2: LINEAR FUNCTIONS AND MATRICES II

1. The rank-nullity theorem

Definition. Let V and W be vector spaces and $T : V \longrightarrow W$ a linear function.
(a) The *kernel* of T is subset of V:

$$\ker(T) = \{\mathbf{v} \in V \,|\, T(\mathbf{v}) = \mathbf{0}_W\} \subseteq V.$$

(b) The *image* of W is the subset of W:

$$\operatorname{im}(T) = \{\mathbf{w} \in W \,|\, \text{there is an } \mathbf{v} \in V \text{ with } T(\mathbf{v}) = \mathbf{w}\}.$$

Remark. $\ker(T)$ is a subspace of V and $\operatorname{im}(T)$ is a subspace of W. Kernel and image of a linear function are important subspaces which carry a lot of information about the linear function.

Theorem 3.1. *Let V and W be vector spaces and $T : V \longrightarrow W$ a linear function.*
(a) *T is one-to-one if and only if $\ker(T) = \{\mathbf{0}_V\}$.*
(b) *T is onto if and only if $\operatorname{im}(T) = W$.*

Proof. (a) For the forward direction $\Rightarrow$ assume that T is a one-to-one linear function and $\mathbf{v} \in \ker(T)$. Then $T(\mathbf{v}) = \mathbf{0}_W = T(\mathbf{0}_V)$ and $\mathbf{v} = \mathbf{0}_V$. Conversely, if $\ker(T) = \{\mathbf{0}_V\}$ let $\mathbf{u}, \mathbf{v} \in V$ with $T(\mathbf{v}) = T(\mathbf{u})$. Then, since T is linear,

$$\mathbf{0}_W = T(\mathbf{v}) - T(\mathbf{u}) = T(\mathbf{v} - \mathbf{u})$$

and $\mathbf{v} - \mathbf{u} \in \ker(T) = \{\mathbf{0}_V\}$. Hence $\mathbf{v} - \mathbf{u} = \mathbf{0}_V$ and T is one-to-one.
(b) If T is onto, then for every vector $\mathbf{w} \in W$ there is a $\mathbf{v} \in V$ with $T(\mathbf{v}) = \mathbf{w}$. Thus, $\operatorname{im}(T) = W$. Conversely, if $\operatorname{im}(T) = W$, then for every $\mathbf{w} \in W$ there is a $\mathbf{v} \in V$ with $T(\mathbf{v}) = \mathbf{w}$. T is onto.

The following theorem relates the dimensions of kernel and image of a linear function, provided that the dimension of the vector space V is finite.

Theorem 3.2. *Let V and W be vector spaces and $T : V \longrightarrow W$ a linear function. Suppose that V is finite dimensional. Then:*

$$\dim(\ker(T)) + \dim(\operatorname{im}(T)) = \dim(V).$$

Proof. Suppose that $\dim(\ker(T)) = r$ and let $\{\mathbf{v}_1, \dots, \mathbf{v}_r\}$ be a basis of $\ker(T)$. Since every linearly independent sequence can be extended to a basis of the vector space, we can extend $\mathbf{v}_1, \dots, \mathbf{v}_r$ to a basis of V, say, $\{\mathbf{v}_1, \dots, \mathbf{v}_r, \mathbf{v}_{r+1}, \dots, \mathbf{v}_n\}$ is a basis of V. The formula follows if we can show that the set $\{T(\mathbf{v}_{r+1}), \dots, T(\mathbf{v}_n)\}$ is a basis of $\operatorname{im}(T)$.

