1. Systems of Linear Algebraic Equations

1.1 Matrices:

Let V, W be finite dimensional vector spaces. Let T be a linear transformation from V into W.

$$T : V \rightarrow W$$

Let $B_V = \{v_1, \dots, v_n\}$ be an ordered basis of V, so that $\dim V = n$. And $B_W = \{w_1, \dots, w_m\}$ is an ordered basis for W, so that $\dim W = m$. We are going to define the matrix of the linear transformation T. It is done as follows:-

A linear transformation is completely determined by its action on the basis vectors. If we know $Tv_1, Tv_2, \dots, Tv_n$ it is enough to completely determine T.

Each of the vectors $Tv_1, Tv_2, \dots, Tv_n$ is determined by the m scalars:

$$\left.\begin{aligned} Tv_1 &= a_{11} w_1 + a_{21} w_2 + \dots + a_{m1} w_m \\ Tv_2 &= a_{12} w_2 + a_{22} w_2 + \dots + a_{m2} w_m \\ &\vdots \\ Tv_n &= a_{1n} w_1 + a_{2n} w_2 + \dots + a_{mn} w_m. \end{aligned}\right\}$$

That is,

$$T(v_j) = \sum_{i=1}^{m} a_{ij} w_i$$

On the right hand side, i is the running index, j is the free index that corresponds to Tv_j. The matrix of the vector $T(v_j)$ relative to the basis B_W is the column vector where entries are the coordinates of Tv_j with respect to B_W.

$$[Tv_j]_{B_W} = \begin{bmatrix} a_{1j} \\ a_{2j} \\ a_{3j} \\ \vdots \\ a_{mj} \end{bmatrix}$$

The matrix of the linear transformation T, that sends $x \in V$ having coordinates $x = (x_1, x_2, \dots, x_n)$ with respect to B_W to $T(x)$ in W with basis B_W is defined as -

$$A = (a_{ij}) = [T]_{B_V}^{B_W} = \begin{array}{c} \begin{array}{cccc} T(v_1) & T(v_2) & \cdots & T(v_n) \end{array} \\ \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ a_{31} & a_{32} & \cdots & a_{3n} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix} \end{array} \begin{array}{c} \\ w_1 \\ w_2 \\ w_3 \\ \vdots \\ w_m \end{array}$$

As an aid to remembering, how $[T]_{B_V}^{B_W}$ is constructed from T, you might write the basis vectors $v_1, v_2, \dots, v_n$ for the domain across the top and the basis vectors $w_1, w_2, \dots, w_m$ for the target space along the right.

In the matrix above, the jth column of $[T]_{\mathcal{B}_V}^{\mathcal{B}_W}$ consists of scalars needed to write Tv_j as a linear combination of the w's. Thus, the picture should remind you that Tv_j is retrieved by multiplying each entry in the jth column, by the corresponding w from the right. and then adding up the resulting vectors.

This is in conformity ~~of~~ with the usual notation of writing a matrix.

<u>Examples</u>

1) Let $T: \mathbb{R}^2 \to \mathbb{R}^3$ be defined as

$$T((x_1, x_2)) = (x_1 + x_2, x_1 - x_2, 2x_2)$$

Determine the matrix of T, $[T]_{\mathcal{B}_V}^{\mathcal{B}_W}$ relative to the standard basis.

<u>Solution</u>

Here $\mathcal{B}_V = \{(1,0), (0,1)\}$

$\mathcal{B}_W = \{(1,0,0), (0,1,0), (0,0,1)\}$.

$$T((1,0)) = (1,1,0) = 1\cdot(1,0,0) + 1\cdot(0,1,0) + 0\cdot(0,0,1)$$
$$T((0,1)) = (1,-1,2) = 1\cdot(1,0,0) - 1(0,1,0) + 2(0,0,1)$$

$$[T]_{\mathcal{B}_V}^{\mathcal{B}_W} = \begin{bmatrix} 1 & 1 \\ 1 & -1 \\ 0 & 2 \end{bmatrix}.$$

Let's redo this example by changing the basis from $\mathcal{B}_W = \{e_1, e_2, e_3\}$ to $\mathcal{B}_W' = \{(1,1,0), (1,-1,0), (0,0,1)\}$.

$$T((1,0)) = (1,1,0) = 1\cdot(1,1,0) + 0(1,-1,0) + 0\cdot(0,0,1)$$
$$T((0,1)) = (1,-1,2) = 0(1,1,0) + 1\cdot(1,-1,0) + 2(0,0,1)$$

$$[T]_{\mathcal{B}_V}^{\mathcal{B}_W'} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 2 \end{bmatrix}$$

2) Let D be the differentiation transformation ~~defined by~~

$$D: \mathcal{P}_3(\mathbb{R}) \to \mathcal{P}_2(\mathbb{R})$$

defined by

$$Dp = p', \quad p \in \mathcal{P}_3(\mathbb{R}).$$

p is a polynomial. p' is the derivative of this polynomial.

$\mathcal{B}_1 = \{1, t, t^2, t^3\}$, $\mathcal{B}_2 = \{1, t, t^2\}$

What is $[D]_{\mathcal{B}_1}^{\mathcal{B}_2}$?

<u>Solution</u>

$$D(1) = 0 = 0\cdot 1 + 0\cdot t + 0\cdot t^2$$
$$D(t) = 1 = 1\cdot 1 + 0\cdot t + 0\cdot t^2$$
$$D(t^2) = 2t = 0\cdot 1 + 2\cdot t + 0\cdot t^2$$
$$D(t^3) = 3t^2 = 0\cdot 1 + 0\cdot t + 3\cdot t^2$$

$$[D]_{\mathcal{B}_1}^{\mathcal{B}_2} = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 3 \end{bmatrix}.$$

Note that, D sends elements from a 4-dimensional vector space to a 3-dimensional vector space.

Let V and W be finite dimensional vector spaces. Assume $\mathcal{B}_V = \{v_1, \dots, v_n\}$, so dim $V = n$ and $\mathcal{B}_W = \{w_1, \dots, w_m\}$, so dim $W = m$. Suppose T is a linear transformation.

$$T : V \longrightarrow W.$$

We have really speaking defined three matrices:

1) The matrix of the vector x with respect to the basis $\mathcal{B}_V$.

$$[x]_{\mathcal{B}_V}$$

2) The matrix of the vector T(x) with respect to basis $\mathcal{B}_W$.

$$[T(x)]_{\mathcal{B}_W}.$$

3) The matrix of the linear transformation T, with respect to the basis $\mathcal{B}_V$ and $\mathcal{B}_W$.

$$[T]_{\mathcal{B}_V}^{\mathcal{B}_W}.$$

How are these matrices related? We will derive a relationship between these three matrices. And if you look at that relationship it will tell you, why the statement, "~~why~~ any linear transformation between between finite dimensional vector spaces is like multiplying the vector x by the matrix A."

Theorem. 1

Let V, W be finite dimensional vector spaces. $\mathcal{B}_V = \{v_1, \dots, v_n\}$ is an ordered basis of V, so dim $V = n$. $\mathcal{B}_W = \{w_1, \dots, w_m\}$ is an ordered basis of W, so dim $W = m$.

Let T be a linear transformation from V into W.

$$T : V \rightarrow W.$$

Then,

$$[T(x)]_{\mathcal{B}_W} = [T]_{\mathcal{B}_V}^{\mathcal{B}_W} [x]_{\mathcal{B}_V}.$$

The matrix of Tx is A times the matrix of x. The matrix of T(x) is equal to the matrix of T times the matrix of x. So any linear transformation between finite dimensional vectors spaces is like the multiplication by a matrix.

Proof.

Let us recall

$$A = [T] = (a_{ij})$$

where

$$T(v_j) = \sum a_{ij} w_i \qquad \text{(the jth column)}$$

Note that, $\mathcal{B}_V = \{v_1, \dots, v_n\}$ is the basis of V., $\mathcal{B}_W = \{w_1, \dots, w_m\}$ is a basis of W. Let start with

$$x = x_1 v_1 + x_2 v_2 + \dots + x_n v_n.$$

Then,

$$[x]_{\mathcal{B}_V} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}$$

since, $x = x_1 v_1 + x_2 v_2 + \dots + x_n v_n$

$$\begin{aligned} T(x) &= T(x_1 v_1 + x_2 v_2 + \dots + x_n v_n) \\ &= x_1 Tv_1 + x_2 Tv_2 + \dots + x_n Tv_n \\ &= \sum_{j=1}^{n} x_j Tv_j \\ &= \sum_{j=1}^{n} x_j \left(\sum_{i=1}^{m} a_{ij} w_i \right) \\ &= \sum_{i=1}^{m} \left(\sum_{j=1}^{n} a_{ij} x_j \right) w_i \\ &= \sum_{i=1}^{m} \beta_i w_i \end{aligned}$$

where $\beta_i = \sum_j a_{ij} x_j$.

To summarize, Tx has been written as a linear combination of $w_1, w_2, \dots, w_m$, where the coefficients are $\beta_1, \beta_2, \dots, \beta_m$.

$T(x)$ is a vector in W. If I want to know the matrix of $T(x)$ relative to $\mathcal{B}_W$, then I must look at the unique representation of $T(x)$ in terms of the basis vectors $w_1, \dots, w_m$, which is precisely

$$T(x) = \sum_{i=1}^{m} \beta_i w_i$$

This means that the matrix of $T(x)$ has these β's as the column entries.

$$[T(x)]_{\mathcal{B}_W} = \begin{bmatrix} \beta_1 \\ \beta_2 \\ \beta_3 \\ \vdots \\ \beta_m \end{bmatrix} = \begin{bmatrix} \sum a_{1j} x_j \\ \sum a_{2j} x_j \\ \sum a_{3j} x_j \\ \vdots \\ \sum a_{mj} x_j \end{bmatrix}$$

$$= \begin{bmatrix} a_{11} x_1 + a_{12} x_2 + \dots + a_{1n} x_n \\ a_{21} x_1 + a_{22} x_2 + \dots + a_{2n} x_n \\ a_{31} x_1 + a_{32} x_2 + \dots + a_{3n} x_n \\ \vdots \\ a_{m1} x_1 + a_{m2} x_2 + \dots + a_{mn} x_n \end{bmatrix}$$

$$= \begin{bmatrix} a_{11} & a_{12} & a_{13} & \dots & a_{1n} \\ a_{21} & a_{22} & a_{23} & \dots & a_{2n} \\ \vdots & & & & \\ a_{m1} & a_{m2} & a_{m3} & \dots & a_{mn} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_m \end{bmatrix}$$

$$= [T]_{\mathcal{B}_V}^{\mathcal{B}_W} [x]_{\mathcal{B}_V}.$$

To summarize, matrices are concrete realizations of ~~f~~ linear transformations ~~between~~ finite dimensional vector spaces. There is in fact a one-to-one correspondence between a matrix and its linear transformation relative to the fixed basis $\mathcal{B}_V, \mathcal{B}_W$.

