1.7 Linear morphisms II

In this section we will discuss the relationship between linear morphisms (of finite dimensional $\mathbb{K}$-vector spaces) and matrices. This material should be familiar to you from your first linear algebra course.

Throughout this section all $\mathbb{K}$-vector spaces will be assumed finite dimensional.

Definition 1.7.1. Let $f : V \to W$ be a linear morphism of $\mathbb{K}$-vector spaces, $\mathcal{B} = (b_1, \dots, b_n) \subset V, \mathcal{C} = (c_1, \dots, c_m) \subset W$ ordered bases of V, W. Then, the *matrix of f with respect to $\mathcal{B}$ and $\mathcal{C}$* is the $m \times n$ matrix

$$[f]_{\mathcal{B}}^{\mathcal{C}} = [[f(b_1)]_{\mathcal{C}}\ [f(b_2)]_{\mathcal{C}} \cdots [f(b_n)]_{\mathcal{C}}],$$

so that the i^{th} column of $[f]_{\mathcal{B}}^{\mathcal{C}}$ is the $\mathcal{C}$-coordinate vector of $f(b_i) \in W$.

If $V = W$ and $\mathcal{B} = \mathcal{C}$ then we write $[f]_{\mathcal{B}} \stackrel{def}{=} [f]_{\mathcal{B}}^{\mathcal{B}}$.

Lemma 1.7.2. *Let $f : V \to W$ be a linear morphism of $\mathbb{K}$-vector spaces, $\mathcal{B} \subset V, \mathcal{C} \subset W$ ordered bases of V, W. Then, for every $v \in V$, we have*

$$[f(v)]_{\mathcal{C}} = [f]_{\mathcal{B}}^{\mathcal{C}}[v]_{\mathcal{B}}.$$

Moreover, if A is an $m \times n$ matrix such that

$$[f(v)]_{\mathcal{C}} = A[v]_{\mathcal{B}}, \quad \textit{for every } v \in V,$$

then $A = [f]_{\mathcal{B}}^{\mathcal{C}}$.

This result should be familiar to you. Note that the standard matrix A_f we defined previously for a linear morphism $f \in \mathrm{Hom}_{\mathbb{K}}(\mathbb{K}^n, \mathbb{K}^m)$ is just

$$A_f = [f]_{\mathcal{S}^{(n)}}^{\mathcal{S}^{(m)}}.$$

We can record the conclusion of the Lemma 1.7.2 in a diagram in a similar fashion as we did in the previous section for $P_{\mathcal{C} \leftarrow \mathcal{B}}$. We have the commutative diagram

$$\begin{tikzcd}
V \arrow[r, "f"] \arrow[d, "{[-]_{\mathcal{B}}}"'] \arrow[dr, phantom, "\circlearrowleft"] & W \arrow[d, "{[-]_{\mathcal{C}}}"] \\
\mathbb{K}^n \arrow[r, "T_{[f]_{\mathcal{B}}^{\mathcal{C}}}"'] & \mathbb{K}^m
\end{tikzcd}$$

where $T_{[f]_{\mathcal{B}}^{\mathcal{C}}} : \mathbb{K}^n \to \mathbb{K}^m$ is the 'multiplication by $[f]_{\mathcal{B}}^{\mathcal{C}}$' morphism and the symbol '$\circlearrowleft$' is translated to mean

'the composite morphism $[-]_{\mathcal{C}} \circ f : V \to \mathbb{K}^m$ equals the composite morphism $T_{[f]_{\mathcal{B}}^{\mathcal{C}}} \circ [-]_{\mathcal{B}} : V \to \mathbb{K}^m$';

this is precisely the statement of Lemma 1.7.2.

So, given ordered bases $\mathcal{B} = (b_1, \dots, b_n) \subset V$ and $\mathcal{C} = (c_1, \dots, c_m) \subset W$ of $\mathbb{K}$-vector space V and W, we have just defined a function

$$[-]_{\mathcal{B}}^{\mathcal{C}} : \mathrm{Hom}_{\mathbb{K}}(V, W) \to Mat_{m,n}(\mathbb{K})\ ;\ f \mapsto [f]_{\mathcal{B}}^{\mathcal{C}}.$$

In fact, this correspondence obeys some particularly nice properties:

Theorem 1.7.3. *The function*

$$[-]_{\mathcal{B}}^{\mathcal{C}} : \mathrm{Hom}_{\mathbb{K}}(V, W) \to Mat_{m,n}(\mathbb{K})\ ;\ f \mapsto [f]_{\mathcal{B}}^{\mathcal{C}},$$

satisfies the following properties:

a) $[-]^{\mathcal{C}}_{\mathcal{B}}$ *is an isomorphism of* $\mathbb{K}$*-vector spaces,*

b) *if* $f \in \mathrm{Hom}_{\mathbb{K}}(U, V), g \in \mathrm{Hom}_{\mathbb{K}}(V, W)$ *and* $\mathcal{A} \subset U, \mathcal{B} \subset V, \mathcal{C} \subset W$ *are bases of* U, V, W*, then*

$$[g \circ f]^{\mathcal{C}}_{\mathcal{A}} = [g]^{\mathcal{C}}_{\mathcal{B}}[f]^{\mathcal{B}}_{\mathcal{A}}.$$

Here

$$g \circ f : U \to V \to W \in \mathrm{Hom}_{\mathbb{K}}(U, W),$$

is the composite morphism and on the RHS of the equation we are considering multiplication of matrices.

c) *for the identity morphism* $\mathrm{id}_V \in \mathrm{Hom}_{\mathbb{K}}(V, V)$ *we have*

$$[\mathrm{id}_V]^{\mathcal{C}}_{\mathcal{B}} = I_n,$$

where I_n *is the* $n \times n$ *identity matrix.*

d) *if* $V = W$ *then*

$$[\mathrm{id}_V]^{\mathcal{C}}_{\mathcal{B}} = P_{\mathcal{C} \leftarrow \mathcal{B}}.$$

e) *If* $A \in Mat_{m,n}(\mathbb{K})$ *and*

$$T_A : \mathbb{K}^n \to \mathbb{K}^m \ ; \ \underline{x} \mapsto A\underline{x},$$

so that $T_A \in \mathrm{Hom}_{\mathbb{K}}(\mathbb{K}^n, \mathbb{K}^m)$*. Then, if* $\mathcal{S}^{(i)} = (e_1, \dots, e_i)$ *is the standard basis of* $\mathbb{K}^i$*, then*

$$[T_A]^{\mathcal{S}^{(m)}}_{\mathcal{S}^{(n)}} = A.$$

We will now show how we can translate properties of morphisms into properties of matrices:

Theorem 1.7.4. *Let* $f \in \mathrm{Hom}_{\mathbb{K}}(V, W)$ *be a linear morphism of* $\mathbb{K}$*-vector spaces* V, W *and let* $\mathcal{B} \subset V, \mathcal{C} \subset W$ *be ordered bases of* V, W*. Then,*

a) f *is injective if and only if* $[f]^{\mathcal{C}}_{\mathcal{B}}$ *has a pivot in every column,*

b) f *is surjective if and only if* $[f]^{\mathcal{C}}_{\mathcal{B}}$ *has a pivot in every row,*

c) f *is an isomorphism if and only if* $[f]^{\mathcal{C}}_{\mathcal{B}}$ *is a square matrix and has a pivot in every row/column,*

d) *Suppose* $\dim V = \dim W$*. Then,* f *is injective if and only if* f *is surjective. In particular,*

$$\text{‘}f \text{ injective} \implies f \text{ surjective} \implies f \text{ bijective} \implies f \text{ injective’}.$$

Remark 1.7.5. 1. Theorem 1.7.3 states various properties that imply that the association of a linear morphism to its matrix (with respect to some ordered bases) behaves well and obeys certain desirable properties.

- *a*) implies that any question concerning the linear algebra properties of the set of $\mathbb{K}$-linear morphisms can be translated into a question concerning matrices. In particular, since it can be easily seen that $\dim_{\mathbb{K}} Mat_{m,n}(\mathbb{K}) = mn$ and isomorphic $\mathbb{K}$-vector spaces have the same dimension, we must therefore have that

$$\dim_{\mathbb{K}} \mathrm{Hom}_{\mathbb{K}}(V, W) = mn,$$

so that $\mathrm{Hom}_{\mathbb{K}}(V, W)$ is finite dimensional.

- *b*) can be summarised by the slogan

$$\boxed{\text{‘}\textbf{matrix multiplication is composition of morphisms}\text{’}.}$$

Together with *e*) this implies that, for an $m \times n$ matrix A and an $n \times p$ matrix B, we have $T_A \circ T_B = T_{AB}$.

2. Theorem 1.7.4 provides a way to show that a linear morphism satisfies certain properties, assuming we have found bases of the domain and codomain.