Typeset by $\mathcal{A}\mathcal{M}\mathcal{S}$-TEX

We first show that $\{T(\mathbf{v}_{r+1}), \ldots, T(\mathbf{v}_n)\}$ is a spanning set. Let $\mathbf{w} \in \operatorname{im}(T)$. Then there is a vector $\mathbf{v} \in V$ with $T(\mathbf{v}) = \mathbf{w}$. Since $\{\mathbf{v}_1, \ldots, \mathbf{v}_r, \mathbf{v}_{r+1}, \ldots, \mathbf{v}_n\}$ is a basis of V, there are scalars $t_i \in \mathbb{R}$ so that

$$\mathbf{v} = t_1\mathbf{v}_1 + \ldots + t_r\mathbf{v}_r + t_{r+1}\mathbf{v}_{r+1} + \ldots + t_n\mathbf{v}_n.$$

This implies

$$\begin{aligned}\mathbf{w} &= T(\mathbf{v})\\ &= T(t_1\mathbf{v}_1 + \ldots + t_r\mathbf{v}_r + t_{r+1}\mathbf{v}_{r+1} + \ldots + t_n\mathbf{v}_n)\\ &= t_1T(\mathbf{v}_1) + \ldots + t_rT(\mathbf{v}_r) + t_{r+1}T(\mathbf{v}_{r+1}) + \ldots + t_nT(\mathbf{v}_n)\\ &= t_{r+1}T(\mathbf{v}_{r+1}) + \ldots + t_nT(\mathbf{v}_n)\end{aligned}$$

The last equation follows since by assumption $\mathbf{v}_i \in \ker(T)$ for all $1 \leq i \leq r$ and therefore $T(\mathbf{v}_i) = \mathbf{0}_W$ for $1 \leq i \leq r$. This shows that $\{T(\mathbf{v}_{r+1}), \ldots, T(\mathbf{v}_n)\}$ is a spanning set of $\operatorname{im}(T)$.

In order to show that $\{T(\mathbf{v}_{r+1}), \ldots, T(\mathbf{v}_n)\}$ is linearly independent, suppose that $t_i \in \mathbb{R}$ with

$$t_{r+1}T(\mathbf{v}_{r+1}) + \ldots + t_nT(\mathbf{v}_n) = \mathbf{0}_W.$$

Since T is linear we have that

$$t_{r+1}T(\mathbf{v}_{r+1}) + \ldots + t_nT(\mathbf{v}_n) = \mathbf{0}_W = T(t_{r+1}\mathbf{v}_{r+1} + \ldots + t_n\mathbf{v}_n)$$

and

$$t_{r+1}\mathbf{v}_{r+1} + \ldots + t_n\mathbf{v}_n \in \ker(T).$$

Using that $\{\mathbf{v}_1, \ldots, \mathbf{v}_r\}$ is a basis of $\ker(T)$, there are scalars $s_i \in \mathbb{R}$ so that

$$t_{r+1}\mathbf{v}_{r+1} + \ldots + t_n\mathbf{v}_n = s_1\mathbf{v}_1 + \ldots + s_r\mathbf{v}_r.$$

Thus

$$s_1\mathbf{v}_1, \ldots, s_r\mathbf{v}_n - t_{r+1}\mathbf{v}_{r+1} - \ldots - t_n\mathbf{v}_n = \mathbf{0}_V$$

and $s_i = t_j = 0$ for all i and j, since $\{\mathbf{v}_1, \ldots, \mathbf{v}_r, \mathbf{v}_{r+1}, \ldots, \mathbf{v}_n\}$ is a basis of V. This shows that the set $\{T(\mathbf{v}_{r+1}), \ldots, T(\mathbf{v}_n)\}$ is linearly independent.

Let A be an $m \times n$ matrix and $\mu_A : \mathbb{R}^n \longrightarrow \mathbb{R}^m$ the linear function defined by $\mu_A(\mathbf{u}) = A\mathbf{u}$ for all $\mathbf{u} \in \mathbb{R}^n$. Let $\mu_A(\mathbf{e}_i) = \mathbf{A}_i$ denote the ith column of A. We write

$$A = [\mathbf{A}_1, \ldots, \mathbf{A}_n]$$

where $\mathbf{A}_i \in \mathbb{R}^m$ is the ith column of A.

