§ 4.6 Change of Basis, Transition matrix

Let V be a finite dimensional vector space and $B=\{v_1,\dots,v_n\}$ a basis for V. Then every $v \in V$ can be written in a unique way as

$$v = k_1 v_1 + k_2 v_2 + \dots + k_n v_n \,.$$

The scalars $k_1, k_2, \dots, k_n$ or the vector $(k_1, k_2, \dots, k_n) \in R^n$ are the coordinates of v with respect to the basis B.

we write

$$\boxed{(k_1, k_2, \dots, k_n) = [v]_B}$$

The change of basis problem: If we change the basis B into a different basis $B'=\{v_1',\dots,v_n'\}$ then the coordinates of v with respect to B' will be different

$$(k_1', k_2', \dots, k_n') = [v]_{B'} \iff v = k_1' v_1' + \dots k_n' v_n'$$

The Question then is: How are the coordinates $[v]_{B'} = (k_1',\dots,k_n')$ related to $[v]_B = (k_1, \dots, k_n)$?

The answer is: there exists an $n \times n$ matrix called the transition matrix $B \to B'$ and denoted by $P_{B\to B'}$ such that

$$\boxed{[v]_{B'} = P_{B\to B'} [v]_B}$$

It turns out that the transition matrix $P_{B\to B'}$ is rather easy to construct. Its columns are the coordinates of the basis elements $v_1, \dots, v_n$ in terms of the basis elements $v_1', v_2', \dots, v_n'$ of B'.

$$P_{B\to B'} = \left[[v_1]_{B'} \;\Big|\; [v_2]_{B'} \;\Big|\; \cdots \;\Big|\; [v_n]_{B'} \right] \qquad (*)$$

In detail,

$$v_1 = \beta_1^{(1)} v_1' + \beta_2^{(1)} v_2' + \cdots + \beta_n^{(1)} v_n' \qquad [v_1]_{B'} = (\beta_1^{(1)}, \beta_2^{(1)}, \ldots, \beta_n^{(1)})$$
$$v_2 = \beta_1^{(2)} v_1' + \beta_2^{(2)} v_2' + \cdots + \beta_n^{(2)} v_n' \qquad [v_2]_{B'} = (\beta_1^{(2)}, \beta_2^{(2)}, \ldots, \beta_n^{(2)})$$
$$\vdots \qquad\qquad \vdots$$
$$v_n = \beta_1^{(n)} v_1' + \beta_2^{(n)} v_2' + \cdots + \beta_n^{(n)} v_n' \qquad [v_n]_{B'} = (\beta_1^{(n)}, \beta_2^{(n)}, \ldots, \beta_n^{(n)})$$

Then

$$P_{B \to B'} = \begin{bmatrix} \beta_1^{(1)} & \beta_1^{(2)} & \cdots & \beta_1^{(n)} \\ \beta_2^{(1)} & \beta_2^{(2)} & \cdots & \beta_2^{(n)} \\ \vdots & \vdots & & \vdots \\ \beta_n^{(1)} & \beta_n^{(2)} & \cdots & \beta_n^{(n)} \end{bmatrix}.$$

It should be clear that the $n \times n$ matrix is well-defined. Indeed, since B' is a basis the scalars $\beta_i^{(j)}$ are all well-defined.

<u>proof of (*)</u>

$$\begin{aligned} v &= k_1 v_1 + k_2 v_2 + \cdots + k_n v_n \\ &= k_1(\beta_1^{(1)} v_1' + \beta_2^{(1)} v_2' + \cdots + \beta_n^{(1)} v_n') \Leftarrow v_1 = \beta_1^{(1)} v_1' + \beta_2^{(1)} v_2' + \cdots + \beta_n^{(1)} v_n' \\ &+ k_2(\beta_1^{(2)} v_1' + \beta_2^{(2)} v_2' + \cdots + \beta_n^{(2)} v_n') \Leftarrow v_2 = \beta_1^{(2)} v_1' + \beta_2^{(2)} v_2' + \cdots + \beta_n^{(2)} v_n' \\ &+ \\ &\vdots \qquad\qquad \vdots \\ &+ k_n(\beta_1^{(n)} v_1' + \beta_2^{(n)} v_2' + \cdots + \beta_n^{(n)} v_n') \Leftarrow v_n = \beta_1^{(n)} v_1' + \beta_2^{(n)} v_2' + \cdots + \beta_n^{(n)} v_n' \end{aligned}$$

Gathering in terms of $v_1', v_2', \ldots, v_n'$,

$$\begin{aligned} v = {} & (k_1 \beta_1^{(1)} + k_2 \beta_1^{(2)} + \cdots + k_n \beta_1^{(n)}) v_1' \\ & + \\ & (k_1 \beta_2^{(1)} + k_2 \beta_2^{(2)} + \cdots + k_n \beta_2^{(n)}) v_2' \\ & + \\ & \vdots \\ & + \\ & (k_1 \beta_n^{(1)} + k_2 \beta_n^{(2)} + \cdots + k_n \beta_n^{(n)}) v_n'. \\ = {} & k_1' v_1' + k_2' v_2' + \cdots + k_n' v_n'. \end{aligned}$$

Since the coordinates of v with respect to any basis are uniquely defined, we must have

$$k_1' = k_1 \beta_1^{(1)} + k_2 \beta_1^{(2)} + \cdots + k_n \beta_1^{(n)}$$
$$k_2' = k_1 \beta_2^{(1)} + k_2 \beta_2^{(2)} + \cdots + k_n \beta_2^{(n)}$$
$$\vdots$$
$$k_n' = k_1 \beta_n^{(1)} + k_2 \beta_n^{(2)} + \cdots + k_n \beta_n^{(n)}.$$