Conversely, Theorem 1.7.4 is also useful in determining properties of matrices by translating to a property of morphisms. For example, suppose that A, B are $n \times n$ matrices such that $AB = I_n$. By definition, a square matrix P is invertible if and only if there is a square matrix Q such that $PQ = QP = I_n$. Thus, even though we know that $AB = I_n$, in order to show that A (or B) is invertible, we would need to show also that $BA = I_n$. This is difficult to show directly (ie, only using matrices) if you only know that $AB = I_n$. However, if we consider the linear maps T_A and T_B then

$$AB = I_n \implies T_A \circ T_B = T_{AB} = id_{\mathbb{K}^n}. \quad \text{(Theorem 1.7.3, b), c), e))}$$

Now, by Lemma 0.2.4, since $\mathrm{id}_{\mathbb{K}^n}$ is injective then T_B is injective. Thus, T_B is an isomorphism by Theorem 1.7.4 so there exists a morphism $g \in \mathrm{Hom}_{\mathbb{K}}(\mathbb{K}^n, \mathbb{K}^n)$ such that $g \circ T_B = T_B \circ g = \mathrm{id}_{\mathbb{K}^n}$. Since $T_A \circ T_B = \mathrm{id}_{\mathbb{K}^n}$, then

$$g = \mathrm{id}_{\mathbb{K}^n} \circ g = (T_A \circ T_B) \circ g = T_A \circ (T_B \circ g) = T_A \circ \mathrm{id}_{\mathbb{K}^n} = T_A,$$

because $f \circ \mathrm{id}_{\mathbb{K}^n} = f$, for any function f with domain $\mathbb{K}^n$, and $\mathrm{id}_{\mathbb{K}^n} \circ f = f$, for any function f with codomain $\mathbb{K}^n$. Hence, we have shown that $g = T_A$ so that $\mathrm{id}_{\mathbb{K}^n} = T_B \circ g = T_B \circ T_A = T_{BA}$. Therefore, $I_n = BA$. Note that we have repeatedly used (various parts of) Theorem 1.7.3 in this last collection of justifications.

Suppose that we have distinct ordered bases $\mathcal{B}_1, \mathcal{B}_2 \subset V$ and $\mathcal{C}_1, \mathcal{C}_2 \subset W$ of the $\mathbb{K}$-vector spaces V, W and $f \in \mathrm{Hom}_{\mathbb{K}}(V, W)$. How are the matrices $[f]_{\mathcal{B}_1}^{\mathcal{C}_1}$ and $[f]_{\mathcal{B}_2}^{\mathcal{C}_2}$ related?

Proposition 1.7.6. *The matrices $[f]_{\mathcal{B}_1}^{\mathcal{C}_1}$ and $[f]_{\mathcal{B}_2}^{\mathcal{C}_2}$ satsify the following relation*

$$[f]_{\mathcal{B}_2}^{\mathcal{C}_2} = P_{\mathcal{C}_2 \leftarrow \mathcal{C}_1} [f]_{\mathcal{B}_1}^{\mathcal{C}_1} P_{\mathcal{B}_1 \leftarrow \mathcal{B}_2},$$

where we are considering multiplication of matrices on the RHS of this equation. Moreover, if there exists a matrix B such that

$$B = P_{\mathcal{C}_2 \leftarrow \mathcal{C}_1} [f]_{\mathcal{B}_1}^{\mathcal{C}_1} P_{\mathcal{B}_1 \leftarrow \mathcal{B}_2},$$

then $B = [f]_{\mathcal{B}_2}^{\mathcal{C}_2}$.

Proof: The proof is trivial once we have Theorem 1.7.3. If we consider that $P_{\mathcal{B}_1 \leftarrow \mathcal{B}_2} = [\mathrm{id}_V]_{\mathcal{B}_2}^{\mathcal{B}_1}$ and $P_{\mathcal{C}_2 \leftarrow \mathcal{C}_1} = [\mathrm{id}_W]_{\mathcal{C}_1}^{\mathcal{C}_2}$ then the RHS of the desired relation is

$$P_{\mathcal{C}_2 \leftarrow \mathcal{C}_1} [f]_{\mathcal{B}_1}^{\mathcal{C}_1} P_{\mathcal{B}_1 \leftarrow \mathcal{B}_2} = [id_W]_{\mathcal{C}_1}^{\mathcal{C}_2} [f]_{\mathcal{B}_1}^{\mathcal{C}_1} [\mathrm{id}_V]_{\mathcal{B}_2}^{\mathcal{B}_1} = [\mathrm{id}_V \circ f \circ \mathrm{id}_V]_{\mathcal{B}_2}^{\mathcal{C}_2} = [f]_{\mathcal{B}_2}^{\mathcal{C}_2}.$$