30-11-2019

Examples.

1) Consider the linear transformation $T: \mathbb{R}^3 \to \mathbb{R}^2$

defined by:

$$T((x_1, x_2, x_3)) = (x_1, x_2).$$

Choose as basis the standard basis vectors in $\mathbb{R}^3$ and $\mathbb{R}^2$ respectively. Then, obtain $[T]_{B_1}^{B_2}$.

Solution.

Let $B_1 = \{(1,0,0), (0,1,0), (0,0,1)\}$ and $B_2 = \{(1,0), (0,1)\}$ be the standard basis of $\mathbb{R}^3$ and $\mathbb{R}^2$ respectively.

$$T((1,0,0)) = (1,0) = 1\cdot(1,0) + 0\cdot(0,1)$$
$$T((0,1,0)) = (0,1) = 0\cdot(0,0) + 1\cdot(0,1)$$
$$T((0,0,1)) = (0,0) = 0\cdot(1,0) + 0\cdot(0,1)$$

$$[T]_{B_1}^{B_2} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix}.$$

2) Write the matrix of the linear transformation $T: \mathbb{R}^3 \to \mathbb{R}^2$

defined by -

$$T((x,y,z)) = (x+2y+2z, 2x+3y+4z)$$

Solution

$B_1 = 6\ \{e_1, e_2, e_3\}$

$B_2 = \{f_1, f_2\}$

$$T(e_1) = (1,2) = 1\cdot f_1 + 2\cdot f_2$$
$$T(e_2) = (2,3) = 2\cdot f_1 + 3\cdot f_2$$
$$T(e_3) = (2,4) = 2\cdot f_1 + 4\cdot f_2$$

$$[T]_{B_1}^{B_2} = \begin{bmatrix} 1 & 2 & 2 \\ 2 & 3 & 4 \end{bmatrix}.$$

3) What is the matrix of T in the above example with respect to the bases

$$B_1' = \{(1,0,0), (0,1,0), (1,2,-1)\}$$
$$B_2' = \{(1,2), (2,3)\}.$$

Solution

$$T((1,0,0)) = (1,2) = 1\cdot(1,2) + 0\cdot(2,3)$$
$$T((0,1,0)) = (2,3) = 0\cdot(1,2) + 1\cdot(2,3)$$
$$T((1,2,-1)) = (3,4) = -1\cdot(1,2) + 2\cdot(2,3)$$

$$[T]_{B_1'}^{B_2'} = \begin{bmatrix} 1 & 0 & -1 \\ 0 & 1 & 2 \end{bmatrix}.$$

4) Find a formula for the linear transformation $T(x)$ where $T: \mathbb{R}^3 \to \mathbb{R}^3$

is described by

$$[T]_{B} = \begin{bmatrix} 1 & 2 & 4 \\ 2 & 3 & 1 \\ 3 & 1 & 2 \end{bmatrix}.$$

where B is the standard basis of $\mathbb{R}^3$.

Solution.

$$[T(x)]_{\mathcal{B}} = [T]_{\mathcal{B}} [x]_{\mathcal{B}}$$

$$= \begin{bmatrix} 1 & 2 & 4 \\ 2 & 3 & 1 \\ 3 & 1 & 2 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$$

$$= \begin{bmatrix} x_1 + 2x_2 + 4x_3 \\ 2x_1 + 3x_2 + x_3 \\ 3x_1 + x_2 + 2x_3 \end{bmatrix}$$

so, $T((x_1, x_2, x_3)) = (x_1 + 2x_2 + 4x_3,\ 2x_1 + 3x_2 + x_3,\ 3x_1 + x_2 + 2x_3)$.

Theorem 2

Let V and W be finite dimensional vector spaces with dim $V = n$ and dim $W = m$ and let $\mathcal{B}_V$ and $\mathcal{B}_W$ be basis for V and W respectively. Then every linear transformation ~~A~~ $T \in \mathcal{L}(V, W)$ determines a unique $m \times n$ matrix with respect to $\mathcal{B}_V$ and $\mathcal{B}_W$ and conversely every such matrix determines a unique linear transformation from V to W. In fact, there is an invertible linear map ϕ between $L(V, W)$ and $\mathbb{F}^{m \times n}$, and $\mathcal{L}(V, W) \cong \mathbb{F}^{m \times n}$.

Proof.

Let $T \in \mathcal{L}(V, W)$.

a) Firstly, let's prove that ϕ is a linear map. ϕ ~~is a map~~ maps a linear transformation T to its matrix $A = [T]_{\mathcal{B}_V}^{\mathcal{B}_W}$. That is,

$$\phi(T) = [T]_{\mathcal{B}_V}^{\mathcal{B}_W}.$$

Let S, T be two linear transformations in $\mathcal{L}(V, W)$. Any linear transformation is completely determined by its action on the basis vectors.

Consider the vectors $S(v_j)$, $T(v_j)$ and $(S+T)(v_j)$. These belong to W. Therefore, we can write them as a linear combination the basis vectors $\mathcal{B}_W$.

6 Suppose $S(v_j) = a_{1j} w_1 + a_{2j} w_2 + \dots + a_{mj} w_m$
$T(v_j) = b_{1j} w_1 + b_{2j} w_2 + \dots + b_{mj} w_m$.

Now, $(S+T)(v_j) = S(v_j) + T(v_j)$ by the usual definition of vector addition in the space of functions or linear transformations.

Thus, $(S+T)(v_j) = \sum_i a_{ij} w_i + \sum_i b_{ij} w_i$

$$= \sum_i (a_{ij} + b_{ij}) w_i$$

$$\phi\, [(S+T)(v_j)]_{\mathcal{B}_W} = \begin{bmatrix} a_{1j} + b_{1j} \\ a_{2j} + b_{2j} \\ a_{3j} + b_{3j} \\ \vdots \\ a_{mj} + b_{mj} \end{bmatrix}$$

$$[S(v_j)]_{\mathcal{B}_W} + [T(v_j)]_{\mathcal{B}_W} = \begin{bmatrix} a_{1j} \\ a_{2j} \\ a_{3j} \\ \vdots \\ a_{mj} \end{bmatrix} + \begin{bmatrix} b_{1j} \\ b_{2j} \\ b_{3j} \\ \vdots \\ b_{mj} \end{bmatrix} = \begin{bmatrix} a_{1j} + b_{1j} \\ a_{2j} + b_{2j} \\ a_{3j} + b_{3j} \\ \vdots \\ a_{mj} + b_{mj} \end{bmatrix}$$

$\therefore \ \phi(S+T) = \phi(S) + \phi(T).$

Also,

$$\alpha T(v_j) = \sum_i \alpha a_{ij} w_i$$

$$[\alpha T(v_j)]_{\mathcal{B}_W} = \begin{bmatrix} \alpha a_{1j} \\ \alpha a_{2j} \\ \alpha a_{3j} \\ \vdots \\ \alpha a_{mj} \end{bmatrix} = \alpha \begin{bmatrix} a_{1j} \\ a_{2j} \\ a_{3j} \\ \vdots \\ a_{mj} \end{bmatrix} = \alpha \, [T(v_j)]_{\mathcal{B}_W}.$$

$\therefore \ \phi(\alpha T) = \alpha \phi(T).$

Hence, ϕ is a linear map.

b) ϕ is injective.

null $\phi = \{ T : \phi(T) = 0_{m\times n} \}$.

$\therefore$ null $\phi = \{0\}$

The null space of ϕ contains the zero transformation; that $0 \in \mathcal{L}(V, W)$ defined as $0(v) = 0$ for all $v \in V$.

c) We are interested to show that ϕ is surjective. Let A be an arbitrary matrix in $\mathbb{F}^{m\times n}$.

$$\text{Define } A := \begin{bmatrix} a_{11} & a_{12} & a_{13} & \cdots & a_{1n} \\ a_{21} & a_{22} & a_{23} & \cdots & a_{2n} \\ a_{31} & a_{32} & a_{33} & \cdots & a_{3n} \\ \vdots \\ a_{m1} & a_{m2} & a_{m3} & \cdots & a_{mn} \end{bmatrix}$$

Then, there exists a linear map $T \in \mathcal{L}(V, W)$, such that $\phi(T) = A$.

Let
$$T(v_j) = \sum_{i=1}^{m} a_{ij} w_i$$
for $j \in \{1, 2, \dots, n\}$.

so, range $\phi = \mathbb{F}^{m\times n}$. ϕ is surjective.

Therefore, ϕ is an ~~one-to-one or~~ invertible linear map as desired. There is one-to-one correspondence between a linear transformation $T \in \mathcal{L}(V, W)$ and its matrix $[T]_{\mathcal{B}_V}^{\mathcal{B}_W}$.

Theorem 3 (Addition of two matrices)

Let S, T be two linear transformations from V into W. $S, T \in \mathcal{L}(V, W)$.

Then,
$$[S+T]_{\mathcal{B}_V}^{\mathcal{B}_W} = [S]_{\mathcal{B}_V}^{\mathcal{B}_W} + [T]_{\mathcal{B}_V}^{\mathcal{B}_W}.$$

Proof.