In matrix-vector form, The above is exactly

$$[v]_{B'} = \begin{bmatrix} k_1' \\ k_2' \\ \vdots \\ k_n' \end{bmatrix} = \begin{bmatrix} \beta_1^{(1)} & \beta_1^{(2)} & \cdots & \beta_1^{(n)} \\ \beta_2^{(1)} & \beta_2^{(2)} & \cdots & \beta_2^{(n)} \\ \vdots & \vdots & & \vdots \\ \beta_n^{(1)} & \beta_n^{(2)} & \cdots & \beta_n^{(n)} \end{bmatrix} \begin{bmatrix} k_1 \\ k_2 \\ \vdots \\ k_n \end{bmatrix} = P_{B \to B'} [v]_B \quad \blacksquare$$

we argued above that the transition matrix $P_{B \to B'}$ is well-defined. we shall next show that it is not only invertible, but it conveniently transpires that its inverse is the transition matrix from B' to B

<u>Theorem 4.6.1</u> (a) The transition matrix $P_{B \to B'}$ is invertible.
(b) $P^{-1}_{B \to B'} = P_{B' \to B}$.

<u>proof</u> (a) Suppose $P_{B \to B'}$ is not invertible. Then There must exist a <u>nonzero</u> $n \times 1$ matrix K such that $P_{B \to B'} K = 0$. <u>Consider the vector v whose coordinates</u> with respect to the basis B are K, i.e. $[v]_B = K$

$$v = k_1 v_1 + k_2 v_2 + \cdots + k_n v_n.$$

Clearly <u>v is nonzero since</u> K is nonzero. on the other hand, we saw above that $[v]_{B'} = P_{B \to B'} [v]_B$ and hence

$$K' \equiv [v]_{B'} = P_{B \to B'} K = 0.$$

However, this means that $v = k_1' v_1' + k_2' v_2' + \cdots + k_n' v_n' = 0$, a clear contradiction to v being nonzero.

(b) From $[v]_{B'} = P_{B\to B'}[v]_B$ we have $[v]_B = P^{-1}_{B\to B'}[v]_{B'}$.
This holds for any v in V and is all we needed as proof. ■

Ex Find the transition (change of basis) matrix from the basis $B=\{1, 2t, -2+4t^2, -12t+8t^3\}$ to the basis $B'=\{1, 1-t, 2-4t+t^2, 6-18t+9t^2-t^3\}$

$$(*)\begin{cases} 1 = \beta_1^{(1)}(1) + \beta_2^{(1)}(1-t) + \beta_3^{(1)}(2-4t+t^2) + \beta_4^{(1)}(6-18t+9t^2-t^3) \\ 2t = \beta_1^{(2)}(1) + \beta_2^{(2)}(1-t) + \beta_3^{(2)}(2-4t+t^2) + \beta_4^{(2)}(6-18t+9t^2-t^3) \\ -2+4t^2 = \beta_1^{(3)}(1) + \beta_2^{(3)}(1-t) + \beta_3^{(3)}(2-4t+t^2) + \beta_4^{(3)}(6-18t+9t^2-t^3) \\ -12t+8t^3 = \beta_1^{(4)}(1) + \beta_2^{(4)}(1-t) + \beta_3^{(4)}(2-4t+t^2) + \beta_4^{(4)}(6-18t+9t^2-t^3) \end{cases}$$

The solution of this problem consists in finding the 16 coefficients (unknowns) $\beta_i^{(j)}$ $i,j=1,2,3,4$. To do this we equate the coefficients of $1, t, t^2, t^3$ on both sides.

From the first equation in (*) we get

$$1 = (\beta_1^{(1)} + \beta_2^{(1)} + 2\beta_3^{(1)} + 6\beta_4^{(1)}) + (-\beta_2^{(1)} - 4\beta_3^{(1)} - 18\beta_4^{(1)})t$$
$$+(\beta_3^{(1)} + 9\beta_4^{(1)})t^2 + (-\beta_4^{(1)})t^3$$

From this, we get 4 equations for $\beta_1^{(1)}, \beta_2^{(1)}, \beta_3^{(1)}, \beta_4^{(1)}$

$$\begin{aligned} \beta_1^{(1)} + \beta_2^{(1)} + 2\beta_3^{(1)} + 6\beta_4^{(1)} &= 1 \\ -\beta_2^{(1)} - 4\beta_3^{(1)} - 18\beta_4^{(1)} &= 0 \\ \beta_3^{(1)} + 9\beta_4^{(1)} &= 0 \\ -\beta_4^{(1)} &= 0 \end{aligned} \iff \begin{bmatrix} 1 & 1 & 2 & 6 \\ 0 & -1 & -4 & -18 \\ 0 & 0 & 1 & 9 \\ 0 & 0 & 0 & -1 \end{bmatrix} \begin{bmatrix} \beta_1^{(1)} \\ \beta_2^{(1)} \\ \beta_3^{(1)} \\ \beta_4^{(1)} \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \end{bmatrix}$$

From the 2nd equation in ⊛, we get

$$2t = (\beta_1^{(2)} + \beta_2^{(2)} + 2\beta_3^{(2)} + 6\beta_4^{(2)}) + t(-\beta_2^{(2)} - 4\beta_3^{(2)} - 18\beta_4^{(2)})$$
$$+ t^2(\beta_3^{(2)} + 9\beta_4^{(2)}) + t^3(-\beta_4^{(2)})$$

This yields 4 equations for $\beta_1^{(2)}, \beta_2^{(2)}, \beta_3^{(2)}, \beta_4^{(2)}$

$$\begin{aligned} \beta_1^{(2)} + \beta_2^{(2)} + 2\beta_3^{(2)} + 6\beta_4^{(2)} &= 0 \\ -\beta_2^{(2)} - 4\beta_3^{(2)} - 18\beta_4^{(2)} &= 2 \\ \beta_3^{2} + 9\beta_4^{(2)} &= 0 \\ -\beta_4^{(2)} &= 0 \end{aligned} \iff \begin{bmatrix} 1 & 1 & 2 & 6 \\ 0 & -1 & -4 & -18 \\ 0 & 0 & 1 & 9 \\ 0 & 0 & 0 & -1 \end{bmatrix} \begin{bmatrix} \beta_1^{(2)} \\ \beta_2^{(2)} \\ \beta_3^{(2)} \\ \beta_4^{(2)} \end{bmatrix} = \begin{bmatrix} 0 \\ 2 \\ 0 \\ 0 \end{bmatrix}$$

Note that this is the same coefficient matrix that appeared in the first set of 4 equations.