$\square$

Corollary 1.7.7. *Let $f \in \mathrm{End}_{\mathbb{K}}(V)$ be an endomorphism of a $\mathbb{K}$-vector space V (recall Definition 1.4.1), $\mathcal{B}, \mathcal{C} \subset V$ ordered bases of V. Then, if we denote $P = P_{\mathcal{B} \leftarrow \mathcal{C}}$, we have*

$$(*) \qquad [f]_{\mathcal{C}} = P^{-1} [f]_{\mathcal{B}} P.$$

Example 1.7.8. 1. Consider the linear morphism

$$f : \mathbb{Q}^3 \to \mathbb{Q}^2 \ ; \ \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} \mapsto \begin{bmatrix} x_1 + 2x_2 \\ -\frac{1}{2}x_2 + 3x_3 \end{bmatrix}.$$

Then, f is linear since we can write

$$f(\underline{x}) = \begin{bmatrix} 1 & 2 & 0 \\ 0 & -\frac{1}{2} & 3 \end{bmatrix} \underline{x}, \quad \text{for every } \underline{x} \in \mathbb{Q}^3.$$

Here, we have

$$A_f = [f]_{\mathcal{S}^{(3)}}^{\mathcal{S}^{(2)}} = \begin{bmatrix} 1 & 2 & 0 \\ 0 & -\frac{1}{2} & 3 \end{bmatrix}.$$

Consider the ordered bases

$$\mathcal{B} = \left(\begin{bmatrix} 0 \\ 2 \\ 1 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ -1 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} \right) \subset \mathbb{Q}^3, \quad \mathcal{C} = \left(\begin{bmatrix} 1 \\ -1 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \end{bmatrix} \right) \subset \mathbb{Q}^2.$$

Then, we can use Proposition 1.7.6 to determine $[f]_{\mathcal{B}}^{\mathcal{C}}$.

We have seen in Example 1.6.3 that

$$P_{\mathcal{S}^{(3)} \leftarrow \mathcal{B}} = \begin{bmatrix} 0 & 0 & 1 \\ 2 & 1 & 1 \\ 1 & -1 & 1 \end{bmatrix}, \quad P_{\mathcal{S}^{(2)} \leftarrow \mathcal{C}} = \begin{bmatrix} 1 & 1 \\ -1 & 1 \end{bmatrix},$$

so that

$$P_{\mathcal{C} \leftarrow \mathcal{S}^{(2)}} = P_{\mathcal{S}^{(2)} \leftarrow \mathcal{C}}^{-1} = \begin{bmatrix} \frac{1}{2} & -\frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} \end{bmatrix}.$$

Hence,

$$[f]_{\mathcal{B}}^{\mathcal{C}} = P_{\mathcal{C} \leftarrow \mathcal{S}^{(2)}} [f]_{\mathcal{S}^{(3)}}^{\mathcal{S}^{(2)}} P_{\mathcal{S}^{(3)} \leftarrow \mathcal{B}} = \begin{bmatrix} \frac{1}{2} & -\frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} \end{bmatrix} \begin{bmatrix} 1 & 2 & 0 \\ 0 & -\frac{1}{2} & 3 \end{bmatrix} \begin{bmatrix} 0 & 0 & 1 \\ 2 & 1 & 1 \\ 1 & -1 & 1 \end{bmatrix} = \begin{bmatrix} 1 & \frac{11}{4} & \frac{1}{4} \\ 3 & -\frac{3}{4} & 4 \end{bmatrix}.$$

2. Consider the linear morphism

$$g : Mat_2(\mathbb{R}) \to Mat_2(\mathbb{R}) \ ; \ A \mapsto A - A^t,$$

where A^t is the transpose of A. It is an exercise to check that g is linear.