Let $S(v_j) = \sum_{i=1}^{m} a_{ij} w_i$

$T(v_j) = \sum_{i=1}^{m} b_{ij} w_i$

$$S(v_j) + T(v_j) = \sum_{i=1}^{m} a_{ij} w_i + \sum_{i=1}^{m} b_{ij} w_i$$

$$= \sum_{i=1}^{m} (a_{ij} + b_{ij}) w_i$$

$$[S(v_j)]_{\mathcal{B}_W} + [T(v_j)]_{\mathcal{B}_W} = \begin{bmatrix} a_{1j} \\ a_{2j} \\ \vdots \\ a_{mj} \end{bmatrix} + \begin{bmatrix} b_{1j} \\ b_{2j} \\ \vdots \\ b_{mj} \end{bmatrix} = \begin{bmatrix} a_{1j} + b_{1j} \\ a_{2j} + b_{2j} \\ \vdots \\ a_{mj} + b_{mj} \end{bmatrix} = [(S+T)(v_j)]_{\mathcal{B}_W}.$$

Thus, $[S+T]_{\mathcal{B}_V}^{\mathcal{B}_W} = [S]_{\mathcal{B}_V}^{\mathcal{B}_W} + [T]_{\mathcal{B}_V}^{\mathcal{B}_W}$.

Theorem 4. Scalar multiplication with a matrix.

Let ϕ be a linear map that sends a linear transformation T to its unique matrix A, relative to the basis $\mathcal{B}_V$ and $\mathcal{B}_W$. We claim that-

$$\phi(cT) = c\,\phi(T).$$

Proof.

$$cT(v_j) = \sum c\, a_{ij} w_i$$

$$[cT(v_j)]_{\mathcal{B}_W} = \begin{bmatrix} c a_{1j} \\ c a_{2j} \\ c a_{3j} \\ \vdots \\ c a_{mj} \end{bmatrix}$$

$$= c \begin{bmatrix} a_{1j} \\ a_{2j} \\ a_{3j} \\ \vdots \\ a_{mj} \end{bmatrix} = c\,[T(v_j)]_{\mathcal{B}_W}.$$

$\therefore\ \phi(cT) = c\,\phi(T)$.

Theorem 5. (Product of two matrices).

Consider the linear maps $S: U \to V$ and $T: V \to W$. The composition TS is a linear map from U to W. How can $\phi(TS)$ be computed from $\phi(T)$ and $\phi(S)$? The nicest solution to this question would be to have the following pretty relationship:

$$\phi(TS) = \phi(T) \cdot \phi(S)$$

Theorem 5. The matrix of the composition of two linear transformations is the product of their respective matrices.

$$\phi(TS) = \phi(T) \cdot \phi(S)$$

Proof.

Let $\mathcal{B}_U = \{u_1, \dots, u_n\}$ be a basis of U. $\dim U = n$

Let $\mathcal{B}_V = \{v_1, \dots, v_m\}$ be a basis of V. $\dim V = m$

Let $\mathcal{B}_W = \{w_1, \dots, w_p\}$ be a basis of W $\dim W = p$.

For $k \in \{1, \dots, n\}$,

suppose
$$TS(u_k) = T(S(u_k))$$
$$= T\left(\sum_{j=1}^{m} b_{jk} v_j\right)$$
$$= \sum_{j=1}^{m} b_{jk} T(v_j)$$
$$= \sum_{j=1}^{m} b_{jk}\left(\sum_{i=1}^{p} a_{ij} w_i\right)$$

Thus, $\phi(TS)$ is a $p\times n$ matrix. The matrix of $TS(u_k)$ is given by:

$$[TS(u_k)]_{B_W} = \begin{array}{c} \begin{bmatrix} \sum_{j=1}^{m} a_{1j} b_{jk} \\ \sum_{j=1}^{m} a_{2j} b_{jk} \\ \vdots \\ \sum_{j=1}^{m} a_{pj} b_{jk} \end{bmatrix} \end{array} \begin{array}{c} w_1 \\ w_2 \\ \vdots \\ w_p \end{array}$$

$$= \begin{bmatrix} a_{11} b_{1k} + a_{12} b_{2k} + \dots + a_{1m} b_{mk} \\ a_{21} b_{1k} + a_{22} b_{2k} + \dots + a_{2m} b_{mk} \\ a_{31} b_{1k} + a_{32} b_{2k} + \dots + a_{3m} b_{mk} \\ \vdots \\ a_{p1} b_{1k} + a_{p2} b_{2k} + \dots + a_{pm} b_{mk} \end{bmatrix}$$

~~The~~ for $k \in \{1, 2, \dots, n\}$.

The ith entry in this column vector is-

$$c_{ik} = a_{i1} b_{1k} + a_{i2} b_{2k} + \dots + a_{im} b_{mk}.$$
$$c_{ik} = \sum_j a_{ij} b_{jk}$$

This is conventional way of the multiplying matrices.

From theorem 1, the matrix of the vector $T(x)$ with respect to a basis B_W can be obtained as-

$$[T(x)]_{B_W} = [T]_{B_V}^{B_W} [x]_{B_V}.$$
$$\therefore [T(S(u_k))]_{B_W} = [T]_{B_V}^{B_W} [S(u_k)]_{B_V}$$
$$\therefore [T(S(u_k))]_{B_W} = [T]_{B_V}^{B_W} [S]_{B_U}^{B_V} [u_k]_{B_U}.$$

We can express the left hand sides as-

$$[TS(u_k)]_{B_W} = [TS]_{B_U}^{B_W} [u_k]_{B_U}$$
$$\Rightarrow [TS]_{B_U}^{B_W} [u_k]_{B_U} = [T]_{B_V}^{B_W} [S]_{B_U}^{B_V} [u_k]_{B_U}$$
$$\therefore [TS]_{B_U}^{B_W} = [T]_{B_V}^{B_W} [S]_{B_U}^{B_V}.$$

Properties of matrix multiplication

1. Matrix multiplication is associative.
$$P(QR) = (PQ)R.$$
2. As a law, commutativity fails to be true for matrix multiplication. For special pairs PQ, it may be true that $PQ = QP$, but it does not hold in general.
3. Matrix multiplication is distributive.
$$A(B+C) = AB + AC$$
$$(A+B)C = AC + BC.$$

Proof.

1. Let $P = (p_{ij})$, $Q = (q_{jk})$, $R = (r_{ml})$.
The proof is a direct computation.
The element in the (i,k) place of $PQ = (c_{ik})$ is:
$$c_{ik} = \sum_j p_{ij} q_{jk}$$
The element in the (i,l) place of $(PQ)R = (d_{il})$ is:
$$d_{il} = \sum_k c_{ik} \cdot r_{kl}$$
$$= \sum_k \sum_j p_{ij} q_{jk} \cdot r_{kl}$$
The element in the (i,l) place of $P(QR)$ may be computed similarly.
First, the element in the (j,l) place of $QR = (c_{jl})$ is-
$$c_{jl} = \sum_k q_{jk} \cdot r_{kl}$$
Hence, the element in the (i,l) place of $P(QR) = (d_{il})$ is -
$$d_{il} = \sum_j p_{ij} c_{jl}$$
$$= \sum_j \sum_k p_{ij} \cdot q_{jk} \cdot r_{kl}.$$
This completes the proof and shows that the triple product PQR is unambiguous even when the parenthesis are omitted.

(2) We shall study at length later in the course, when the product of two matrices P and Q is commutative, but for now, we easily see that-
$$\sum_j p_{ij} q_{jk} \neq \sum_j q_{ij} p_{jk}$$
That is:
$$p_{i1} q_{1k} + p_{i2} q_{2k} + \dots \neq p_{1i} q_{i1} + p_{2i} q_{i2} + \dots$$

(3) Matrix multiplication is distributive.
$A(B+C) = AB + AC.$
Let $A = (a_{ij})$, $B = (b_{jk})$, $C = (c_{jk})$
The element in the (i,k) place of $A(B+C)$ is given by:
$$\sum_j a_{ij} \cdot (b_{jk} + c_{jk})$$
$$= \sum_j a_{ij} b_{jk} + \sum_j a_{ij} c_{jk}$$
So, the element in (i,k) place $A(B+C)$ is sum of the elements in (i,k) place of AB and AC.
Thus, $A(B+C) = AB + AC$.

Claim. Let $A = (a_{ij})$, $B = (b_{ij})$, $C = (c_{jk})$.
Let d_{ik} be the element in the (i,k) place of $(A+B)C$.

Then, $d_{ik} = \sum_j (a_{ij} + b_{ij}) \cdot c_{jk}$

$= \sum_j a_{ij} \cdot c_{jk} + \sum_j b_{ij} \cdot c_{jk}$

$=$ Element in (i, k) place of $AC +$ Element in (i, k) place of BC.

$= AC + BC$.

Partitioned Matrices.

We are interested to find an expression for the ith row and the kth column of $C = AB$. The element in the (i, k) place of $C = (c_{ik})$ is found by taking the inner product of row i of A with the column k of B. Let $A \in \mathbb{R}^{m \times n}$, $B \in \mathbb{R}^{n \times p}$.

$$c_{ik} = (a_{i1}, a_{i2}, \dots, a_{in}) \begin{bmatrix} b_{1k} \\ b_{2k} \\ \vdots \\ b_{nk} \end{bmatrix}$$

Let $C = (C_1, C_2, \dots, C_p)$ and $A = (A_1, A_2, \dots, A_n)$.
Let us find an expression for C_k.

$$C_k = \begin{bmatrix} a_{11} & a_{12} & a_{13} & \dots & a_{1n} \\ a_{21} & a_{22} & a_{23} & \dots & a_{2n} \\ a_{31} & a_{32} & a_{33} & \dots & a_{3n} \\ \vdots \\ a_{m1} & a_{m2} & a_{m3} & \dots & a_{mn} \end{bmatrix} \begin{bmatrix} b_{1k} \\ b_{2k} \\ b_{3k} \\ \vdots \\ b_{nk} \end{bmatrix} = A(:, j) \times B(j, k)$$

$$= A_1 b_{1k} + A_2 b_{2k} + \dots + A_n b_{nk}.$$

Similarly, if

$$C = \begin{bmatrix} C_1 \\ C_2 \\ \vdots \\ C_m \end{bmatrix}$$

$$C_i = \begin{bmatrix} a_{i1} & a_{i2} & \dots & a_{in} \end{bmatrix} \begin{bmatrix} b_{11} & b_{12} & b_{13} & \dots & b_{1p} \\ b_{21} & b_{22} & b_{23} & \dots & b_{2p} \\ b_{31} & b_{32} & b_{33} & \dots & b_{3p} \\ \vdots \\ b_{n1} & b_{n2} & b_{n3} & \dots & b_{np} \end{bmatrix}$$

$= a_{i1} (b_{11}, b_{12}, \dots, b_{1p})$
$+ a_{i2} (b_{21}, b_{22}, \dots, b_{2p})$
$+ \dots$
$+ a_{in} (b_{n1}, b_{n2}, \dots, b_{np})$

$= a_{i1} B_1 + a_{i2} B_2 + \dots + a_{in} B_n$

Thus, any column ~~k can be expressed as the linear~~ of $C = AB$ can be expressed as a linear combination of the columns of A.
Any row i of the product $C = AB$ can be expressed as the linear combination of the rows of B.

