|  |
| --- |
| **Definition of a Matrix**  **DEFINITION 1.1.1 (Matrix)*****A rectangular array of numbers is called a matrix.***  We shall mostly be concerned with matrices having real numbers as entries.  The horizontal arrays of a matrix are called its ROWS and the vertical arrays are called its COLUMNS. A matrix having $ m$ rows and $ n$ columns is said to have the order $ m \times n.$  A matrix $ A$ of ORDER $ m \times n$ can be represented in the following form:  $\displaystyle A = \begin{bmatrix}a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & ... ...& \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix},$  where $ a_{ij}$ is the entry at the intersection of the $ i^{\mbox{th}}$ row and $ j^{\mbox{th}}$ column.  In a more concise manner, we also denote the matrix $ A$ by $ [a_{ij}]$ by suppressing its order.  **Remark 1.1.2**   *Some books also use $ \begin{pmatrix}a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots &... ... & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix}$ to represent a matrix.*  Let $ A = \begin{bmatrix}1 & 3 & 7 \\ 4 & 5 & 6 \end{bmatrix}.$ Then $ a_{11} = 1, \; a_{12} = 3, \; a_{13} = 7, \; a_{21} = 4, \; a_{22} = 5, \; $ and $ \; a_{23} = 6.$  A matrix having only one column is called a COLUMN VECTOR; and a matrix with only one row is called a ROW VECTOR.  WHENEVER A VECTOR IS USED, IT SHOULD BE UNDERSTOOD FROM THE CONTEXT WHETHER IT IS A ROW VECTOR OR A COLUMN VECTOR.  **DEFINITION 1.1.3 (Equality of two Matrices)*****Two matrices $ A=[a_{ij}]$ and $ B= [b_{ij}]$ having the same order $ m \times n$ are equal if $ a_{ij} = b_{ij}$ for each $ i =1, 2, \ldots, m$ and $ j =1, 2, \ldots, n.$***  In other words, two matrices are said to be equal if they have the same order and their corresponding entries are equal.  **EXAMPLE 1.1.4   *The linear system of equations $ 2 x + 3 y = 5$ and $ 3 x + 2 y = 5$ can be identified with the matrix $ \begin{bmatrix}2 & 3 &: & 5 \\ 3 & 2 & : & 5\end{bmatrix}.$*** |
| Special Matrices **DEFINITION 1.1.5**   1. A matrix in which each entry is zero is called a zero-matrix, denoted by $ {\mathbf 0}.$ For example,   $\displaystyle {\mathbf 0}_{2 \times 2} = \begin{bmatrix}0 & 0 \\ 0 & 0 \end{bma... ... {\mathbf 0}_{2 \times 3} = \begin{bmatrix}0 & 0 & 0\\ 0 & 0 & 0 \end{bmatrix}.$   1. A matrix for which the number of rows equals the number of columns, is called a square matrix. So, if $ A$ is an $ n \times n$ matrix then $ A$ is said to have order $ n$ . 2. In a square matrix, $ A=[a_{ij}],$ of order $ n$ , the entries $ a_{11}, a_{22}, \ldots, a_{nn}$ are called the diagonal entries and form the principal diagonal of $ A.$ 3. A square matrix $ A=[a_{ij}]$ is said to be a diagonal matrix if $ \; a_{ij} = 0$ for $ i \neq j.$ In other words, the non-zero entries appear only on the principal diagonal. For example, the zero matrix $ {{\mathbf 0}}_n$ and$ \begin{bmatrix}4&0\\ 0&1    \end{bmatrix}$ are a few diagonal matrices.   A diagonal matrix $ D$ of order $ n$ with the diagonal entries $ d_1, d_2, \ldots, d_n$ is denoted by $ D= {\mbox{diag}}(d_1, \ldots, d_n).$  If $ d_i=d$ for all $ i = 1, 2, \ldots, n$ then the diagonal matrix $ D$ is called a **scalar matrix**.   1. A diagonal matrix $ A$ of order $ n$ is called an IDENTITY MATRIX if $ d_i = 1$ for all $ i = 1, 2, \ldots, n$ . This matrix is denoted by $ I_n$ .   For example, $ I_2 = \begin{bmatrix}1&0\\ 0&1 \end{bmatrix}$ and $ I_3 = \begin{bmatrix}1&0&0\\ 0&1&0 \\ 0&0&1 \end{bmatrix}.$  The subscript $ n$ is suppressed in case the order is clear from the context or if no confusion arises.   1. A square matrix $ A=[a_{ij}]$ is said to be an upper triangular matrix if $ a_{ij} = 0$ for $ i > j.$   A square matrix $ A=[a_{ij}]$ is said to be a lower triangular matrix if $ a_{ij} = 0$ for $ i < j.$  A square matrix $ A$ is said to be triangular if it is an upper or a lower triangular matrix.  For example $ \begin{bmatrix}2 & 1 & 4\\ 0&3&-1\\ 0&0&-2 \end{bmatrix}$ is an upper triangular matrix. An upper triangular matrix will be represented by $ \begin{bmatrix}a_{11} & a_{12} & \cdots & a_{1n} \\ 0 & a_{22} & \cdots & a_{... ...\\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & a_{nn} \end{bmatrix}.$ |
| **Operations on Matrices**  **DEFINITION 1.2.1 (Transpose of a Matrix)*****The transpose of an $ m \times n$ matrix $ A=[a_{ij}]$ is defined as the $ n \times m$ matrix $ B = [b_{ij}],$ with $ b_{ij} = a_{ji}$ for $ 1 \leq i \leq m$ and $ 1 \leq j \leq n.$ The transpose of $ A$ is denoted by $ A^t.$***  That is, by the transpose of an $ m \times n$ matrix $ A,$ we mean a matrix of order $ n \times m$ having the rows of $ A$ as its columns and the columns of $ A$ as its rows.  For example, if $ A=\begin{bmatrix}1 & 4 & 5 \\ 0 & 1 & 2 \end{bmatrix}$ then $ A^t =\begin{bmatrix}1 & 0 \\ 4 & 1 \\ 5 & 2 \end{bmatrix}.$  Thus, the transpose of a row vector is a column vector and vice-versa.  **THEOREM 1.2.2   *For any matrix $ A, \;$ $ \; (A^t)^t = A.$***  *Proof*. Let $ A = [a_{ij}], \; A^t = [b_{ij}]$ and $ (A^t)^t = [c_{ij}].$ Then, the definition of transpose gives  $\displaystyle c_{ij} = b_{ji} = a_{ij} \;\; {\mbox{ for all }} \;\; i,j$  and the result follows. height6pt width 6pt depth 0pt  **DEFINITION 1.2.3 (Addition of Matrices)*****let $ A=[a_{ij}]$ and $ B= [b_{ij}]$ be are two $ m \times n$ matrices. Then the sum $ A + B$ is defined to be the matrix $ C = [c_{ij}]$ with $ c_{ij} = a_{ij} + b_{ij}. $***  Note that, we define the sum of two matrices only when the order of the two matrices are same.  **DEFINITION 1.2.4 (Multiplying a Scalar to a Matrix)*****Let $ A=[a_{ij}]$ be an $ m \times n$ matrix. Then for any element $ k \in {\mathbb{R}},$ we define $ k A = [k a_{ij}].$***  For example, if $ A=\begin{bmatrix}1 & 4 & 5 \\ 0 & 1 & 2 \end{bmatrix}$ and $ k = 5,$ then $ 5 A = \begin{bmatrix}5 & 20 & 25 \\ 0 & 5 & 10 \end{bmatrix}.$  **THEOREM 1.2.5   *Let $ A, B$ and $ C$ be matrices of order $ m \times n,$ and let $ k, \ell \in {\mathbb{R}}.$ Then***   1. $ A + B = B + A \; \hspace{1.5in} {\mbox{ (commutativity)}}.$ 2. $ ( A + B ) + C = A + (B + C) \; \hspace{0.635in} {\mbox{    (associativity)}}.$ 3. $ k ( \ell A) = (k \ell) A.$ 4. $ ( k+ \ell) A = k A + \ell A.$   *Proof*. Part [1](http://nptel.iitm.ac.in/courses/Webcourse-contents/IIT-KANPUR/mathematics-2/node6.html#thm:sum:one).  Let $ A=[a_{ij}]$ and $ B= [b_{ij}].$ Then  $\displaystyle A + B = [a_{ij}] + [b_{ij}] = [ a_{ij} + b_{ij} ] = [ b_{ij} + a_{ij}] = [b_{ij}] + [a_{ij}]= B + A$  as real numbers commute.  The reader is required to prove the other parts as all the results follow from the properties of real numbers. height6pt width 6pt depth 0pt  **EXERCISE 1.2.6**   1. Suppose $ A + B = A.$ Then show that $ B = {\mathbf 0}.$ 2. Suppose $ A + B = {\mathbf 0}.$ Then show that $ B = (-1) A = [- a_{ij}].$   **DEFINITION 1.2.7 (Additive Inverse)   *Let $ A$ be an $ m \times n$ matrix.***   1. Then there exists a matrix $ B$ with $ A + B = {\mathbf 0}.$ This matrix $ B$ is called the additive inverse of $ A,$ and is denoted by $ - A = (-1) A. $ 2. Also, for the matrix $ {\mathbf 0}_{m \times n},$ $ A + {\mathbf 0}= {\mathbf 0}+ A = A.$ Hence, the matrix $ {\mathbf 0}_{m \times n}$ is called the additive identity. |
| Multiplication of Matrices **DEFINITION 1.2.8 (Matrix Multiplication / Product)*****Let $ A=[a_{ij}]$ be an $ m \times n$ matrix and $ B= [b_{ij}]$ be an $ n \times r$ matrix. The product $ A B$ is a matrix $ C = [c_{ij}]$ of order $ m \times r,$ with***  $\displaystyle c_{ij} = \sum\limits_{k=1}^n a_{ik} b_{kj} = a_{i1} b_{1j} + a_{i2} b_{2j} + \cdots + a_{in} b_{nj}.$  That is, if $ A_{m \times n} = \left[\begin{array}{cccc} & & \cdots & \\ & & \cdots & \\ a_... ...a_{i2} & \cdots & a_{in} \\ & & \cdots & \\ & & \cdots & \\ \end{array}\right]$ and $ B_{n \times r} = \left[\begin{array}{ccc} \cdots & b_{1j} & \cdots \\ \cdots ... ... \\ \vdots & \vdots & \vdots \\ \cdots & b_{mj} & \cdots \\ \end{array}\right]$ then  $\displaystyle A B = [(AB)_{ij}]_{m \times r} {\mbox{ and }} (AB)_{ij} = a_{i1} b_{1j} + a_{i2} b_{2j} + \cdots + a_{in} b_{nj}.$  Observe that the product $ A B$ is defined if and only if  THE NUMBER OF COLUMNS OF MATHEND000# THE NUMBER OF ROWS OF MATHEND000#  For example, if $ A= \begin{bmatrix}1 & 2 & 3 \\ 2 & 4 & 1 \end{bmatrix}$ and $ B = \begin{bmatrix}1 & 2 & 1\\ 0 & 0 & 3 \\ 1 & 0 & 4 \end{bmatrix}$ then   |  |  |  |  | | --- | --- | --- | --- | | $\displaystyle A B$ | $\displaystyle =$ | $\displaystyle \begin{bmatrix}1 +0+ 3 & 2+0+0 & 1 + 6 + 12 \\ 2 + 0+1 & 4+0+0 & 2 + 12 + 4 \end{bmatrix} = \begin{bmatrix}4 & 2 & 19 \\ 3 & 4 & 18 \end{bmatrix}$ |  | |  | $\displaystyle =$ | $\displaystyle \begin{bmatrix}1 \cdot (1 \ 2 \ 1) + 2 \cdot (0 \ 0 \ 3 ) + 3 \cd... ... 2 \cdot (1 \ 2 \ 1) + 4 \cdot (0 \ 0 \ 3 ) + 1 \cdot (1 \ 0 \ 4) \end{bmatrix}$ | (1.2.1) | |  | $\displaystyle =$ | $\displaystyle \left[\begin{array}{c} \left(\begin{array}{c} 1 \\ 2 \end{array}\... ... 0 + \left(\begin{array}{c} 3 \\ 1 \end{array}\right) \cdot 1\end{array}\right.$ |  | |  |  | $\displaystyle \hspace{.75in} \left(\begin{array}{c} 1 \\ 2 \end{array}\right) \... ...rray}\right) \cdot 0 + \left(\begin{array}{c} 3 \\ 1 \end{array}\right) \cdot 0$ |  | |  |  | $\displaystyle \hspace{1.5in} \left. \begin{array}{c} \left(\begin{array}{c} 1 \... ... + \left(\begin{array}{c} 3 \\ 1 \end{array}\right) \cdot 4 \end{array}\right].$ | (1.2.2) |   Observe the following:   1. In this example, while $ A B$ is defined, the product $ B A$ is not defined.   However, for square matrices $ A$ and $ B$ of the same order, both the product $ A B$ and $ B A$ are defined.   1. The product $ A B$ corresponds to operating on the rows of the matrix*$ B$*(see [1.2.1](http://nptel.iitm.ac.in/courses/Webcourse-contents/IIT-KANPUR/mathematics-2/node7.html#mul:1)), and 2. The product $ A B$ also corresponds to operating on the columns of the matrix*$ A$*(see [1.2.2](http://nptel.iitm.ac.in/courses/Webcourse-contents/IIT-KANPUR/mathematics-2/node7.html#mul:2)).   **DEFINITION 1.2.9   *Two square matrices $ A$ and $ B$ are said to commute if $ A B = B A.$***  **Remark 1.2.10**   1. Note that if $ A$ is a square matrix of order $ n$ then $ A    I_n = I_n A.$ Also, a scalar matrix of order $ n$ commutes with any square matrix of order $ n$ . 2. In general, the matrix product is not commutative. For example, consider the following two matrices $ A = \begin{bmatrix}1 & 1 \\ 0 & 0 \end{bmatrix}$ and $ B = \begin{bmatrix}1 & 0 \\ 1 & 0 \end{bmatrix}$ . Then check that the matrix product   $\displaystyle A B = \begin{bmatrix}2 & 0 \\ 0 & 0 \end{bmatrix} \neq \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix} = B A.$  **THEOREM 1.2.11   *Suppose that the matrices $ A, \;B$ and $ C$ are so chosen that the matrix multiplications are defined.***   1. Then $ (A B) C = A (B C).$ That is, the matrix multiplication is associative. 2. For any $ k \in {\mathbb{R}}, \; (k A) B = k ( A B) = A ( k B).$ 3. Then $ A(B + C) = A B + A C.$ That is, multiplication distributes over addition. 4. If $ A$ is an $ n \times n$ matrix then $ A I_n = I_n A = A.$ 5. For any square matrix $ A$ of order $ n$ and $ D={\mbox{diag}}(d_1, d_2, \ldots, d_n),$ we have    * the first row of $ D A$ is $ d_1$ times the first row of $ A;$    * for $ 1 \leq i \leq n,$ the $ i^{\mbox{th}}$ row of $ D A$ is $ d_i$ times the $ i^{\mbox{th}}$ row of $ A.$   A similar statement holds for the columns of $ A$ when $ A$ is multiplied on the right by $ D.$  *Proof*. Part [1](http://nptel.iitm.ac.in/courses/Webcourse-contents/IIT-KANPUR/mathematics-2/node7.html#thm:product:one). $ \;\;$ Let $ A= [a_{ij}]_{m \times n}, \; B= [b_{ij}]_{n \times p}$ and $ C = [c_{ij}]_{p \times q}.$ Then  $\displaystyle (BC)_{kj} = \sum_{\ell=1}^p b_{k \ell} c_{\ell j} \;\; {\mbox{ and }} \;\; (AB)_{i\ell} = \sum_{k=1}^n a_{i k} b_{k \ell}.$  Therefore,   |  |  |  |  | | --- | --- | --- | --- | | $\displaystyle \bigl( A(BC) \bigr)_{ij}$ | $\displaystyle =$ | $\displaystyle \sum_{k=1}^n a_{ik} \bigl(BC\bigr)_{kj} = \sum_{k=1}^n a_{ik} \bigl( \sum_{\ell=1}^p b_{k\ell} c_{\ell j}\bigr)$ |  | |  | $\displaystyle =$ | $\displaystyle \sum_{k=1}^n \sum_{\ell=1}^p a_{ik} \bigl( b_{k\ell}c_{\ell j}\bigr) = \sum_{k=1}^n \sum_{\ell=1}^p \bigl( a_{ik} b_{k\ell} \bigr) c_{\ell j}$ |  | |  | $\displaystyle =$ | $\displaystyle \sum_{\ell=1}^p \bigl(\sum_{k=1}^n a_{ik} b_{k\ell} \bigr) c_{\ell j} = \sum_{\ell=1}^t \bigl(AB \bigr)_{i\ell}c_{\ell j}$ |  | |  | $\displaystyle =$ | $\displaystyle \bigl( (AB)C \bigr)_{ij}.$ |  |     Part [5](http://nptel.iitm.ac.in/courses/Webcourse-contents/IIT-KANPUR/mathematics-2/node7.html#thm:product:five). $ \;\;$ For all $ j = 1, 2, \ldots, n,$ we have  $\displaystyle (D A)_{ij} = \sum_{k=1}^n d_{ik} a_{kj} = d_i a_{ij}$  as $ d_{ik} = 0$ whenever $ i \neq k.$ Hence, the required result follows. |
| Inverse of a Matrix **DEFINITION 1.2.13 (Inverse of a Matrix)*****Let $ A$ be a square matrix of order $ n.$***   1. A square matrix $ B$ is said to be a LEFT INVERSE of $ A$ if $ B A = I_n.$ 2. A square matrix $ C$ is called a RIGHT INVERSE of $ A,$ if $ A C = I_n.$ 3. A matrix $ A$ is said to be INVERTIBLE (or is said to have an INVERSE) if there exists a matrix $ B$ such that $ A B = B A = I_n.$   **LEMMA 1.2.14   *Let $ A$ be an $ n \times n$ matrix. Suppose that there exist $ n \times n$ matrices $ B$ and $ C$ such that $ A B = I_n$ and $ C A = I_n,$ then $ B = C.$***  *Proof*. Note that  $\displaystyle C = C I_n = C( A B) = (C A) B = I_n B = B.$  height6pt width 6pt depth 0pt  **Remark 1.2.15**   1. From the above lemma, we observe that if a matrix $ A$ is invertible, then the inverse is unique. 2. As the inverse of a matrix $ A$ is unique, we denote it by $ A^{-1}.$ That is, $ A A^{-1} = A^{-1} A = I.$   **THEOREM 1.2.16   *Let $ A$ and $ B$ be two matrices with inverses $ A^{-1}$ and $ B^{-1},$ respectively. Then***   1. $ (A^{-1})^{-1}= A.$ 2. $ ( A B )^{-1} = B^{-1} A^{-1}.$ 3. $ (A^t)^{-1} =    (A^{-1})^t.$   *Proof*. Proof of Part [1](http://nptel.iitm.ac.in/courses/Webcourse-contents/IIT-KANPUR/mathematics-2/node8.html#inverse:inverse).  By definition $ A A^{-1} = A^{-1} A = I.$ Hence, if we denote $ A^{-1}$ by $ B,$ then we get $ A B = B A = I.$ Thus, the definition, implies $ B^{-1} = A,$ or equivalently $ (A^{-1})^{-1}= A.$  Proof of Part [2](http://nptel.iitm.ac.in/courses/Webcourse-contents/IIT-KANPUR/mathematics-2/node8.html#inverse:product).  Verify that $ (A B) (B^{-1} A^{-1}) = I = (B^{-1} A^{-1}) (A B).$  Proof of Part [3](http://nptel.iitm.ac.in/courses/Webcourse-contents/IIT-KANPUR/mathematics-2/node8.html#inverse:transpose).  We know $ A A^{-1} = A^{-1} A = I.$ Taking transpose, we get  $\displaystyle (A A^{-1})^t = (A^{-1} A)^t = I^t \Longleftrightarrow (A^{-1})^t A^t = A^t (A^{-1})^t = I.$  Hence, by definition $ (A^t)^{-1} = (A^{-1})^t.$ |
| **Some More Special Matrices**  **DEFINITION 1.3.1**   1. A matrix $ A$ over $ {\mathbb{R}}$ is called symmetric if $ A^{t} = A$ and skew-symmetric if $ A^{t} = -A.$ 2. A matrix $ A$ is said to be orthogonal if $ A A^t = A^t A = I.$   **EXAMPLE 1.3.2**   1. Let $ A = \begin{bmatrix}1 & 2 & 3 \\ 2 & 4 & -1 \\    3 & -1 & 4 \end{bmatrix}$ and $ B = \begin{bmatrix}0 & 1 & 2 \\ -1 & 0    & -3 \\ -2 & 3 & 0 \end{bmatrix}.$ Then $ A$ is a symmetric matrix and $ B$ is a skew-symmetric matrix. 2. Let $ A = \begin{bmatrix}\frac{1}{\sqrt{3}} &\frac{1}{\sqrt{3}}    & \frac{1}{\sqrt{3}}...    ...\frac{1}{\sqrt{6}} & \frac{1}{\sqrt{6}} & - \frac{2}{ \sqrt{6} } \end{bmatrix}.$ Then $ A$ is an orthogonal matrix. 3. Let $ A=[a_{ij}]$ be an $ n \times n$ matrix with $ a_{ij} = \begin{cases}1 & {\mbox{ if }} i= j+1 \\ 0 &{\mbox{ otherwise }}    \end{cases}.$ Then $ A^n = {\mathbf 0}$ and $ A^{\ell} \neq {\mathbf 0}$ for $ 1 \leq \ell \leq n-1.$ The matrices $ A$ for which a positive integer $ k$ exists such that $ A^k = {\mathbf 0}$ are called NILPOTENT matrices. The least positive integer $ k$ for which $ A^k = {\mathbf 0}$ is called the ORDER OF NILPOTENCY. 4. Let $ A = \begin{bmatrix}1 & 0 \\ 0 & 0 \end{bmatrix}.$ Then $ A^2 = A.$ The matrices that satisfy the condition that $ A^2 = A$ are called IDEMPOTENT matrices.   For any square matrix $ A,$ $ S = \frac{1}{2} (A+A^{t})$ is symmetric, $ T = \frac{1}{2} (A - A^{t})$ is skew-symmetric, and $ A = S + T.$ |
| Submatrix of a Matrix **DEFINITION 1.3.4*****A matrix obtained by deleting some of the rows and/or columns of a matrix is said to be a submatrix of the given matrix.***  For example, if $ A=\begin{bmatrix}1 & 4 & 5 \\ 0 & 1 & 2 \end{bmatrix},$ a few submatrices of $ A$ are  $\displaystyle [ 1 ], [ 2 ], \begin{bmatrix}1 \\ 0 \end{bmatrix}, [ 1 \; 5 ], \begin{bmatrix}1 & 5 \\ 0 & 2 \end{bmatrix}, \; A.$  But the matrices $ \begin{bmatrix}1& 4\\ 1& 0 \end{bmatrix}$ and $ \begin{bmatrix}1& 4\\ 0& 2 \end{bmatrix}$ are not submatrices of $ A.$ |