We have the standard ordered basis of $Mat_2(\mathbb{R})$, $\mathcal{S} = (e_{11}, e_{12}, e_{21}, e_{22})$, where e_{ij} is the 2×2 matrix with 0 everywhere except a 1 in the ij-entry. Also, we have the ordered bases[31]

$$\mathcal{B} = (e_{12}, e_{21}, e_{11} - e_{22}, e_{11} + e_{22}), \ \mathcal{C} = (e_{11}, e_{22}, e_{12} + e_{21}, e_{12} - e_{21}) \subset Mat_2(\mathbb{R}).$$

Now, we see that

$$g(e_{11}) = 0, \ g(e_{12}) = e_{12} - e_{21}, \ g(e_{21}) = -e_{12} + e_{21}, \ g(e_{22}) = 0,$$

so that

$$[g]_{\mathcal{S}} = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 1 & -1 & 0 \\ 0 & -1 & 1 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}.$$

We use Proposition 1.7.6 to determine $[g]_{\mathcal{B}}^{\mathcal{C}}$. We have

$$P_{\mathcal{S} \leftarrow \mathcal{B}} = \begin{bmatrix} 0 & 0 & 1 & 1 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & -1 & 1 \end{bmatrix}, \ P_{\mathcal{S} \leftarrow \mathcal{C}} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 1 & -1 \\ 0 & 1 & 0 & 0 \end{bmatrix}.$$

[31] Check that these are bases of $Mat_2(\mathbb{R})$.

Then,

$$P_{\mathcal{C}\leftarrow\mathcal{S}} = P_{\mathcal{S}\leftarrow\mathcal{C}}^{-1} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & \frac{1}{2} & \frac{1}{2} & 0 \\ 0 & \frac{1}{2} & -\frac{1}{2} & 0 \end{bmatrix},$$

and we have

$$\begin{aligned} [g]_{\mathcal{B}}^{\mathcal{C}} = P_{\mathcal{C}\leftarrow\mathcal{S}}[g]_{\mathcal{S}}P_{\mathcal{S}\leftarrow\mathcal{B}} &= \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & \frac{1}{2} & \frac{1}{2} & 0 \\ 0 & \frac{1}{2} & -\frac{1}{2} & 0 \end{bmatrix} \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 1 & -1 & 0 \\ 0 & -1 & 1 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} 0 & 0 & 1 & 1 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & -1 & 1 \end{bmatrix} \\ &= \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 1 & -1 & 0 & 0 \end{bmatrix} \end{aligned}$$

Indeed, we have that $g(e_{12}) = e_{12} - e_{21}$, which gives

$$[g(e_{12})]_{\mathcal{C}} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 1 \end{bmatrix}.$$

Using the definition of $[g]_{\mathcal{B}}^{\mathcal{C}}$ this should be the first column, so that the matrix we have obtained above corroborates this.

Remark 1.7.9. The relationship established in Proposition 1.7.6 can be indicated in the following 'rooftop' or 'prism' diagram (think of the arrow $V \to W$ as the top of the rooftop)

$$\begin{tikzcd}
 & V \arrow[rrr, "f"] \arrow[dl, "{[-]_{\mathcal{B}_1}}"'] \arrow[dd, "{[-]_{\mathcal{B}_2}}"] & & & W \arrow[dl, "{[-]_{\mathcal{C}_1}}"'] \arrow[dd, "{[-]_{\mathcal{C}_2}}"] \\
\mathbb{K}^n \arrow[rrr, crossing over, "T_{[f]_{\mathcal{B}_1}^{\mathcal{C}_1}}"'] \arrow[dr, "T_{P_{\mathcal{B}_2\leftarrow\mathcal{B}_1}}"'] & & & \mathbb{K}^m \arrow[dr, "T_{P_{\mathcal{C}_2\leftarrow\mathcal{C}_1}}"'] & \\
 & \mathbb{K}^n \arrow[rrr, "T_{[f]_{\mathcal{B}_2}^{\mathcal{C}_2}}"'] & & & \mathbb{K}^m
\end{tikzcd}$$

Here we are assuming that all squares that appear are the commutative squares appearing after Lemma 1.7.2, and that the triangles that appear at the end of the prism are the commutative triangles that appeared in Remark 1.6.2. So, Proposition 1.7.6 corresponds to the 'bottom square' being a commutative diagram.

This diagram can be confusing at first but the more you try and understand it the better you will understand the relationship between linear morphisms, matrices and change of coordinates.

Note that in the rooftop diagram all arrows which have some vertical component are isomorphisms; this means that we can go forward and backwards along these arrows.

For example, suppose we start at V and go along the sequence of arrows $(\downarrow, \rightarrow)$. Then, the commutativity of the bottom square and the fact that that the arrows $\searrow$ are isomorphisms means we have

$$\rightarrow \; = \; (\nwarrow, \rightarrow, \searrow),$$

where $\nwarrow$ denotes the inverse morphism to $\searrow$.