Transpose of a matrix

The transpose of a matrix $A = (a_{ij})$ is defined as:

$$A^T = (a_{ji})$$

Theorem 6. Let $A, B \in \mathbb{R}^{m \times n}$ and $\alpha \in \mathbb{R}$. Then

1) $(A+B)^T = A^T + B^T$
2) $(\alpha A)^T = \alpha A^T$
3) $(A^T)^T = A$
4) $(AB)^T = B^T A^T$.

Proof.

1) $$\begin{aligned}(A+B)^T &= (a_{ij} + b_{ij})^T \\ &= (a_{ji} + b_{ji}) \\ &= (a_{ji}) + (b_{ji}) \\ &= (a_{ij})^T + (b_{ij})^T \\ &= A^T + B^T.\end{aligned}$$

2) $$\begin{aligned}(\alpha A^T) &= (\alpha a_{ij})^T \\ &= (\alpha a_{ji}) \\ &= \alpha (a_{ji}) \\ &= \alpha (a_{ij})^T \\ &= \alpha A^T.\end{aligned}$$

3) $$\begin{aligned}(A^T)^T &= ((a_{ij})^T)^T \\ &= (a_{ji})^T \\ &= (a_{ij}) \\ &= A\end{aligned}$$

4) ~~(AB)~~ We are interested to prove the reverse order law.
Let $C = (c_{ik})$.

$$C^T = (c_{ik})^T = c_{ki}$$

$$= \sum_j a_{kj} b_{ji}$$

This ~~the~~ is the inner product of the k-th row of A with the ith column of B.

~~Let $A = (a_{jk})$, k~~

But, b_{ji} is the element in the (i,j) place of B^T. a_{kj} is ~~(j,k)~~ element of the (j,k) place of A^T.

Let $b_{ji} = b'_{ij}$ ~~where~~ then $B^T = (b'_{ij})$
and $a_{kj} = a'_{jk}$ then $A^T = (a'_{jk})$.

$$\therefore (C^T_{ik}) = \sum_j b'_{ij} a'_{jk}$$

$$\Rightarrow C^T = B^T \cdot A^T.$$

Square Matrices.
A square matrix A is a matrix with the same number of rows and columns. $\mathbb{R}^{n\times n}$ is the vector space of all $n\times n$ square matrices.

Symmetric Matrices.
A square matrix $A \in \mathbb{R}^{n\times n}$ is said to be symmetric, if and only if
$$A = A^T.$$

Skew-symmetric Matrix.
Let A be a square matrix of order $n\times n$. $A \in \mathbb{R}^{n\times n}$. The matrix A is said to be skew-symmetric, if and only if
$$A = -A^T.$$

Hermitian and Skew-Hermitian Matrices.
Let A be a matrix over the field $\mathbb{C}$ of complex numbers.

(1) A is said to be a Hermitian matrix, if and only if:
$$A = (\bar{A})^T$$
that is $$A = A^*$$

(2) A is said to be skew-hermitian if and only if:
$$A = -(\bar{A})^T.$$

Diagonal, upper and lower triangular matrices.
Let $A\ T \in \mathcal{L}(V,W)$ be a linear transformation. suppose $\dim V = \dim W = n$. The transformation T is defined as follows—

(1) $Tv_j = \sum_{i=1} a_{ij} w_i$ where $a_{ij} = 0$ when $i \neq j$.

Then, $Tv_1 = a_{11}w_1 + 0\cdot w_2 + \ldots + 0\cdot w_n$
$Tv_2 = 0\cdot w_1 + a_{22}w_2 + \ldots + 0\cdot w_n$
$\vdots$
$Tv_n = 0\cdot w_1 + 0\cdot w_2 + \ldots + a_{nn}w_n.$

Therefore,
$$[T]_{B_2} = \begin{bmatrix} a_{11} & 0 & 0 & \ldots & 0 \\ 0 & a_{22} & 0 & \ldots & 0 \\ 0 & 0 & a_{33} & \ldots & 0 \\ & & & \ddots & \vdots \\ 0 & 0 & 0 & \ldots & a_{nn} \end{bmatrix}$$

Such a matrix is called a diagonal matrix. All elements off the principal diagonal are zero.

(2) $Tv_j = \sum_{i=1}^{n} a_{ij} w_i$ where $a_{ij} = 0$ for $i > j$.

Then, $Tv_1 = a_{11}w_1 + 0\cdot w_2 + 0\cdot w_3 + \ldots + 0\cdot w_n$
$Tv_2 = a_{12}w_1 + a_{22}w_2 + 0\cdot w_3 + \ldots + 0\cdot w_n$
$Tv_3 = a_{13}w_1 + a_{23}w_2 + a_{33}w_3 + \ldots + 0\cdot w_n$
$\vdots$
$Tv_n = a_{1n}w_1 + a_{2n}w_2 + a_{3n}w_3 + \ldots + a_{nn}w_n.$

$$[T]_{\mathcal{B}_W} = \begin{bmatrix} a_{11} & a_{12} & a_{13} & \cdots & a_{1n} \\ 0 & a_{22} & a_{23} & \cdots & a_{2n} \\ 0 & 0 & a_{33} & & a_{3n} \\ \vdots & & & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & a_{nn} \end{bmatrix}$$

Such a matrix is called an upper triangular matrix.

(3) Similarly, if

$$T(v_j) = \sum_{i=1}^{m} a_{ij} w_i \text{ and } a_{ij} = 0 \text{ when } i < j.$$

Then

$$T(v_1) = a_{11} \cdot w_1 + a_{21} w_2 + a_{31} w_3 + \dots + a_{m1} w_n$$
$$T(v_2) = 0 \cdot w_1 + a_{22} w_2 + a_{32} w_3 + \dots + a_{n2} w_n$$
$$T(v_3) = 0 \cdot w_1 + 0 \cdot w_2 + a_{33} w_3 + \dots + a_{n3} w_n$$
$$\vdots$$
$$T(v_n) = 0 \cdot w_1 + 0 \cdot w_2 + 0 \cdot w_3 + \dots + a_{nn} w_n$$

$$[T]_{\mathcal{B}} = \begin{bmatrix} a_{11} & 0 & 0 & \cdots & 0 \\ a_{21} & a_{22} & 0 & \cdots & 0 \\ a_{31} & a_{32} & a_{33} & \cdots & 0 \\ & & & \ddots & \\ a_{n1} & a_{n2} & a_{n3} & \cdots & a_{nn} \end{bmatrix}$$

such a matrix is called a lower triangular matrix.

Inverse of a matrix

Theorem 7: Let $T: V \to V$ be an invertible linear operator and $\mathcal{B}$ be a basis of V. Then, the inverse of a matrix is given by —

$$[T^{-1}]_{\mathcal{B}} = [T]_{\mathcal{B}}^{-1}.$$

Proof.

Since T is invertible, there is and inverse operator S, such that $ST = I$ and $TS = I$. Remember, this holds true, if a linear transformation has an inverse. There are no matrices here.

$$[TS]_{\mathcal{B}} = [I]_{\mathcal{B}}.$$

Invoking the formula $[TS]_{\mathcal{B}} = [T]_{\mathcal{B}} [S]_{\mathcal{B}}$, we have:

$$[T]_{\mathcal{B}} [S]_{\mathcal{B}} = [I]_{\mathcal{B}}.$$

We can also do this for the matrix in the second equation ST and write down —

$$[S]_{\mathcal{B}} [T]_{\mathcal{B}} = [I]_{\mathcal{B}}.$$

What this means is, $[S]_{\mathcal{B}}$ is the inverse of the matrix $[T]_{\mathcal{B}}$.

That is, $[S]_{\mathcal{B}} = [T]_{\mathcal{B}}^{-1}$.

But, $S = T^{-1}$

$$\therefore \quad [T^{-1}]_{\mathcal{B}} = [T]_{\mathcal{B}}^{-1}.$$

Definition.

A square matrix $A \in \mathbb{R}^{n\times n}$ is said to be invertible, iff and only if there exists a unique matrix B such that:

$$A \cdot B = I = B \cdot A.$$

The matrix B is called the inverse of the matrix A.

$$B = A^{-1}.$$

Note that: I_n is the inverse of the identity matrix itself, since

$$I_n \cdot I_n = I_n.$$

The zero matrix O has no inverse, since

$$O \cdot A = O$$

for all matrices A.

Properties of the inverse of a matrix.

Let A, B be invertible matrices of order ~~n~~ n.

(1) $(A^{-1})^{-1} = A$.

(2) $(A^{T})^{-1} = (A^{-1})^{T}$

(3) $(AB)^{-1} = B^{-1}A^{-1}$.

Proof.

(1) By definition $A \cdot A^{-1} = I = A^{-1} \cdot A$.

Thus, the matrix A is the inverse of A^{-1}.

So, $(A^{-1})^{-1} = A$.