Definition. (a) The *column space* $C(A)$ of A is the subspace of $\mathbb{R}^m$ which is generated by the columns of A, i.e. $C(A) = \operatorname{span}\{\mathbf{A}_1, \ldots, \mathbf{A}_n\} \subseteq \mathbb{R}^m$.
(b) The number $\dim(C(A))$ is called the *column rank* of A and is denoted by $\operatorname{Crk}(A)$.

Remark. (a) The column space of A is the image of the linear function μ_A, that is,

$$\operatorname{im}(\mu_A) = \operatorname{span}\{\mathbf{A}_1, \ldots, \mathbf{A}_n\} = C(A).$$

Therefore $\dim(\operatorname{im}(A)) = \dim(C(A)) = \operatorname{Crk}(A)$.
(b) Note that the kernel of μ_A is the solution set of the homogeneous linear system $A\mathbf{x} = \mathbf{0}$.

Definition. The number $\dim(\ker(\mu_A))$ is called the *nullity* of A and is denoted by $\operatorname{null}(A)$.

Now Theorem 3.2 yields:

Theorem 3.3 (Rank-Nullity-Theorem). *Let A be an $m \times n$ matrix. Then:*

$$\mathrm{Crk}(A) + \mathrm{null}(A) = n.$$

Remark. Suppose that

$$A = \begin{bmatrix} \mathbf{a}_1 \\ \mathbf{a}_2 \\ \vdots \\ \mathbf{a}_m \end{bmatrix}$$

where $\mathbf{a}_i$ is the ith row of A. In the previous chapter we defined the row space of A as the subspace of $\mathbb{R}^n$ spanned by the rows of A:

$$R(A) = \mathrm{span}\{\mathbf{a}_1, \ldots, \mathbf{a}_n\}.$$

The *row rank* of A is the dimension of the row space of A:

$$\mathrm{Rrk}(A) = \dim(R(A)) = \dim(\mathrm{span}\{\mathbf{a}_1, \ldots, \mathbf{a}_n\}).$$

The row rank of a matrix is easy to compute, since it does not change under elementary row operations. So, if A is an $m \times n$ matrix, we bring A into reduced echelon form D. Then the row rank of A is the number of nonzero rows of D. Since the row space does not change under elementary row operations, the nonzero rows of D provide a basis of the row space $R(A)$. Note that the column space changes under elementary row operations and the linearly independent columns of D are NOT a bais of $C(A)$. The following statement, however, is true:

Theorem 3.4. *If A is an $m \times n$ matrix, then*

$$Rrk(A) = Crk(A).$$

Proof. Suppose that D is the row reduced echelon form of A. The kernel of μ_A, that is, the solution set of the linear system $A\mathbf{x} = \mathbf{0}$, equals the solution set of the homogeneous system $D\mathbf{x} = \mathbf{0}$. Thus the dimension of $\ker(\mu_A)$ is the number of free variables of the system $D\mathbf{x} = \mathbf{0}$ which is the number of columns of D without a pivot one. On the other hand, the number of rows of D with pivot ones is exactly the dimension of $R(A)$. This gives:

$$\begin{aligned} \mathrm{Rrk}(A) &= n - \dim(\ker(A)) \\ &= n - \mathrm{null}(A). \end{aligned}$$

By Theorem 3.3:

$$n - \mathrm{null}(A) = \mathrm{Crk}(A)$$

and we have shown that $\mathrm{Rrk}(A) = \mathrm{Crk}(A)$.

Definition. If A is an $m \times n$ matrix the *rank* of A, denoted $\mathrm{rk}(A)$, is the row (or column) rank of A.