| **Matrices over Complex Numbers**  Here the entries of the matrix are complex numbers. All the definitions still hold. One just needs to look at the following additional definitions.  **DEFINITION 1.4.1 (Conjugate Transpose of a Matrix)**   1. Let $ A$ be an $ m \times n$ matrix over $ {\mathbb{C}}.$ If $ A=[a_{ij}]$ then the Conjugate of $ A,$ denoted by $ \overline{A},$ is the matrix $ B= [b_{ij}]$ with $ b_{ij} = \overline{a_{ij}}.$   For example, Let $ A=\begin{bmatrix}1 & 4 + 3 i & i \\ 0 & 1 & i - 2 \end{bmatrix}.$ Then  $\displaystyle \overline{A} =\begin{bmatrix}1 & 4 - 3 i & - i \\ 0 & 1 & -i - 2 \end{bmatrix}.$   1. Let $ A$ be an $ m \times n$ matrix over $ {\mathbb{C}}.$ If $ A=[a_{ij}]$ then the Conjugate Transpose of $ A,$ denoted by $ A^*,$ is the matrix $ B= [b_{ij}]$ with $ b_{ij} = \overline{a_{ji}}.$   For example, Let $ A=\begin{bmatrix}1 & 4 + 3 i & i \\ 0 & 1 & i - 2 \end{bmatrix}.$ Then  $\displaystyle A^* =\begin{bmatrix}1 & 0 \\ 4 - 3 i & 1 \\ - i & -i - 2 \end{bmatrix}.$   1. A square matrix $ A$ over $ {\mathbb{C}}$ is called Hermitian if $ A^*    = A.$ 2. A square matrix $ A$ over $ {\mathbb{C}}$ is called skew-Hermitian if $ A^{*} = -A.$ 3. A square matrix $ A$ over $ {\mathbb{C}}$ is called unitary if $ A^{*}A = A A^{*} = I.$ 4. A square matrix $ A$ over $ {\mathbb{C}}$ is called Normal if $ A A^{*} = A^{*} A.$   **Remark 1.4.2**   *If $ A=[a_{ij}]$ with $ a_{ij} \in {\mathbb{R}},$ then $ A^* = A^t.$*   1. For any square matrix $ A,$ $ S =    \frac{A+A^*}{2} $ is Hermitian, $ T = \frac{A- A^*}{2}$ is skew-Hermitian, and $ A = S + T.$ 2. If $ A$ is a complex triangular matrix and $ A A^* =    A^* A$ then $ A$ is a diagonal matrix. |
| **Introduction**  Let us look at some examples of linear systems.   1. Suppose $ a, b \in {\mathbb{R}}.$ Consider the system $ a x    = b.$    1. If $ a \neq 0$ then the system has a UNIQUE SOLUTION $ x = \frac{b}{a}.$    2. If $ a = 0$ and       1. $ b \neq 0$ then the system has NO SOLUTION.       2. $ b = 0$ then the system has INFINITE NUMBER OF SOLUTIONS, namely all $ x \in {\mathbb{R}}.$ 2. We now consider a system with $ 2$ equations in $ 2$ unknowns.  Consider the equation $ a x + b y = c.$ If one of the coefficients, $ a$ or $ b$ is non-zero, then this linear equation represents a line in $ {\mathbb{R}}^2.$ Thus for the system   $\displaystyle a_1 x + b_1 y = c_1 \; {\mbox{ and }} \; a_2 x + b_2 y = c_2, $  the set of solutions is given by the points of intersection of the two lines. There are three cases to be considered. Each case is illustrated by an example.   * 1. UNIQUE SOLUTION  $ x+ 2 y = 1 $ and $ x + 3 y = 1.$ The unique solution is $ (x, y)^t =      (1, 0)^t.$  Observe that in this case, $ a_1 b_2 - a_2 b_1 \neq 0.$   2. INFINITE NUMBER OF SOLUTIONS  $ x+ 2 y = 1 $ and $ 2x + 4 y = 2.$ The set of solutions is $ (x, y)^t = (1 - 2y, y)^t= (1, 0)^t + y      (-2, 1)^t$ with $ y$ arbitrary. In other words, both the equations represent the same line.  Observe that in this case, $ a_1 b_2 - a_2 b_1 = 0,\; a_1 c_2 - a_2 c_1 = 0$ and $ b_1 c_2 - b_2 c_1 = 0.$   3. NO SOLUTION  $ x+ 2 y = 1 $ and $ 2x + 4 y = 3.$ The equations represent a pair of parallel lines and hence there is no point of intersection.  Observe that in this case, $ a_1 b_2 - a_2 b_1 = 0$ but $ a_1 c_2 - a_2 c_1 \neq 0.$  1. As a last example, consider $ 3$ equations in $ 3$ unknowns.  A linear equation $ a x + b y + c z = d$ represent a plane in $ {\mathbb{R}}^3$ provided $ (a, b, c) \neq (0, 0, 0).$ As in the case of $ 2$ equations in $ 2$ unknowns, we have to look at the points of intersection of the given three planes. Here again, we have three cases. The three cases are illustrated by examples.    1. UNIQUE SOLUTION  Consider the system $ x + y + z = 3, \;\; x + 4 y + 2 z =       7$ and $ 4 x + 10 y - z = 13.$ The unique solution to this system is $ (x, y, z)^t = (1, 1, 1)^t;$ *i.e.* THE THREE PLANES INTERSECT AT A POINT.    2. INFINITE NUMBER OF SOLUTIONS  Consider the system $ x + y + z = 3, \;\; x + 2 y + 2 z = 5$ and $ 3 x + 4 y + 4 z = 11.$ The set of solutions to this system is $ (x, y, z)^t = (1, 2-z, z)^t = (1, 2, 0)^t + z (0, -1, 1)^t,$ with $ z$arbitrary: THE THREE PLANES INTERSECT ON A LINE.    3. NO SOLUTION  The system $ x + y + z = 3, \;\; x + 2 y + 2 z = 5$ and $ 3 x + 4 y + 4 z = 13$ has no solution. In this case, we get three parallel lines as intersections of the above planes taken two at a time.   The readers are advised to supply the proof.  **DEFINITION 2.1.1 (Linear System)*****A linear system of $ m$ equations in $ n$ unknowns $ x_1, x_2, \ldots, x_n$ is a set of equations of the form***   |  |  |  |  | | --- | --- | --- | --- | | $\displaystyle a_{11} x_1 + a_{12}x_2 + \cdots + a_{1n}x_n$ | $\displaystyle =$ | $\displaystyle b_1$ |  | | $\displaystyle a_{21} x_1 + a_{22}x_2 + \cdots + a_{2n}x_n$ | $\displaystyle =$ | $\displaystyle b_2$ |  | |  |  |  | (2.1.1) | | $\displaystyle a_{m1} x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n$ | $\displaystyle =$ | $\displaystyle b_m$ |  |   ***where for $ 1 \leq i \leq n,$ and $ 1 \leq j \leq m; \; a_{ij}, b_i \in {\mathbb{R}}.$ Linear System (***[***2.1.1***](http://nptel.iitm.ac.in/courses/Webcourse-contents/IIT-KANPUR/mathematics-2/node15.html#eqn:sys)***) is called HOMOGENEOUS******if $ b_1 = 0 = b_2= \cdots = b_m$ and NON-HOMOGENEOUS******otherwise.***  We rewrite the above equations in the form $ A {\mathbf x}= {\mathbf b},$ where  $ A = \begin{bmatrix}a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdo... ...}, \; \; {\mathbf x}=\begin{bmatrix}x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}, $ and $ {\mathbf b}= \begin{bmatrix}b_1\\ b_2 \\ \vdots \\ b_m \end{bmatrix}$  The matrix $ A$ is called the COEFFICIENT matrix and the block matrix $ \left[ A \; \; {\mathbf b}\right],$ is the AUGMENTED matrix of the linear system ([2.1.1](http://nptel.iitm.ac.in/courses/Webcourse-contents/IIT-KANPUR/mathematics-2/node15.html#eqn:sys)).  **Remark 2.1.2**   *Observe that the $ i^{\mbox{th}}$ row of the augmented matrix $ [A \;\; {\mathbf b}]$ represents the $ i^{\mbox{th}}$ equation and the $ j^{\mbox{th}}$ column of the coefficient matrix $ A$ corresponds to coefficients of the $ j^{\mbox{th}}$variable $ x_j.$ That is, for $ 1 \leq i \leq m$ and $ 1 \leq j \leq n,$ the entry $ a_{ij}$ of the coefficient matrix $ A$ corresponds to the $ i^{\mbox{th}}$ equation and $ j^{\mbox{th}}$ variable $ x_j..$*  For a system of linear equations $ A {\mathbf x}= {\mathbf b},$ the system $ A {\mathbf x}= {\mathbf 0}$ is called the ASSOCIATED HOMOGENEOUS SYSTEM.  **DEFINITION 2.1.3 (Solution of a Linear System)*****A solution of the linear system $ A {\mathbf x}= {\mathbf b}$ is a column vector $ {\mathbf y}$ with entries $ y_1, y_2, \ldots, y_n$ such that the linear system (***[***2.1.1***](http://nptel.iitm.ac.in/courses/Webcourse-contents/IIT-KANPUR/mathematics-2/node15.html#eqn:sys)***) is satisfied by substituting $ y_i$ in place of $ x_i.$***  That is, if $ {\mathbf y}^t = [ y_1, y_2, \ldots, y_n ]$ then $ A {\mathbf y}= {\mathbf b}$ holds.  **Note:** The zero $ n$ -tuple $ {\mathbf x}={\mathbf 0}$ is always a solution of the system $ A {\mathbf x}= {\mathbf 0},$ and is called the TRIVIAL solution. A non-zero $ n$ -tuple $ {\mathbf x},$ if it satisfies $ A {\mathbf x}= {\mathbf 0},$ is called a NON-TRIVIAL solution. |
| **Row Operations and Equivalent Systems**  **DEFINITION 2.2.1 (Elementary Operations)*****The following operations***[***1***](http://nptel.iitm.ac.in/courses/Webcourse-contents/IIT-KANPUR/mathematics-2/node17.html#operation:one)***,***[***2***](http://nptel.iitm.ac.in/courses/Webcourse-contents/IIT-KANPUR/mathematics-2/node17.html#operation:two)***and***[***3***](http://nptel.iitm.ac.in/courses/Webcourse-contents/IIT-KANPUR/mathematics-2/node17.html#operation:three)***are called elementary operations.***   1. interchange of two equations, say ``interchange the $ i^{\mbox{th}}$ and $ j^{\mbox{th}}$ equations";   (compare the system ([2.1.2](http://nptel.iitm.ac.in/courses/Webcourse-contents/IIT-KANPUR/mathematics-2/node16.html#step1)) with the original system.)   1. multiply a non-zero constant throughout an equation, say ``multiply the $ k^{\mbox{th}}$ equation by $ c \neq 0$ ";   (compare the system ([2.1.5](http://nptel.iitm.ac.in/courses/Webcourse-contents/IIT-KANPUR/mathematics-2/node16.html#step4)) and the system ([2.1.4](http://nptel.iitm.ac.in/courses/Webcourse-contents/IIT-KANPUR/mathematics-2/node16.html#step3)).)   1. replace an equation by itself plus a constant multiple of another equation, say ``replace the $ k^{\mbox{th}}$ equation by $ k^{\mbox{th}}$ equation plus $ c$ times the $ j^{\mbox{th}}$ equation".   (compare the system ([2.1.3](http://nptel.iitm.ac.in/courses/Webcourse-contents/IIT-KANPUR/mathematics-2/node16.html#step2)) with ([2.1.2](http://nptel.iitm.ac.in/courses/Webcourse-contents/IIT-KANPUR/mathematics-2/node16.html#step1)) or the system ([2.1.4](http://nptel.iitm.ac.in/courses/Webcourse-contents/IIT-KANPUR/mathematics-2/node16.html#step3)) with ([2.1.3](http://nptel.iitm.ac.in/courses/Webcourse-contents/IIT-KANPUR/mathematics-2/node16.html#step2)).)  **Remark 2.2.2**   1. In Example [2.1.4](http://nptel.iitm.ac.in/courses/Webcourse-contents/IIT-KANPUR/mathematics-2/node16.html#exa:gauss), observe that the elementary operations helped us in getting a linear system ([2.1.5](http://nptel.iitm.ac.in/courses/Webcourse-contents/IIT-KANPUR/mathematics-2/node16.html#step4)), which was easily solvable. 2. Note that at Step [1](http://nptel.iitm.ac.in/courses/Webcourse-contents/IIT-KANPUR/mathematics-2/node16.html#exa:step1), if we interchange the first and the second equation, we get back to the linear system from which we had started. This means the operation at Step [1](http://nptel.iitm.ac.in/courses/Webcourse-contents/IIT-KANPUR/mathematics-2/node16.html#exa:step1), has an inverse operation. In other words, INVERSE OPERATION sends us back to the step where we had precisely started.   So, in Example [2.1.4](http://nptel.iitm.ac.in/courses/Webcourse-contents/IIT-KANPUR/mathematics-2/node16.html#exa:gauss), the application of a finite number of elementary operations helped us to obtain a simpler system whose solution can be obtained directly. That is, after applying a finite number of elementary operations, a simpler linear system is obtained which can be easily solved. Note that the three elementary operations defined above, have corresponding INVERSE operations, namely,   1. ``interchange the $ i^{\mbox{th}}$ and $ j^{\mbox{th}}$ equations", 2. ``divide the $ k^{\mbox{th}}$ equation by $ c \neq 0$ "; 3. ``replace the $ k^{\mbox{th}}$ equation by $ k^{\mbox{th}}$ equation minus $ c$ times the $ j^{\mbox{th}}$ equation".   It will be a useful exercise for the reader to IDENTIFY THE INVERSE OPERATIONS at each step in Example [2.1.4](http://nptel.iitm.ac.in/courses/Webcourse-contents/IIT-KANPUR/mathematics-2/node16.html#exa:gauss).  **DEFINITION 2.2.3 (Equivalent Linear Systems)*****Two linear systems are said to be equivalent if one can be obtained from the other by a finite number of elementary operations.***  The linear systems at each step in Example [2.1.4](http://nptel.iitm.ac.in/courses/Webcourse-contents/IIT-KANPUR/mathematics-2/node16.html#exa:gauss) are equivalent to each other and also to the original linear system.  **LEMMA 2.2.4   *Let $ C {\mathbf x}= {\mathbf d}$ be the linear system obtained from the linear system $ A {\mathbf x}= {\mathbf b}$ by a single elementary operation. Then the linear systems $ A {\mathbf x}= {\mathbf b}$ and $ C {\mathbf x}= {\mathbf d}$ have the same set of solutions.***  *Proof*. We prove the result for the elementary operation ``the $ k^{\mbox{th}}$ equation is replaced by $ k^{\mbox{th}}$ equation plus $ c$ times the $ j^{\mbox{th}}$ equation." The reader is advised to prove the result for other elementary operations.  In this case, the systems $ A {\mathbf x}= {\mathbf b}$ and $ C {\mathbf x}= {\mathbf d}$ vary only in the $ k^{\mbox{th}}$ equation. Let $ ({\alpha}_1, {\alpha}_2, \ldots, {\alpha}_n)$ be a solution of the linear system $ A {\mathbf x}= b.$ Then substituting for $ {\alpha}_i$ 's in place of $ x_i$ 's in the $ k^{\mbox{th}}$ and$ j^{\mbox{th}}$ equations, we get  $\displaystyle a_{k1} {\alpha}_1 + a_{k2} {\alpha}_2 + \cdots a_{kn} {\alpha}_n ... ...d }} \; a_{j1} {\alpha}_1 + a_{j2} {\alpha}_2 + \cdots a_{jn} {\alpha}_n = b_j.$  Therefore,   |  |  | | --- | --- | | $\displaystyle (a_{k1} + c a_{j1}) {\alpha}_1 + (a_{k2} + c a_{j2}) {\alpha}_2 + \cdots + (a_{kn} + c a_{jn}) {\alpha}_n = b_k + c b_j.$ | (2.2.1) |     But then the $ k^{\mbox{th}}$ equation of the linear system $ C {\mathbf x}= {\mathbf d}$ is   |  |  | | --- | --- | | $\displaystyle (a_{k1} + c a_{j1}) x_1 + (a_{k2} + c a_{j2}) x_2 + \cdots + (a_{kn} + c a_{jn}) x_n = b_k + c b_j.$ | (2.2.2) |     Therefore, using Equation ([2.2.1](http://nptel.iitm.ac.in/courses/Webcourse-contents/IIT-KANPUR/mathematics-2/node17.html#ajk)), $ ({\alpha}_1, {\alpha}_2, \ldots, {\alpha}_n)$ is also a solution for the $ k^{\mbox{th}}$ Equation ([2.2.2](http://nptel.iitm.ac.in/courses/Webcourse-contents/IIT-KANPUR/mathematics-2/node17.html#cjk)).  Use a similar argument to show that if $ (\beta_1, \beta_2, \ldots, \beta_n)$ is a solution of the linear system $ C {\mathbf x}= {\mathbf d}$ then it is also a solution of the linear system $ A {\mathbf x}= {\mathbf b}.$  Hence, we have the proof in this case. height6pt width 6pt depth 0pt  Lemma [2.2.4](http://nptel.iitm.ac.in/courses/Webcourse-contents/IIT-KANPUR/mathematics-2/node17.html#lem:equivalent) is now used as an induction step to prove the main result of this section (Theorem [2.2.5](http://nptel.iitm.ac.in/courses/Webcourse-contents/IIT-KANPUR/mathematics-2/node17.html#thm:equivalent)).  **THEOREM 2.2.5   *Two equivalent systems have the same set of solutions.***  *Proof*. Let $ n$ be the number of elementary operations performed on $ A {\mathbf x}= {\mathbf b}$ to get $ C {\mathbf x}= {\mathbf d}.$ We prove the theorem by induction on $ n.$  If $ n = 1,$ Lemma [2.2.4](http://nptel.iitm.ac.in/courses/Webcourse-contents/IIT-KANPUR/mathematics-2/node17.html#lem:equivalent) answers the question. If $ n > 1,$ assume that the theorem is true for $ n = m.$ Now, suppose $ n = m+1.$ Apply the Lemma [2.2.4](http://nptel.iitm.ac.in/courses/Webcourse-contents/IIT-KANPUR/mathematics-2/node17.html#lem:equivalent) again at the ``last step" (that is, at the $ (m+1)^{\mbox{th}}$step from the $ m^{\mbox{th}}$ step) to get the required result using induction. height6pt width 6pt depth 0pt  Let us formalise the above section which led to Theorem [2.2.5](http://nptel.iitm.ac.in/courses/Webcourse-contents/IIT-KANPUR/mathematics-2/node17.html#thm:equivalent). For solving a linear system of equations, we applied elementary operations to equations. It is observed that in performing the elementary operations, the calculations were made on the COEFFICIENTS (numbers). The variables $ x_1, x_2, \ldots, x_n$ and the sign of equality (that is, $ \lq\lq =''$ ) are not disturbed. Therefore, in place of looking at the system of equations as a whole, we just need to work with the coefficients. These coefficients when arranged in a rectangular array gives us the augmented matrix $ [ A \; \; {\mathbf b}].$  **DEFINITION 2.2.6 (Elementary Row Operations)*****The elementary row operations are defined as:***   1. interchange of two rows, say ``interchange the $ i^{\mbox{th}}$ and $ j^{\mbox{th}}$ rows", denoted $ R_{ij};$ 2. multiply a non-zero constant throughout a row, say ``multiply the $ k^{\mbox{th}}$ row by $ c \neq 0$ ", denoted $ R_k(c);$ 3. replace a row by itself plus a constant multiple of another row, say ``replace the $ k^{\mbox{th}}$ row by $ k^{\mbox{th}}$ row plus $ c$ times the $ j^{\mbox{th}}$ row", denoted $ R_{kj}(c).$   **EXERCISE 2.2.7   *Find the INVERSE row operations corresponding to the elementary row operations that have been defined just above.***  **DEFINITION 2.2.8 (Row Equivalent Matrices)*****Two matrices are said to be row-equivalent if one can be obtained from the other by a finite number of elementary row operations.***  **EXAMPLE 2.2.9   *The three matrices given below are row equivalent.  $ \begin{bmatrix}0 & 1 & 1 & 2 \\ 2 & 0 & 3 & 5 \\ 1 & 1 & 1 & 3 \end{bmatrix} \... ... 0 & \frac{3}{2} & \frac{5}{2} \\ 0 & 1 & 1 & 2 \\ 1 & 1 & 1 & 3 \end{bmatrix}.$***  ***Whereas the matrix $ \begin{bmatrix}0 & 1 & 1 & 2 \\ 2 & 0 & 3 & 5 \\ 1 & 1 & 1 & 3 \end{bmatrix}$ is not row equivalent to the matrix $ \begin{bmatrix}1 & 0 & 1 & 2 \\ 0 & 2 & 3 & 5 \\ 1 & 1 & 1 & 3 \end{bmatrix}.$*** |