From the 3rd equation in ⊛ we get

$$-2 + 4t^2 = (\beta_1^{(3)} + \beta_2^{(3)} + 2\beta_3^{(3)} + 6\beta_4^{(3)}) + t(-\beta_2^{(3)} - 4\beta_3^{(3)} - 18\beta_4^{(3)}$$
$$+ t^2(\beta_3^{(3)} + 9\beta_4^{(3)}) + t^3(-\beta_4^{(3)}).$$

The next set of 4 equations are

$$\begin{aligned} \beta_1^{(3)} + \beta_2^{(3)} + 2\beta_3^{(3)} + 6\beta_4^{(3)} &= -2 \\ -\beta_2^{(3)} - 4\beta_2^{(3)} - 18\beta_4^{(3)} &= 0 \\ \beta_3^{(3)} + 9\beta_4^{(3)} &= 4 \\ -\beta_4^{(3)} &= 0 \end{aligned} \iff \begin{bmatrix} 1 & 1 & 2 & 6 \\ 0 & -1 & -4 & -18 \\ 0 & 0 & 1 & 9 \\ 0 & 0 & 0 & -1 \end{bmatrix} \begin{bmatrix} \beta_1^{(3)} \\ \beta_2^{(3)} \\ \beta_3^{(3)} \\ \beta_4^{(3)} \end{bmatrix} = \begin{bmatrix} -2 \\ 0 \\ 4 \\ 0 \end{bmatrix}$$

From the 4th equation in ⊛ we get

$$-12t + 8t^3 = (\beta_1^{(4)} + \beta_2^{(4)} + 2\beta_3^{(4)} + 6\beta_4^{(4)}) + t(-\beta_2^{(4)} - 4\beta_3^{(4)} - 18\beta_4^{(4)})$$
$$+ t^2(\beta_3^{(4)} + 9\beta_4^{(4)}) + t^3(-\beta_4^{(4)})$$

$$\Rightarrow \begin{aligned} \beta_1^{(4)} + \beta_2^{(4)} + 2\beta_3^{(4)} + 6\beta_4^{(4)} &= 0 \\ -\beta_2^{(4)} - 4\beta_3^{(4)} - 18\beta_4^{(4)} &= -12 \\ \beta_3^{(4)} + 9\beta_4^{(4)} &= 0 \\ -\beta_4^{(4)} &= 8 \end{aligned} \iff \begin{bmatrix} 1 & 1 & 2 & 6 \\ 0 & -1 & -4 & -18 \\ 0 & 0 & 1 & 9 \\ 0 & 0 & 0 & -1 \end{bmatrix} \begin{bmatrix} \beta_1^{(4)} \\ \beta_2^{(4)} \\ \beta_3^{(4)} \\ \beta_4^{(4)} \end{bmatrix} = \begin{bmatrix} 0 \\ -12 \\ 0 \\ 8 \end{bmatrix}$$

we have 4 systems of equations to solve. we noticed already that the coefficient matrix is the same, only the right-hand-sides being different. we combine all 4 systems into

$$(**)\quad \begin{bmatrix} 1 & 1 & 2 & 6 \\ 0 & -1 & -4 & -18 \\ 0 & 0 & 1 & 9 \\ 0 & 0 & 0 & -1 \end{bmatrix} \begin{bmatrix} \beta_1^{(1)} & \beta_1^{(2)} & \beta_1^{(3)} & \beta_1^{(4)} \\ \beta_2^{(1)} & \beta_2^{(2)} & \beta_2^{(3)} & \beta_2^{(4)} \\ \beta_3^{(1)} & \beta_3^{(2)} & \beta_3^{(3)} & \beta_3^{(4)} \\ \beta_4^{(1)} & \beta_4^{(2)} & \beta_4^{(3)} & \beta_4^{(4)} \end{bmatrix} = \begin{bmatrix} 1 & 0 & -2 & 0 \\ 0 & 2 & 0 & -12 \\ 0 & 0 & 4 & 0 \\ 0 & 0 & 0 & 8 \end{bmatrix}.$$

$$A' \quad \cdot \quad P_{B\to B'} \quad = A$$

$$\Rightarrow \underline{P_{B\to B'} = (A')^{-1}A} = \begin{bmatrix} 1 & 2 & 6 & 36 \\ 0 & -2 & -16 & -132 \\ 0 & 0 & 4 & 72 \\ 0 & 0 & 0 & -8 \end{bmatrix}. \quad \text{using Matlab}$$

~~Remark 1 The matrices A and A' are upper triangular. This is due to the fact that the degrees of the basis elements v_i and v_i' are less than or equal to i, $i=1,\dots,4$.~~

Remark 2 we see that finding transition matrices can involve tedious calculations devoted mainly to setting up the systems of equations (*) leading to (**). However, all that work can be bypassed if we notice that

The columns of A are the coefficients of the basis elements of B in terms of the standard basis
The columns of A' " " " " " " "
of B' " " " " " ".