Example. Consider the matrix:

$$A = \begin{bmatrix} 0 & 3 & -6 & 6 & 4 & -5 \\ 3 & -7 & 8 & -5 & 8 & 9 \\ 3 & -9 & 12 & -9 & 6 & 15 \\ 3 & -4 & 2 & 1 & 12 & 4 \end{bmatrix}$$

with row reduced echelon form

$$D = \begin{bmatrix} 1 & 0 & -2 & 3 & 0 & -24 \\ 0 & 1 & -2 & 2 & 0 & -7 \\ 0 & 0 & 0 & 0 & 1 & 4 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}.$$

From D we obtain that

$$\{(1,0,-2,3,0,-24),(0,1,-22,0,-7),(0,0,0,0,1,4)\} \subseteq \mathbb{M}(1,6)$$

is a basis of $R(A)$ and hence $\operatorname{rk}(A) = 3$. Therefore the associated linear function $\mu_A : \mathbb{R}^6 \longrightarrow \mathbb{R}^4$ maps $\mathbb{R}^4$ into a 3-dimensional subspace of $\mathbb{R}^4$. From the reduced echelon form D we see that

$$\left\{ \begin{bmatrix} 2 \\ 2 \\ 1 \\ 0 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} -3 \\ -2 \\ 0 \\ 1 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 24 \\ 7 \\ 0 \\ 0 \\ -4 \\ 1 \end{bmatrix} \right\}$$

is a basis of $\ker(\mu_A)$ and $\operatorname{null}(A) = 3$. In order to find a basis of $C(A)$ we use a theorem we have not shown in class:

Theorem. *Let A be an $m \times n$ matrix and D its reduced echelon form. The columns of A which correspond to the columns of D with pivot ones form a basis of the column space of A.*

In the example the first, second, and fifth column of D are the columns of D with pivot ones. Thus the first, second, and fifth column of A form a basis of $C(A)$, i.e.

$$\left\{ \begin{bmatrix} 0 \\ 3 \\ 3 \\ 3 \end{bmatrix}, \begin{bmatrix} 3 \\ -7 \\ -9 \\ -4 \end{bmatrix}, \begin{bmatrix} 4 \\ 8 \\ 6 \\ 12 \end{bmatrix} \right\}.$$

is a basis of $C(A)$.

2. Coordinate functions and the matrix of a linear function

Let V be a finite dimensional vector space and $B = \{\mathbf{v}_1, \dots, \mathbf{v}_n\}$ a basis of V. Usually in a set it does not matter in which order the elements of the set are listed, for example, $\{1,2\} = \{2,1\}$. In the following we need to fix an order on B. We will say that $B = \{\mathbf{v}_1, \dots, \mathbf{v}_n\}$ is an *ordered basis* of V and mean that now the order of the basis vectors is fixed, that is, vector $\mathbf{v}_1$ comes first, vector $\mathbf{v}_2$ is second and so on. In particular, we consider the ordered basis $B = \{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n\}$ different from the ordered basis $B' = \{\mathbf{v}_2, \mathbf{v}_1, \ldots, \mathbf{v}_n\}$ although the sets are the same.

We fix the order of the vectors in B and assume that $B = \{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ is an ordered basis of V. Then there is a unique linear function

$$[\ ]_B : V \longrightarrow \mathbb{R}^n$$

defined by $[\ ]_B(\mathbf{v}_i) = [\mathbf{v}_i]_B = \mathbf{e}_i$ where $\mathbf{e}_i$ is the ith standard basis vector of $\mathbb{R}^n$. If $\mathbf{v} \in V$ with $\mathbf{v} = r_1\mathbf{v}_1 + \ldots + r_n\mathbf{v}_n$ then

$$[\mathbf{v}]_B = [r_1\mathbf{v}_1 + \ldots + r_n\mathbf{v}_n]_B = \begin{bmatrix} r_1 \\ r_2 \\ \vdots \\ r_n \end{bmatrix}.$$

This vector $[\mathbf{v}]_B$ is called the *coordinate vector* of the vector $\mathbf{v}$ with respect to the ordered basis B. The linear function $[\ ]_B : V \longrightarrow \mathbb{R}^n$ is called the *coordinate function* of V with respect to the ordered basis B.