| Gauss Elimination Method **DEFINITION 2.2.10 (Forward/Gauss Elimination Method)*****Gaussian elimination is a method of solving a linear system $ A {\mathbf x}= {\mathbf b}$ (consisting of $ m$ equations in $ n$ unknowns) by bringing the augmented matrix***  $\displaystyle [A \;\; {\mathbf b}] = \left[\begin{array}{cccc\vert c} a_{11} & ... ... \vdots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} & b_m \end{array}\right]$  ***to an upper triangular form***  $\displaystyle \left[\begin{array}{cccc\vert c} c_{11} & c_{12} & \cdots & c_{1n... ... \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & c_{mn} & d_m \end{array}\right].$  ***This elimination process is also called the forward elimination method.***  The following examples illustrate the Gauss elimination procedure.  **EXAMPLE 2.2.11   *Solve the linear system by Gauss elimination method.***   |  |  |  |  | | --- | --- | --- | --- | | $\displaystyle y + z$ | $\displaystyle =$ | $\displaystyle 2$ |  | | $\displaystyle 2 x + 3 z$ | $\displaystyle =$ | $\displaystyle 5$ |  | | $\displaystyle x+ y + z$ | $\displaystyle =$ | $\displaystyle 3$ |  |     **Solution:** In this case, the augmented matrix is $ \begin{bmatrix}0 & 1 & 1 & 2 \\ 2 & 0 & 3 & 5 \\ 1 & 1 & 1 & 3 \end{bmatrix}.$ The method proceeds along the following steps.   1. Interchange $ 1^{\mbox{st}}$ and $ 2^{\mbox{nd}}$ equation (or $ R_{12}$ ).   $\displaystyle \begin{array}{cr} 2 x + 3 z &= 5 \\ y + z &= 2 \\ x + y + z &= 3 ... ...} \begin{bmatrix}2 & 0 & 3 & 5 \\ 0 & 1 & 1 & 2 \\ 1 & 1 & 1 & 3 \end{bmatrix}.$   1. Divide the $ 1^{\mbox{st}}$ equation by $ 2$ (or $ R_1(1/2)$ ).   $\displaystyle \begin{array}{cr} x + \frac{3}{2} z &= \frac{5}{2} \\ y + z &= 2... ...0 & \frac{3}{2} & \frac{5}{2} \\ 0 & 1 & 1 & 2 \\ 1 & 1 & 1 & 3 \end{bmatrix}. $   1. Add $ -1$ times the $ 1^{\mbox{st}}$ equation to the $ 3^{\mbox{rd}}$ equation (or $ R_{31}(-1)$ ).   \begin{displaymath}\begin{array}{cr} x + \frac{3}{2} z &= \frac{5}{2} \\ y + z &... ...1 & 1 & 2 \\ 0 & 1 & -\frac{1}{2} & \frac{1}{2} \end{bmatrix}. \end{displaymath}   1. Add $ -1$ times the $ 2^{\mbox{nd}}$ equation to the $ 3^{\mbox{rd}}$ equation (or $ R_{32}(-1)$ ).   $\displaystyle \begin{array}{cr} x + \frac{3}{2} z &= \frac{5}{2} \\ y + z &= 2 ... ...ac{5}{2} \\ 0 & 1 & 1 & 2 \\ 0 & 0 & -\frac{3}{2} & -\frac{3}{2} \end{bmatrix}.$   1. Multiply the $ 3^{\mbox{rd}}$ equation by $ \frac{-2}{3}$ (or $ R_3(-\frac{2}{3})$ ).   $\displaystyle \begin{array}{cr} x + \frac{3}{2} z &= \frac{5}{2} \\ y + z &= 2 ... ... 0 & \frac{3}{2} & \frac{5}{2} \\ 0 & 1 & 1 & 2 \\ 0 & 0 & 1 & 1 \end{bmatrix}.$  The last equation gives $ z=1,$ the second equation now gives $ y=1.$ Finally the first equation gives $ x = 1.$ Hence the set of solutions is $ (x, y, z)^t = (1, 1, 1)^t,$ A UNIQUE SOLUTION.  **EXAMPLE 2.2.12   *Solve the linear system by Gauss elimination method.***   |  |  |  |  | | --- | --- | --- | --- | | $\displaystyle x+ y + z$ | $\displaystyle =$ | $\displaystyle 3$ |  | | $\displaystyle x + 2 y + 2 z$ | $\displaystyle =$ | $\displaystyle 5$ |  | | $\displaystyle 3 x + 4 y + 4 z$ | $\displaystyle =$ | $\displaystyle 11$ |  |     **Solution:** In this case, the augmented matrix is $ \begin{bmatrix}1 & 1 & 1 & 3 \\ 1 & 2 & 2 & 5 \\ 3 & 4 & 4 & 11 \end{bmatrix}$ and the method proceeds as follows:   1. Add $ -1$ times the first equation to the second equation.   $\displaystyle \begin{array}{cr} x + y + z &= 3 \\ y + z &= 2 \\ 3 x +4 y +4 z ... ...begin{bmatrix} 1 & 1 & 1 & 3 \\ 0 & 1 & 1 & 2 \\ 3 & 4 & 4 & 11 \end{bmatrix}. $   1. Add $ -3$ times the first equation to the third equation.   $\displaystyle \begin{array}{cr} x + y + z &= 3 \\ y + z &= 2 \\ y + z &= 2 \en... ...\begin{bmatrix} 1 & 1 & 1 & 3 \\ 0 & 1 & 1 & 2 \\ 0 & 1 & 1 & 2 \end{bmatrix}. $   1. Add $ -1$ times the second equation to the third equation   $\displaystyle \begin{array}{cr} x + y+ z &= 3 \\ y + z &= 2 \end{array} \hspace... ... \begin{bmatrix}1 & 1 & 1 & 3 \\ 0 & 1 & 1 & 2 \\ 0 & 0 & 0 & 0 \end{bmatrix}. $  Thus, the set of solutions is $ (x, y, z)^t = (1, 2-z, z)^t = (1, 2, 0)^t + z (0, -1, 1)^t,$ with $ z$ arbitrary. In other words, the system has INFINITE NUMBER OF SOLUTIONS.  **EXAMPLE 2.2.13   *Solve the linear system by Gauss elimination method.***   |  |  |  |  | | --- | --- | --- | --- | | $\displaystyle x+ y + z$ | $\displaystyle =$ | $\displaystyle 3$ |  | | $\displaystyle x + 2 y + 2 z$ | $\displaystyle =$ | $\displaystyle 5$ |  | | $\displaystyle 3 x + 4 y + 4 z$ | $\displaystyle =$ | $\displaystyle 12$ |  |     **Solution:** In this case, the augmented matrix is $ \begin{bmatrix}1 & 1 & 1 & 3 \\ 1 & 2 & 2 & 5 \\ 3 & 4 & 4 & 12 \end{bmatrix}$ and the method proceeds as follows:   1. Add $ -1$ times the first equation to the second equation.   $\displaystyle \begin{array}{cr} x + y + z &= 3 \\ y + z &= 2 \\ 3 x +4 y +4 z ... ...begin{bmatrix} 1 & 1 & 1 & 3 \\ 0 & 1 & 1 & 2 \\ 3 & 4 & 4 & 12 \end{bmatrix}. $   1. Add $ -3$ times the first equation to the third equation.   $\displaystyle \begin{array}{cr} x + y + z &= 3 \\ y + z &= 2 \\ y + z &= 3 \en... ...\begin{bmatrix} 1 & 1 & 1 & 3 \\ 0 & 1 & 1 & 2 \\ 0 & 1 & 1 & 3 \end{bmatrix}. $   1. Add $ -1$ times the second equation to the third equation   $\displaystyle \begin{array}{cr} x + y+ z &= 3 \\ y + z &= 2 \\ 0 & = 1 \end{arr... ... \begin{bmatrix}1 & 1 & 1 & 3 \\ 0 & 1 & 1 & 2 \\ 0 & 0 & 0 & 1 \end{bmatrix}. $  The third equation in the last step is  $\displaystyle 0 x + 0 y + 0 z = 1.$  This can never hold for any value of $ x, y, z.$ Hence, the system has NO SOLUTION.  **Remark 2.2.14**   *Note that to solve a linear system, $ A {\mathbf x}= {\mathbf b},$ one needs to apply only the elementary row operations to the augmented matrix $ [ A \; \; {\mathbf b}].$* |
| **Rank of a Matrix**  In previous sections, we solved linear systems using Gauss elimination method or the Gauss-Jordan method. In the examples considered, we have encountered three possibilities, namely   1. existence of a unique solution, 2. existence of an infinite number of solutions, and 3. no solution.   Based on the above possibilities, we have the following definition.  **DEFINITION 2.4.1 (Consistent, Inconsistent)*****A linear system is called CONSISTENT if it admits a solution and is called INCONSISTENT if it admits no solution.***  The question arises, as to whether there are conditions under which the linear system $ A {\mathbf x}= {\mathbf b}$ is consistent. The answer to this question is in the affirmative. To proceed further, we need a few definitions and remarks.  Recall that the row reduced echelon form of a matrix is unique and therefore, the number of non-zero rows is a unique number. Also, note that the number of non-zero rows in either the row reduced form or the row reduced echelon form of a matrix are same.  **DEFINITION 2.4.2 (Row rank of a Matrix)*****The number of non-zero rows in the row reduced form of a matrix is called the row-rank of the matrix.***  By the very definition, it is clear that row-equivalent matrices have the same row-rank. For a matrix $ A,$ we write ` $ {\mbox{row-rank }}(A)$ ' to denote the row-rank of $ A.$  **EXAMPLE 2.4.3**   1. Determine the row-rank of $ A = \begin{bmatrix}1 &    2 & 1 \\ 2 & 3 & 1 \\ 1 & 1 & 2 \end{bmatrix}.$  **Solution:** To determine the row-rank of $ A,$ we proceed as follows.    1. $ \begin{bmatrix}1 & 2 & 1 \\ 2 & 3 & 1 \\ 1       & 1 & 2 \end{bmatrix} \overrightarr...       ..._{31}(-1)}       \begin{bmatrix}1 & 2 & 1 \\ 0 & -1 & -1 \\ 0 & -1 & 1       \end{bmatrix}.$    2. $ \begin{bmatrix}1 & 2 & 1 \\ 0 & -1 & -1 \\ 0 & -1 & 1       \end{bmatrix} \overright...       ..., R_{32}(1) }       \begin{bmatrix}1 & 2 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & 2       \end{bmatrix}.$    3. $ \begin{bmatrix}1 & 2 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & 2       \end{bmatrix}\overrightarro...       ...R_{12}(-2) }       \begin{bmatrix}1 & 0 & -1 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{bmatrix}.$    4. $ \begin{bmatrix}1 & 0 & -1 \\ 0 & 1 & 1 \\ 0 & 0 & 1       \end{bmatrix}\overrightarr...       ...1), R_{13}(1)}\begin{bmatrix}1 & 0 & 0       \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} $   The last matrix in Step [1d](http://nptel.iitm.ac.in/courses/Webcourse-contents/IIT-KANPUR/mathematics-2/node22.html#last:1) is the row reduced form of $ A$ which has $ 3$ non-zero rows. Thus, $ {\mbox{row-rank}}(A)~=~3.$ This result can also be easily deduced from the last matrix in Step [1b](http://nptel.iitm.ac.in/courses/Webcourse-contents/IIT-KANPUR/mathematics-2/node22.html#last:but:1).   1. Determine the row-rank of $ A = \begin{bmatrix}1 & 2 & 1    \\ 2 & 3 & 1 \\ 1 & 1 & 0 \end{bmatrix}.$  **Solution:** Here we have    1. $ \begin{bmatrix}1 & 2 & 1 \\ 2 & 3 & 1 \\ 1       & 1 & 0 \end{bmatrix} \overrightarr...       ...31}(-1) }       \begin{bmatrix}1 & 2 & 1 \\ 0 & -1 & -1 \\ 0 & -1 & -1       \end{bmatrix}.$    2. $ \begin{bmatrix}1 & 2 & 1 \\ 0 & -1 & -1 \\ 0 & -1 & -1       \end{bmatrix} \overrigh...       ..., R_{32}(1) }       \begin{bmatrix}1 & 2 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & 0       \end{bmatrix}.$   From the last matrix in Step [2b](http://nptel.iitm.ac.in/courses/Webcourse-contents/IIT-KANPUR/mathematics-2/node22.html#last:2), we deduce $ {\mbox{row-rank}}(A)=2.$  **Remark 2.4.4**   *Let $ A {\mathbf x}= {\mathbf b}$ be a linear system with $ m$ equations and $ n$ unknowns. Then the row-reduced echelon form of $ A$ agrees with the first $ n$ columns of $ [A \; \; {\mathbf b}],$ and hence*  $\displaystyle {\mbox{row-rank}} (A) \leq {\mbox{row-rank}} ([A \; \; {\mathbf b}]).$  *The reader is advised to supply a proof.*  **Remark 2.4.5**   *Consider a matrix $ A.$ After application of a finite number of elementary column operations (see Definition*[*2.3.16*](http://nptel.iitm.ac.in/courses/Webcourse-contents/IIT-KANPUR/mathematics-2/node21.html#defn:column:trans)*) to the matrix $ A,$ we can have a matrix, say $ B,$ which has the following properties:*   1. The first nonzero entry in each column is $ 1.$ 2. A column containing only 0 's comes after all columns with at least one non-zero entry. 3. The first non-zero entry (the leading term) in each non-zero column moves down in successive columns.   Therefore, we can define **column-rank** of $ A$ as the number of non-zero columns in $ B.$ It will be proved later that  $\displaystyle {\mbox{row-rank}} (A) = {\mbox{column-rank}} (A).$  Thus we are led to the following definition.  **DEFINITION 2.4.6   *The number of non-zero rows in the row reduced form of a matrix $ A$ is called the rank of $ A,$ denoted $ {\mbox{rank }} (A).$***  **THEOREM 2.4.7   *Let $ A$ be a matrix of rank $ r.$ Then there exist elementary matrices $ E_{1},E_{2},\ldots ,E_{s}$ and $ F_{1},F_{2},\ldots ,F_{\ell }$ such that***  $\displaystyle E_{1}E_{2}\ldots E_{s} \; A \; F_{1}F_{2}\ldots F_{\ell } = \begin{bmatrix}I_{r}& {\mathbf 0}\\ {\mathbf 0}& {\mathbf 0}\end{bmatrix}. $  *Proof*. Let $ C$ be the row reduced echelon matrix obtained by applying elementary row operations to the given matrix $ A.$ As $ {\mbox{rank}}(A) = r,$ the matrix $ C$ will have the first $ r$ rows as the non-zero rows. So by Remark [2.3.5](http://nptel.iitm.ac.in/courses/Webcourse-contents/IIT-KANPUR/mathematics-2/node19.html#rem:leading), $ C$ will have $ r$ leading columns, say $ i_1, i_2, \ldots, i_r.$ Note that, for $ 1 \leq s \leq r, $ the $ i_s^{\mbox{th}}$ column will have $ 1$ in the $ s^{\mbox{th}}$ row and zero elsewhere.  We now apply column operations to the matrix $ C.$ Let $ D$ be the matrix obtained from $ C$ by successively interchanging the $ s^{\mbox{th}}$ and $ i_s^{\mbox{th}}$ column of $ C$ for $ 1 \leq s \leq r.$ Then the matrix $ D$ can be written in the form $ \begin{bmatrix}I_r & B \\ {\mathbf 0}& {\mathbf 0}\end{bmatrix},$ where $ B$ is a matrix of appropriate size. As the $ (1,1)$ block of $ D$ is an identity matrix, the block $ (1,2)$ can be made the zero matrix by application of column operations to $ D.$ This gives the required result. height6pt width 6pt depth 0pt  **COROLLARY 2.4.8   *Let $ A$ be a $ n \times n$ matrix of rank $ r<n.$ Then the system of equations $ A {\mathbf x}= {\mathbf 0}$ has infinite number of solutions.***  *Proof*. By Theorem [2.4.7](http://nptel.iitm.ac.in/courses/Webcourse-contents/IIT-KANPUR/mathematics-2/node22.html#thm:pqr), there exist elementary matrices $ E_{1},E_{2},\ldots ,E_{s}$ and $ F_{1},F_{2},\ldots ,F_{\ell }$ such that $ E_{1}E_{2}\ldots E_{s} \; A \; F_{1}F_{2}\ldots F_{\ell } = \begin{bmatrix}I_{r}&0\\ 0&0\end{bmatrix}. $ Define $ Q= F_{1}F_{2}\ldots F_{\ell }$ . Then the matrix  $\displaystyle A Q = \left[\begin{array}{c\vert c} & \\ \star & {\mathbf 0}\\ & \end{array}\right]$  as the elementary martices $ E_i$ 's are being multiplied on the left of the matrix $ \begin{bmatrix}I_{r}& {\mathbf 0}\\ {\mathbf 0}& {\mathbf 0}\end{bmatrix}.$ Let $ Q_1, Q_2, \ldots, Q_n$ be the columns of the matrix $ Q$ . Then check that $ A Q_i = {\mathbf 0}$ for $ i = r+1, r+2, \ldots, n$ . Hence, we can use the $ Q_i$ 's which are non-zero (Use Exercise [1.2.17](http://nptel.iitm.ac.in/courses/Webcourse-contents/IIT-KANPUR/mathematics-2/node8.html#exe:inver).[2](http://nptel.iitm.ac.in/courses/Webcourse-contents/IIT-KANPUR/mathematics-2/node8.html#exe:inver:3)) to generate infinite number of solutions. height6pt width 6pt depth 0pt  **EXERCISE 2.4.9**   1. Determine the ranks of the coefficient and the augmented matrices that appear in Part [1](http://nptel.iitm.ac.in/courses/Webcourse-contents/IIT-KANPUR/mathematics-2/node20.html#linear) and Part [2](http://nptel.iitm.ac.in/courses/Webcourse-contents/IIT-KANPUR/mathematics-2/node20.html#echelon) of Exercise [2.3.12](http://nptel.iitm.ac.in/courses/Webcourse-contents/IIT-KANPUR/mathematics-2/node20.html#exe:sys:soln). 2. Let $ A$ be an $ n \times n$ matrix with $ {\mbox{rank}}(A) =n.$ Then prove that $ A$ is row-equivalent to $ I_n.$ 3. If $ P$ and $ Q$ are invertible matrices and $ P A Q$ is defined then show that $ {\mbox{rank }}(P A Q) = {\mbox{ rank }} (A).$ 4. Find matrices $ P$ and $ Q$ which are product of elementary matrices such that $ B = P A Q$ where $ A = \begin{bmatrix}    2 & 4 & 8\\ 1 & 3 & 2 \end{bmatrix}$ and $ B =    \begin{bmatrix}1 & 0 & 0\\ 0 & 1 & 0    \end{bmatrix}.$ 5. Let $ A$ and $ B$ be two matrices. Show that    1. if $ A + B$ is defined, then $ {\mbox{rank}}(A+B) \leq {\mbox{rank}}(A) + {\mbox{rank}}(B),$    2. if $ A B$ is defined, then $ {\mbox{rank}}(AB) \leq {\mbox{rank}}(A)$ and $ {\mbox{rank}}(AB) \leq {\mbox{rank}}(B).$ 6. Let $ A$ be any matrix of rank $ r.$ Then show that there exists invertible matrices $ B_i, C_i$ such that  $ B_1 A =    \begin{bmatrix}R_1 & R_2 \\ {\mathbf 0}& {\mathbf 0}\end{bmatrix},    \;\...    ...C_2 = \begin{bmatrix}A_1 & {\mathbf 0}\\ {\mathbf 0}& {\mathbf 0}\end{bmatrix},$ and $ B_3 A C_3 =    \begin{bmatrix}I_r & {\mathbf 0}\\ {\mathbf 0}& {\mathbf 0}\end{bmatrix}.$ Also, prove that the matrix $ A_1$ is an $ r \times r$ invertible matrix. 7. Let $ A$ be an $ m \times n$ matrix of rank $ r.$ Then $ A$ can be written as $ A = B C,$ where both $ B$ and $ C$ have rank $ r$ and $ B$ is a matrix of size $ m \times r$ and $ C$ is a matrix of size $ r \times n.$ 8. Let $ A$ and $ B$ be two matrices such that $ A B$ is defined and $ {\mbox{rank }}(A) = {\mbox{rank }}(A B).$ Then show that $ A = A B X$ for some matrix $ X.$ Similarly, if $ B A$ is defined and $ {\mbox{rank }}(A) = {\mbox{rank }}(B A),$then $ A =    Y B A$ for some matrix $ Y.$ *[Hint: Choose non-singular matrices $ P, Q $ and $ R$ such that $ P A Q = \begin{bmatrix}A_1 &    {\mathbf 0}\\ {\mathbf 0}& {\mathbf 0}    \end{bmatrix}$ and $ P (A B) R = \begin{bmatrix}C & {\mathbf 0}\\ {\mathbf 0}& {\mathbf 0}    \end{bmatrix}.$ Define $ X = R \begin{bmatrix}C^{-1} A_1 & {\mathbf 0}\\ {\mathbf 0}& {\mathbf 0}    \end{bmatrix} Q^{-1}.$ ]* 9. If matrices $ B$ and $ C$ are invertible and the involved partitioned products are defined, then show that   $\displaystyle \begin{bmatrix}A&B \\ C&{\mathbf 0}\end{bmatrix}^{-1} = \begin{bmatrix}{\mathbf 0}&C^{-1}\\ B^{-1}&-B^{-1}AC^{-1}\end{bmatrix}. $   1. Suppose $ A$ is the inverse of a matrix $ B.$ Partition $ A$ and $ B$ as follows:   $\displaystyle A = \begin{bmatrix}A_{11} & A_{12} \\ A_{21} & A_{22}\end{bmatrix}, \;\; B = \begin{bmatrix}B_{11}& B_{12} \\ B_{21} & B_{22} \end{bmatrix}. $  If $ A_{11}$ is invertible and $ P = A_{22} - A_{21}(A^{-1}_{11} A_{12}),$ then show that  $\displaystyle B_{11} = A^{-1}_{11} + (A^{-1}_{11}A_{12})P^{-1}(A_{21} A^{-1}_{1... ... = - P^{-1}(A_{21} A^{-1}_{11}), \;\; B_{12} = - (A^{-1}_{11} A_{12})P^{-1}, $  and $ B_{22} = P^{-1}.$ |