Ex. Find the transition matrix $P_{B\to B'}$ for the bases $B=\{(1,-1,2),(4,5,-1),(3,5,4)$ and
$B'=\{(2,1,4),(7,-2,3),(4,-3,1)\}$ of R^3.

Using Remark 2 as guide, we set

$$\underbrace{\begin{bmatrix} 2 & 7 & 4 \\ 1 & -2 & -3 \\ 4 & 3 & 1 \end{bmatrix}}_{A'} P_{B \to B'} = \underbrace{\begin{bmatrix} 1 & 4 & 3 \\ -1 & 5 & 5 \\ 2 & -1 & 4 \end{bmatrix}}_{A}$$

$$\Rightarrow P_{B \to B'} = (A')^{-1} A = \frac{1}{11}\begin{bmatrix} 8 & -22 & 2 \\ 7 & 44 & 23 \\ 11 & -55 & -33 \end{bmatrix}.$$

This is exactly the procedure described on page 233 of the text, except that our remark is much more general and applies to any two bases B and B' in a general vector space V.

We can generalize the above procedures as follows: Suppose we have three bases B, B', B'' in a vector space V. Then it is rather straightforward showing that

$$\boxed{P_{B \to B''} = P_{B' \to B''} P_{B \to B'}}$$

Ex. Let $V = \text{span}\{B\}$ $\quad B = \{\sin x, \cos x\}$.

(a) Show that $B' = \{2\sin x + \cos x, 3\cos x\}$ is a basis for V

(b) Find the transition matrix $P_{B \to B'}$.

(c) Find the transition matrix $P_{B' \to B}$.

(d) Compute the coordinate vectors $[f]_B$ and $[f]_{B'}$ for $f = 2\sin x - 5\cos x$

(a) $\sin x$ and $\cos x$ are linearly independent since neither is a constant scalar multiple of the other $\forall x$. Hence B is a basis for V and $\dim V = 2$.

Clearly $\text{span}(B')$ is a subspace of V since each element of B' is a linear combination of the elements of B. Also, since $\dim V = 2$ and B' contains 2 elements, if we show B' is a linearly independent set, then B' will be a basis for V.

we can do this directly!

Suppose

$$k_1(2\sin x + \cos x) + k_2(3\cos x) = 0 \quad \forall x \in \mathbb{R}$$

$$\Leftrightarrow \quad (2k_1)\sin x + (k_1 + 3k_2)\cos x = 0 \quad \forall x \in \mathbb{R}.$$

Since $\sin x$ and $\cos x$ are linearly independent, we must have

$$2k_1 = 0 \quad \Rightarrow k_1 = 0$$

$$k_1 + 3k_2 = 0 \quad \Rightarrow k_2 = -\frac{1}{3}k_1 = 0 \quad \checkmark.$$

Thus B' is lin. indep. $\Rightarrow$ B' is a basis for V.

(b) we use formula (**). Indeed B can be considered as the standard basis for V

$$\Rightarrow A = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}. \quad \text{Also, } A' = \begin{bmatrix} 2 & 0 \\ 1 & 3 \end{bmatrix}.$$

$$\Rightarrow \quad A' P_{B\to B'} = A \Rightarrow P_{B\to B'} = (A')^{-1} A$$

$$= \frac{1}{6}\begin{bmatrix} 3 & 0 \\ -1 & 2 \end{bmatrix}\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = \frac{1}{6}\begin{bmatrix} 3 & 0 \\ -1 & 2 \end{bmatrix}.$$

(c) we know $P_{B'\to B} = (P_{B\to B'})^{-1} = \frac{1}{6}\begin{bmatrix} 3 & 0 \\ -1 & 2 \end{bmatrix} = \begin{bmatrix} 2 & 0 \\ 1 & 3 \end{bmatrix}.$

(d) Obviously $[f]_B = (2, -5)$

$$[f]_{B'} = P_{B\to B'}[f]_B = \frac{1}{6}\begin{bmatrix} 3 & 0 \\ -1 & 2 \end{bmatrix}\begin{bmatrix} 2 \\ -5 \end{bmatrix} = \frac{1}{6}\begin{bmatrix} 6 \\ -12 \end{bmatrix} = \begin{bmatrix} 1 \\ -2 \end{bmatrix}$$

<u>check</u> $1\cdot(2\sin x + \cos x) - 2(3\cos x) = 2\sin x - 5\cos x = f \quad \checkmark$

§4.7 Row space, column space and Null space of a matrix

In This section and the next, we will introduce and study some important vector spaces that are associated with matrices. This will provide a deeper understanding of "what a matrix is", and will help solving a host of problems that arise in Linear Algebra.

Recall that an $m \times n$ matrix is a rectangular array of elements

$$A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix}.$$

we shall also think of A as an (ordered) set of m "rows" which are $1 \times n$ matrices:

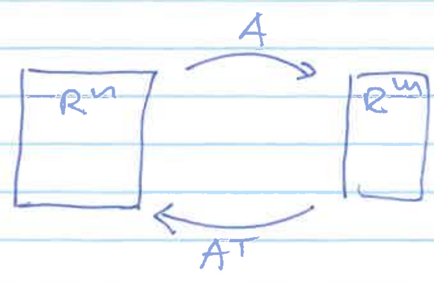
$$[a_{11}, a_{12}, \ldots, a_{1n}], [a_{21}, a_{22}, \ldots, a_{2n}], \ldots\ldots, [a_{m1}, a_{m2}, \ldots, a_{mn}].$$

Of course we should also think of the rows of A as vectors in R^n.