(2) Claim. $(A^{T})^{-1} = (A^{-1})^{T}$

$A \cdot A^{-1} = I$.

$(A \cdot A^{-1})^{T} = I^{T}$

$(A^{-1})^{T} \cdot A^{T} = I$.

Thus, A^{T} is the inverse of $(A^{-1})^{T}$.

In other words: $(A^{T})^{-1} = (A^{-1})^{T}$.

(3) Reverse order law. Let ~~B~~ A and B ~~as~~ be any two matrices.

~~AB~~ $A \cdot A^{-1} = I = A^{-1}A$

$B \cdot B^{-1} = I = B^{-1}B$.

Therefore,

$$A^{-1}A = I.$$
$$A^{-1}AB = IB.$$
$$A^{-1}AB = B.$$
$$B^{-1}A^{-1}AB = B^{-1}B$$
$$(B^{-1}A^{-1})(AB) = I. \qquad (a)$$

And

$$\cancel{AA^{-1} = I}$$
$$BB^{-1} = I$$
$$ABB^{-1} = AI$$
$$ABB^{-1} = A$$
$$ABB^{-1}A^{-1} = AA^{-1}$$
$$(AB)(B^{-1}A^{-1}) = I \qquad (b)$$

So, $(AB)(B^{-1}A^{-1}) = I = (B^{-1}A^{-1})(AB)$.

Thus, $(AB)^{-1} = B^{-1}A^{-1}$.

How do matrices corresponding to different bases behave? The answer to this given in the next result.

Theorem 8. Let $T: V \to V$ be a linear operator and $\mathcal{B}_1, \mathcal{B}_2$ be the bases of V. Look at the matrix of the identity transformation with respect to the bases $\mathcal{B}_1$ and $\mathcal{B}_2$.

$$M = [I]_{\mathcal{B}_1}^{\mathcal{B}_2}.$$

Note that, this is not the identity matrix. Then, we have the following -

$$[x]_{\mathcal{B}_2} = M\,[x]_{\mathcal{B}_1}.$$

And further

$$[T]_{\mathcal{B}_2} = M\,[T]_{\mathcal{B}_1}\,M^{-1}.$$

Remember that this involves M^{-1}, so we must show that M is invertible. We prove this at length later.

Proof.

1) $$\begin{aligned}[x]_{\mathcal{B}_2} &= [I(x)]_{\mathcal{B}_2} \\ &= [I]_{\mathcal{B}_1}^{\mathcal{B}_2}\,[x]_{\mathcal{B}_1}. \quad \text{Recalling that } [T(x)]_{\mathcal{B}_W} = [T]_{\mathcal{B}_V}^{\mathcal{B}_W}\,[x]_{\mathcal{B}_V}. \\ &= M\,[x]_{\mathcal{B}_1}\end{aligned}$$

2) Consider $[T(x)]_{\mathcal{B}_2}$.

Let $y = T(x)$. Appealing to the previous formula, y relative to $\mathcal{B}_2$ is M times y relative to $\mathcal{B}_1$.

$$[Tx]_{\mathcal{B}_2} = M\,[Tx]_{\mathcal{B}_1}$$

$$[T]_{\mathcal{B}_2}[x]_{\mathcal{B}_2} = M\,[T]_{\mathcal{B}_1}[x]_{\mathcal{B}_1} \qquad \text{since } [T(x)]_{\mathcal{B}} = [T]_{\mathcal{B}}\,[x]_{\mathcal{B}}.$$

$$\therefore [T]_{\mathcal{B}_2}\,M\,[x]_{\mathcal{B}_1} = M\,[T]_{\mathcal{B}_1}\,[x]_{\mathcal{B}_1}$$

Let $\mathcal{B}_1 = \{e_1, e_2, \dots, e_n\}$.
If we replace x by the basis vectors $e_1, e_2, \dots, e_j, \dots, e_n$, then the left hand side $[T]_{\mathcal{B}_2} M e_j$ represents the jth column of the matrix $[T]_{\mathcal{B}_2} M$. Similarly, $M\,[T]_{\mathcal{B}_1} e_j$ represents the jth column of the matrix $M\,[T]_{\mathcal{B}_1}$. Thus, the columns ~~are~~ of the matrix $[T]_{\mathcal{B}_2} M$ and $M\,[T]_{\mathcal{B}_1}$ are equal.

Therefore,
$$\begin{aligned}[T]_{\mathcal{B}_2} M &= M\,[T]_{\mathcal{B}_1}. \\ [T]_{\mathcal{B}_2} M \cdot M^{-1} &= M\,[T]_{\mathcal{B}_1} M^{-1} \\ [T]_{\mathcal{B}_2} I &= M\,[T]_{\mathcal{B}_1} M^{-1} \\ [T]_{\mathcal{B}_2} &= M\,[T]_{\mathcal{B}_1} M^{-1}.\end{aligned}$$

This closes the proof.

Similar matrices. (Definition).

If B and A are $n \times n$ matrices over $\mathbb{F}$, B is said to be similar to A, if there is a non-singular matrix P with elements in $\mathbb{F}$ such that
$$B = PAP^{-1}.$$

Lemma. The relation of similarity is an R,S,T relation.

Proof.

(1) A matrix is similar to itself, since
$$A = IAI^{-1}.$$

(2) If A is similar to B, B is similar to A.
Assume that,
$$\begin{aligned}
A &= MBM^{-1} \\
AM &= MBM^{-1}M \\
AM &= MB(M^{-1}M) \\
AM &= MBI \\
AM &= MB \\
M^{-1}AM &= M^{-1}MB \\
&= (M^{-1}M)B \\
&= IB \\
M^{-1}AM &= B.
\end{aligned}$$
Hence, B is similar to A.

(3) If A is similar to B, and B is similar to C, then A is similar to C.
Suppose $A = MBM^{-1}$
$B = NCN^{-1}$

$$\begin{aligned}
\Rightarrow A &= MBM^{-1} \\
&= M(NCN^{-1})M^{-1} \\
&= (MN)C(N^{-1}M^{-1}) \\
&= (MN)C(MN)^{-1}
\end{aligned}$$
$\therefore$ A is similar to C.

Hence, similarity is an equivalence relation.

If two matrices are ~~a~~ similar (but not equal), they share many common properties. ~~We shall study these at length later in the course.~~ What does it preserve? It preserves what are called eigen-values. That's something which we shall discuss at length later in the course.

Example.
In $\mathbb{R}^2$, $\mathcal{B}_1 = \{(1,0),(0,1)\}$ is a standard basis. We can show that $\mathcal{B}_2 = \{(\cos\theta, \sin\theta), (-\sin\theta, \cos\theta)\}$ obtained by the rotation of the standard basis vectors through an angle θ also forms a basis of $\mathbb{R}^2$. Find the matrix M.

~~Proof~~
Solution $M = [I]_{\mathcal{B}_1}^{\mathcal{B}_2}$

$$\begin{aligned}
I(e_1) &= e_1 = (1,0) \\
&= (\cos\theta)(\cos\theta, \sin\theta) + (-\sin\theta)(-\sin\theta, \cos\theta)
\end{aligned}$$

$$\begin{aligned}
I(e_2) &= e_2 = (0,1) \\
&= (\sin\theta)(\cos\theta, \sin\theta) + (\cos\theta)(-\sin\theta, \cos\theta)
\end{aligned}$$

$$[I]_{\mathcal{B}_1}^{\mathcal{B}_2} = \begin{bmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{bmatrix}.$$

Change of Basis Matrix (Definition). The matrix $M = [I]_{\mathcal{B}_1}^{\mathcal{B}_2}$ is called the matrix of the change of basis from $\mathcal{B}_1$ to $\mathcal{B}_2$.

Let V be a finite dimensional vector space and let

$$\mathcal{B}_1 = \{u_1, \dots, u_n\}$$
$$\mathcal{B}_2 = \{w_1, \dots, w_n\}$$

be bases of V.

Suppose $u_j = \sum a_{ij} w_i$. Let the coordinates of the vector x w.r.t $\mathcal{B}_1$ be $(x_1, x_2, \dots, x_n)$. Then,

$$x = x_1 u_1 + x_2 u_2 + x_3 u_3 + \dots + x_n u_n$$
$$= \sum_j x_j u_j$$
$$= \sum_j x_j \left(\sum_i a_{ij} w_i \right) = \sum_i \left(\sum_j a_{ij} x_j \right) w_i$$

$$\therefore [x]_{\mathcal{B}_2} = \begin{bmatrix} a_{11} x_1 + a_{12} x_2 + \dots + a_{1n} x_n \\ a_{21} x_1 + a_{22} x_2 + \dots + a_{2n} x_n \\ \vdots \\ a_{n1} x_1 + a_{n2} x_2 + \dots + a_{nn} x_n \end{bmatrix} = \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & & & \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}$$

$$[x]_{\mathcal{B}_2} = M \cdot [x]_{\mathcal{B}_1}$$

$$M = [I]_{\mathcal{B}_1}^{\mathcal{B}_2} = \begin{bmatrix} a_{11} & a_{12} & a_{13} & \dots & a_{1n} \\ a_{21} & a_{22} & a_{23} & & a_{2n} \\ a_{31} & a_{32} & a_{33} & & a_{3n} \\ \vdots & & & \ddots & \\ a_{n1} & a_{n2} & a_{n3} & \dots & a_{nn} \end{bmatrix}$$

Example

1) Let $\mathcal{B}_1$ be a standard basis for $\mathbb{R}^3$ and $\mathcal{B}_2$ be another basis such that:

$$M = [I]_{\mathcal{B}_1}^{\mathcal{B}_2} = \begin{bmatrix} 0 & 1 & 1 \\ 1 & 1 & 0 \\ 0 & 0 & 3 \end{bmatrix}$$

What are elements of $\mathcal{B}_2$?