| Main Theorem **THEOREM 2.5.1   *[Existence and Non-existence] Consider a linear system $ A {\mathbf x}= {\mathbf b},$ where $ A$ is a $ m \times n$ matrix, and $ \; {\mathbf x}, \; {\mathbf b}$ are vectors with orders $ n \times 1,$ and $ m \times 1,$ respectively. Suppose$ {\mbox{rank }}(A) = r$ and $ {\mbox{rank}} ([A \; \;{\mathbf b}]) = r_a.$ Then exactly one of the following statement holds:***   1. if $ \; r_a = r < n,$ the set of solutions of the linear system is an infinite set and has the form   $\displaystyle \{ {\mathbf u}_0 + k_1 {\mathbf u}_1 + k_2 {\mathbf u}_2 + \cdots... ...n-r} {\mathbf u}_{n-r} \; : \;\; k_i \in {\mathbb{R}}, \; 1 \leq i \leq n-r \},$  where $ {\mathbf u}_0, {\mathbf u}_1, \ldots, {\mathbf u}_{n-r}$ are $ n \times 1$ vectors satisfying $ A {\mathbf u}_0 = {\mathbf b}$ and $ A {\mathbf u}_i = {\mathbf 0}$ for $ 1 \leq i \leq n-r.$   1. if $ \; r_a = r = n,$ the solution set of the linear system has a unique $ n \times 1$ vector $ {\mathbf x}_0$ satisfying $ A {\mathbf x}_0 = {\mathbf b}.$ 2. If $ \; r < r_a,$ the linear system has no solution.   **Remark 2.5.2**   *Let $ A$ be an $ m \times n$ matrix and consider the linear system $ A {\mathbf x}= {\mathbf b}.$ Then by Theorem*[*2.5.1*](http://nptel.iitm.ac.in/courses/Webcourse-contents/IIT-KANPUR/mathematics-2/node25.html#thm:solns)*, we see that the linear system $ A {\mathbf x}= {\mathbf b}$ is consistent if and only if*  $\displaystyle {\mbox{rank }}(A) = {\mbox{rank}} ([A \; \;{\mathbf b}]).$  The following corollary of Theorem [2.5.1](http://nptel.iitm.ac.in/courses/Webcourse-contents/IIT-KANPUR/mathematics-2/node25.html#thm:solns) is a very important result about the homogeneous linear system $ A {\mathbf x}= {\mathbf 0}.$  **COROLLARY 2.5.3   *Let $ A$ be an $ m \times n$ matrix. Then the homogeneous system $ A {\mathbf x}= {\mathbf 0}$ has a non-trivial solution if and only if rank$ (A) < n.$***  *Proof*. Suppose the system $ A {\mathbf x}= {\mathbf 0}$ has a non-trivial solution, $ {\mathbf x}_0.$ That is, $ A {\mathbf x}_0 = {\mathbf 0}$ and $ {\mathbf x}_0 \neq {\mathbf 0}.$ Under this assumption, we need to show that $ {\mbox{rank}} (A) < n.$ On the contrary, assume that rank$ (A) = n.$ So,  $\displaystyle n = {\mbox{rank}}(A) = {\mbox{rank}} \bigl([A \;\; {\mathbf 0}]\bigr) = r_a.$  Also $ A {\mathbf 0}= {\mathbf 0}$ implies that $ {\mathbf 0}$ is a solution of the linear system $ A {\mathbf x}= {\mathbf 0}.$ Hence, by the uniqueness of the solution under the condition $ r = r_a = n$ (see Theorem [2.5.1](http://nptel.iitm.ac.in/courses/Webcourse-contents/IIT-KANPUR/mathematics-2/node25.html#thm:solns)), we get $ {\mathbf x}_0 = {\mathbf 0}.$ A contradiction to the fact that $ {\mathbf x}_0$ was a given non-trivial solution.  Now, let us assume that rank$ (A) < n.$ Then  $\displaystyle r_a = {\mbox{rank}} \bigl( [A \;\; {\mathbf 0}] \bigr) = {\mbox{rank}} (A) < n.$  So, by Theorem [2.5.1](http://nptel.iitm.ac.in/courses/Webcourse-contents/IIT-KANPUR/mathematics-2/node25.html#thm:solns), the solution set of the linear system $ A {\mathbf x}= {\mathbf 0}$ has infinite number of vectors $ {\mathbf x}$ satisfying $ A {\mathbf x}= {\mathbf 0}.$ From this infinite set, we can choose any vector $ {\mathbf x}_0$ that is different from $ {\mathbf 0}.$ Thus, we have a solution $ {\mathbf x}_0 \neq {\mathbf 0}.$ That is, we have obtained a non-trivial solution $ {\mathbf x}_0.$ height6pt width 6pt depth 0pt  We now state another important result whose proof is immediate from Theorem [2.5.1](http://nptel.iitm.ac.in/courses/Webcourse-contents/IIT-KANPUR/mathematics-2/node25.html#thm:solns) and Corollary [2.5.3](http://nptel.iitm.ac.in/courses/Webcourse-contents/IIT-KANPUR/mathematics-2/node25.html#cor:system).  **PROPOSITION 2.5.4   *Consider the linear system $ A {\mathbf x}= {\mathbf b}.$ Then the two statements given below cannot hold together.***   1. The system $ A {\mathbf x}= {\mathbf b}$ has a unique solution for every $ {\mathbf b}.$ 2. The system $ A {\mathbf x}= {\mathbf 0}$ has a non-trivial solution.   **Remark 2.5.5**   1. Suppose $ {\mathbf x}_1, {\mathbf x}_2$ are two solutions of $ A {\mathbf x}= {\mathbf 0}.$ Then $ k_1 {\mathbf x}_1 + k_2 {\mathbf x}_2$ is also a solution of $ A {\mathbf x}= {\mathbf 0}$ for any $ k_1, k_2 \in {\mathbb{R}}.$ 2. If $ {\mathbf u}, {\mathbf v}$ are two solutions of $ A {\mathbf x}= {\mathbf b}$ then $ {\mathbf u}- {\mathbf v}$ is a solution of the system $ A {\mathbf x}= {\mathbf 0}.$ That is, $ {\mathbf u}- {\mathbf v}= {\mathbf x}_h$ for some solution $ {\mathbf x}_h$ of $ A {\mathbf x}= {\mathbf 0}.$ That is, any two solutions of $ A {\mathbf x}= {\mathbf b}$ differ by a solution of the associated homogeneous system $ A {\mathbf x}= {\mathbf 0}.$   In conclusion, for $ {\mathbf b}\neq {\mathbf 0},$ the set of solutions of the system $ A {\mathbf x}= {\mathbf b}$ is of the form, $ \{{\mathbf x}_0 + {\mathbf x}_h\};$ where $ {\mathbf x}_0$ is a particular solution of $ A {\mathbf x}= {\mathbf b}$ and $ {\mathbf x}_h$ is a solution $ A {\mathbf x}= {\mathbf 0}.$ |
| **THEOREM 2.5.8   *For a square matrix $ A$ of order $ n,$ the following statements are equivalent.***   1. $ A$ is invertible. 2. $ A$ is of full rank. 3. $ A$ is row-equivalent to the identity matrix. 4. $ A$ is a product of elementary matrices.   *Proof*. [1](http://nptel.iitm.ac.in/courses/Webcourse-contents/IIT-KANPUR/mathematics-2/node26.html#equi:1) $ \Longrightarrow $ [2](http://nptel.iitm.ac.in/courses/Webcourse-contents/IIT-KANPUR/mathematics-2/node26.html#equi:2)  Let if possible rank $ (A) = r < n.$ Then there exists an invertible matrix $ P$ (a product of elementary matrices) such that $ P A = \begin{bmatrix}B_1 & B_2 \\ {\mathbf 0}& {\mathbf 0}\end{bmatrix},$ where $ B_1$ is an $ r \times r$ matrix. Since $ A$ is invertible, let$ A^{-1} = \begin{bmatrix}C_1 \\ C_2 \end{bmatrix},$ where $ C_1$ is an $ r \times n$ matrix. Then   |  |  | | --- | --- | | $\displaystyle P = P I_n = P (A A^{-1})= (P A ) A^{-1} = \begin{bmatrix}B_1 & B_... ...2 \end{bmatrix} = \begin{bmatrix}B_1 C_1 + B_2 C_2 \\ {\mathbf 0}\end{bmatrix}.$ | (2.5.1) |     Thus the matrix $ P$ has $ n-r$ rows as zero rows. Hence, $ P$ cannot be invertible. A contradiction to $ P$ being a product of invertible matrices. Thus, $ A$ is of full rank.  [2](http://nptel.iitm.ac.in/courses/Webcourse-contents/IIT-KANPUR/mathematics-2/node26.html#equi:2) $ \Longrightarrow $ [3](http://nptel.iitm.ac.in/courses/Webcourse-contents/IIT-KANPUR/mathematics-2/node26.html#equi:3)  Suppose $ A$ is of full rank. This implies, the row reduced echelon form of $ A$ has all non-zero rows. But $ A$ has as many columns as rows and therefore, the last row of the row reduced echelon form of $ A$ will be $ (0, 0, \ldots, 0, 1).$ Hence, the row reduced echelon form of $ A$ is the identity matrix.  [3](http://nptel.iitm.ac.in/courses/Webcourse-contents/IIT-KANPUR/mathematics-2/node26.html#equi:3) $ \Longrightarrow $ [4](http://nptel.iitm.ac.in/courses/Webcourse-contents/IIT-KANPUR/mathematics-2/node26.html#equi:4)  Since $ A$ is row-equivalent to the identity matrix there exist elementary matrices $ E_1, E_2, \ldots, E_k$ such that $ A = E_1 E_2 \cdots E_k I_n.$ That is, $ A$ is product of elementary matrices.  [4](http://nptel.iitm.ac.in/courses/Webcourse-contents/IIT-KANPUR/mathematics-2/node26.html#equi:4) $ \Longrightarrow $ [1](http://nptel.iitm.ac.in/courses/Webcourse-contents/IIT-KANPUR/mathematics-2/node26.html#equi:1)  Suppose $ A = E_1 E_2 \cdots E_k; $ where the $ E_i$ 's are elementary matrices. We know that elementary matrices are invertible and product of invertible matrices is also invertible, we get the required result. height6pt width 6pt depth 0pt  The ideas of Theorem [2.5.8](http://nptel.iitm.ac.in/courses/Webcourse-contents/IIT-KANPUR/mathematics-2/node26.html#thm:equi) will be used in the next subsection to find the inverse of an invertible matrix. The idea used in the proof of the first part also gives the following important Theorem. We repeat the proof for the sake of clarity.  **THEOREM 2.5.9   *Let $ A$ be a square matrix of order $ n.$***   1. Suppose there exists a matrix $ B$ such that $ A B = I_n.$ Then $ A^{-1}$ exists. 2. Suppose there exists a matrix $ C$ such that $ C A = I_n.$ Then $ A^{-1}$ exists.   *Proof*. Suppose that $ A B = I_n.$ We will prove that the matrix $ A$ is of full rank. That is, $ {\mbox{rank}}(A) =n.$  Let if possible, rank $ (A) = r < n.$ Then there exists an invertible matrix $ P$ (a product of elementary matrices) such that $ P A = \begin{bmatrix}C_1 & C_2 \\ {\mathbf 0}& {\mathbf 0}\end{bmatrix}.$ Let $ B = \begin{bmatrix}B_1 \\ B_2 \end{bmatrix},$ where $ B_1$ is an $ r \times n$ matrix. Then   |  |  | | --- | --- | | $\displaystyle P = P I_n = P (A B)= (P A ) B = \begin{bmatrix}C_1 & C_2 \\ {\mat... ...2 \end{bmatrix} = \begin{bmatrix}C_1 B_1 + C_2 B_2 \\ {\mathbf 0}\end{bmatrix}.$ | (2.5.2) |     Thus the matrix $ P$ has $ n-r$ rows as zero rows. So, $ P$ cannot be invertible. A contradiction to $ P$ being a product of invertible matrices. Thus, $ {\mbox{rank}}(A) =n.$ That is, $ A$ is of full rank. Hence, using Theorem [2.5.8](http://nptel.iitm.ac.in/courses/Webcourse-contents/IIT-KANPUR/mathematics-2/node26.html#thm:equi), $ A$ is an invertible matrix. That is, $ BA = I_n$ as well.  Using the first part, it is clear that the matrix $ C$ in the second part, is invertible. Hence  $\displaystyle A C = I_n = C A.$  Thus, $ A$ is invertible as well. height6pt width 6pt depth 0pt  **Remark 2.5.10**   *This theorem implies the following: ``if we want to show that a square matrix $ A$ of order $ n$ is invertible, it is enough to show the existence of*   1. either a matrix $ B$ such that $ A B = I_n$ 2. or a matrix $ C$ such that $ C A = I_n.$   **THEOREM 2.5.11   *The following statements are equivalent for a square matrix $ A$ of order $ n.$***   1. $ A$ is invertible. 2. $ A {\mathbf x}= {\mathbf 0}$ has only the trivial solution $ {\mathbf x}= {\mathbf 0}.$ 3. $ A {\mathbf x}= {\mathbf b}$ has a solution $ {\mathbf x}$ for every $ {\mathbf b}.$   *Proof*. [1](http://nptel.iitm.ac.in/courses/Webcourse-contents/IIT-KANPUR/mathematics-2/node26.html#equi:soln:1) $ \Longrightarrow $ [2](http://nptel.iitm.ac.in/courses/Webcourse-contents/IIT-KANPUR/mathematics-2/node26.html#equi:soln:2)  Since $ A$ is invertible, by Theorem [2.5.8](http://nptel.iitm.ac.in/courses/Webcourse-contents/IIT-KANPUR/mathematics-2/node26.html#thm:equi) $ A$ is of full rank. That is, for the linear system $ A {\mathbf x}= {\mathbf 0},$ the number of unknowns is equal to the rank of the matrix $ A.$ Hence, by Theorem [2.5.1](http://nptel.iitm.ac.in/courses/Webcourse-contents/IIT-KANPUR/mathematics-2/node25.html#thm:solns) the system $ A {\mathbf x}= {\mathbf 0}$has a unique solution $ {\mathbf x}= {\mathbf 0}.$  [2](http://nptel.iitm.ac.in/courses/Webcourse-contents/IIT-KANPUR/mathematics-2/node26.html#equi:soln:2) $ \Longrightarrow $ [1](http://nptel.iitm.ac.in/courses/Webcourse-contents/IIT-KANPUR/mathematics-2/node26.html#equi:soln:1)  Let if possible $ A$ be non-invertible. Then by Theorem [2.5.8](http://nptel.iitm.ac.in/courses/Webcourse-contents/IIT-KANPUR/mathematics-2/node26.html#thm:equi), the matrix $ A$ is not of full rank. Thus by Corollary [2.5.3](http://nptel.iitm.ac.in/courses/Webcourse-contents/IIT-KANPUR/mathematics-2/node25.html#cor:system), the linear system $ A {\mathbf x}= {\mathbf 0}$ has infinite number of solutions. This contradicts the assumption that $ A {\mathbf x}= {\mathbf 0}$ has only the trivial solution $ {\mathbf x}= {\mathbf 0}.$  [1](http://nptel.iitm.ac.in/courses/Webcourse-contents/IIT-KANPUR/mathematics-2/node26.html#equi:soln:1) $ \Longrightarrow $ [3](http://nptel.iitm.ac.in/courses/Webcourse-contents/IIT-KANPUR/mathematics-2/node26.html#equi:soln:3)  Since $ A$ is invertible, for every $ {\mathbf b},$ the system $ A {\mathbf x}= {\mathbf b}$ has a unique solution $ {\mathbf x}= A^{-1} {\mathbf b}.$  [3](http://nptel.iitm.ac.in/courses/Webcourse-contents/IIT-KANPUR/mathematics-2/node26.html#equi:soln:3) $ \Longrightarrow $ [1](http://nptel.iitm.ac.in/courses/Webcourse-contents/IIT-KANPUR/mathematics-2/node26.html#equi:soln:1)  For $ 1 \leq i \leq n,$ define $ {\mathbf e}_i = (0, \ldots, 0, \underbrace{1}_{i^{\mbox{th position}}}, 0, \ldots, 0)^t,$ and consider the linear system $ A {\mathbf x}= {\mathbf e}_i.$ By assumption, this system has a solution $ {\mathbf x}_i$ for each $ i, \; 1 \leq i \leq n.$ Define a matrix$ B = [{\mathbf x}_1, {\mathbf x}_2, \ldots, {\mathbf x}_n ].$ That is, the $ i^{\mbox{th}}$ column of $ B$ is the solution of the system $ A {\mathbf x}= {\mathbf e}_i.$ Then  $\displaystyle A B = A [{\mathbf x}_1, {\mathbf x}_2 \ldots, {\mathbf x}_n] = [A... ..., A {\mathbf x}_n]= [{\mathbf e}_1, {\mathbf e}_2 \ldots, {\mathbf e}_n] = I_n.$  Therefore, by Theorem [2.5.9](http://nptel.iitm.ac.in/courses/Webcourse-contents/IIT-KANPUR/mathematics-2/node26.html#thm:left:right:inverse), the matrix $ A$ is invertible. |
| **Determinant**  Notation: For an $ n \times n$ matrix $ A,$ by $ A({\alpha}\vert \beta),$ we mean the submatrix $ B$ of $ A,$ which is obtained by deleting the $ {\alpha}^{\mbox{th}}$ row and $ \beta^{\mbox{th}}$ column.  **EXAMPLE 2.6.1   *Consider a matrix $ A = \begin{bmatrix}1 & 2 & 3\\ 1 &3&2 \\ 2 & 4&7 \end{bmatrix}.$ Then $ A(1\vert 2) = \begin{bmatrix}1 & 2 \\ 2 & 7 \end{bmatrix},$ $ A(1\vert 3) = \begin{bmatrix}1 & 3 \\ 2 & 4 \end{bmatrix},$ and $ A(1,2\vert 1,3) = [4].$***  **DEFINITION 2.6.2 (Determinant of a Square Matrix)*****Let $ A$ be a square matrix of order $ n.$ With $ A,$ we associate inductively (on $ n$ ) a number, called the determinant of $ A,$ written $ \det (A)$(or $ \vert A\vert$ ) by***  \begin{displaymath}\det(A) = \left \{ \begin{array}{lc} a & {\mbox{if }} A = [a]... ...l(A(1\vert j)\bigr), & {\mbox{ otherwise}}. \end{array} \right.\end{displaymath}  **EXAMPLE 2.6.3**   1. Let $ A= \begin{bmatrix}a_{11} & a_{12} \\ a_{21}    & a_{22}    \end{bmatrix}.$ Then,   $\displaystyle \det(A) = \vert A\vert = a_{11} \det(A{(1\vert 1)}) - a_{12} \det(A{(1\vert 2)}) = a_{11} a_{22} - a_{12} a_{21}.$  For example, for $ A = \begin{bmatrix}1 & 2 \\ 3 & 5 \end{bmatrix},$ $ \; \det(A) = \begin{vmatrix}1 & 2 \\ 3 & 5 \end{vmatrix} = 1\cdot 5 - 2 \cdot 3 = -1.$   1. Let $ A= \begin{bmatrix}a_{11} & a_{12} &    a_{13}\\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33}    \end{bmatrix}.$ Then,  |  |  |  |  | | --- | --- | --- | --- | | $\displaystyle \det(A)$ | $\displaystyle =$ | $\displaystyle \vert A\vert = a_{11} \det(A{(1\vert 1)}) - a_{12} \det(A{(1\vert 2)}) + a_{13} \det(A{(1\vert 3)})$ |  | |  | $\displaystyle =$ | $\displaystyle a_{11} \begin{vmatrix}a_{22} & a_{23} \\ \nonumber a_{32} & a_{33... ..._{13} \begin{vmatrix}a_{21} & a_{22} \\ \nonumber a_{31} & a_{32} \end{vmatrix}$ |  | |  | $\displaystyle =$ | $\displaystyle a_{11} ( a_{22} a_{33} - a_{23} a_{32} ) - a_{12} ( a_{21} a_{33} - a_{31} a_{23} ) + a_{13} ( a_{21} a_{32} - a_{31} a_{22} )$ |  | |  | $\displaystyle =$ | $\displaystyle a_{11} a_{22} a_{33} - a_{11} a_{23} a_{32} - a_{12} a_{21} a_{33} + a_{12} a_{23} a_{31} + a_{13} a_{21}a_{32} - a_{13} a_{22} a_{31}$ | (2.6.1) |  2. For example, if $ A = \begin{bmatrix}1 & 2 & 3 \\ 2 & 3 & 1 \\ 1    & 2 & 2 \end{bmatrix}$ then  $ \det(A) = \begin{vmatrix}1 & 2 & 3 \\ 2 & 3 & 1 \\ 1    & 2 & 2 \end{vmatrix} = 1...    ...x} + 3 \cdot \begin{vmatrix}2 &    3 \\ 1 & 2 \end{vmatrix} = 4 - 2(3) + 3(1) = 1.$   **EXERCISE 2.6.4**   1. Find the determinant of the following matrices.  $ i) \; \begin{bmatrix}1 & 2 & 7 & 8 \\ 0 & 4 & 3 & 2 \\ 0 & 0 & 2 & 3\\    0 & 0 ...    ...ii) \;    \begin{bmatrix}1 & a & a^2 \\ 1 & b & b^2 \\    1 & c & c^2 \end{bmatrix}.$ 2. Show that the determinant of a triangular matrix is the product of its diagonal entries.   **DEFINITION 2.6.5   *A matrix $ A$ is said to be a singular matrix if $ \det (A) = 0.$ It is called non-singular if $ \det (A) \neq 0.$***  The proof of the next theorem is omitted. The interested reader is advised to go through Appendix [14.3](http://nptel.iitm.ac.in/courses/Webcourse-contents/IIT-KANPUR/mathematics-2/node134.html#app:determinant).  **THEOREM 2.6.6   *Let $ A$ be an $ n \times n$ matrix. Then***   1. if $ B$ is obtained from $ A$ by interchanging two rows, then $ \det (B) = - \det (A)$ , 2. if $ B$ is obtained from $ A$ by multiplying a row by $ c$ then $ \det (B) = c \det (A)$ , 3. if all the elements of one row or column of $ A$ are 0 then $ \det (A) = 0$ , 4. if $ B$ is obtained from $ A$ by replacing the $ j$ th row by itself plus $ k$ times the $ i$ th row, where $ i \neq j$ then $ \det (B) = \det (A)$ , 5. if $ A$ is a square matrix having two rows equal then $ \det (A) = 0$ .   **Remark 2.6.7**   1. Many authors define the determinant using ``Permutations." It turns out that THE WAY WE HAVE DEFINED DETERMINANT is usually called the expansion of the determinant along the first row. 2. Part [1](http://nptel.iitm.ac.in/courses/Webcourse-contents/IIT-KANPUR/mathematics-2/node28.html#lem:interchange) of Lemma [2.6.6](http://nptel.iitm.ac.in/courses/Webcourse-contents/IIT-KANPUR/mathematics-2/node28.html#lem:det:basic) implies that ``one can also calculate the determinant by expanding along any row." Hence, for an $ n \times n$ matrix $ A,$ for every $ k, \; 1 \le k \le n$ , one also has   $\displaystyle \det (A) = \sum_{j=1}^n (-1)^{k+j} a_{kj} \det\bigl(A(k\vert j)\bigr).$  **Remark 2.6.8**   1. Let $ {\mathbf u}^t = (u_1, u_2)$ and $ {\mathbf v}^t = (v_1, v_2)$ be two vectors in $ {\mathbb{R}}^2.$ Then consider the parallelogram, $ PQRS,$ formed by the vertices $ P = (0,0)^t, Q = {\mathbf u}, R={\mathbf u}+{\mathbf v}$ and $ S= {\mathbf v}.$ We   $\displaystyle {\mbox{Claim: }} \hspace{.5in} {\mbox{Area}} \; (PQRS) = \vert u_1 v_2 - u_2 v_1\vert = \begin{vmatrix}u_1 & v_1 \\ u_2 & v_2 \end{vmatrix}.$  Recall that the dot product, $ {\mathbf u}\bullet {\mathbf v}= u_1 v_1 + u_2 v_2,$ and $ \sqrt{{\mathbf u}\bullet{\mathbf u}} = \sqrt{(u_1^2 + u_2^2)},$ is the length of the vector $ {\mathbf u}.$ We denote the length by $ \ell({\mathbf u}).$ With the above notation, if $ \theta$ is the angle between the vectors $ {\mathbf u}$ and $ {\mathbf v},$ then  $\displaystyle \cos(\theta) = \frac{ {\mathbf u}\bullet {\mathbf v}}{\ell({\mathbf u}) \ell({\mathbf v})}.$  Which tells us,   |  |  |  |  | | --- | --- | --- | --- | | $\displaystyle {\mbox{Area}}(PQRS)$ | $\displaystyle =$ | $\displaystyle \ell({\mathbf u}) \ell({\mathbf v}) \sin(\theta) = \ell({\mathbf ... ...{\mathbf u}\bullet {\mathbf v}}{\ell({\mathbf u}) \ell({\mathbf v})} \right)^2}$ |  | |  | $\displaystyle =$ | $\displaystyle \sqrt{ \ell({\mathbf u})^2 + \ell(v)^2 - ({\mathbf u}\bullet{\mathbf v})^2} = \sqrt{(u_1 v_2 - u_2 v_1)^2}$ |  | |  | $\displaystyle =$ | $\displaystyle \vert u_1 v_2 - u_2 v_1\vert.$ |  |   Hence, the claim holds. That is, in $ {\mathbb{R}}^2,$ the determinant is $ \pm$ times the area of the parallelogram.   1. Let $ {\mathbf u}= (u_1, u_2, u_3), {\mathbf v}= (v_1, v_2, v_3)$ and $ {\mathbf w}= (w_1, w_2, w_3)$ be three elements of $ {\mathbb{R}}^3.$ Recall that the cross product of two vectors in $ {\mathbb{R}}^3$ is,   $\displaystyle {\mathbf u}\times {\mathbf v}= (u_2 v_3 - u_3 v_2, u_3 v_1 - u_1 v_3, u_1 v_2 - u_2 v_1).$  Note here that if $ A = [ {\mathbf u}^t, {\mathbf v}^t, {\mathbf w}^t],$ then  $\displaystyle \det(A) = \begin{vmatrix}u_1 & v_1 & w_1 \\ u_2 & v_2 & w_2 \\ u... ...hbf w}\times {\mathbf u}) = {\mathbf w}\bullet ({\mathbf u}\times {\mathbf v}).$  Let $ P$ be the parallelopiped formed with $ (0,0,0)$ as a vertex and the vectors $ {\mathbf u}, {\mathbf v}, {\mathbf w}$ as adjacent vertices. Then observe that $ {\mathbf u}\times {\mathbf v}$ is a vector perpendicular to the plane that contains the parallelogram formed by the vectors $ {\mathbf u}$ and $ {\mathbf v}.$ So, to compute the volume of the parallelopiped $ P,$ we need to look at $ \cos (\theta),$ where $ \theta$ is the angle between the vector $ {\mathbf w}$ and the normal vector to the parallelogram formed by $ {\mathbf u}$ and $ {\mathbf v}.$ So,  $\displaystyle {\mbox{volume }} (P) = \vert {\mathbf w}\bullet ({\mathbf u}\times {\mathbf v}) \vert.$  Hence, $ \vert\det (A)\vert = {\mbox{volume }}\; ( P).$   1. Let $ {\mathbf u}_1, {\mathbf u}_2,\ldots, {\mathbf u}_n \in {\mathbb{R}}^{n \times 1}$ and let $ A = [ {\mathbf u}_1, {\mathbf u}_2, \ldots, {\mathbf u}_n]$ be an $ n \times n$ matrix. Then the following properties of $ \det (A)$ also hold for the volume of an $ n$ -dimensional parallelopiped formed with$ {\mathbf 0}\in {\mathbb{R}}^{n\times 1}$ as one vertex and the vectors $ {\mathbf u}_1, {\mathbf u}_2, \ldots, {\mathbf u}_n$ as adjacent vertices:    1. If $ {\mathbf u}_1 = (1,0,\ldots, 0)^t, {\mathbf u}_2 = (0,1,0,\ldots, 0)^t, \ldots, $ and $ {\mathbf u}_n = (0,\ldots, 0, 1)^t,$ then $ \det(A) = 1.$ Also, volume of a unit $ n$ -dimensional cube is $ 1.$    2. If we replace the vector $ {\mathbf u}_i$ by $ {\alpha}{\mathbf u}_i,$ for some $ {\alpha}\in {\mathbb{R}},$ then the determinant of the new matrix is $ {\alpha}\cdot \det(A)$ . This is also true for the volume, as the original volume gets multiplied by $ {\alpha}.$    3. If $ {\mathbf u}_1 = {\mathbf u}_i$ for some $ i, \; 2 \leq i \leq n,$ then the vectors $ {\mathbf u}_1, {\mathbf u}_2, \ldots, {\mathbf u}_n$ will give rise to an $ (n-1)$ -dimensional parallelopiped. So, this parallelopiped lies on an $ (n-1)$ -dimensional hyperplane. Thus, its $ n$ -dimensional volume will be zero. Also, $ \vert\det(A)\vert = \vert 0\vert = 0.$   In general, for any $ n \times n$ matrix $ A,$ it can be proved that $ \vert\det (A)\vert$ is indeed equal to the volume of the $ n$ -dimensional parallelopiped. |
| Adjoint of a Matrix **DEFINITION 2.6.9 (Minor, Cofactor of a Matrix)*****The number $ \det \left(A(i\vert j)\right)$ is called the $ (i,j)^{\mbox{th}}$ minor of $ A$ . We write $ A_{ij} = \det \left(A(i\vert j)\right).$ The $ (i,j)^{\mbox{th}}$ cofactor of $ A,$ denoted $ C_{ij},$ is the number $ (-1)^{i+j} A_{ij}.$***  **DEFINITION 2.6.10 (Adjoint of a Matrix)*****Let $ A$ be an $ n \times n$ matrix. The matrix $ B= [b_{ij}]$ with $ b_{ij} = C_{ji},$ for $ 1 \leq i, j \leq n$ is called the Adjoint of $ A,$ denoted $ Adj (A).$***  **EXAMPLE 2.6.11   *Let $ A = \begin{bmatrix}1 & 2 & 3 \\ 2 & 3 & 1 \\ 1 & 2 & 2 \end{bmatrix}.$ Then $ Adj(A) = \begin{bmatrix}4 & 2 & -7 \\ -3 & -1& 5 \\ 1 & 0 & -1 \end{bmatrix};$  as $ C_{11} = (-1)^{1+1}A_{11} = 4, C_{12} = (-1)^{1+2} A_{12} = -3, C_{13} = (-1)^{1+3} A_{13} = 1, $ and so on.***  **THEOREM 2.6.12   *Let $ A$ be an $ n \times n$ matrix. Then***   1. for $ 1 \leq i \leq n,$ $ \; \sum\limits_{j=1}^n a_{ij} \; C_{ij}    = \sum\limits_{j=1}^n a_{ij} (-1)^{i+j} \; A_{i j} = \det(A),$ 2. for $ i \neq \ell, \; \sum\limits_{j=1}^n a_{ij} \; C_{\ell j}    = \sum\limits_{j=1}^n a_{ij} (-1)^{\ell+j} \; A_{\ell j} = 0,$ and 3. $ \; A (Adj (A) ) = \det(A) I_n.$ Thus,  |  |  | | --- | --- | | $\displaystyle \det (A) \neq 0 \Rightarrow A^{-1} = \frac{1}{\det(A)} Adj (A).$ | (2.6.2) |  1. $ B= [b_{ij}]$ *Proof*. Let  be a square matrix with  * the $ \ell^{\mbox{th}}$ row of $ B$ as the $ i^{\mbox{th}}$ row of $ A,$ * the other rows of $ B$ are the same as that of $ A.$   By the construction of $ B,$ two rows ( $ i^{\mbox{th}}$ and $ \ell^{\mbox{th}}$ ) are equal. By Part [5](http://nptel.iitm.ac.in/courses/Webcourse-contents/IIT-KANPUR/mathematics-2/node28.html#lem:tworows:same) of Lemma [2.6.6](http://nptel.iitm.ac.in/courses/Webcourse-contents/IIT-KANPUR/mathematics-2/node28.html#lem:det:basic), $ \det (B) = 0.$ By construction again, $ \det\bigl(A(\ell\vert j)\bigr) = \det\bigl(B(\ell\vert j)\bigr)$ for $ 1 \leq j \leq n.$ Thus, by Remark [2.6.7](http://nptel.iitm.ac.in/courses/Webcourse-contents/IIT-KANPUR/mathematics-2/node28.html#rem:det:anyrow), we have   |  |  |  |  | | --- | --- | --- | --- | | $\displaystyle 0 = \det (B)$ | $\displaystyle =$ | $\displaystyle \sum_{j=1}^n (-1)^{\ell + j} b_{\ell j} \det\bigl(B(\ell\vert j)\bigr) = \sum_{j=1}^n (-1)^{\ell + j} a_{ij} \det\bigl(B(\ell\vert j)\bigr)$ |  | |  | $\displaystyle =$ | $\displaystyle \sum_{j=1}^n (-1)^{\ell + j} a_{ij} \det\bigl(A(\ell\vert j)\bigr) = \sum_{j=1}^n a_{ij} C_{\ell j}.$ |  |   Now,   |  |  |  |  | | --- | --- | --- | --- | | $\displaystyle \biggl(A\bigl( {\mbox{Adj}}(A) \bigr)\biggr)_{ij}$ | $\displaystyle =$ | $\displaystyle \sum_{k=1}^n a_{ik} \bigl( {\mbox{Adj}}(A)\bigr)_{kj} = \sum_{k=1}^n a_{ik} C_{jk}$ |  | |  | $\displaystyle =$ | $\displaystyle \left\{\begin{array}{cc} 0 & {\mbox{ if }} i \neq j \\ \det(A) & {\mbox{ if }} i = j \end{array}\right.$ |  |   Thus, $ \; A (Adj (A) ) = \det(A) I_n.$ Since, $ \det(A) \neq 0,$ $ \; A \displaystyle \frac{1}{\det(A)} Adj(A) = I_n.$ Therefore, $ A$ has a right inverse. Hence, by Theorem [2.5.9](http://nptel.iitm.ac.in/courses/Webcourse-contents/IIT-KANPUR/mathematics-2/node26.html#thm:left:right:inverse) $ A$ has an inverse and  $\displaystyle A^{-1} = \frac{1}{\det(A)} Adj (A).$  **EXAMPLE 2.6.13   *Let $ A= \begin{bmatrix}1 & -1 & 0 \\ 0 & 1 & 1\\ 1 & 2 & 1 \end{bmatrix}.$ Then***  $\displaystyle Adj (A) = \begin{bmatrix} -1 & 1 & -1 \\ 1 & 1 & -1\\ -1 &-3 & 1 \end{bmatrix}$  ***and $ \det (A) = -2.$ By Theorem***[***2.6.12***](http://nptel.iitm.ac.in/courses/Webcourse-contents/IIT-KANPUR/mathematics-2/node29.html#thm:adj)***.***[***3***](http://nptel.iitm.ac.in/courses/Webcourse-contents/IIT-KANPUR/mathematics-2/node29.html#inverse:adjoint)***, $ A^{-1} = \begin{bmatrix}1/2 & -1/2 & 1/2 \\ -1/2 & -1/2 & 1/2\\ 1/2 & 3/2 & -1/2 \end{bmatrix}.$***  The next corollary is an easy consequence of Theorem [2.6.12](http://nptel.iitm.ac.in/courses/Webcourse-contents/IIT-KANPUR/mathematics-2/node29.html#thm:adj) (recall Theorem [2.5.9](http://nptel.iitm.ac.in/courses/Webcourse-contents/IIT-KANPUR/mathematics-2/node26.html#thm:left:right:inverse)).  **COROLLARY 2.6.14   *If $ A$ is a non-singular matrix, then  $ \bigl(Adj(A) \bigr) A = \det(A) I_n \;\; $ and $ \;\; \sum\limits_{i=1}^n a_{ij} \; C_{ik} = \left\{\begin{array}{cc} \det (A) & {\mbox{ if }} j = k \\ 0 & {\mbox{if }} j \neq k \end{array}\right..$***  **THEOREM 2.6.15   *Let $ A$ and $ B$ be square matrices of order $ n.$ Then $ \;\det (A B) = \det (A) \det (B).$***  *Proof*. **Step 1.** Let $ \det (A) \neq 0.$  This means, $ A$ is invertible. Therefore, either $ A$ is an elementary matrix or is a product of elementary matrices (see Theorem [2.5.8](http://nptel.iitm.ac.in/courses/Webcourse-contents/IIT-KANPUR/mathematics-2/node26.html#thm:equi)). So, let $ E_1, E_2, \ldots, E_k$ be elementary matrices such that$ A = E_1 E_2 \cdots E_k.$ Then, by using Parts [1](http://nptel.iitm.ac.in/courses/Webcourse-contents/IIT-KANPUR/mathematics-2/node28.html#lem:interchange), [2](http://nptel.iitm.ac.in/courses/Webcourse-contents/IIT-KANPUR/mathematics-2/node28.html#lem:rkc) and [4](http://nptel.iitm.ac.in/courses/Webcourse-contents/IIT-KANPUR/mathematics-2/node28.html#lem:rijc) of Lemma [2.6.6](http://nptel.iitm.ac.in/courses/Webcourse-contents/IIT-KANPUR/mathematics-2/node28.html#lem:det:basic) repeatedly, we get   |  |  |  |  | | --- | --- | --- | --- | | $\displaystyle \det(AB)$ | $\displaystyle =$ | $\displaystyle \det (E_1 E_2 \cdots E_k B) =\det (E_1) \det( E_2 \cdots E_k B)$ |  | |  | $\displaystyle =$ | $\displaystyle \det(E_1) \det( E_2) \det(E_3 \cdots E_k B)$ |  | |  | $\displaystyle =$ | $\displaystyle \det(E_1 E_2) \det(E_3 \cdots E_k B)$ |  | |  | $\displaystyle =$ | $\displaystyle \vdots$ |  | |  | $\displaystyle =$ | $\displaystyle \det(E_1 E_2 \cdots E_k) \det( B)$ |  | |  | $\displaystyle =$ | $\displaystyle \det(A) \det(B).$ |  |   Thus, we get the required result in case $ A$ is non-singular.  **Step 2.** Suppose $ \det (A) = 0.$  Then $ A$ is not invertible. Hence, there exists an invertible matrix $ P$ such that $ P A = C,$ where $ C = \begin{bmatrix}C_1 \\ {\mathbf 0} \end{bmatrix}.$ So, $ A = P^{-1} C, $ and therefore   |  |  |  |  | | --- | --- | --- | --- | | $\displaystyle \det(AB)$ | $\displaystyle =$ | $\displaystyle \det ((P^{-1} C) B) = \det (P^{-1} (C B)) = \det \left( P^{-1} \begin{bmatrix}C_1 B \\ {\mathbf 0}\end{bmatrix}\right)$ |  | |  | $\displaystyle =$ | $\displaystyle \det( P^{-1} ) \cdot \det \left( \begin{bmatrix}C_1 B \\ {\mathbf 0} \end{bmatrix}\right) \;\; {\mbox{ as }} P^{-1} {\mbox{ is non-singular}}$ |  | |  | $\displaystyle =$ | $\displaystyle \det (P) \cdot 0 = 0 = 0 \cdot \det (B) = \det(A) \det(B).$ |  |   Thus, the proof of the theorem is complete. height6pt width 6pt depth 0pt  **COROLLARY 2.6.16   *Let $ A$ be a square matrix. Then $ A$ is non-singular if and only if $ A$ has an inverse.***  *Proof*. Suppose $ A$ is non-singular. Then $ \det(A) \neq 0$ and therefore, $ A^{-1} = \displaystyle\frac{1}{\det(A)} Adj(A).$ Thus, $ A$ has an inverse.  Suppose $ A$ has an inverse. Then there exists a matrix $ B$ such that $ A B = I = BA.$ Taking determinant of both sides, we get  $\displaystyle \det(A) \det (B) = \det(AB) = \det(I) = 1.$  This implies that $ \det (A) \neq 0.$ Thus, $ A$ is non-singular. height6pt width 6pt depth 0pt  **THEOREM 2.6.17   *Let $ A$ be a square matrix. Then $ \det (A) = \det(A^t).$***  *Proof*. If $ A$ is a non-singular Corollary [2.6.14](http://nptel.iitm.ac.in/courses/Webcourse-contents/IIT-KANPUR/mathematics-2/node29.html#cor:adj) gives $ \det (A) = \det(A^t).$  If $ A$ is singular, then $ \det (A) = 0.$ Hence, by Corollary [2.6.16](http://nptel.iitm.ac.in/courses/Webcourse-contents/IIT-KANPUR/mathematics-2/node29.html#cor:inverse), $ A$ doesn't have an inverse. Therefore, $ A^t$ also doesn't have an inverse (for if $ A^t$ has an inverse then $ A^{-1} = \bigl((A^t)^{-1}\bigr)^t).$ Thus again by Corollary [2.6.16](http://nptel.iitm.ac.in/courses/Webcourse-contents/IIT-KANPUR/mathematics-2/node29.html#cor:inverse), $ \det(A^t) = 0.$ Therefore, we again have $ \det(A) = 0 = \det(A^t).$  Hence, we have $ \det (A) = \det(A^t).$ |
| Cramer's Rule Recall the following:   * The linear system $ A {\mathbf x}= {\mathbf b}$ has a unique solution for every $ {\mathbf b}$ if and only if $ A^{-1}$ exists. * $ A$ has an inverse if and only if $ \det (A) \neq 0.$   Thus, $ A {\mathbf x}= {\mathbf b}$ has a unique solution FOR EVERY $ {\mathbf b}$ if and only if $ \det (A) \neq 0.$  The following theorem gives a direct method of finding the solution of the linear system $ A {\mathbf x}= {\mathbf b}$ when $ \det (A) \neq 0.$  **THEOREM 2.6.18 (Cramer's Rule)   *Let $ A {\mathbf x}= {\mathbf b}$ be a linear system with $ n$ equations in $ n$ unknowns. If $ \det(A) \neq 0,$ then the unique solution to this system is***  $\displaystyle x_j = \frac{ \det(A_j)}{\det(A)}, \;\; {\mbox{ for }} j=1, 2, \ldots, n,$  ***where $ A_j$ is the matrix obtained from $ A$ by replacing the $ j$ th column of $ A$ by the column vector $ {\mathbf b}.$***  *Proof*. Since $ \det(A) \neq 0,\; $ $ A^{-1} = \displaystyle\frac{1}{\det(A)} Adj(A).$ Thus, the linear system $ A {\mathbf x}= {\mathbf b}$ has the solution $ {\mathbf x}= \displaystyle\frac{1}{\det(A)} Adj(A) {\mathbf b}.$ Hence, $ x_j,$ the $ j$ th coordinate of $ {\mathbf x}$ is given by  $\displaystyle x_j=\frac{b_1 C_{1j} + b_2 C_{2j} + \cdots + b_n C_{nj}}{\det(A)} = \frac{\det (A_j)}{\det(A)}.$  height6pt width 6pt depth 0pt  The theorem implies that  $\displaystyle x_1 = \frac{1}{\det (A)} \; \begin{vmatrix}b_1 & a_{12} & \cdots ... ...ts & \vdots & \ddots & \vdots \\ b_n & a_{n2} & \cdots & a_{nn} \end{vmatrix}, $  and in general  $\displaystyle x_j = \frac{1}{\det (A)} \; \begin{vmatrix}a_{11} & \cdots & a_{1... ... a_{1n} & \cdots & a_{n j-1} & b_n & a_{n j+1} & \cdots & a_{nn} \end{vmatrix} $  for $ j=2, 3, \ldots, n.$  **EXAMPLE 2.6.19   *Suppose that $ A = \begin{bmatrix}1 & 2 & 3 \\ 2 & 3 & 1 \\ 1 & 2 & 2 \end{bmatrix}$ and $ {\mathbf b}= \begin{bmatrix}1\\ 1 \\ 1\end{bmatrix}.$ Use Cramer's rule to find a vector $ {\mathbf x}$ such that $ A {\mathbf x}= {\mathbf b}.$  Solution: Checkthat $ \det(A) = 1.$ Therefore $ x_1 = \begin{vmatrix}1 &2 & 3 \\ 1 & 3 & 1 \\ 1 & 2 & 2\end{vmatrix}= -1,$  $ x_2 = \begin{vmatrix}1 & 1 & 3 \\ 2 & 1 & 1 \\ 1 & 1 & 2\end{vmatrix}= 1,$ and $ x_3 = \begin{vmatrix}1 &2 & 1 \\ 2 & 3 & 1 \\ 1 & 2 & 1 \end{vmatrix}=0.$ That is, $ {\mathbf x}^t = (-1,1,0).$*** |