Similarly, A can be thought of as an ordered set of n columns which are $m \times 1$ matrices (or vectors in R^m)

$$\begin{bmatrix} a_{11} \\ a_{21} \\ \vdots \\ a_{m1} \end{bmatrix}, \begin{bmatrix} a_{12} \\ a_{22} \\ \vdots \\ a_{m2} \end{bmatrix}, \ldots, \begin{bmatrix} a_{1n} \\ a_{2n} \\ \vdots \\ a_{mn} \end{bmatrix}.$$

Also, it is very important to develop the point of view that a matrix is in essence a linear transformation. Indeed, if A is an $m \times n$ matrix, then the map $x \longmapsto Ax$ is a linear transformation from R^n to R^m.



Conversely, given a linear transformation $T: R^n \to R^m$, there is a unique $m \times n$ matrix that can be associated with T. This is done as follows:

Let $\{v_1, \ldots, v_n\}$ be a basis for R^n and $\{w_1, \ldots, w_m\}$ be a basis for R^m. Since the vectors $T(v_1), \ldots, T(v_n)$ belong to R^m, we can expand them in terms of the basis $\{w_1, \ldots, w_m\}$

$$T(v_1) = a_{11} w_1 + a_{21} w_2 + \cdots + a_{m1} w_m$$
$$T(v_2) = a_{12} w_1 + a_{22} w_2 + \cdots + a_{m2} w_m$$
$$\vdots$$
$$T(v_n) = a_{1n} w_1 + a_{2n} w_2 + \cdots + a_{mn} w_m .$$

We define an $m \times n$ matrix A by

$$A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix} = \left[[T(v_1)] \,\middle|\, [T(v_2)] \,\middle|\, \cdots \,\middle|\, [T(v_n)] \right]$$

where $[T(v_1)], \ldots, [T(v_n)]$ are the coordinates of $T(v_1), \ldots, T(v_n)$ with respect to the basis $\{w_1, \ldots, w_m\}$.

Now let $x \in R^n$, $y \in R^m$ be such that $y = T(x)$. We will show that in fact $[y] = A[x]$ where $[x]$ is the coordinate vector of x with respect to $\{v_1, \ldots, v_n\}$ and $[y]$ is the coordinate vector of y with respect to $\{w_1, \ldots, w_m\}$.

Indeed,

$$T(x) = T\left(\sum_{j=1}^{n} [x]_j v_j\right) = \sum_{j=1}^{n} [x]_j \, T(v_j) \qquad \text{Linearity of } T$$
$$= \sum_{j=1}^{n} [x_j] \sum_{i=1}^{m} a_{ij} w_i$$
$$= \sum_{i=1}^{m} \left(\sum_{j=1}^{n} a_{ij} [x]_j \right) w_i .$$

Since $T(x) = y$, $\sum_{j=1}^{n} a_{ij} [x]_j = y_i$, $i = 1, \ldots, m$

by uniqueness of the coordinates y_i. But this is exactly what we wanted to show, i.e. $A[x] = [y]$.

Defn. The row space of a matrix is the span of its rows considered as vectors in R^n. It is therefore a subspace of R^n.

Defn. The column space of a matrix is the span of its columns considered as vectors in R^m. It is therefore a subspace of R^m.

It is important to note that the column space of a matrix A is the same as the range of A, with A considered as a linear transformation from R^n to R^m.

Defn. The Null space or Kernel of a matrix A is

$$\mathrm{Ker}(A) = \{ x \in R^n \mid Ax = 0 \}.$$

It is a subspace of R^n.

Defn. Since the row space is a subspace of R^n, it has finite dimension which is called the row rank of A. We have row-rank$(A) \leq n$.

Defn. Since the column space is a subspace of R^m, it has finite dimension called the column-rank of A. Clearly column-rank$(A) \leq m$.

Defn. Ker(A) is a subspace of R^n. Its dimension is the nullity of A with nullity of $A \leq n$.

It makes sense to also introduce the same 3 spaces of A^T. However, note that only $\mathrm{Ker}(A^T)$ is "new" here since

row space of A^T = column space of A

and

column space of A^T = row space of A.

This motivates us to speak of the 4 fundamental spaces associated with a matrix A.

The row space of A
The column space of A
The null space of A
The null space of A^T

Our next important task will be to

1) Construct/identify a basis for each of the 4 fundamental spaces. This in particular will allow us to determine their dimensions. A crucial role is played by reduction to reduced row echelon form using elementary row operations.

2) Establish that
row rank of A = column rank of A = rank(A).

recall that early on we defined rank(A) as the number of nonzero rows in the reduced row echelon form of A.

3) Establish the fundamental identities $\boxed{\text{rank}(A) = \text{rank}(A^T)}$

$\boxed{n = \text{nullity of } A + \text{rank}(A)}$, $\boxed{m = \text{nullity of } A^T + \text{rank}(A).}$

Before we accomplish this for a general matrix A, let us do so for a matrix which is in reduced row echelon form. It turns out that the same information can be obtained from the row echelon form or even in upper triangular form.

Lemma let R be in rre form.
(a) A basis for the row space of R is given by the non-zero rows of R.

(b) A basis for the column space of R is given by the pivot columns of R.

(c) A basis for Ker(A) can be constructed as follows:
If there are no free variables, then Ker(A) = {0}.
let $x_{j_1}, x_{j_2}, \dots, x_{j_\nu}$ be the free variables.

(Recall that the free variables are those whose columns are not pivot.)

$Rz_1 = 0$

put $x_{j_1} = 1$, $x_{j_2} = 0$, ..., $x_{j_\nu} = 0$, solve for pivot variables $\Rightarrow z_{j_1}$

put $x_{j_1} = 0$, $x_{j_2} = 1$, $x_{j_3} = 0, \ldots, x_{j_\nu} = 0$ " , , " $\Rightarrow z_{j_2}$

put $x_{j_1} = \cdots = x_{j_{\nu-1}} = 0$, $x_{j_\nu} = 1$, " " " " $\Rightarrow z_{j_\nu}$

Then $\{z_1, \ldots, z_\nu\}$ is a basis for $\ker(R)$.