Solution

Let $\mathcal{B}_1 = \{e_1, e_2, e_3\}$, $\mathcal{B}_2 = \{f_1, f_2, f_3\}$.

$$e_1 = 0 \cdot f_1 + 1 \cdot f_2 + 0 \cdot f_3$$
$$e_2 = 1 \cdot f_1 + 1 \cdot f_2 + 0 \cdot f_3$$
$$e_3 = 1 \cdot f_1 + 0 \cdot f_2 + 3 f_3$$

$$\therefore f_2 = (1,0,0)$$
$$f_1 = (0,1,0) - (1,0,0) = (-1,1,0)$$

and

$$3f_3 = (0,0,1) - (-1,1,0)$$
$$= (1,-1,1)$$
$$f_3 = \frac{1}{3}(1,-1,1).$$

$$\mathcal{B}_2 = \left\{ (-1,1,0), (1,0,0), \frac{1}{3}(1,-1,1) \right\}.$$

2) Consider the following two bases of $\mathbb{R}^2$:

$$S = \{u_1, u_2\} = \{(1,2), (3,5)\}$$
$$S' = \{v_1, v_2\} = \{(1,-1), (1,-2)\}.$$

Find the change-of-basis matrix $[I]_S^{S'}$ and $[I]_{S'}^{S}$.

Solution.

Let the change of basis matrix from S to S' be

$$[I]_{S_1}^{S_2} = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}.$$

We have:

$$v_1 = a_{11} u_1 + a_{21} u_2$$

$$\begin{pmatrix} 1 \\ -1 \end{pmatrix} = a_{11}\begin{pmatrix} 1 \\ 2 \end{pmatrix} + a_{21}\begin{pmatrix} 3 \\ 5 \end{pmatrix}$$

$$\therefore \quad 1 = a_{11} + 3a_{21}$$
$$-1 = 2a_{11} + 5a_{21}$$

yielding $a_{11} = -8$, $a_{21} = 3$.

And $v_2 = a_{12} u_1 + a_{22} u_2$

$$\begin{pmatrix} 1 \\ -2 \end{pmatrix} = a_{12}\begin{pmatrix} 1 \\ 2 \end{pmatrix} + a_{22}\begin{pmatrix} 3 \\ 5 \end{pmatrix}$$

$$1 = a_{12} + 3a_{22}$$
$$-2 = 2a_{12} + 5a_{22}$$

yielding $a_{22} = 4$, $a_{12} = -11$.

$$[I]^{\mathcal{B}_2}_{\mathcal{B}_1} = \begin{bmatrix} -8 & -11 \\ 3 & 4 \end{bmatrix}$$ is the change of basis matrix from S to S'.

Let the change of basis matrix from S' to S be

$$[I]^{S}_{S'} = \begin{bmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{bmatrix}$$

$$u_1 = b_{11} v_1 + b_{21} v_2$$

$$\begin{pmatrix} 1 \\ 2 \end{pmatrix} = b_{11} \begin{pmatrix} 1 \\ -1 \end{pmatrix} + b_{21} \begin{pmatrix} 1 \\ -2 \end{pmatrix}$$

$$1 = b_{11} + b_{21}$$
$$2 = -b_{11} - 2b_{21}$$

yielding $b_{11} = 4$, $b_{21} = -3$

$$u_2 = b_{12} v_1 + b_{22} v_2$$

$$\begin{pmatrix} 3 \\ 5 \end{pmatrix} = b_{12} \begin{pmatrix} 1 \\ -1 \end{pmatrix} + b_{22} \begin{pmatrix} 1 \\ -2 \end{pmatrix}$$

yielding $b_{12} = 11$, $b_{22} = -8$

$$[I]^{S}_{S'} = \begin{bmatrix} 4 & 11 \\ -3 & -8 \end{bmatrix}.$$

Theorem 9

Let $\mathcal{B}_1, \mathcal{B}_2, \mathcal{B}_3$ be the three bases of V. Then,

$$[I]^{\mathcal{B}_3}_{\mathcal{B}_2} \cdot [I]^{\mathcal{B}_2}_{\mathcal{B}_1} = [I]^{\mathcal{B}_3}_{\mathcal{B}_1}.$$

Proof.

$$[x]_{\mathcal{B}_2} = [I]^{\mathcal{B}_2}_{\mathcal{B}_1} [x]_{\mathcal{B}_1}$$

$$[x]_{\mathcal{B}_3} = [I]^{\mathcal{B}_3}_{\mathcal{B}_2} [x]_{\mathcal{B}_2}$$

$$= [I]^{\mathcal{B}_3}_{\mathcal{B}_2} \cdot [I]^{\mathcal{B}_2}_{\mathcal{B}_1} [x]_{\mathcal{B}_1} \qquad (a)$$

But, $[x]_{\mathcal{B}_3} = [I]^{\mathcal{B}_3}_{\mathcal{B}_1} [x]_{\mathcal{B}_1}$ (b)

From (a) and (b),

$$[I]^{\mathcal{B}_3}_{\mathcal{B}_1} [x]_{\mathcal{B}_1} = [I]^{\mathcal{B}_3}_{\mathcal{B}_2} [I]^{\mathcal{B}_2}_{\mathcal{B}_1} [x]_{\mathcal{B}_1}.$$

If we replace $x = e_j$, both the left hand side and right side represent the jth columns of a matrix. Therefore,

$$[I]^{\mathcal{B}_3}_{\mathcal{B}_1} = [I]^{\mathcal{B}_3}_{\mathcal{B}_2} [I]^{\mathcal{B}_2}_{\mathcal{B}_1}.$$

Corollary.

Let $\mathcal{B}_1, \mathcal{B}_2$ be bases of V. Then,

$$[I]^{\mathcal{B}_1}_{\mathcal{B}_2} [I]^{\mathcal{B}_2}_{\mathcal{B}_1} = I = [I]^{\mathcal{B}_2}_{\mathcal{B}_1} [I]^{\mathcal{B}_1}_{\mathcal{B}_2}.$$

Proof.

consider

$$[x]_{\mathcal{B}_1} = [I]_{\mathcal{B}_2}^{\mathcal{B}_1} [x]_{\mathcal{B}_2}$$

$$= [I]_{\mathcal{B}_2}^{\mathcal{B}_1} [I]_{\mathcal{B}_1}^{\mathcal{B}_2} [x]_{\mathcal{B}_1}.$$

$$\therefore \; [I]_{\mathcal{B}_2}^{\mathcal{B}_1} [I]_{\mathcal{B}_1}^{\mathcal{B}_2} = I.$$

Also consider

$$[x]_{\mathcal{B}_2} = [I]_{\mathcal{B}_1}^{\mathcal{B}_2} [x]_{\mathcal{B}_1}$$

$$= [I]_{\mathcal{B}_1}^{\mathcal{B}_2} [I]_{\mathcal{B}_2}^{\mathcal{B}_1} [x]_{\mathcal{B}_2}$$

$$\therefore \; [I]_{\mathcal{B}_1}^{\mathcal{B}_2} [I]_{\mathcal{B}_2}^{\mathcal{B}_1} = I$$

Thus, $[I]_{\mathcal{B}_2}^{\mathcal{B}_1} [I]_{\mathcal{B}_1}^{\mathcal{B}_2} = I = [I]_{\mathcal{B}_1}^{\mathcal{B}_2} [I]_{\mathcal{B}_2}^{\mathcal{B}_1}$

How does the change of basis affect a linear transformation on finite dimensional vector spaces?

Theorem 10.

Let $T \in \mathcal{L}(V, W)$ be a linear transformation from V into W.
Let $\mathcal{B}_1 = \{e_1, e_2, \dots, e_n\}$ and $\mathcal{B}_2 = \{f_1, f_2, \dots, f_n\}$ be a pair of basis of V and W respectively. Let $\mathcal{B}_1' = \{e_1', e_2', \dots, e_n'\}$ and $\mathcal{B}_2' = \{f_1', f_2', \dots, f_n'\}$ be another pair of basis of V and W.
Then,

$$[T]_{\mathcal{B}_1'}^{\mathcal{B}_2'} = [I]_{\mathcal{B}_2}^{\mathcal{B}_2'} [T]_{\mathcal{B}_1}^{\mathcal{B}_2} [I]_{\mathcal{B}_1'}^{\mathcal{B}_1}$$

Proof.

consider

$$[T]_{\mathcal{B}_1'}^{\mathcal{B}_2'} = [I_W T \cdot I_V]_{\mathcal{B}_1'}^{\mathcal{B}_2'}$$

$$= [I_W] [T]_{\mathcal{B}_1'}^{\mathcal{B}_2'} [I_V]$$

$$= [I_W]_{\mathcal{B}_2'}^{\mathcal{B}_2'} [I_W]_{\mathcal{B}_2'}^{\mathcal{B}_2} [T]_{\mathcal{B}_1'}^{\mathcal{B}_2'} [I]_{\mathcal{B}_1}^{\mathcal{B}_1'} [I]_{\mathcal{B}_1'}^{\mathcal{B}_1}$$

Now, consider the product $[I]_{\mathcal{B}_2'}^{\mathcal{B}_2} [T]_{\mathcal{B}_1'}^{\mathcal{B}_2'} [I]_{\mathcal{B}_1}^{\mathcal{B}_1'}$.

we can write:

$$[T(x)]_{\mathcal{B}_2} = [I_W]_{\mathcal{B}_2'}^{\mathcal{B}_2} [T(x)]_{\mathcal{B}_2'}$$

$$= [I_W]_{\mathcal{B}_2'}^{\mathcal{B}_2} [T]_{\mathcal{B}_1'}^{\mathcal{B}_2'} [x]_{\mathcal{B}_1'}$$

$$[T]_{\mathcal{B}_1}^{\mathcal{B}_2} [x]_{\mathcal{B}_1} = [I_W]_{\mathcal{B}_2'}^{\mathcal{B}_2} [T]_{\mathcal{B}_1'}^{\mathcal{B}_2'} [I_V]_{\mathcal{B}_1}^{\mathcal{B}_1'} [x]_{\mathcal{B}_1}$$

$$\therefore \; [T]_{\mathcal{B}_1}^{\mathcal{B}_2} = [I_W]_{\mathcal{B}_2'}^{\mathcal{B}_2} [T]_{\mathcal{B}_1'}^{\mathcal{B}_2'} [I_V]_{\mathcal{B}_1}^{\mathcal{B}_1'}$$

so, the middle term can be replaced by $[T]_{\mathcal{B}_1}^{\mathcal{B}_2}$.

$$\therefore \; [T]_{\mathcal{B}_1'}^{\mathcal{B}_2'} = [I_W]_{\mathcal{B}_2}^{\mathcal{B}_2'} [T]_{\mathcal{B}_1}^{\mathcal{B}_2} [I]_{\mathcal{B}_1'}^{\mathcal{B}_1}.$$

Example. Let $\mathcal{P}_3(\mathbb{R})$ the space of all polynomials over the field $\mathbb{R}$ of real numbers, which have degree lesser than or equal to 3. Let D be the differentiation operator on $\mathcal{P}_3(\mathbb{R})$ defined by

$$Dp = p' \quad \text{for all } p \in \mathcal{P}_3(\mathbb{R}).$$

Let $\mathcal{B} = \{f_1, f_2, f_3, f_4\}$ be an ordered basis of defined by:

$f_1(x) = 1$
$f_2(x) = x$
$f_3(x) = x^2$
$f_4(x) = x^3$

Let c be a fixed real constant and define the polynomials $g_i(x) = (x+c)^{i-1}$ (all basis functions shifted by a distance c). Suppose $\mathcal{B}' = \{g_1, g_2, g_3, g_4\}$.