Since $[\ ]_B$ maps a basis of V onto a basis of $\mathbb{R}^n$, $[\ ]_B$ is an isomorphism of vector spaces and has an inverse function $L_B : \mathbb{R}^n \longrightarrow V$. L_B is determined by $L_B(\mathbf{e}_i) = \mathbf{v}_i$ for all $1 \leq i \leq n$ and thus is given by:

$$L_B\left(\begin{bmatrix} t_1 \\ t_2 \\ \vdots \\ t_n \end{bmatrix}\right) = t_1\mathbf{v}_1 + t_2\mathbf{v}_2 + \ldots + t_n\mathbf{v}_n.$$

Suppose now that V and W are finite dimensional vector spaces and that $B_V = \{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ is an ordered basis of V and $B_W = \{\mathbf{w}_1, \ldots, \mathbf{w}_m\}$ an ordered basis of W. If $T : V \longrightarrow W$ is a linear function we obtain a diagram:

$$\begin{array}{ccc} V & \xrightarrow{T} & W \\ {\scriptstyle [\,]_{B_V}}\downarrow & & \downarrow{\scriptstyle [\,]_{B_W}} \\ \mathbb{R}^n & & \mathbb{R}^m \end{array}$$

The question is if we can complete this diagram in a nice way, this means, if there is a linear function $S : \mathbb{R}^n \longrightarrow \mathbb{R}^m$ so that we have a diagram

$$\begin{array}{ccc} V & \xrightarrow{T} & W \\ {\scriptstyle [\ ]_{B_V}}\downarrow & & \downarrow{\scriptstyle [\ ]_{B_W}} \\ \mathbb{R}^n & \xrightarrow{S} & \mathbb{R}^m \end{array}$$

with $S \circ [\ ]_{B_V} = [\ ]_{B_W} \circ T$. Actually there is. Since coordinate functions are isomorphisms, there is a linear function

$$S = [\ ]_{B_W} \circ T \circ L_B : \mathbb{R}^n \longrightarrow \mathbb{R}^m$$

with

$$S \circ [\ ]_{B_V} = [\ ]_{B_W} \circ T \circ L_{B_V} \circ [\ ]_{B_V} = [\ ]_{B_W} \circ T \circ \mathrm{id}_V = [\ ]_{B_W} \circ T.$$

In general we define: A diagram of sets and functions

$$\begin{array}{ccc} X & \xrightarrow{f} & Y \\ {\scriptstyle h}\big\downarrow & & \big\downarrow{\scriptstyle g} \\ S & \xrightarrow{\ell} & T \end{array}$$

is called a *commutative diagram* if $g \circ f = \ell \circ h$.

We have just proved:

Theorem 3.5. *Let V and W are vector spaces with ordered basis $B_V = \{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ and $B_W = \{\mathbf{w}_1, \ldots, \mathbf{w}_m\}$ and $T : V \longrightarrow W$ a linear function. Then there is a linear function $S = \mu_A : \mathbb{R}^n \longrightarrow \mathbb{R}^m$ so that the diagram*

$$\begin{array}{ccc} V & \xrightarrow{T} & W \\ {\scriptstyle [\ ]_{B_V}}\big\downarrow & & \big\downarrow{\scriptstyle [\ ]_{B_W}} \\ \mathbb{R}^n & \xrightarrow{\mu_A} & \mathbb{R}^m \end{array}$$

commutes, that is, $\mu_A \circ [\ ]_{B_V} = [\ ]_{B_W} \circ T$.