| Eigenvalues, Eigenvectors and Diagonalisation **Introduction and Definitions**  In this chapter, the linear transformations are from a given finite dimensional vector space $ V$ to itself. Observe that in this case, the matrix of the linear transformation is a square matrix. So, in this chapter, all the matrices are square matrices and *a vector* $ {\mathbf x}$ means $ {\mathbf x}=(x_1,x_2,\ldots,x_n)^t$ for some positive integer $ n.$  **EXAMPLE 6.1.1   *Let $ A$ be a real symmetric matrix. Consider the following problem:***  $\displaystyle {\mbox{ Maximize (Minimize)}} \;\;{\mathbf x}^t A {\mathbf x}{\mb... ... }} {\mathbf x} \in {\mathbb{R}}^n {\mbox{ and }} {\mathbf x}^t {\mathbf x}= 1.$  ***To solve this, consider the Lagrangian***  $\displaystyle L({\mathbf x}, \lambda) = {\mathbf x}^t A {\mathbf x}- \lambda ( ... ...\sum_{i=1}^n\sum_{j=1}^n a_{ij} x_i x_j - \lambda (\sum_{i=1}^n x_i^2 \;\; -1).$  ***Partially differentiating $ L({\mathbf x}, \lambda)$ with respect to $ x_i$ for $ 1 \leq i \leq n,$ we get***  $\displaystyle \frac{\partial L}{\partial x_1} = 2 a_{11} x_1 + 2 a_{12} x_2 + \cdots + 2 a_{1n} x_n - 2 \lambda x_1, $  $\displaystyle \frac{\partial L}{\partial x_2} = 2 a_{21} x_1 + 2 a_{22} x_2 + \cdots + 2 a_{2n} x_n - 2 \lambda x_2, $  ***and so on, till***  $\displaystyle \frac{\partial L}{\partial x_n} = 2 a_{n1} x_1 + 2 a_{n2} x_2 + \cdots + 2 a_{nn} x_n - 2 \lambda x_n. $  ***Therefore, to get the points of extrema, we solve for***  $\displaystyle (0,0,\ldots,0)^t = (\frac{\partial L}{\partial x_1}, \frac{\parti... ...rac{\partial L}{\partial {\mathbf x}} = 2 (A {\mathbf x}- \lambda {\mathbf x}).$  ***We therefore need to find a $ \lambda \in {\mathbb{R}}$ and $ {\mathbf 0}\neq {\mathbf x}\in {\mathbb{R}}^n$ such that $ A {\mathbf x}= \lambda {\mathbf x}$ for the extremal problem.***  **EXAMPLE 6.1.2   *Consider a system of $ n$ ordinary differential equations of the form***   |  |  | | --- | --- | | $\displaystyle \frac{d \; {\mathbf y}(t)}{d t} = A {\mathbf y}, \; t \geq 0;$ | (6.1.1) |     ***where $ A$ is a real $ n \times n$ matrix and $ {\mathbf y}$ is a column vector.  To get a solution, let us assume that***   |  |  | | --- | --- | | $\displaystyle {\mathbf y}(t) = {\mathbf c}e^{ {\lambda}t}$ | (6.1.2) |     ***is a solution of (***[***6.1.1***](http://nptel.iitm.ac.in/courses/Webcourse-contents/IIT-KANPUR/mathematics-2/node53.html#sys:diff)***) and look into what $ {\lambda}$ and $ {\mathbf c}$ has to satisfy, i.e., we are investigating for a necessary condition on $ {\lambda}$ and $ {\mathbf c}$ so that (***[***6.1.2***](http://nptel.iitm.ac.in/courses/Webcourse-contents/IIT-KANPUR/mathematics-2/node53.html#eqn:diff)***) is a solution of (***[***6.1.1***](http://nptel.iitm.ac.in/courses/Webcourse-contents/IIT-KANPUR/mathematics-2/node53.html#sys:diff)***). Note here that (***[***6.1.1***](http://nptel.iitm.ac.in/courses/Webcourse-contents/IIT-KANPUR/mathematics-2/node53.html#sys:diff)***) has the zero solution, namely $ y(t) \equiv 0$ and so we are looking for a non-zero $ {\mathbf c}.$ Differentiating (***[***6.1.2***](http://nptel.iitm.ac.in/courses/Webcourse-contents/IIT-KANPUR/mathematics-2/node53.html#eqn:diff)***) with respect to $ t$ and substituting in (***[***6.1.1***](http://nptel.iitm.ac.in/courses/Webcourse-contents/IIT-KANPUR/mathematics-2/node53.html#sys:diff)***), leads to***   |  |  | | --- | --- | | $\displaystyle {\lambda}e^{{\lambda}t} {\mathbf c}= A e^{{\lambda}t} {\mathbf c}\;\; {\mbox{or equivalently}} \;\; (A - {\lambda}I) {\mathbf c}= {\mathbf 0}.$ | (6.1.3) |     ***So, (***[***6.1.2***](http://nptel.iitm.ac.in/courses/Webcourse-contents/IIT-KANPUR/mathematics-2/node53.html#eqn:diff)***) is a solution of the given system of differential equations if and only if $ {\lambda}$ and $ {\mathbf c}$ satisfy (***[***6.1.3***](http://nptel.iitm.ac.in/courses/Webcourse-contents/IIT-KANPUR/mathematics-2/node53.html#eqn:diff:1)***). That is, given an $ n \times n$ matrix $ A,$ we are this lead to find a pair $ ({\lambda}, {\mathbf c})$ such that$ {\mathbf c}\neq {\mathbf 0}$ and (***[***6.1.3***](http://nptel.iitm.ac.in/courses/Webcourse-contents/IIT-KANPUR/mathematics-2/node53.html#eqn:diff:1)***) is satisfied.***  Let $ A$ be a matrix of order $ n.$ In general, we ask the question:  For what values of $ {\lambda}\in {\mathbb{F}},$ there exist a non-zero vector $ {\mathbf x}\in {\mathbb{F}}^n$ such that   |  |  | | --- | --- | | $\displaystyle A {\mathbf x}= \lambda {\mathbf x}?$ | (6.1.4) |     Here, $ {\mathbb{F}}^n$ stands for either the vector space $ {\mathbb{R}}^n$ over $ {\mathbb{R}}$ or $ {\mathbb{C}}^n$ over $ {\mathbb{C}}.$ Equation ([6.1.4](http://nptel.iitm.ac.in/courses/Webcourse-contents/IIT-KANPUR/mathematics-2/node53.html#eqn:e)) is equivalent to the equation  $\displaystyle (A - \lambda I) {\mathbf x}= {\mathbf 0}.$  By Theorem [2.5.1](http://nptel.iitm.ac.in/courses/Webcourse-contents/IIT-KANPUR/mathematics-2/node25.html#thm:solns), this system of linear equations has a non-zero solution, if  $\displaystyle {\mbox{rank }} (A - \lambda I) < n, \;\; {\mbox{ or equivalently }} \;\; \det(A - {\lambda}I) = 0.$  So, to solve ([6.1.4](http://nptel.iitm.ac.in/courses/Webcourse-contents/IIT-KANPUR/mathematics-2/node53.html#eqn:e)), we are forced to choose those values of $ \lambda \in {\mathbb{F}}$ for which $ \det (A - \lambda I) = 0.$ Observe that $ \det(A - {\lambda}I)$ is a polynomial in $ {\lambda}$ of degree $ n.$ We are therefore lead to the following definition.  **DEFINITION 6.1.3 (Characteristic Polynomial)*****Let $ A$ be a matrix of order $ n.$ The polynomial $ \det(A - {\lambda}I)$ is called the characteristic polynomial of $ A$ and is denoted by $ p({\lambda}).$ The equation $ p({\lambda}) = 0$ is called the characteristic equation of $ A.$ If $ {\lambda}\in {\mathbb{F}}$ is a solution of the characteristic equation $ p({\lambda}) = 0,$ then $ {\lambda}$ is called a characteristic value of $ A.$***  ***Some books use the term EIGENVALUE in place of characteristic value.***  **THEOREM 6.1.4   *Let $ A= [a_{ij}]; \; a_{ij} \in {\mathbb{F}}, \; {\mbox{ for }} 1 \leq i, j \leq n.$ Suppose $ \lambda = \lambda_0 \in {\mathbb{F}}$ is a root of the characteristic equation. Then there exists a non-zero $ {\mathbf v}\in {{\mathbb{F}}}^n$ such that $ A {\mathbf v}= \lambda_0 {\mathbf v}.$***  *Proof*. Since $ \lambda_0$ is a root of the characteristic equation, $ \det (A - \lambda_0 I) = 0.$ This shows that the matrix $ A - \lambda_0 I$ is singular and therefore by Theorem [2.5.1](http://nptel.iitm.ac.in/courses/Webcourse-contents/IIT-KANPUR/mathematics-2/node25.html#thm:solns) the linear system  $\displaystyle (A - \lambda_0 I_n) {\mathbf x}= {\mathbf 0}$  has a non-zero solution. height6pt width 6pt depth 0pt  **Remark 6.1.5**   *Observe that the linear system $ A {\mathbf x}= {\lambda}{\mathbf x}$ has a solution $ {\mathbf x}={\mathbf 0}$ for every $ {\lambda}\in {\mathbb{F}}.$ So, we consider only those $ {\mathbf x}\in {\mathbb{F}}^n$ that are non-zero and are solutions of the linear system$ A {\mathbf x}= {\lambda}{\mathbf x}.$*  **DEFINITION 6.1.6 (Eigenvalue and Eigenvector)*****If the linear system $ A {\mathbf x}= {\lambda}{\mathbf x}$ has a non-zero solution $ {\mathbf x}\in {\mathbb{F}}^n$ for some $ {\lambda}\in {\mathbb{F}},$ then***   1. $ \lambda \in {\mathbb{F}}$ is called an eigenvalue of $ A,$ 2. $ {\mathbf 0}\neq {\mathbf x}\in {\mathbb{F}}^n$ is called an eigenvector corresponding to the eigenvalue $ {\lambda}$ of $ A,$ and 3. the tuple $ (\lambda, {\mathbf x})$ is called an eigenpair.   **Remark 6.1.7**   *To understand the difference between a characteristic value and an eigenvalue, we give the following example.*  *Consider the matrix $ A = \begin{bmatrix}0 & 1 \\ -1 & 0 \end{bmatrix}.$ Then the characteristic polynomial of $ A$ is*  $\displaystyle p({\lambda}) = {\lambda}^2 + 1.$  *Given the matrix $ A,$ recall the linear transformation $ T_A: {\mathbb{F}}^2 {\longrightarrow}{\mathbb{F}}^2$ defined by*  $\displaystyle T_A({\mathbf x}) = A {\mathbf x}\;\; {\mbox{ for every }} \;\; {\mathbf x}\in {\mathbb{F}}^2.$   1. If $ \; {\mathbb{F}}= {\mathbb{C}},$ that is, if $ A$ is considered a COMPLEX matrix, then the roots of $ p({\lambda}) = 0$ in $ {\mathbb{C}}$ are $ \pm i.$ So, $ A$ has $ (i, (1,i)^t)$ and $ (-i, (i, 1)^t)$ as eigenpairs. 2. If $ \; {\mathbb{F}}= {\mathbb{R}},$ that is, if $ A$ is considered a REAL matrix, then $ p({\lambda}) = 0$ has no solution in $ {\mathbb{R}}.$ Therefore, if $ {\mathbb{F}}= {\mathbb{R}},$ then $ A$ has no eigenvalue but it has $ \pm i$ as characteristic values.   **Remark 6.1.8**   *Note that if $ (\lambda, {\mathbf x})$ is an eigenpair for an $ n \times n$ matrix $ A$ then for any non-zero $ c \in {\mathbb{F}},\;c \neq 0, \; (\lambda, c {\mathbf x})$ is also an eigenpair for $ A.$ Similarly, if $ {\mathbf x}_1, {\mathbf x}_2, \ldots, {\mathbf x}_r$ are eigenvectors of $ A$ corresponding to the eigenvalue $ {\lambda},$ then for any non-zero $ (c_1, c_2, \ldots, c_r) \in {\mathbb{F}}^r,$ it is easily seen that if $ \sum\limits_{i=1}^r c_i {\mathbf x}_i \neq {\mathbf 0}$ , then $ \sum\limits_{i=1}^r c_i {\mathbf x}_i$ is also an eigenvector of $ A$ corresponding to the eigenvalue $ {\lambda}.$ Hence, when we talk of eigenvectors corresponding to an eigenvalue $ \lambda,$ we mean LINEARLY INDEPENDENT EIGENVECTORS.*  *Suppose $ {\lambda}_0 \in {\mathbb{F}}$ is a root of the characteristic equation $ \det(A - {\lambda}_0 I) = 0.$ Then $ A - {\lambda}_0 I$ is singular and $ {\mbox{rank }}(A - {\lambda}_0 I) < n.$ Suppose $ {\mbox{rank }}(A - {\lambda}_0 I)=r < n.$ Then by Corollary*[*4.3.9*](http://nptel.iitm.ac.in/courses/Webcourse-contents/IIT-KANPUR/mathematics-2/node45.html#cor:rank:nullity)*, the linear system $ (A - {\lambda}_0 I) {\mathbf x}= {\mathbf 0}$ has $ n-r$ linearly independent solutions. That is, $ A$ has $ n-r$ linearly independent eigenvectors corresponding to the eigenvalue $ {\lambda}_0$ whenever$ {\mbox{rank }}(A - {\lambda}_0 I)=r < n.$*  **EXAMPLE 6.1.9**   1. Let $ A = {\mbox{diag}}(d_1, d_2, \ldots, d_n)$ with $ d_i \in {\mathbb{R}}$ for $ 1 \leq i \leq n.$ Then $ p({\lambda}) = \prod_{i=1}^n ({\lambda}- d_i)$ is the characteristic equation. So, the eigenpairs are   $\displaystyle (d_1, (1,0,\ldots, 0)^t), (d_2, (0,1,0,\ldots,0)^t), \ldots, (d_n, (0,\ldots, 0,1)^t).$   1. Let $ A = \begin{bmatrix}1 & 1    \\ 0 & 1 \end{bmatrix}.$ Then $ \det (A - \lambda I_2) =    (1-\lambda)^2.$ Hence, the characteristic equation has roots $ 1,    1.$ That is $ 1$ is a repeated eigenvalue. Now check that the equation $ (A - I_2) {\mathbf x}= {\mathbf 0}$ for$ {\mathbf x}= (x_1, x_2)^{t}$ is equivalent to the equation $ x_2 = 0.$ And this has the solution $ {\mathbf x}= (x_1, 0)^{t}.$ Hence, from the above remark, $ (1, 0)^{t}$ is a representative for the eigenvector. Therefore, HERE WE HAVE TWO EIGENVALUES MATHEND000# BUT ONLY ONE EIGENVECTOR. 2. Let $ A    = \begin{bmatrix}1 & 0 \\ 0 & 1 \end{bmatrix}.$ Then $ \det (A - \lambda I_2) =    (1-\lambda)^2.$ The characteristic equation has roots $ 1,    1.$ Here, the matrix that we have is $ I_2$ and we know that $ I_2 {\mathbf x}= {\mathbf x}$ for every $ {\mathbf x}^{t} \in {\mathbb{R}}^2$ and we canCHOOSE ANY TWO LINEARLY INDEPENDENT VECTORS $ {\mathbf x}^{t},    {\mathbf y}^{t} $ from $ {\mathbb{R}}^2$ to get $ (1, {\mathbf x})$ and $ (1, {\mathbf y})$ as the two eigenpairs.   In general, if $ {\mathbf x}_1, {\mathbf x}_2, \ldots, {\mathbf x}_n$ are linearly independent vectors in $ {\mathbb{R}}^n,$ then $ (1, {\mathbf x}_1), \; (1, {\mathbf x}_2), \; \ldots, (1, {\mathbf x}_n)$ are eigenpairs for the identity matrix, $ I_n.$   1. Let $ A = \begin{bmatrix}1 & 2 \\ 2 & 1 \end{bmatrix}.$ Then $ \det (A - \lambda I_2) = (\lambda- 3)(\lambda + 1).$ The characteristic equation has roots $ 3, -1.$ Now check that the eigenpairs are $ (3, (1,1)^{t}), $ and $ (-1, (1, -1)^{t}).$ In this case, we haveTWO DISTINCT EIGENVALUES AND THE CORRESPONDING EIGENVECTORS ARE ALSO LINEARLY INDEPENDENT. The reader is required to prove the linear independence of the two eigenvectors. 2. Let $ A = \begin{bmatrix}1 & -1 \\ 1 & 1 \end{bmatrix}.$ Then $ \det (A - \lambda I_2) = \lambda^2 - 2 \lambda + 2.$ The characteristic equation has roots $ 1 + i, 1 - i.$ Hence, over $ {\mathbb{R}},$ the matrix $ A$ has no eigenvalue. Over $ {\mathbb{C}},$ the reader is required to show that the eigenpairs are $ (1+i, (i,1)^t)$ and $ (1-i, (1,i)^t).$   **EXERCISE 6.1.10**   1. Find the eigenvalues of a triangular matrix. 2. Find eigenpairs over $ {\mathbb{C}},$ for each of the following matrices:  $ \begin{bmatrix}1 & 0 \\ 0 & 0 \end{bmatrix}, \hspace{0.1in}    \begin{bmatrix}1 &...    ...bmatrix}\cos \theta & - \sin \theta \\ \sin \theta &    \cos \theta \end{bmatrix},$ and $ \;\;\begin{bmatrix}\cos \theta & \sin \theta \\ \sin \theta &    - \cos \theta \end{bmatrix}.$ 3. Let $ A$ and $ B$ be similar matrices.    1. Then prove that $ A$ and $ B$ have the same set of eigenvalues.    2. Let $ ({\lambda}, {\mathbf x})$ be an eigenpair for $ A$ and $ ({\lambda}, {\mathbf y})$ be an eigenpair for $ B.$ What is the relationship between the vectors $ {\mathbf x}$ and $ {\mathbf y}$ ?   [*Hint: Recall that if the matrices $ A$ and $ B$ are similar, then there exists a non-singular matrix $ P$ such that $ B = P A P^{-1}.$*]   1. Let $ A=(a_{ij})$ be an $ n \times n$ matrix. Suppose that for all $ i, \; 1 \leq i \leq n, \; \sum\limits_{j=1}^n    a_{ij} = a.$ Then prove that $ a$ is an eigenvalue of $ A.$ What is the corresponding eigenvector? 2. Prove that the matrices $ A$ and $ A^t$ have the same set of eigenvalues. Construct a $ 2 \times 2$ matrix $ A$ such that the eigenvectors of $ A$ and $ A^t$ are different. 3. Let $ A$ be a matrix such that $ A^2 = A$ ($ A$ is called an idempotent matrix). Then prove that its eigenvalues are either 0 or $ 1$ or both. 4. Let $ A$ be a matrix such that $ A^k = {\mathbf 0}$ ($ A$ is called a nilpotent matrix) for some positive integer $ k \ge 1$ . Then prove that its eigenvalues are all 0 .   **THEOREM 6.1.11   *Let $ A=[a_{ij}]$ be an $ n \times n$ matrix with eigenvalues $ \lambda_1, \lambda_2, \ldots, \lambda_n,$ not necessarily distinct. Then $ \det (A) = \prod\limits_{i=1}^n \lambda_i$ and $ {\mbox{ tr}}(A) = \sum\limits_{i=1}^n a_{ii} = \sum\limits_{i=1}^n \lambda_i.$***  *Proof*. Since $ \lambda_1, \lambda_2, \ldots, \lambda_n$ are the $ n$ eigenvalues of $ A,$ by definition,   |  |  | | --- | --- | | $\displaystyle \det (A - \lambda I_n) = p({\lambda}) = (-1)^n (\lambda - \lambda_1) (\lambda - \lambda_2) \cdots (\lambda - \lambda_n).$ | (6.1.5) |     ([6.1.5](http://nptel.iitm.ac.in/courses/Webcourse-contents/IIT-KANPUR/mathematics-2/node53.html#factors)) is an identity in $ {\lambda}$ as polynomials. Therefore, by substituting $ \lambda = 0$ in ([6.1.5](http://nptel.iitm.ac.in/courses/Webcourse-contents/IIT-KANPUR/mathematics-2/node53.html#factors)), we get  $\displaystyle \det (A) = (-1)^n (-1)^n \prod_{i=1}^n \lambda_i = \prod_{i=1}^n \lambda_i.$  Also,   |  |  |  |  | | --- | --- | --- | --- | | $\displaystyle \det (A - \lambda I_n)$ | $\displaystyle =$ | $\displaystyle \begin{bmatrix}a_{11} - \lambda & a_{12} & \cdots & a_{1n} \\ a_{... ... & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} - \lambda \end{bmatrix}$ | (6.1.6) | |  | $\displaystyle =$ | $\displaystyle a_0 - \lambda a_1 + \lambda^2 a_2 + \cdots$ |  | |  |  | $\displaystyle +(-1)^{n-1} {\lambda}^{n-1} a_{n-1} + (-1)^n \lambda^n$ | (6.1.7) |   for some $ a_0, a_1, \ldots, a_{n-1} \in {\mathbb{F}}.$ Note that $ a_{n-1},$ the coefficient of $ (-1)^{n-1} {\lambda}^{n-1},$ comes from the product  $\displaystyle (a_{11} - {\lambda})(a_{22} - {\lambda}) \cdots (a_{nn} - {\lambda}).$  So, $ a_{n-1} = \sum\limits_{i=1}^n a_{ii} = {\mbox{tr}}(A)$ by definition of trace.  But , from ([6.1.5](http://nptel.iitm.ac.in/courses/Webcourse-contents/IIT-KANPUR/mathematics-2/node53.html#factors)) and ([6.1.7](http://nptel.iitm.ac.in/courses/Webcourse-contents/IIT-KANPUR/mathematics-2/node53.html#polynomial)), we get   |  |  |  |  | | --- | --- | --- | --- | |  |  | $\displaystyle a_0 - \lambda a_1 + \lambda^2 a_2 + \cdots + (-1)^{n-1} {\lambda}^{n-1} a_{n-1} + (-1)^n \lambda^n$ |  | |  | $\displaystyle =$ | $\displaystyle (-1)^n (\lambda - \lambda_1) (\lambda - \lambda_2) \cdots (\lambda - \lambda_n).$ | (6.1.8) |   Therefore, comparing the coefficient of $ (-1)^{n-1} \lambda^{n-1},$ we have  $\displaystyle {\mbox{tr}}(A)= a_{n-1} = (-1) \{ (-1)\sum\limits_{i=1}^n \lambda_i \} = \sum\limits_{i=1}^n \lambda_i.$  Hence, we get the required result. height6pt width 6pt depth 0pt  **EXERCISE 6.1.12**   1. Let $ A$ be a skew symmetric matrix of order $ 2n + 1.$ Then prove that 0 is an eigenvalue of $ A.$ 2. Let $ A$ be a $ 3 \times 3$ orthogonal matrix $ (A A^t = I)$ .If $ \det(A) = 1$ , then prove that there exists a non-zero vector $ {\mathbf v}\in {\mathbb{R}}^3$ such that $ A {\mathbf v}= {\mathbf v}.$   Let $ A$ be an $ n \times n$ matrix. Then in the proof of the above theorem, we observed that the characteristic equation $ \det(A - {\lambda}I) = 0$ is a polynomial equation of degree $ n$ in $ {\lambda}.$ Also, for some numbers$ a_0, a_1, \ldots, a_{n-1} \in {\mathbb{F}},$ it has the form  $\displaystyle {\lambda}^n + a_{n-1} {\lambda}^{n-1} + a_{n-2} {\lambda}^2 + \cdots a_1 {\lambda}+ a_0 = 0.$  Note that, in the expression $ \det(A - {\lambda}I)= 0, \;\; {\lambda}$ is an element of $ {\mathbb{F}}.$ Thus, we can only substitute $ {\lambda}$ by elements of $ {\mathbb{F}}.$  It turns out that the expression  $\displaystyle A^n + a_{n-1} A^{n-1} + a_{n-2} A^2 + \cdots a_1 A + a_0 I = {\mathbf 0}$  holds true as a matrix identity. This is a celebrated theorem called the Cayley Hamilton Theorem. We state this theorem without proof and give some implications.  **THEOREM 6.1.13 (Cayley Hamilton Theorem)*****Let $ A$ be a square matrix of order $ n.$ Then $ A$ satisfies its characteristic equation. That is,***  $\displaystyle A^n + a_{n-1} A^{n-1} + a_{n-2} A^2 + \cdots a_1 A + a_0 I = {\mathbf 0}$  ***holds true as a matrix identity.***  Some of the implications of Cayley Hamilton Theorem are as follows.  **Remark 6.1.14**   1. Let $ A = \left[\begin{array}{cc}0&1 \\ 0 & 0    \end{array}\right].$ Then its characteristic polynomial is $ p({\lambda}) = {\lambda}^2.$ Also, for the function, $ f(x) = x,$ $ \;f(0) = 0,$ and $ \; f(A) = A \neq {\mathbf 0}.$ This shows that the condition $ f({\lambda}) = 0$ for each eigenvalue $ {\lambda}$ of $ A$ does not imply that $ f(A) = {\mathbf 0}.$ 2. Suppose we are given a square matrix $ A$ of order $ n$ and we are interested in calculating $ A^{\ell}$ where $ \ell$ is large compared to $ n.$ Then we can use the division algorithm to find numbers$ {\alpha}_0, {\alpha}_1, \ldots,    {\alpha}_{n-1}$ and a polynomial $ f({\lambda})$ such that  |  |  |  |  | | --- | --- | --- | --- | | $\displaystyle {\lambda}^{\ell}$ | $\displaystyle =$ | $\displaystyle f({\lambda}) \bigl( {\lambda}^n + a_{n-1} {\lambda}^{n-1} + a_{n-2} {\lambda}^2 + \cdots a_1 {\lambda}+ a_0 \bigr)$ |  | |  |  | $\displaystyle + {\alpha}_0 + {\lambda}{\alpha}_1 + \cdots + {\lambda}^{n-1} {\alpha}_{n-1}.$ |  |  1. Hence, by the Cayley Hamilton Theorem, 2. $\displaystyle A^{\ell} = {\alpha}_0 I + {\alpha}_1 A + \cdots + {\alpha}_{n-1} A^{n-1}.$ 3. That is, we just need to compute the powers of $ A$ till $ n-1.$ 4. In the language of graph theory, it says the following:  ``Let $ G$ be a graph on $ n$ vertices. Suppose there is no path of length $ n-1$ or less from a vertex $ v$ to a vertex $ u$ of $ G.$ Then there is no path from $ v$ to $ u$ of any length. That is, the graph $ G$ is disconnected and $ v$ and $ u$are in different components." 5. Let $ A$ be a non-singular matrix of order $ n.$ Then note that $ a_n = \det (A) \neq 0$ and   $\displaystyle A^{-1} = \frac{-1}{a_n} [ A^{n-1} + a_{n-1} A^{n-2} + \cdots + a_1 I ].$  This matrix identity can be used to calculate the inverse.  Note that the vector $ A^{-1}$ (as an element of the vector space of all $ n \times n$ matrices) is a linear combination of the vectors $ I, A, \ldots, A^{n-1}.$  **EXERCISE 6.1.15   *Find inverse of the following matrices by using the Cayley Hamilton Theorem***  $\displaystyle i)\;\; \begin{bmatrix}2&3&4 \\ 5&6&7\\ 1&1&2 \end{bmatrix}\;\;\;\... ...\; iii) \begin{bmatrix} 1 & -2 & -1 \\ -2 & 1 & -1 \\ 0 & -1 & 2 \end{bmatrix}.$  **THEOREM 6.1.16   *If $ {\lambda}_1, {\lambda}_2, \ldots, {\lambda}_k$ are distinct eigenvalues of a matrix $ A$ with corresponding eigenvectors $ {\mathbf x}_1, {\mathbf x}_2, \ldots, {\mathbf x}_k,$ then the set $ \{{\mathbf x}_1, {\mathbf x}_2, \ldots, {\mathbf x}_k\}$ is linearly independent.***  *Proof*. The proof is by induction on the number $ m$ of eigenvalues. The result is obviously true if $ m = 1$ as the corresponding eigenvector is non-zero and we know that any set containing exactly one non-zero vector is linearly independent.  Let the result be true for $ m, \; 1 \leq m < k.$ We prove the result for $ m+1.$ We consider the equation   |  |  | | --- | --- | | $\displaystyle c_1 x_1 + c_2 x_2 + \cdots + c_{m+1} x_{m+1} = {\mathbf 0}$ | (6.1.9) |     for the unknowns $ c_1, c_2, \ldots, c_{m+1}.$ We have   |  |  |  |  | | --- | --- | --- | --- | | $\displaystyle {\mathbf 0}= A {\mathbf 0}$ | $\displaystyle =$ | $\displaystyle A ( c_1 x_1 + c_2 x_2 + \cdots + c_{m+1} x_{m+1} )$ |  | |  | $\displaystyle =$ | $\displaystyle c_1 A x_1 + c_2 A x_2 + \cdots + c_{m+1} A x_{m+1}$ |  | |  | $\displaystyle =$ | $\displaystyle c_1 \lambda_1 x_1 + c_2 \lambda_2 x_2 + \cdots + c_{m+1} \lambda_{m+1} x_{m+1}.$ | (6.1.10) |     From Equations ([6.1.9](http://nptel.iitm.ac.in/courses/Webcourse-contents/IIT-KANPUR/mathematics-2/node53.html#eigen:li)) and ([6.1.10](http://nptel.iitm.ac.in/courses/Webcourse-contents/IIT-KANPUR/mathematics-2/node53.html#eigen:li2)), we get  $\displaystyle c_2 (\lambda_2 - \lambda_1 ) {\mathbf x}_2 + c_3 (\lambda_3 - \la... ...+ \cdots + c_{m+1} (\lambda_{m+1} - \lambda_1) {\mathbf x}_{m+1} = {\mathbf 0}.$  This is an equation in $ m$ eigenvectors. So, by the induction hypothesis, we have  $\displaystyle c_i (\lambda_i - \lambda_1) = 0 \;\; {\mbox{ for }} \;\; 2 \leq i \leq m+1.$  But the eigenvalues are distinct implies $ \lambda_i - {\lambda}_1 \neq 0$ for $ 2 \leq i \leq m+1.$ We therefore get $ c_i = 0$ for $ 2 \leq i \leq m+1.$ Also, $ {\mathbf x}_1 \neq {\mathbf 0}$ and therefore ([6.1.9](http://nptel.iitm.ac.in/courses/Webcourse-contents/IIT-KANPUR/mathematics-2/node53.html#eigen:li)) gives $ c_1 = 0.$  Thus, we have the required result. height6pt width 6pt depth 0pt  We are thus lead to the following important corollary.  **COROLLARY 6.1.17   *The eigenvectors corresponding to distinct eigenvalues of an $ n \times n$ matrix $ A$ are linearly independent.***  **EXERCISE 6.1.18**   1. For an $ n \times n$ matrix $ A,$ prove the following.    1. $ A$ and $ A^{t}$ have the same set of eigenvalues.    2. If $ {\lambda}$ is an eigenvalue of an invertible matrix $ A$ then $ \displaystyle\frac{1}{{\lambda}}$ is an eigenvalue of $ A^{-1}.$    3. If $ {\lambda}$ is an eigenvalue of $ A$ then $ {\lambda}^{k}$ is an eigenvalue of $ A^k$ for any positive integer $ k.$    4. If $ A$ and $ B$ are $ n \times n$ matrices with $ A$ nonsingular then $ B A^{-1}$ and $ A^{-1} B$ have the same set of eigenvalues.   In each case, what can you say about the eigenvectors?   1. Let $ A$ and $ B$ be $ 2 \times 2$ matrices for which $ \det(A) =    \det(B)$ and $ {\mbox{tr}}(A) = {\mbox{tr}}(B).$    1. Do $ A$ and $ B$ have the same set of eigenvalues?    2. Give examples to show that the matrices $ A$ and $ B$ need not be similar. 2. Let $ ({\lambda}_1, {\mathbf u})$ be an eigenpair for a matrix $ A$ and let $ ({\lambda}_2, {\mathbf u})$ be an eigenpair for another matrix $ B.$    1. Then prove that $ ({\lambda}_1+{\lambda}_2, {\mathbf u})$ is an eigenpair for the matrix $ A+B.$    2. Give an example to show that if $ {\lambda}_1, {\lambda}_2$ are respectively the eigenvalues of $ A$ and $ B,$ then $ {\lambda}_1+{\lambda}_2$ need not be an eigenvalue of $ A+B.$ 3. Let $ {\lambda}_i, 1 \leq i \leq n$ be distinct non-zero eigenvalues of an $ n \times n$ matrix $ A.$ Let $ {\mathbf u}_i, 1 \leq i \leq n$ be the corresponding eigenvectors. Then show that $ {\cal {B}} =    \{{\mathbf u}_1, {\mathbf u}_2, \ldots, {\mathbf u}_n \}$ forms a basis of $ {\mathbb{F}}^n ({\mathbb{F}}).$ If$ [{\mathbf b}]_{{\cal B}} = (c_1, c_2, \ldots, c_n)^t$ then show that $ A {\mathbf x}= {\mathbf b}$ has the unique solution   $\displaystyle {\mathbf x}= \frac{c_1}{{\lambda}_1} {\mathbf u}_1 + \frac{c_2}{{\lambda}_2} {\mathbf u}_2 + \cdots + \frac{c_n}{{\lambda}_n} {\mathbf u}_n.$ |
| **Diagonalisation**  Let $ A$ be a square matrix of order $ n$ and let $ T_A: {\mathbb{F}}^n {\longrightarrow}{\mathbb{F}}^n$ be the corresponding linear transformation. In this section, we ask the question ``does there exist a basis $ {\cal B}$ of $ {\mathbb{F}}^n$ such that $ T_A[{\cal B},{\cal B}],$ the matrix of the linear transformation $ T_A,$ is in the simplest possible form."  We know that, the simplest form for a matrix is the identity matrix and the diagonal matrix. In this section, we show that for a certain class of matrices $ A,$ we can find a basis $ {\cal B}$ such that $ T_A[{\cal B},{\cal B}]$ is a diagonal matrix, consisting of the eigenvalues of $ A.$ This is equivalent to saying that $ A$ is similar to a diagonal matrix. To show the above, we need the following definition.  **DEFINITION 6.2.1 (Matrix Diagonalisation)*****A matrix $ A$ is said to be diagonalisable if there exists a non-singular matrix $ P$ such that $ P^{-1} A P$ is a diagonal matrix.***  **Remark 6.2.2**   *Let $ A$ be an $ n \times n$ diagonalisable matrix with eigenvalues $ {\lambda}_1, {\lambda}_2, \ldots, {\lambda}_n.$ By definition, $ A$ is similar to a diagonal matrix $ D.$ Observe that $ D = {\mbox{diag}}({\lambda}_1, {\lambda}_2, \ldots, {\lambda}_n)$ as similar matrices have the same set of eigenvalues and the eigenvalues of a diagonal matrix are its diagonal entries.*  **EXAMPLE 6.2.3   *Let $ A= \left[\begin{array}{cc} 0 & 1 \\ -1 & 0 \end{array}\right].$ Then we have the following:***   1. Let $ V = {\mathbb{R}}^2.$ Then $ A$ has no real eigenvalue (see Example [6.1.8](http://nptel.iitm.ac.in/courses/Webcourse-contents/IIT-KANPUR/mathematics-2/node53.html#rem:no:eigenvalue) and hence $ A$ doesn't have eigenvectors that are vectors in $ {\mathbb{R}}^2.$ Hence, there does not exist any non-singular $ 2 \times 2$ real matrix $ P$such that $ P^{-1} A P$ is a diagonal matrix. 2. In case, $ V = {\mathbb{C}}^2 ({\mathbb{C}}),$ the two complex eigenvalues of $ A$ are $ -i, i$ and the corresponding eigenvectors are $ (i, 1)^t$ and $ (-i, 1)^t,$ respectively. Also, $ (i, 1)^t$ and $ (-i, 1)^t$ can be taken as a basis of$ {\mathbb{C}}^2 ({\mathbb{C}}).$ Define a $ 2 \times 2$ complex matrix by $ U = \frac{1}{\sqrt{2}}\left[\begin{array}{cc}    i & -i \\ 1 & 1 \end{array}\right].$ Then   $\displaystyle U^* A U = \left[\begin{array}{cc} -i & 0 \\ 0 & i \end{array}\right].$  **THEOREM 6.2.4   *let $ A$ be an $ n \times n$ matrix. Then $ A$ is diagonalisable if and only if $ A$ has $ n$ linearly independent eigenvectors.***  *Proof*. Let $ A$ be diagonalisable. Then there exist matrices $ P$ and $ D$ such that  $\displaystyle P^{-1} A P = D = {\mbox{diag}}({\lambda}_1, {\lambda}_2, \ldots, {\lambda}_n). $  Or equivalently, $ A P = P D.$ Let $ P = [{\mathbf u}_1, {\mathbf u}_2, \ldots, {\mathbf u}_n].$ Then $ A P = P D$ implies that  $\displaystyle A {\mathbf u}_i = d_i {\mathbf u}_i \;\; {\mbox{ for }} \;\; 1 \leq i \leq n.$  Since $ {\mathbf u}_i$ 's are the columns of a non-singular matrix $ P,$ they are non-zero and so for $ 1 \leq i \leq n,$ we get the eigenpairs $ (d_i, {\mathbf u}_i)$ of $ A.$ Since, $ {\mathbf u}_i$ 's are columns of the non-singular matrix $ P,$ using Corollary[4.3.9](http://nptel.iitm.ac.in/courses/Webcourse-contents/IIT-KANPUR/mathematics-2/node45.html#cor:rank:nullity), we get $ {\mathbf u}_1, {\mathbf u}_2, \ldots, {\mathbf u}_n$ are linearly independent.  Thus we have shown that if $ A$ is diagonalisable then $ A$ has $ n$ linearly independent eigenvectors.  Conversely, suppose $ A$ has $ n$ linearly independent eigenvectors $ {\mathbf u}_i, \; 1 \leq i \leq n$ with eigenvalues $ \lambda_i.$ Then $ A {\mathbf u}_i = \lambda_i {\mathbf u}_i.$ Let $ P = [{\mathbf u}_1, {\mathbf u}_2, \ldots, {\mathbf u}_n].$ Since $ {\mathbf u}_1, {\mathbf u}_2, \ldots, {\mathbf u}_n$ are linearly independent, by Corollary [4.3.9](http://nptel.iitm.ac.in/courses/Webcourse-contents/IIT-KANPUR/mathematics-2/node45.html#cor:rank:nullity), $ P$ is non-singular. Also,   |  |  |  |  | | --- | --- | --- | --- | | $\displaystyle A P$ | $\displaystyle =$ | $\displaystyle [A {\mathbf u}_1, A {\mathbf u}_2, \ldots, A {\mathbf u}_n] = [{\... ..._1 {\mathbf u}_1, {\lambda}_2 {\mathbf u}_2, \ldots, {\lambda}_n {\mathbf u}_n]$ |  | |  | $\displaystyle =$ | $\displaystyle [{\mathbf u}_1, {\mathbf u}_2, \ldots, {\mathbf u}_n] \begin{bmat... ...a}_2 & 0 \\ \vdots & \ddots & \vdots \\ 0 & 0 & {\lambda}_n \end{bmatrix}= P D.$ |  |   Therefore the matrix $ A$ is diagonalisable. height6pt width 6pt depth 0pt  **COROLLARY 6.2.5   *let $ A$ be an $ n \times n$ matrix. Suppose that the eigenvalues of $ A$ are distinct. Then $ A$ is diagonalisable.***  *Proof*. As $ A$ is an $ n \times n$ matrix, it has $ n$ eigenvalues. Since all the eigenvalues of $ A$ are distinct, by Corollary [6.1.17](http://nptel.iitm.ac.in/courses/Webcourse-contents/IIT-KANPUR/mathematics-2/node53.html#cor:distinct:eigenvalues), the $ n$ eigenvectors are linearly independent. Hence, by Theorem [6.2.4](http://nptel.iitm.ac.in/courses/Webcourse-contents/IIT-KANPUR/mathematics-2/node54.html#thm:diagonal), $ A$ is diagonalisable. height6pt width 6pt depth 0pt  **COROLLARY 6.2.6   *Let $ A$ be an $ n \times n$ matrix with $ {\lambda}_1, {\lambda}_2, \ldots, {\lambda}_k$ as its distinct eigenvalues and $ p({\lambda}) $ as its characteristic polynomial. Suppose that for each $ i, \; 1 \le i \le k, \; (x - {\lambda}_i)^{m_i}$divides $ p({\lambda}) $ but $ (x - {\lambda}_i)^{m_i+1}$ does not divides $ p({\lambda}) $ for some positive integers $ m_i$ . Then***  $\displaystyle A \;{\mbox{ is diagonalisable if and only if }} \; \dim\bigl(\ker(A - {\lambda}_i I)\bigr) = m_i \;{\mbox{ for each }} \; i, \; 1 \le i \le k.$  ***Or equivalently $ A \;{\mbox{ is diagonalisable if and only if }} \; {\mbox{rank}}(A - {\lambda}_i I) = n - m_i \;{\mbox{ for each }} \; i, \; 1 \le i \le k.$***  *Proof*. As $ A$ is diagonalisable, by Theorem [6.2.4](http://nptel.iitm.ac.in/courses/Webcourse-contents/IIT-KANPUR/mathematics-2/node54.html#thm:diagonal), $ A$ has $ n$ linearly independent eigenvalues. Also, $ \sum\limits_{i=1}^k m_i = n$ as $ \deg( p({\lambda})) = n$ . Hence, for each eigenvalue $ {\lambda}_i, \; 1 \le i \le k$ , $ A$ has exactly $ m_i$linearly independent eigenvectors. Thus, for each $ i, \; 1 \le i \le k$ , the homogeneous linear system $ (A - {\lambda}_i I) {\mathbf x}= {\mathbf 0}$ has exactly $ m_i$ linearly independent vectors in its solution set. Therefore,$ \dim\bigl(\ker(A - {\lambda}_i I)\bigr) \ge m_i$ . Indeed $ \dim\bigl(\ker(A - {\lambda}_i I)\bigr) = m_i$ for $ 1 \le i \le k$ follows from a simple counting argument.  Now suppose that for each $ i, \; 1 \le i \le k, \;\dim\bigl(\ker(A - {\lambda}_i I)\bigr) = m_i$ . Then for each $ i, \; 1 \le i \le k$ , we can choose $ m_i$ linearly independent eigenvectors. Also by Corollary [6.1.17](http://nptel.iitm.ac.in/courses/Webcourse-contents/IIT-KANPUR/mathematics-2/node53.html#cor:distinct:eigenvalues), the eigenvectors corresponding to distinct eigenvalues are linearly independent. Hence $ A$ has $ n = \sum\limits_{i=1}^k m_i$ linearly independent eigenvectors. Hence by Theorem [6.2.4](http://nptel.iitm.ac.in/courses/Webcourse-contents/IIT-KANPUR/mathematics-2/node54.html#thm:diagonal), $ A$ is diagonalisable. height6pt width 6pt depth 0pt  **EXAMPLE 6.2.7**   1. Let $ A=\left[\begin{array}{ccc}2 & 1 & 1\\ 1 & 2 & 1\\ 0 & -1    & 1 \end{array}\right].$ Then $ \det ( A - {\lambda}I) = (2 -    {\lambda})^2 (1 - {\lambda}).$ Hence, $ A$ has eigenvalues $ 1, 2, 2.$ It is easily seen that $ \bigl(1, (1,0, -1)^t \bigr)$ and $ (\bigl( 2, (1,1,-1)^t \bigr)$ are the only eigenpairs. That is, the matrix $ A$ has exactly one eigenvector corresponding to the repeated eigenvalue $ 2.$ Hence, by Theorem [6.2.4](http://nptel.iitm.ac.in/courses/Webcourse-contents/IIT-KANPUR/mathematics-2/node54.html#thm:diagonal), the matrix $ A$ is not diagonalisable. 2. Let $ A=\left[\begin{array}{ccc}2 & 1 & 1\\ 1 & 2 & 1\\ 1 & 1    & 2 \end{array}\right].$ Then $ \det ( A - {\lambda}I) = (4 -    {\lambda})(1 - {\lambda})^2.$ Hence, $ A$ has eigenvalues $ 1, 1, 4.$ It can be easily verified that $ (1,-1,0)^t$ and $ (1,0,-1)^t$ correspond to the eigenvalue $ 1$ and $ (1,1,1)^t$ corresponds to the eigenvalue $ \; 4.$ Note that the set $ \{ (1, -1, 0)^t, (1, 0, -1)^t \}$ consisting of eigenvectors corresponding to the eigenvalue $ 1$ are not orthogonal. This set can be replaced by the orthogonal set $ \{(1,0,-1)^t, (1,-2,1)^t\}$ which still consists of eigenvectors corresponding to the eigenvalue $ 1$ as $ (1, -2, 1) = 2 (1,-1,0) - (1,0,-1)$ . Also, the set$ \{(1,1,1), (1,0,-1), (1,-2,1)\}$ forms a basis of $ {\mathbb{R}}^3.$ So, by Theorem [6.2.4](http://nptel.iitm.ac.in/courses/Webcourse-contents/IIT-KANPUR/mathematics-2/node54.html#thm:diagonal), the matrix $ A$ is diagonalisable. Also, if $ U = \left[\begin{array}{ccc} \frac{1}{\sqrt{3}} &    \frac{1}{\sqrt{2}} & \frac{1...    ...frac{1}{\sqrt{3}} & -\frac{1}{\sqrt{2}}    & \frac{1}{\sqrt{6}} \end{array}\right]$ is the corresponding unitary matrix then$ U^* A U = {\mbox{diag}}(4,1,1).$   Observe that the matrix $ A$ is a symmetric matrix. In this case, the eigenvectors are mutually orthogonal. In general, for any $ n \times n$ real symmetric matrix $ A,$ there always exist $ n$ eigenvectors and they are mutually orthogonal |