$g_1(x) = 1$
$g_2(x) = x + c$
$g_3(x) = x^2 + 2cx + c^2$
$g_4(x) = x^3 + 3cx^2 + 3c^2x + c^3$.

The change of basis matrix can be found as follows:

$g_1 = f_1$
$g_2 = c \cdot f_1 + 1 \cdot f_2$
$g_3 = c^2 \cdot f_1 + 2c f_2 + 1 \cdot f_3$.
$g_4 = c^3 f_1 + 3c^2 f_2 + 3c f_3 + f_4$

$$M = [I]_{\mathcal{B}_1}^{\mathcal{B}_2} = \begin{bmatrix} 1 & c & c^2 & c^3 \\ 0 & 1 & 2c & 3c^2 \\ 0 & 0 & 1 & 3c \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

This matrix is seen to be easily invertible and its inverse is –

$$M^{-1} = \begin{bmatrix} 1 & -c & +c^2 & -c^3 \\ 0 & 1 & -2c & 3c^2 \\ 0 & 0 & 1 & -3c \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

~~Let~~ The action of a linear transformation is completely determined by its action on the basis vectors.

$Df_1 = (1)' = 0 = 0 \cdot f_1 + 0 \cdot f_2 + 0 \cdot f_3 + 0 \cdot f_4$
$Df_2 = (x)' = 1 = 1 \cdot f_1 + 0 \cdot f_2 + 0 \cdot f_3 + 0 \cdot f_4$
$Df_3 = (x^2)' = 2x = 0 \cdot f_1 + 2 \cdot f_2 + 0 \cdot f_3 + 0 \cdot f_4$
$Df_4 = (x^3)' = 3x^2 = 0 \cdot f_1 + 0 \cdot f_2 + 3 f_3 + 0 \cdot f_4$.

$$[D]_{\mathcal{B}} = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 3 \\ 0 & 0 & 0 & 0 \end{bmatrix}.$$

The matrix of the differentiation of operator relative to the ordered basis $\mathcal{B}'$ is, therefore:

$$[D]_{\mathcal{B}'} = M^{-1} [D]_{\mathcal{B}} M$$

$$= \begin{bmatrix} 1 & -c & c^2 & -c^3 \\ 0 & 1 & -2c & 3c^2 \\ 0 & 0 & 1 & -3c \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 3 \\ 0 & 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} 1 & c & c^2 & c^3 \\ 0 & 1 & 2c & 3c^2 \\ 0 & 0 & 1 & 3c \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

$$= \begin{bmatrix} 0 & 1 & -2c & 3c^2 \\ 0 & 0 & 2 & -6c \\ 0 & 0 & 0 & 3 \\ 0 & 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} 1 & c & c^2 & c^3 \\ 0 & 1 & 2c & 3c^2 \\ 0 & 0 & 1 & 3c \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

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

Thus, D is represented by the same matrix relative to the ordered bases $\mathcal{B}$ and $\mathcal{B}'$.

Row space and column space of a matrix

~~Definition. Let $A \in \mathbb{R}^{m\times n}$ be a matrix over the field of real numbers. The matrix A has m rows, which are The subspace of $\mathbb{R}^n$ generated by the row vectors of A is called the row-space of A.~~

The column space of A.

The most important subspaces in linear algebra are tied to the matrix A. ~~If A~~ We are trying to solve $Ax = b$. Suppose A is not invertible, the system is solvable for some b and not solvable for other ~~b~~ right hand side vectors b. We want to describe the good right sides b - the vectors that can be written as A times some vector x. Those b's form the column space of A.

Definition. The column space of A is the subspace generated by all linear combinations of the columns of a matrix A.

Remember that, $Ax = b$ is solvable, if and only if b is in the column space of A. Since, $b \in \mathbb{R}^m$, the column space of A is a subspace of $\mathbb{R}^m$.

Examples.

1) Consider the system of linear algebraic equations

$$Ax = \begin{bmatrix} 1 & 0 \\ 4 & 3 \\ 2 & 3 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} b_1 \\ b_2 \\ b_3 \end{bmatrix}.$$

The column space of A is $x_1 \begin{pmatrix} 1 \\ 4 \\ 2 \end{pmatrix} + x_2 \begin{pmatrix} 0 \\ 3 \\ 3 \end{pmatrix}$. This fills up a plane in $\mathbb{R}^3$, since it is a linear combination of two vectors.

This plane has zero thickness, so very few right sides b in $\mathbb{R}^3$ are in the column space. In fact, the dimension of the column space in this case is 2.

<u>Definition (Row space of a matrix)</u>. Let A be a matrix of order $m \times n$ over the field of real numbers $\mathbb{R}$. $A \in \mathbb{R}^{m \times n}$. The subspace of $\mathbb{R}^n$ generated by the row-vectors of A is called the row space of A.

<u>Examples.</u>

1) If $A = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix}$, does $(0,0,1)$ belong to the row space of A?

<u>Solution.</u>

The row space of A is the subspace of $\mathbb{R}^3$ generated by the vectors $(1,0,0)$, and $(0,1,0)$.

$RS(A) := \{(x_1, x_2, 0) : x_1, x_2 \in \mathbb{R}\}$.

Thus, $(0,0,1)$ does not belong to the row-space of A.

2) If $A = \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 2 & 0 \end{bmatrix}$, what is the dimension of the row space of A?

<u>Solution.</u>

The row space of A is a subspace of $\mathbb{R}^2$ generated by the row vectors $(1,0)$, $(0,1)$, $(2,0)$. But, $(2,0)$ already lies in the span of $\{(1,0), (0,1)\}$. Therefore, the row space of A equals $\mathbb{R}^2$.

dim rowspace of $A = 2$.

<u>Row rank of a Matrix A. (Definition)</u>. The dimension of the row space of A is called the row rank of A, denoted by $\rho_r(A)$.

<u>Column rank of a matrix (Definition)</u>. The column rank of a matrix A, is the dimension of the column space of A.

The row space is a subspace of $\mathbb{R}^n$. The column space is a subspace of $\mathbb{R}^m$. Each row vector has n coordinates, each column vector has m coordinates.

rowspace $(A) \subseteq \mathbb{R}^n$
columnspace $(A) \subseteq \mathbb{R}^m$.

An interesting fact is that the row rank of A equals the column rank of A. It is a very important result. These subspace may lie in ~~a~~ different vector space, but their dimension is the same.

<u>Properties of row rank and column rank.</u>

<u>Lemma.</u> The row rank of $A \leq \min(m, n)$.

<u>Proof</u> Firstly, the row rank of A must be equal to the number of linearly independent rows of A. Thus, the row rank of $A \leq$ rows of A. Thus, $\rho_r(A) \leq m$.

Next, the row-vectors in A each have n-coordinates. row space $(A) \subseteq \mathbb{R}^n$. Thus, $\rho_r(A) \leq n$.

So, we must have $\rho_r(A) \leq \min(m, n)$.

Theorem 11. (Invertible matrix theorem). Let $A \in \mathbb{R}^{n \times n}$. Then, the following statements are equivalent:

(1) The columns of A are linearly independent.
(2) The rows of A are linearly independent.
(3) A is invertible, that is A has an inverse, is non-singular, or is non-degenerate. $AB = I_n = BA$.
(4) The null space of A is trivial, that is, it contains only the zero vector as an element. $\text{null}(A) = \{0\}$.
(5) The columns of A span $\mathbb{R}^n$.

Proof.

(1) Claim — The columns of A are linearly independent.
A is a square matrix over the field of real numbers of order $n \times n$.
Let $T: \mathbb{R}^n \to \mathbb{R}^n$ be the linear transformation such that
$$A = [T]_{\mathcal{B}}.$$

A is an invertible matrix. T is invertible. So, T is an isomorphism.
Let $\mathcal{B} = \{e_1, e_2, \ldots, e_n\}$ be a basis of $\mathbb{R}^n$.

As T is an isomorphism, $\{Te_1, Te_2, \ldots, Te_n\}$ must ~~be~~ form a basis. ~~of~~
This is because, —

- T is a bijection. T is both an injection and surjection.
- Let y be an arbitrary element in the codomain.
  $\exists x \in \mathbb{R}^n$, such that $y = T(x)$. $\forall \cancel{x}$
- $y = T(x)$
  $= T(x_1 e_1 + x_2 e_2 + x_3 e_3 + \ldots + x_n e_n)$
  $= x_1 Te_1 + x_2 Te_2 + x_3 Te_3 + \ldots + x_n Te_n$.
- Therefore, $\{Te_1, Te_2, \ldots, Te_n\}$ span $\mathbb{R}^n$

- Also, $\alpha_1 Te_1 + \alpha_2 Te_2 + \ldots + \alpha_n Te_n = 0$
  $\Rightarrow T(\alpha_1 e_1 + \alpha_2 e_2 + \ldots + \alpha_n e_n) = 0$.
- T is injective. So, $\text{null}(T) = 0$.
- Thus, $\alpha_1 e_1 + \alpha_2 e_2 + \ldots + \alpha_n e_n = 0$.
- But, $\{e_1, e_2, \ldots, e_n\}$ are a basis of $\mathbb{R}^n$ and are linearly independent.
- $\alpha_1 = \alpha_2 = \ldots = \alpha_n = 0$.
- So, $\{Te_1, Te_2, \ldots, Te_n\}$ is a linearly independent set in $\mathbb{R}^n$.