Before proving these facts, let us look at an example

ex.

$$R = \begin{bmatrix} 1 & 0 & 5 & 0 & 3 \\ 0 & 1 & 3 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

rank(R) = 3 number of nonzero rows of R

pivot columns are 1, 2, 4

free " " 3, 5

Basis for row space of R is $\{(1,0,5,0,3), (0,1,3,0,0), (0,0,0,1,0)\}$.

Basis for col. space of R: $\{(1,0,0,0), (0,1,0,0), (0,0,1,0)\}$

Basis for $\ker(R)$ the free variables are x_3, x_5

put $x_3 = 1$, $x_5 = 0$, solve for $x_1, x_2, x_4 \Rightarrow (-5, -3, 1, 0, 0) = z_1$

put $x_3 = 0$, $x_5 = 1$, " " $\Rightarrow (-3, 0, 0, 0, 1) = z_2$

$$\ker(R) = \operatorname{span}\{(-5,-3,1,0,0), (-3,0,0,0,1)\}.$$

proof of (a) let $r = \operatorname{rank}(R)$ be the number of nonzero rows of R. Hence, it suffices to show that the nonzero rows of R are linearly independent.

Let $j_1, j_2, \ldots, j_r$ be the pivot columns. Note that j_i is the column that contains the leading 1 (or first nonzero element) of row i. Now consider

$$\alpha_1 \text{ row } 1 + \alpha_2 \text{ row } 2 + \cdots + \alpha_r \text{ row } r = \mathbf{0} = (0,0,\ldots,0) \in R^n$$

$\Downarrow$

$$(\gamma, \ldots, \gamma, d_1, \gamma, \ldots, \gamma, d_2, \gamma, \ldots, \gamma, d_r, \gamma, \ldots, \gamma) = 0 = (0,0,\ldots,0) \in R^n$$

pivot col. j_1 ... pivot col. j_2 ... pivot col. j_r

where γ is a real number not necessarily the same in any two places. It is clear that we must have $\alpha_1 = \cdots = \alpha_r = 0$.

Thus row space of R = span of nonzero rows of R.

and rank(r) = row rank(R).

proof of (b) Given the upper triangular structure of R, it is easily shown that any free (non-pivot) column of R can be expressed as a linear combination of the pivot columns of R. Hence the column space of R is equal to the span of the pivot columns of R (*). Hence, it suffices to show that the pivot columns of R are linearly independent. So consider the equality

$$\alpha_1 \text{ col. } j_1 + \alpha_2 \text{ col. } j_2 + \cdots + \alpha_r \text{ col. } j_r = 0 = (0,0,\ldots,0) \in R^m$$

or

$$\alpha_1 e_1 + \alpha_2 e_2 + \cdots + \alpha_r e_r = 0$$

This translates into a system of the form

$$\left.\begin{array}{l} \alpha_1 = 0 \\ \vdots \\ \alpha_r = 0 \end{array}\right\} r$$

$$\left.\begin{array}{l} 0 = 0 \\ \vdots \\ 0 = 0 \end{array}\right\} m-r$$

It is immediately obvious that $\alpha_1 = \cdots = \alpha_r = 0$. Thus the pivot columns of R are linearly independent. We have in fact shown that

column space of R = span of $\{e_1, \ldots, e_r\}$

$\therefore$ column rank of $R = r$ = rank(r) = row rank of R.

(*) On one hand span{pivot columns} subspace of column space
on the other column space is a subspace of span{pivot columns}

proof of (c) If $\ker(A) = \{0\}$ Then there is no basis for it.

Now suppose $\ker(A) \neq \{0\}$, i.e. there are free variables. To simplify the exposition, let us assume that the pivot columns/variables are numbered $1, 2, \ldots, r$ whereas the free columns/variables are $r+1, \ldots, n$. There is no loss of generality in this assumption as it can easily be achieved by a suitable permutation of the columns of R.

In view of this assumption, R can be partitioned as follows:

$$R = \begin{array}{c|cc|l} & r & \nu & \\ \hline r & I_r & B & r \\ & 0 & 0 & m-r \end{array}$$

I_r is the $r \times r$ identity matrix, B is $\nu \times r$ matrix with $\nu = n - r$.

Let $\{e_1, e_2, \ldots, e_\nu\}$ be the canonical/standard basis of R^ν. Define the set of ν vectors $\{z_1, \ldots, z_\nu\}$ of R^n by

$$z_i = \begin{array}{l} \begin{bmatrix} -Be_i \\ \hline e_i \end{bmatrix} \begin{matrix} r \\ \nu \end{matrix} \end{array}, \quad i = 1, \ldots, \nu.$$

Clearly the set $S = \{z_1, \ldots, z_\nu\}$ is linearly independent and its elements belong to $\ker(R)$. We will show that S is indeed a basis for $\ker(R)$. To see this, let w satisfy $Rw = 0$. With

$$w = \begin{bmatrix} \tilde{w} \\ \hat{w} \end{bmatrix} \begin{matrix} r \\ \nu \end{matrix}, \quad \text{There holds } \tilde{w} + B\hat{w} = 0.$$

Now

$$\hat{w}_1 z_1 + \cdots + \hat{w}_\nu z_\nu = \hat{w}_1 \begin{bmatrix} -Be_1 \\ \hline e_1 \end{bmatrix} + \cdots + \hat{w}_\nu \begin{bmatrix} -Be_\nu \\ \hline e_\nu \end{bmatrix}$$

$$= \begin{bmatrix} -B(\hat{w}_1 e_1 + \cdots + \hat{w}_\nu e_\nu) \\ \hline \hat{w}_1 e_1 + \cdots + \hat{w}_\nu e_\nu \end{bmatrix} = \begin{bmatrix} -B\hat{w} \\ \hline \hat{w} \end{bmatrix} = \begin{bmatrix} \tilde{w} \\ \hline \hat{w} \end{bmatrix}.$$

Thus $w \in \operatorname{span}\{z_1, \ldots, z_\nu\}$. This also shows that nullity of $R = \nu = n - r$, from which the fundamental relation

$$\text{nullity}(A) + \text{rank}(A) = n$$

follows. $\square$

The case of a general matrix

Since every matrix can be transformed into reduced row echelon form using elementary row operations, a good way to treat the general case will be feasible if we can relate the 3 fundamental subspaces of A to those of its reduced row echelon form.