Definition. The $m \times n$ matrix of Theorem 3.5 is called the *matrix of T with respect to ordered basis B_V and B_W.*

Suppose as above that V and W are vector spaces with ordered basis $B_V = \{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ and $B_W = \{\mathbf{w}_1, \ldots, \mathbf{w}_m\}$ and that $T : V \longrightarrow W$ is a linear function. We want to compute the matrix of T with respect to B_V and B_W. Remember that the function $S : \mathbb{R}^n \longrightarrow \mathbb{R}^m$ is defined by:

$$S = [\ ]_{B_W} \circ T \circ L_{B_V}.$$

Hence $S = \mu_A$ and the ith column of A equals $S(\mathbf{e}_i)$ which is given by:

$$\begin{aligned} \mathbf{A}_i &= S(\mathbf{e}_i) \\ &= ([\ ]_{B_W} \circ T \circ L_{B_V})(\mathbf{e}_i) \\ &= ([\ ]_{B_W} \circ T)(L_{B_V}(\mathbf{e}_i) \\ &= ([\ ]_{B_W} \circ T)(\mathbf{v}_i) \\ &= [\ ]_{B_W}(T(\mathbf{v}_i)) \\ &= [T(\mathbf{v}_i)]_{B_W}. \end{aligned}$$

Thus the ith column of A is $[T(\mathbf{v}_i)]_{B_W}$ and therefore

$$A = [[T(\mathbf{v}_1)]_{B_W}, \ldots, [T(\mathbf{v}_n)]_{B_W}].$$

Example. Let $T : \mathbb{P}_3 \longrightarrow \mathbb{P}_2$ be the linear function defined by $T(f(x)) = f'(x)$. If we choose for ordered basis of $\mathbb{P}_3$ the set $B_3 = \{1, x, x^2, x^3\}$ and for $\mathbb{P}_2$ the ordered basis $B_2 = \{1, x, x^2\}$, then we compute

$$T(1) = 0 \text{ and } [T(0)]_{B_2} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

$$T(x) = 1 \text{ and } [T(x)]_{B_2} = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$$

$$T(x^2) = 2x \text{ and } [T(x^2)]_{B_2} = \begin{bmatrix} 0 \\ 2 \\ 0 \end{bmatrix}$$

$$T(x^3) = 3x^2 \text{ and } [T(x^3)]_{B_2} = \begin{bmatrix} 0 \\ 0 \\ 3 \end{bmatrix}$$

Thus the matrix of T with respect to basis B_3 and B_2 is:

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

If we choose different ordered basis of $\mathbb{P}_3$ and $\mathbb{P}_2$ say, $B_3' = \{1, x+1, (x+1)^2, (x+1)^3\}$ for $\mathbb{P}_3$ and $B_2' = \{(x-1)^2, x-1, 1\}$ for $\mathbb{P}_2$ we have:

$$\begin{aligned} T(1) &= 0 \\ T(x+1) &= 1 \\ T((x+1)^2) &= 2(x-1) + 4 \\ T((x+1)^3) &= 3(x-1)^2 + 12(x-1) + 12 \end{aligned}$$

This gives:

$$[T(0)]_{B_2} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

$$[T(x)]_{B_2} = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$$

$$[T(x^2)]_{B_2} = \begin{bmatrix} 0 \\ 2 \\ 4 \end{bmatrix}$$

$$[T(x^3)]_{B_2} = \begin{bmatrix} 3 \\ 12 \\ 12 \end{bmatrix}$$

and the matrix of T with respect to the ordered basis B_3' and B_2' is:

$$A' = \begin{bmatrix} 0 & 0 & 0 & 3 \\ 0 & 0 & 2 & 12 \\ 0 & 1 & 4 & 12 \end{bmatrix}.$$

3. Base Change

Suppose that V is a finite dimensional vector space with ordered bases $B = \{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ and $B' = \{\mathbf{u}_1, \ldots, \mathbf{u}_n\}$. Let $\mathbf{v} \in V$ a vector with coordinates (with respect to B)

$$[\mathbf{v}]_B = \begin{bmatrix} r_1 \\ r_2 \\ \vdots \\ r_n \end{bmatrix}$$

that is, $\mathbf{v} = r_1\mathbf{v}_1 + \ldots + r_n\mathbf{v}_n$. We want to find a formula to get from the coordinate vector $[\mathbf{v}]_B$ with respect to ordered basis B to the coordinate vector $[\mathbf{v}]_{B'}$ with respect to B'.