We conclude that, $\{Te_1, \ldots, Te_n\}$ is a basis of $\mathbb{R}^n$.
But, $[Te_j]_{\mathcal{B}} = [T]_{\mathcal{B}} [e_j]$ is the jth column of the matrix A.

Thus, the columns of A are linearly independent and they span $\mathbb{R}^n$.

(2) The rows of A are linearly independent. ~~As A invertible, A^T is also~~

As A is invertible, A^T is also invertible. Therefore, the columns of A and thus the rows of A are linearly independent.

(3) By definition, A has both a left-inverse

$$AB = I_n$$

and a right inverse

$$BA = I_n.$$

Thus, $AB = I_n = BA$.

(4) Since T is injective, null $(T) = \{0\}$.

Examples.

1) Let

$$A = \begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix}$$

be a 3×3 matrix.

Determine whether or not A is invertible.

Solution.

Let's verify if the columns of A are ~~invertible~~ linearly independent.

Suppose, $Ax = 0$

$$\begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

$$1\cdot x_1 + 0\cdot x_2 + 1\cdot x_3 = 0$$
$$0\cdot x_1 + 1\cdot x_2 + 1\cdot x_3 = 0$$
$$1\cdot x_1 + 1\cdot x_2 + 1\cdot x_3 = 0$$

Thus,

$$1\cdot x_1 + 0\cdot x_2 + 1\cdot x_3 = 0$$
$$0\cdot x_1 + 1\cdot x_2 + 1\cdot x_3 = 0$$
$$1\cdot x_1 + 0\cdot x_2 + 0\cdot x_3 = 0$$

which yields $x_1 = 0$, $x_2 = 0$, $x_3 = 0$.

The columns of A are linearly independent. Invoking invertible matrix theorem, A is invertible.

2) Check if

$$\begin{bmatrix} 2 & 0 & 1 \\ 0 & 0 & 1 \\ 0 & 3 & 0 \end{bmatrix}$$

is invertible.

Solution.

Let us check if the columns of A are linearly independent.

Suppose $Ax = 0$

$$\begin{bmatrix} 2 & 0 & 1 \\ 0 & 0 & 1 \\ 0 & 3 & 0 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

$$2x_1 + 0\cdot x_2 + 1\cdot x_3 = 0$$
$$x_3 = 0$$
$$3x_2 = 0$$

which yields $x_1 = 0$, $x_2 = 0$, $x_3 = 0$.

By invertible matrix theorem, A is invertible.

Elementary Matrix (Definition). Any matrix P obtained by performing a single elementary row operation on I_n is called an elementary matrix.

Given any elementary matrix E, there exists a matrix D, such that $DE = I = ED$. Every elementary matrix E is invertible.

Proof (Informal).
Each column of E is a unique linear combination of the columns of I_n.

Let $E = (E_1, E_2, \dots, E_j, \dots, E_n)$ be an elementary matrix of order n. The jth column of E is E_j.

suppose $E_j = \begin{bmatrix} a_{1j} \\ a_{2j} \\ a_{3j} \\ \vdots \\ a_{nj} \end{bmatrix}$, where atleast one of the a's is non-zero.

Each column vector of E is a unique linear combination of the basis vectors $e_1, e_2, \dots, e_n$.

Therefore, $E_j = a_{1j} \begin{bmatrix} 1 \\ 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix} + a_{2j} \begin{bmatrix} 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix} + a_{3j} \begin{bmatrix} 0 \\ 0 \\ 1 \\ \vdots \\ 0 \end{bmatrix} + \dots + a_{nj} \begin{bmatrix} 0 \\ 0 \\ 0 \\ \vdots \\ 1 \end{bmatrix}$

$$E_j = a_{1j} e_1 + a_{2j} e_2 + a_{3j} e_3 + \dots + a_{nj} e_n$$

$$E_j = \sum_{i=1}^{n} a_{ij} e_i$$

where e_i's are the columns of the identity matrix I_n.

Since each E_j has a unique representation, E_2 is not a scalar multiple of E_1, E_3 is not in the span (E_1, E_2) and in general $E_j \notin$ span $(E_1, E_2, \dots, E_{j-1})$.

Therefore, the columns of E are linearly independent. Invoking inverse matrix theorem, E is an invertible matrix.

Each of the elementary row or column operations on a matrix is as like multiplying a matrix A with an elementary matrix E. For example, if

$$A = \begin{bmatrix} 0 & 1 & 2 \\ 3 & 0 & 0 \\ 2 & 1 & 0 \end{bmatrix}$$

Left multiplying this matrix by E_{12} is as good as interchanging the first and second rows.

$$E_{12} A = \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 0 & 1 & 2 \\ 3 & 0 & 0 \\ 2 & 1 & 0 \end{bmatrix} = \begin{bmatrix} 3 & 0 & 0 \\ 0 & 1 & 2 \\ 2 & 1 & 0 \end{bmatrix}.$$

Right-multiplying the matrix A by E_{12}, is like performing a column operation.

$$AE_{12} = \begin{bmatrix} 0 & 1 & 2 \\ 3 & 0 & 0 \\ 2 & 1 & 0 \end{bmatrix} \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 2 \\ 0 & 3 & 0 \\ 1 & 2 & 0 \end{bmatrix}.$$

Again consider multiplying or dividing a row/column by a scalar: ~~This is a scalar k multi~~ Multiplying a row by a scalar k is akin to left multiplication $E_i(k)A$. Multiplying a column by a scalar is similar to the right multiplication $AE_i(k)$.

As an illustration:

$$E_3(2)A = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 2 \end{bmatrix} \begin{bmatrix} 0 & 1 & 2 \\ 3 & 0 & 0 \\ 2 & 1 & 0 \end{bmatrix} = \begin{bmatrix} 0 & 1 & 2 \\ 3 & 0 & 0 \\ 4 & 2 & 0 \end{bmatrix}.$$

$$AE_3(2) = \begin{bmatrix} 0 & 1 & 2 \\ 3 & 0 & 0 \\ 2 & 1 & 0 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 2 \end{bmatrix} = \begin{bmatrix} 0 & 1 & 4 \\ 3 & 0 & 0 \\ 2 & 1 & 0 \end{bmatrix}.$$

Lastly, if an equation (row) e_i is replaced by the sum of $e_i + ke_j$, where $j \neq i$ and k is any scalar, it is equivalent to multiplication by the elementary matrix $E_{ij}(k)$.

$$E_{13}(5)(A) = \begin{bmatrix} 1 & 0 & 5 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 0 & 1 & 2 \\ 3 & 0 & 0 \\ 2 & 1 & 0 \end{bmatrix} = \begin{bmatrix} 10 & 6 & 2 \\ 3 & 0 & 0 \\ 2 & 1 & 0 \end{bmatrix}$$

$$AE_{13}(5) = \begin{bmatrix} 0 & 1 & 2 \\ 3 & 0 & 0 \\ 2 & 1 & 0 \end{bmatrix} \begin{bmatrix} 1 & 0 & 5 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} 10 & 1 & 2 \\ 3 & 0 & 0 \\ 2 & 1 & 0 \end{bmatrix}.$$

Many problems in linear algebra require that a matrix be converted to the row-echelon form (ref) or its stricter variant reduced row echelon form (rref). Every computer algebra system and most scientific graphing calculators have commands which produce these forms for any matrix.

Row-Echelon Form (ref).

A rectangular matrix in the row-echelon form has the following three defining properties.

1) The first r rows for some $r \geq 0$ are non-zero, and the remaining rows if any are zero.

2) In the ith row $(i = 1, 2, 3, \ldots, r)$, the first non-zero element is equal to unity, the column in which it occurs is c_i.

3) $c_1 < c_2 < c_3 < \cdots < c_r$.

A row-echelon

01·12·2019

A matrix in the row-echelon form is upper triangular.

For example,

$$U = \begin{bmatrix} 1 & 0 & 3 & 3 \\ 0 & 1 & 3 & 4 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

is in the row-echelon form.

The reduced row Echelon form.
If a matrix in row-echelon form satisfies the following conditions then it is said to be in reduced-row echelon form.

1) The matrix is in row-echelon form.

2) Each leading 1 is the only non-zero entry in its column.

The reduced row echelon form of the matrix discussed in the previous section is

$$R = \begin{bmatrix} 1 & 0 & 3 & 0 \\ 0 & 1 & 3 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}.$$

.1·2 Equivalent systems of Linear Equations.
The sole objective of linear algebra is to solve the system of equations:

$$Ax = b$$

where A is a matrix of order $m \times n$ over the field of reals, $x \in \mathbb{R}^n$, the right side vector $b \in \mathbb{R}^m$. We are interested to find $x = (x_1, x_2, \dots, x_n)$ that satisfies the above system of equations.

A solution of a linear system is therefore, an assignment of values to the variables $x_1, x_2, \dots, x_n$ such that each of the equations is satisfied. The set of all possible solutions is called the solution set.

Consider the system of equations

$$2x_1 + 3x_2 = 6$$
$$3x_1 + 2x_2 = 4$$

$$A = \begin{bmatrix} 2 & 3 \\ 3 & 2 \end{bmatrix} \quad x = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \quad b = \begin{bmatrix} 6 \\ 4 \end{bmatrix}.$$

The ~~one~~

$$6x_1 + 9x_2 = 18$$
$$-6x_1 + 4x_2 = 8$$
$$5x_2 = 10$$
$$x_2 = 2$$
$$x_1 = 0.$$