It is quite convenient to recall that the application of an elementary row operation to a matrix A can be duplicated by premultiplying the matrix by the elementary matrix corresponding to the elementary row operation. In essence, we can write

$$\mathrm{rref}(A) = E\,A\ , \quad \text{where } E = E_p \cdots E_1$$

is a product of elementary matrices. The important thing here is that E is invertible.

<u>Theorem</u> Let A be an $m \times n$ matrix and B an $m \times m$ invertible matrix. Then

(a) The row spaces of BA and A are the same.

(b) $\ker(BA) = \ker(A)$.

<u>Proof</u> (a) A linear combination of rows of BA is of the form $w^T(BA)$ for some vector w in $\mathbb{R}^m$.

Now

$w^T(BA) = (w^T B)A$, which is a linear combination of rows of A. In other words

row space of BA is a subspace of the row space of A.

Note that this is always the case, since we did not use invertibility of B.

If on the other hand B is invertible, then writing $A = B^{-1}(BA)$ we see that every linear combination of rows of A is a linear combination of the rows of BA. i.e.

row space of A is a subspace of the row space of BA.

This and the above show that rows of A and BA are the same as vector subspaces of $\mathbb{R}^n$.

(b) Let $z \in Ker(A)$ i.e. $Az = 0$. Then $(BA)z = B0 = 0$. Thus $Ker(A) \subseteq Ker(BA)$. Again note that this is always true since we did not use invertibility of B. Now suppose B is invertible and let $z \in Ker(BA)$.

$$\Rightarrow BAz = 0 \Rightarrow B^{-1}BAz = Az = 0. \quad \blacksquare$$

In view of this result:

A basis for the row space of a general matrix A is given by the nonzero rows of rref(A). It is also given by the nonzero rows of the row echelon form of A. It is also given by the set of nonzero rows of the upper triangular form of A.

The Null space of a general matrix A is equal to the Null space of its reduced row echelon form, it is equal to the Null space of its row echelon form and is also equal to the Null space of its upper triangular form.
Hence a basis for the Null space of A can be constructed as shown above using one of the reduced row echelon, row echelon or upper triangular forms.

As for the column space or range of A, we will show the following

A basis for the column space of a matrix A is given by the <u>columns of A</u> corresponding to pivot columns of rref(A) or the pivot columns of the row echelon form of A or the pivot columns of the upper triangular form of A.

The proof of this result is based on the following lemma

<u>Lemma</u> Let $\{v_1, v_2, \ldots, v_n\}$ be a collection/set of vectors in R^m. Let B be an $p \times m$ matrix. Then

(a) If the set $\{Bv_1, Bv_2, \ldots, Bv_n\}$ is linearly independent, then the set $\{v_1, v_2, \ldots, v_n\}$ must be linearly independent.

(b) Suppose now that B is $m \times m$ and invertible. If the set $\{v_1, v_2, \ldots, v_n\}$ is linearly independent, then the set $\{Bv_1, Bv_2, \ldots, Bv_n\}$ must be linearly independent.

Proof (a) Consider the equation $k_1 v_1 + k_2 v_2 + \cdots + k_n v_n = 0$. We want to show that $k_1 = k_2 = \cdots = k_n = 0$. Multiply the equation by B to get

$$0 = B0 = B(k_1 v_1 + \cdots + k_n v_n)$$
$$= k_1 (Bv_1) + \cdots + k_n (Bv_n).$$

Since the set $\{Bv_1, \ldots, Bv_n\}$ is assumed to be linearly independent, it follows that $k_1 = \cdots = k_n = 0$. ✓

(b) Consider now the equation $k_1(Bv_1) + \cdots + k_n(Bv_n) = 0$. We want to show that $k_1 = k_2 = \cdots = k_n = 0$. Multiply this equation by B^{-1}, to get

$$0 = B^{-1}0 = B^{-1}(k_1(Bv_1) + \cdots + k_n(Bv_n))$$
$$= B^{-1}B(k_1 v_1 + \cdots + k_n v_n)$$
$$= k_1 v_1 + \cdots + k_n v_n.$$

Since it was assumed that the set $\{v_1, \ldots, v_n\}$ is linearly independent, we must have $k_1 = \cdots = k_n = 0$. ✓ ∎

Corollary Let $\{v_1, \ldots, v_n\}$ be a set of vectors in $\mathbb{R}^m$ and let B be an $m \times m$ invertible matrix. Then the set $\{v_1, \ldots, v_n\}$ is linearly independent if and only if the set $\{Bv_1, \ldots, Bv_n\}$ is linearly independ.

We apply this corollary to the problem of finding a basis for the column space of a general matrix A using the fact that

$$\text{rref}(A) = EA$$

where the matrix E is the product of elementary matrices and is invertible. Viewing a matrix as an ordered collection of columns, we have

$$\left[\begin{array}{c|c|c|c} r_1 & r_2 & \cdots & r_n \end{array}\right] = E\left[\begin{array}{c|c|c|c} c_1 & c_2 & \cdots & c_n \end{array}\right]$$

$$= \left[\begin{array}{c|c|c|c} Ec_1 & Ec_2 & \cdots & Ec_n \end{array}\right]$$

where $r_1, \ldots, r_n$ are the columns of rref(A) and $c_1, \ldots, c_n$ are the columns of A. It follows from the Corollary that

(a) column rank of A = column rank of rref(A)

(b) A basis for the column space of A is given by the columns of A that correspond to the pivot columns of rref(A).