The identity function $\mathrm{id}_V : V \longrightarrow V$ is a linear function. According to section 2 there is a linear function $\mu_P : \mathbb{R}^n \longrightarrow \mathbb{R}^n$ so that the following diagram commutes:

$$\begin{array}{ccc} V & \xrightarrow{\mathrm{id}_V} & V \\ {\scriptstyle [\,]_B}\big\downarrow & & \big\downarrow{\scriptstyle [\,]_{B'}} \\ \mathbb{R}^n & \xrightarrow{\mu_P} & \mathbb{R}^n \end{array}$$

We have that

$$\begin{aligned} \mu_P([\mathbf{v}]_B) &= \mu_P \circ [\ ]_B(\mathbf{v}) \\ &= [\ ]_{B'} \circ \mathrm{id}_V(\mathbf{v}) \\ &= [\mathbf{v}]_{B'} \end{aligned}$$

Thus μ_P maps the coordinates of a vector with respect to ordered basis B into the coordinates of the vector with respect to ordered basis B'. This means that there is an $n \times n$ matrix P so that

$$P[\mathbf{v}]_B = [\mathbf{v}]_{B'}.$$

Definition. The $n \times n$ matrix P is called the *change-of-bases-matrix* from ordered basis B to ordered basis B'.

Remark. Note that the function μ_P is the composition of isomorphisms:

$$\mu_P = [\ ]_{B'} \circ \mathrm{id}_V \circ L_B = [\ ]_{B'} \circ L_B.$$

In particular, μ_P is an isomorphism and P is an invertible matrix.

Suppose now that V and W are finite dimensional vector space and that $T : V \longrightarrow W$ is a linear function. Let $B_V = \{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ and $B'_V = \{\mathbf{v}'_1, \ldots, \mathbf{v}'_n\}$ be ordered bases of V and $B_W = \{\mathbf{w}_1, \ldots, \mathbf{w}_m\}$ and $B'_W = \{\mathbf{w}'_1, \ldots, \mathbf{w}'_m\}$ be ordered basis of W. If A is the matrix of T with respect to ordered bases B_V and B_W we have a commutative diagram:

$$\begin{array}{ccc} V & \xrightarrow{T} & W \\ {\scriptstyle [\,]_{B_V}}\big\downarrow & & \big\downarrow{\scriptstyle [\,]_{B_W}} \\ \mathbb{R}^n & \xrightarrow{\mu_A} & \mathbb{R}^m \end{array}$$

We want to compute the matrix B of T with respect to the bases B'_V and B'_W. Let P denote the change-of-bases-matrix for changing from basis B'_V to basis B_V, This gives a commutative diagram:

$$\begin{array}{ccc} V & \xrightarrow{\mathrm{id}_V} & V \\ {\scriptstyle [\,]_{B'_V}}\downarrow & & \downarrow{\scriptstyle [\,]_{B_V}} \\ \mathbb{R}^n & \xrightarrow{\mu_P} & \mathbb{R}^n \end{array}$$

If Q is the change-of-bases-matrix from basis B_W to basis B'_W we obtain a third commutative diagram:

$$\begin{array}{ccc} W & \xrightarrow{\mathrm{id}_W} & W \\ {\scriptstyle [\,]_{B_W}}\downarrow & & \downarrow{\scriptstyle [\,]_{B'_W}} \\ \mathbb{R}^m & \xrightarrow{\mu_Q} & \mathbb{R}^m \end{array}$$