| **Diagonalisable matrices**  In this section, we will look at some special classes of square matrices which are diagonalisable. We will also be dealing with matrices having complex entries and hence for a matrix $ A=[a_{ij}],$ recall the following definitions.  **DEFINITION 6.3.1 (Special Matrices)**   1. $ A^* = ( {\overline{a_{ji}}} ),$ is called the **conjugate transpose** of the matrix $ A.$   Note that $ A^* = {\overline{ A^{t}}} = {\overline {A}}^{t}.$   1. A square matrix $ A$ with complex entries is called    1. a Hermitian matrix if $ A^*       = A.$    2. a unitary matrix if $ A \; A^* = A^* A = I_n.$    3. a skew-Hermitian matrix if $ A^* = - A.$    4. a normal matrix if $ A^* A =       A A^*.$ 2. A square matrix $ A$ with real entries is called    1. a **symmetric** matrix if $ A^{t} = A.$    2. an **orthogonal** matrix if $ A \;A^{t} = A^{t} A = I_n.$    3. a **skew-symmetric** matrix if $ A^{t} = -A.$   Note that a symmetric matrix is always Hermitian, a skew-symmetric matrix is always skew-Hermitian and an orthogonal matrix is always unitary. Each of these matrices are normal. If $ A$ is a unitary matrix then $ A^* = A^{-1}.$  **EXAMPLE 6.3.2**   1. Let $ B= \begin{bmatrix}i & 1 \\ -1 & i    \end{bmatrix}.$ Then $ B$ is skew-Hermitian. 2. Let $ A = \frac{1}{\sqrt{2}}\begin{bmatrix}1 & i \\ i & 1    \end{bmatrix}$ and $ B = \begin{bmatrix}1 & 1 \\ -1 & 1    \end{bmatrix}.$ Then $ A$ is a unitary matrix and $ B$ is a normal matrix. Note that $ \sqrt{2} A$ is also a normal matrix.   **DEFINITION 6.3.3 (Unitary Equivalence)*****Let $ A$ and $ B$ be two $ n \times n$ matrices. They are called unitarily equivalent if there exists a unitary matrix $ U$ such that $ A = U^* B U.$***  ***Note that $ U^* = U^{-1}$ as $ U$ is a unitary matrix. So, $ A$ is unitarily similar to the matrix $ B$ .***  **EXERCISE 6.3.4**   1. Let $ A$ be a square matrix such that $ U A U^*$ is a diagonal matrix for some unitary matrix $ U$ . Prove that $ A$ is a normal matrix. 2. Let $ A$ be any matrix. Then $ A = \frac{1}{2}(A + A^*) + \frac{1}{2}(A - A^*) $ where $ \frac{1}{2}(A + A^*) $ is the Hermitian part of $ A$ and $ \frac{1}{2}(A - A^*) $ is the skew-Hermitian part of $ A.$ 3. Every matrix can be uniquely expressed as $ A = S + i T$ where both $ S$ and $ T$ are Hermitian matrices. 4. Show that $ A - A^*$ is always skew-Hermitian. 5. Does there exist a unitary matrix $ U$ such that $ U^{-1} A U = B$ where  $ A = \begin{bmatrix}1 & 1 & 4\\ 0 &2 & 2\\ 0&0&3 \end{bmatrix}$ and $ B = \begin{bmatrix}2 & -1 & 3 \sqrt{2}\\ 0 &1 & \sqrt{2}\\ 0&0&3    \end{bmatrix}.$   **PROPOSITION 6.3.5   *Let $ A$ be an $ n \times n$ Hermitian matrix. Then all the eigenvalues of $ A$ are real.***  *Proof*. Let $ (\lambda, {\mathbf x})$ be an eigenpair. Then $ A {\mathbf x}= \lambda {\mathbf x}$ and $ A = A^*$ implies  $\displaystyle {\mathbf x}^* A = {\mathbf x}^* A^* = (A {\mathbf x})^* = ({\lambda}{\mathbf x})^* = \overline{{\lambda}} {\mathbf x}^*.$  Hence  $\displaystyle \lambda {\mathbf x}^*{\mathbf x}= {\mathbf x}^* ({\lambda}{\mathb... ...}} {\mathbf x}^*) {\mathbf x}= {\overline{ \lambda}} {\mathbf x}^* {\mathbf x}.$  But $ {\mathbf x}$ is an eigenvector and hence $ {\mathbf x}\neq {\mathbf 0}$ and so the real number $ \Vert{\mathbf x}\Vert^2 = {\mathbf x}^* {\mathbf x}$ is non-zero as well. Thus $ \lambda = {\overline{\lambda}}.$ That is, $ {\lambda}$ is a real number. height6pt width 6pt depth 0pt  **THEOREM 6.3.6   *Let $ A$ be an $ n \times n$ Hermitian matrix. Then $ A$ is unitarily diagonalisable. That is, there exists a unitary matrix $ U$ such that $ U^* A U = D;$ where $ D$ is a diagonal matrix with the eigenvalues of $ A$ as the diagonal entries.***  ***In other words, the eigenvectors of $ A$ form an orthonormal basis of $ {\mathbb{C}}^n.$***  *Proof*. We will prove the result by induction on the size of the matrix. The result is clearly true if $ n=1.$ Let the result be true for $ n = k-1.$ we will prove the result in case $ n = k.$ So, let $ A$ be a $ k \times k$matrix and let $ (\lambda_1, {\mathbf x})$ be an eigenpair of $ A$ with $ \Vert {\mathbf x}\Vert = 1.$ We now extend the linearly independent set $ \{ {\mathbf x}\}$ to form an orthonormal basis $ \{{\mathbf x}, {\mathbf u}_2, {\mathbf u}_3, \ldots, {\mathbf u}_k \}$ (using *Gram-Schmidt Orthogonalisation*) of $ {\mathbb{C}}^k$ .  As $ \{{\mathbf x}, {\mathbf u}_2, {\mathbf u}_3, \ldots, {\mathbf u}_k \}$ is an orthonormal set,  $\displaystyle {\mathbf u}_i^* {\mathbf x}= 0 \;\; {\mbox{ for all }} \; i = 2, 3, \ldots, k.$  Therefore, observe that for all $ i, \; 2 \leq i \leq k,$  $\displaystyle (A {\mathbf u}_i)^* {\mathbf x}= ({\mathbf u}_i* A^*) {\mathbf x}... ..._i^* ({\lambda}_1 {\mathbf x}) = {\lambda}_1 ({\mathbf u}_i^* {\mathbf x}) = 0.$  Hence, we also have $ {\mathbf x}^* (A {\mathbf u}_i) = 0$ for $ 2 \leq i \leq k.$ Now, define $ U_1 = [ {\mathbf x}, \; {\mathbf u}_2, \; \cdots, {\mathbf u}_k ]$ (with $ {\mathbf x}, {\mathbf u}_2, \ldots, {\mathbf u}_k$ as columns of $ U_1$ ). Then the matrix $ U_1$ is a unitary matrix and   |  |  |  |  | | --- | --- | --- | --- | | $\displaystyle U_1^{*} A U_1$ | $\displaystyle =$ | $\displaystyle U_1^* [ A {\mathbf x}\; A {\mathbf u}_2 \; \cdots A {\mathbf u}_k ]$ |  | |  | $\displaystyle =$ | $\displaystyle \begin{bmatrix}{\mathbf x}^* \\ {\mathbf u}_2^* \\ \vdots \\ {\ma... ...mbda}_1 {\mathbf x}) & \cdots & {\mathbf u}_k^* (A {\mathbf u}_k) \end{bmatrix}$ |  | |  | $\displaystyle =$ | $\displaystyle \left[\begin{array}{c\vert c} \lambda_1 & {\mathbf 0}\\ \hline {\mathbf 0}& \\ \vdots & B \\ {\mathbf 0}& \end{array} \right],$ |  |   where $ B$ is a $ (k-1) \times (k-1)$ matrix. As $ A^* = A$ ,we get $ (U_1^{*} A U_1)^* = U_1^{*} A U_1$ . This condition, together with the fact that $ {\lambda}_1$ is a real number (use Proposition [6.3.5](http://nptel.iitm.ac.in/courses/Webcourse-contents/IIT-KANPUR/mathematics-2/node55.html#prop:hermitian)), implies that $ B^* = B$ . That is,$ B$ is also a Hermitian matrix. Therefore, by induction hypothesis there exists a $ (k-1) \times (k-1)$ unitary matrix $ U_2$ such that  $\displaystyle U_2^{*} B U_2 = D_2 = {\mbox{diag}}(\lambda_2, \ldots, \lambda_k).$  Recall that , the entries $ {\lambda}_i, \; $ for $ 2 \leq i \leq k$ are the eigenvalues of the matrix $ B.$ We also know that two similar matrices have the same set of eigenvalues. Hence, the eigenvalues of $ A$ are $ \lambda_1, \lambda_2, \ldots, \lambda_k.$Define $ U= U_1 \begin{bmatrix}1 & {\mathbf 0}\\ {\mathbf 0}& U_2 \end{bmatrix}.$ Then $ U$ is a unitary matrix and   |  |  |  |  | | --- | --- | --- | --- | | $\displaystyle U^{*} A U$ | $\displaystyle =$ | $\displaystyle \left( U_1 \begin{bmatrix}1 & {\mathbf 0}\\ {\mathbf 0}& U_2 \end... ...left(U_1 \begin{bmatrix}1 & {\mathbf 0}\\ {\mathbf 0}& U_2 \end{bmatrix}\right)$ |  | |  | $\displaystyle =$ | $\displaystyle \left(\begin{bmatrix}1 & {\mathbf 0}\\ {\mathbf 0}& U_2^{*} \end{... ...ft( U_1 \begin{bmatrix}1 & {\mathbf 0}\\ {\mathbf 0}& U_2 \end{bmatrix} \right)$ |  | |  | $\displaystyle =$ | $\displaystyle \begin{bmatrix}1 & {\mathbf 0}\\ {\mathbf 0}& U_2^{*} \end{bmatri... ...*} A U_1 \bigr) \begin{bmatrix}1 & {\mathbf 0}\\ {\mathbf 0}& U_2 \end{bmatrix}$ |  | |  | $\displaystyle =$ | $\displaystyle \begin{bmatrix}1 & {\mathbf 0}\\ {\mathbf 0}& U_2^{*} \end{bmatri... ...in{bmatrix}{\lambda}_1 & {\mathbf 0}\\ {\mathbf 0}& U_2^{*} B U_2 \end{bmatrix}$ |  | |  | $\displaystyle =$ | $\displaystyle \begin{bmatrix}{\lambda}_1 & {\mathbf 0}\\ {\mathbf 0}& D_2 \end{bmatrix}.$ |  |   Thus, $ U^{*} A U$ is a diagonal matrix with diagonal entries $ \lambda_1, \lambda_2, \ldots, \lambda_k,$ the eigenvalues of $ A.$ Hence, the result follows. height6pt width 6pt depth 0pt  **COROLLARY 6.3.7   *Let $ A$ be an $ n \times n$ real symmetric matrix. Then***   1. the eigenvalues of $ A$ are all real, 2. the corresponding eigenvectors can be chosen to have real entries, and 3. the eigenvectors also form an orthonormal basis of $ {\mathbb{R}}^n.$   *Proof*. As $ A$ is symmetric, $ A$ is also an Hermitian matrix. Hence, by Proposition [6.3.5](http://nptel.iitm.ac.in/courses/Webcourse-contents/IIT-KANPUR/mathematics-2/node55.html#prop:hermitian), the eigenvalues of $ A$ are all real. Let $ ({\lambda}, \; {\mathbf x})$ be an eigenpair of $ A.$ Suppose $ {\mathbf x}^t \in {\mathbb{C}}^n.$ Then there exist $ {\mathbf y}^t, {\mathbf z}^t \in {\mathbb{R}}^n$such that $ {\mathbf x}= {\mathbf y}+ i {\mathbf z}.$ So,  $\displaystyle A {\mathbf x}= {\lambda}{\mathbf x}\Longrightarrow A ({\mathbf y}+ i {\mathbf z}) = {\lambda}( {\mathbf y}+ i {\mathbf z}).$  Comparing the real and imaginary parts, we get $ A {\mathbf y}= {\lambda}{\mathbf y}$ and $ A {\mathbf z}= {\lambda}{\mathbf z}.$ Thus, we can choose the eigenvectors to have real entries.  To prove the orthonormality of the eigenvectors, we proceed on the lines of the proof of Theorem [6.3.6](http://nptel.iitm.ac.in/courses/Webcourse-contents/IIT-KANPUR/mathematics-2/node55.html#thm:hermitian), Hence, the readers are advised to complete the proof. height6pt width 6pt depth 0pt  **EXERCISE 6.3.8**   1. Let $ A$ be a skew-Hermitian matrix. Then all the eigenvalues of $ A$ are either zero or purely imaginary. Also, the eigenvectors corresponding to distinct eigenvalues are mutually orthogonal.  *[Hint: Carefully study the proof of Theorem*[*6.3.6*](http://nptel.iitm.ac.in/courses/Webcourse-contents/IIT-KANPUR/mathematics-2/node55.html#thm:hermitian)*.]* 2. Let $ A$ be an $ n \times n$ unitary matrix. Then    1. the rows of $ A$ form an orthonormal basis of $ {\mathbb{C}}^n.$    2. the columns of $ A$ form an orthonormal basis of $ {\mathbb{C}}^n.$    3. for any two vectors $ {\mathbf x}, {\mathbf y}\in {\mathbb{C}}^{n \times 1},\;$ $ \langle A {\mathbf x}, A {\mathbf y}\rangle = \langle {\mathbf x}, {\mathbf y}\rangle.$    4. for any vector $ {\mathbf x}\in {\mathbb{C}}^{n \times 1},\; $ $ \Vert A {\mathbf x}\Vert = \Vert {\mathbf x}\Vert.$    5. for any eigenvalue $ \lambda$ $ A, \;$ $ \vert \lambda\vert = 1.$    6. the eigenvectors $ {\mathbf x}, {\mathbf y}$ corresponding to distinct eigenvalues $ {\lambda}$ and $ \mu$ satisfy $ \langle {\mathbf x}, {\mathbf y}\rangle = 0.$ That is, if $ ({\lambda}, {\mathbf x})$ and $ (\mu, {\mathbf y})$ are eigenpairs, with $ {\lambda}\neq \mu,$ then $ {\mathbf x}$ and $ {\mathbf y}$ are mutually orthogonal. 3. Let $ A$ be a normal matrix. Then, show that if $ (\lambda, {\mathbf x})$ is an eigenpair for $ A$ then $ ({\overline{\lambda}}, {\mathbf x})$ is an eigenpair for $ A^*.$ 4. Show that the matrices $ A = \begin{bmatrix}4&4\\ 0&4    \end{bmatrix}$ and $ B = \begin{bmatrix}10&9 \\ -4&-2    \end{bmatrix}$ are similar. Is it possible to find a unitary matrix $ U$ such that $ A = U^* B U?$ 5. Let $ A$ be a $ 2 \times 2$ orthogonal matrix. Then prove the following:    1. if $ \det (A) = 1,$ then $ A = \begin{bmatrix}\cos \theta & - \sin \theta \\ \sin \theta & \cos \theta       \end{bmatrix}$ for some $ \theta, \;\; 0 \leq \theta < 2 \pi.$    2. if $ \det A = -1,$ then there exists a basis of $ {\mathbb{R}}^2$ in which the matrix of $ A$ looks like $ \begin{bmatrix}       1 & 0 \\ 0 & -1 \end{bmatrix}.$   Or equivalently, $ A = \begin{bmatrix}\cos \theta & \sin \theta \\ \sin \theta & - \cos \theta \end{bmatrix}$ for some $ \theta, \;\; 0 \leq \theta < 2 \pi.$ In this case, prove that $ A$ reflects the vectors in $ {\mathbb{R}}^2$ about a line passing through origin. Also, determine this line.   1. Let $ A = \begin{bmatrix}2 & 1 & 1 \\ 1 & 2 & 1 \\ 1 & 1 & 2 \end{bmatrix}.$ Determine $ A^{301}$ . 2. Let $ A$ be a $ 3 \times 3$ orthogonal matrix. Then prove the following:    1. if $ \det (A) = 1,$ then $ A$ is a rotation about a fixed axis, in the sense that $ A$ has an eigenpair $ (1, {\mathbf x})$ such that the restriction of $ A$ to the plane $ {\mathbf x}^{\perp}$ is a two dimensional rotation of $ {\mathbf x}^{\perp}.$    2. if $ \det A = -1,$ then the action of $ A$ corresponds to a reflection through a plane $ P,$ followed by a rotation about the line through the origin that is perpendicular to $ P.$   **Remark 6.3.9**   *In the previous exercise, we saw that the matrices $ A = \begin{bmatrix}4&4\\ 0&4 \end{bmatrix}$ and $ B = \begin{bmatrix}10&9 \\ -4&-2 \end{bmatrix}$ are similar but not unitarily equivalent, whereas unitary equivalence implies similarity equivalence as $ U^* = U^{-1}.$ But in numerical calculations, unitary transformations are preferred as compared to similarity transformations. The main reasons being:*   1. Exercise [6.3.8](http://nptel.iitm.ac.in/courses/Webcourse-contents/IIT-KANPUR/mathematics-2/node55.html#exe:diagonal:2).[2](http://nptel.iitm.ac.in/courses/Webcourse-contents/IIT-KANPUR/mathematics-2/node55.html#exe:unitary:absolute) implies that an orthonormal change of basis leaves unchanged the sum of squares of the absolute values of the entries which need not be true under a non-orthonormal change of basis. 2. As $ U^* = U^{-1}$ for a unitary matrix $ U,$ unitary equivalence is computationally simpler. 3. Also in doing ``conjugate transpose", the loss of accuracy due to round-off errors doesn't occur.   We next prove the Schur's Lemma and use it to show that normal matrices are unitarily diagonalisable.  **LEMMA 6.3.10 (Schur's Lemma)   *Every $ n \times n$ complex matrix is unitarily similar to an upper triangular matrix.***  *Proof*. We will prove the result by induction on the size of the matrix. The result is clearly true if $ n=1.$ Let the result be true for $ n = k-1.$ we will prove the result in case $ n = k.$ So, let $ A$ be a $ k \times k$matrix and let $ (\lambda_1, {\mathbf x})$ be an eigenpair for $ A$ with $ \Vert {\mathbf x}\Vert = 1.$ Then the linearly independent set $ \{ {\mathbf x}\}$ can be extended, using the *Gram-Schmidt Orthogonalisation process,* to get an orthonormal basis$ \{{\mathbf x}, {\mathbf u}_2, {\mathbf u}_3, \ldots, {\mathbf u}_k \}$ of $ {\mathbb{C}}^n({\mathbb{C}})$ . Then $ U_1 = [ {\mathbf x}\; {\mathbf u}_2 \; \cdots {\mathbf u}_k ]$ (with $ {\mathbf x}, {\mathbf u}_2, \ldots, {\mathbf u}_k$ as the columns of the matrix $ U_1$ ) is a unitary matrix and   |  |  |  |  | | --- | --- | --- | --- | | $\displaystyle U_1^{*} A U_1$ | $\displaystyle =$ | $\displaystyle U_1^* [ A {\mathbf x}\; A {\mathbf u}_2 \; \cdots A {\mathbf u}_k ]$ |  | |  | $\displaystyle =$ | $\displaystyle \begin{bmatrix}{\mathbf x}^* \\ {\mathbf u}_2^* \\ \vdots \\ {\ma... ..._1 & * \\ \hline {\mathbf 0}& \\ \vdots & B \\ {\mathbf 0}& \end{array} \right]$ |  |   where $ B$ is a $ (k-1) \times (k-1)$ matrix. By induction hypothesis there exists a $ (k-1) \times (k-1)$ unitary matrix $ U_2$ such that $ U_2^{*} B U_2 $ is an upper triangular matrix with diagonal entries $ \lambda_2, \ldots, \lambda_k,$ the eigen values of the matrix $ B.$ Observe that since the eigenvalues of $ B$ are $ \lambda_2, \ldots, \lambda_k$ the eigenvalues of $ A$ are $ \lambda_1, \lambda_2, \ldots, \lambda_k.$ Define $ U= U_1 \begin{bmatrix}1 & {\mathbf 0}\\ {\mathbf 0}& U_2 \end{bmatrix}.$ Then check that $ U$ is a unitary matrix and $ U^{*} A U$ is an upper triangular matrix with diagonal entries $ \lambda_1, \lambda_2, \ldots, \lambda_k,$ the eigenvalues of the matrix $ A.$ Hence, the result follows. height6pt width 6pt depth 0pt  **EXERCISE 6.3.11**   1. Let $ A$ be an $ n \times n$ real invertible matrix. Prove that there exists an orthogonal matrix $ P$ and a diagonal matrix $ D$ with positive diagonal entries such that $ A A^t = P D P^{-1}$ . 2. Show that matrices $ A = \begin{bmatrix}1 & 1 &    1\\ 0 & 2 & 1\\ 0 & 0 & 3 \end{bmatrix}$ and $ B =    \begin{bmatrix}2 & -1 & \sqrt{2}\\ 0 & 1 & 0\\ 0 & 0 & 3    \end{bmatrix}$ are unitarily equivalent via the unitary matrix $ U = \frac{1}{\sqrt{2}} \begin{bmatrix}1 & 1 & 0\\ 1 & -1    & 0\\ 0 & 0 & \sqrt{2} \end{bmatrix}.$ Hence, conclude that the upper triangular matrix obtained in the "Schur's Lemma" need not be unique. 3. Show that the normal matrices are diagonalisable.  *[Hint: Show that the matrix $ B$ in the proof of the above theorem is also a normal matrix and if $ T$ is an upper triangular matrix with $ T^* T = T T^*$ then $ T$ has to be a diagonal matrix].*   **Remark 6.3.12** (The Spectral Theorem for Normal Matrices)   *Let $ A$ be an $ n \times n$ normal matrix. Then the above exercise shows that there exists an orthonormal basis $ \{{\mathbf x}_1, {\mathbf x}_2, \ldots, {\mathbf x}_n \}$ of$ {\mathbb{C}}^n({\mathbb{C}})$ such that $ A {\mathbf x}_i = \lambda_i {\mathbf x}_i$ for $ 1 \leq i \leq n.$*   1. Let $ A$ be a normal matrix. Prove the following:    1. if all the eigenvalues of $ A$ are $ 0,$ then $ A = {\mathbf 0},$    2. if all the eigenvalues of $ A$ are $ 1,$ then $ A = I.$ 2. Let $ A$ be an $ n \times n$ matrix. Prove that    1. if $ A$ is Hermitian and $ {\mathbf x}A {\mathbf x}^* = 0$ for all $ {\mathbf x}\in {\mathbb{C}}^n$ then $ A = {\mathbf 0}$ .    2. if $ A$ is a real, symmetric matrix and $ {\mathbf x}A {\mathbf x}^* = 0$ for all $ {\mathbf x}\in {\mathbb{R}}^n$ then $ A = {\mathbf 0}$ . |