Ex. Find Bases for each of the row, column and null spaces of

$$A = \begin{bmatrix} 1 & -2 & 0 & 0 & 3 \\ 2 & -5 & -3 & -2 & 6 \\ 0 & 5 & 15 & 10 & 0 \\ 2 & 6 & 18 & 8 & 6 \end{bmatrix} \xrightarrow[-2\times r_1 + r_4]{-2\times r_1 + r_2} \begin{bmatrix} 1 & -2 & 0 & 0 & 3 \\ 0 & -1 & -3 & -2 & 0 \\ 0 & 5 & 15 & 10 & 0 \\ 0 & 10 & 18 & 8 & 0 \end{bmatrix}$$

$$\xrightarrow[10\times r_2 + r_4]{5\times r_2 + r_3} \begin{bmatrix} 1 & -2 & 0 & 0 & 3 \\ 0 & -1 & -3 & -2 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & -12 & -12 & 0 \end{bmatrix} \xrightarrow{r_3 \leftrightarrow r_4} U = \begin{bmatrix} 1 & -2 & 0 & 0 & 3 \\ 0 & -1 & -3 & -2 & 0 \\ 0 & 0 & -12 & -12 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

This is the upper triangular form of A and has all the information we need.

Pivot columns are 1, 2, 3, rank(A) = 3, nullity(A) = 5 − 3 = 2.

Basis for the column space is the set of columns 1, 2, 3 of A.

i.e. $\begin{bmatrix}1\\2\\0\\2\end{bmatrix}, \begin{bmatrix}-2\\-5\\5\\6\end{bmatrix}, \begin{bmatrix}0\\-3\\15\\18\end{bmatrix}$.

Basis for the row space of A is given by the nonzero rows of U

$$\{(1,-2,0,0,3),(0,-1,-3,-2,0),(0,0,-12,-12,0)\}$$

Note Another choice would be rows 1, 2 and 4 of A. This takes into account the row interchange $r_3 \leftrightarrow r_4$. The fact that the third row of A was transformed to zero, shows that it is a linear combination of rows 1, 2, 4.

Basis for Ker(A) We know there will be 2 vectors in it, each corresponding to one of the 2 free variables x_4, x_5.

set $x_4 = 1$, $x_5 = 0$, solve for x_1, x_2, x_3 from $Ux = 0$

$\Rightarrow z_1 = (2, 1, -1, 1, 0)$

set $x_4 = 0$, $x_5 = 1$, solve for x_1, x_2, x_3

$\Rightarrow z_2 = (-3, 0, 0, 0, 1)$

Remark A basis for $Ker(A^T)$ can be obtained by reducing A^T to upper triangular form. From $m = \text{nullity}(A^T) + \text{rank}(A)$, we know that $\text{nullity}(A^T) = 4 - 3 = 1$, i.e. $Ker(A^T)$ is a one-dimensional subspace of R^4.

There is another way to obtain a basis for $Ker(A^T)$, namely through knowledge of the range or column space of A. We will discuss this approach in the next section.

Remark Besides providing useful information about matrices, and linear transformations, the procedures studied in this section can be used to solve a host of problems of general nature in vector spaces.

Problem Given a set of nonzero vectors $S = \{v_1, \ldots, v_k\}$ in R^n;
(a) Find a subset of S that forms a basis for span(S).
(b) Express each vector that is not in that basis as a linear combination of the basis vectors.

Solution

(a) (1) Form the matrix A whose columns are the vectors in S
(2) Reduce/transform A into upper triangular form U or rreform.
(3) Identify the pivot columns of U or of rref(A).
(4) The columns of A corresponding to the pivot columns of U form a basis for span(S).

(b) This part is handled more efficiently if rref(A) is available
(1) Express each free column of rref(A) as a linear combination of the pivot columns of rref(A). This process is quite simple since the pivot columns are nothing but the standard basis of $R^{\text{rank}(A)}$.

(2) A column of A that corresponds to a free column of rref(A) is the same linear combination of the basis columns of A.

Ex. Find a basis for the set of vectors already organized as columns of the matrix

$$A = \begin{bmatrix} 2 & 6 & 0 & 8 & 4 & 12 & 4 \\ 3 & 9 & -2 & 8 & 6 & 18 & 6 \\ 3 & 9 & -7 & -2 & 6 & -3 & -1 \\ 2 & 6 & 5 & 18 & 4 & 33 & 11 \\ 1 & 3 & -2 & 0 & 2 & 6 & 2 \end{bmatrix}$$

From rank$(A) \le \min\{m, n\}$, we know that at most 5 of the columns of A can be linearly independent.
We also know from $n = \text{nullity}(A) + \text{rank}(A)$ that nullity$(A) \ge 2$.

using Matlab, we have

$$\text{rref}(A) = \begin{bmatrix} 1 & 3 & 0 & 4 & 2 & 0 & 0 \\ 0 & 0 & 1 & 2 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 1/3 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

we see that actually the rank is 3. Also, the 1st, third and 6th columns of A, i.e.

$$\underset{C_1}{(2,3,3,2,1)},\ \underset{C_3}{(0,-2,-7,5,-2)},\ \underset{C_6}{(12,18,-3,33,6)}$$

form a basis for the column space of A.

Furthermore

$C_2 = 3C_1$

$C_4 = 4C_1 + 2C_3$

$C_5 = 2C_1$

$C_7 = \frac{1}{3} C_6$