Using these three commutative diagrams we obtain a big diagram:

$$\begin{array}{ccccccc} V & \xrightarrow{\mathrm{id}_V} & V & \xrightarrow{T} & W & \xrightarrow{\mathrm{id}_W} & W \\ {\scriptstyle [\,]_{B'_V}}\downarrow & & {\scriptstyle [\,]_{B_V}}\downarrow & & {\scriptstyle [\,]_{B_W}}\downarrow & & \downarrow{\scriptstyle [\,]_{B'_W}} \\ \mathbb{R}^n & \xrightarrow{\mu_P} & \mathbb{R}^n & \xrightarrow{\mu_A} & \mathbb{R}^m & \xrightarrow{\mu_Q} & \mathbb{R}^m \end{array}$$

All squares in this diagram commute again. Thus the big (exterior diagram) gives a commutative diagram:

$$\begin{array}{ccc} V & \xrightarrow{T} & W \\ {\scriptstyle [\,]_{B'_V}}\downarrow & & \downarrow{\scriptstyle [\,]_{B_W}} \\ \mathbb{R}^n & \xrightarrow{\mu_{QAP}} & \mathbb{R}^m \end{array}$$

where $\mu_{QAP} = \mu_Q \circ \mu_A \circ \mu_P$. We have just shown:

Theorem 3.6. *Let $T : V \longrightarrow W$ be a linear function of finite dimensional vector spaces. If A is the matrix of T with respect to ordered bases B_V and B_W, then the matrix of T with respect to ordered bases B'_V and B'_W is*

$$QAP$$

where P is the change-of-bases-matrix from B'_V to B_V and Q is the change-of-bases-matrix from B_W to B'_W.

Example. In the previous section we studied the differentiation function $T : \mathbb{P}_3 \longrightarrow \mathbb{P}_2$. We found that the matrix of T with respect to ordered bases $B_3 = \{1, x, x^2, x^3\}$ and $B_2 = \{1, x, x^2\}$ is:

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

With respect to the ordered bases $B'_3 = \{1, x+1, (x+1)^2, (x+1)^3\}$ and $B'_2 = \{(x-1)^2, x-1, 1\}$ the matrix of T is:

$$A' = \begin{bmatrix} 0 & 0 & 0 & 3 \\ 0 & 0 & 2 & 12 \\ 0 & 1 & 4 & 12 \end{bmatrix}.$$

Since

$$\begin{aligned}
1 &= 1 \\
x+1 &= x+1 \\
(x+1)^2 &= x^2+2x+1 \\
(x+1)^3 &= x^3+3x^2+3x+1
\end{aligned}$$

the change-of-bases-matrix from B_3' to B_3 is:

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

Similarly,

$$\begin{aligned}
1 &= 1 \\
x &= (x-1)+1 \\
x^2 &= (x-1)^2+2(x-1)+1
\end{aligned}$$

and

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

is the change-of-bases-matrix from ordered basis B_2 to ordered basis B_2'. Then

$$\begin{aligned}
QAP &= \begin{bmatrix} 0 & 0 & 1 \\ 0 & 1 & 2 \\ 1 & 1 & 1 \end{bmatrix} \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 3 \end{bmatrix} \begin{bmatrix} 1 & 1 & 1 & 1 \\ 0 & 1 & 2 & 3 \\ 0 & 0 & 1 & 3 \\ 0 & 0 & 0 & 1 \end{bmatrix} \\
&= \begin{bmatrix} 0 & 0 & 1 \\ 0 & 1 & 2 \\ 1 & 1 & 1 \end{bmatrix} \begin{bmatrix} 0 & 1 & 2 & 3 \\ 0 & 0 & 2 & 6 \\ 0 & 0 & 0 & 3 \end{bmatrix} \\
&= \begin{bmatrix} 0 & 0 & 0 & 3 \\ 0 & 0 & 2 & 12 \\ 0 & 1 & 4 & 12 \end{bmatrix} \\
&= A'.
\end{aligned}$$