Then, because we write composition of functions in the reverse order ($g \circ f$ means 'do f first, then g') we have
$$T_{[f]^{\mathcal{C}_2}_{\mathcal{B}_2}} \circ [-]_{\mathcal{B}_2} = T_{P_{\mathcal{C}_2 \leftarrow \mathcal{C}_1}} \circ T_{[f]^{\mathcal{C}_1}_{\mathcal{B}_1}} \circ T_{P_{\mathcal{B}_1 \leftarrow \mathcal{B}_2}} \circ [-]_{\mathcal{B}_2};$$
that is, for every $v \in V$, we have
$$[f]^{\mathcal{C}_2}_{\mathcal{B}_2}[v]_{\mathcal{B}_2} = P_{\mathcal{C}_2 \leftarrow \mathcal{C}_1}[f]^{\mathcal{C}_1}_{\mathcal{B}_1} P_{\mathcal{B}_1 \leftarrow \mathcal{B}_2}[v]_{\mathcal{B}_2},$$
and this is Proposition 1.7.6.

Definition 1.7.10 (Similar matrices)**.** Let $A, B \in Mat_n(\mathbb{K})$. We say that A *is similar to* B if and only if there exists an invertible matrix Q such that
$$A = Q^{-1}BQ.$$
This definition is symmetric with respect to A and B: namely, A is similar to B if and only if B is similar to A, since
$$A = Q^{-1}BQ \implies QAQ^{-1} = B,$$
so that if we let $P = Q^{-1}$ then we have a relation
$$B = P^{-1}AP.$$
Here we have used the (assumed known) fact that $(P^{-1})^{-1} = P$, for any invertible matrix P.

Moreover, if A is similar to B and B is similar to C, so that
$$A = Q^{-1}BQ, \quad \text{and} \quad B = P^{-1}CP,$$
then
$$A = Q^{-1}BQ = Q^{-1}P^{-1}CPQ = (PQ)^{-1}C(PQ),$$
so that A is similar to C.[32]

Corollary 1.7.7 states that matrices of linear endomorphisms with respect to different bases are similar. There is a converse to this result.

Proposition 1.7.11. *Let $A, B \in Mat_n(\mathbb{K})$ be similar matrices, so that $A = P^{-1}BP$, where $P \in \mathrm{GL}_n(\mathbb{K})$ is an invertible $n \times n$ matrix. Then, there exists a linear endomorphism $f \in \mathrm{End}_{\mathbb{K}}(\mathbb{K}^n)$ and ordered bases $\mathcal{B}, \mathcal{C} \subset \mathbb{K}^n$ such that*
$$[f]_{\mathcal{B}} = A, \quad \textit{and} \quad [f]_{\mathcal{C}} = B.$$

Proof: We take $\mathcal{C} = \mathcal{S}^{(n)} = (e_1, \dots, e_n)$, $\mathcal{B} = (b_1, \dots, b_n)$, where b_i is the i^{th} column of P, and $f = T_B \in \mathrm{End}_{\mathbb{K}}(\mathbb{K}^n)$. The details are left to the reader. □

Hence, Proposition 1.7.11 tells us that we can think of similar matrices A and B as being the matrices of the same linear morphism with respect to different ordered bases. As such, we expect that similar matrices should have certain equivalent properties; namely, those properties that can arise by considering the linear morphism T_A (or, equivalently, T_B), for example, *rank, diagonalisability, invertibility*.

1.7.1 Rank, classification of linear morphisms

Let $f \in \mathrm{Hom}_{\mathbb{K}}(V, W)$ be a linear morphism and recall the definition of the kernel of f and the image of f (Definition 1.4.4).

[32] These facts, along with the trivial statement that A is similar to A, imply that the notion of similarity defines an *equivalence relation* on $Mat_n(\mathbb{K})$.

Definition 1.7.12. We define the *rank of f*, denoted rank f, to be the number

$$\text{rank } f = \dim \text{im} f.$$

We define the *nullity of f*, denoted nul f, to be the number

$$\text{nul } f = \dim \ker f.$$

If A is an $m \times n$ matrix then we define the *rank of A*, denoted rank A, to be the rank of the linear morphism T_A determined by A. Similary, we define the *nullity of A*, denoted nul A, to be the nullity of the linear morphism T_A.

There exists a basic relationship between rank and nullity.

Theorem 1.7.13 (Rank Theorem)**.** *Let $f \in \text{Hom}_{\mathbb{K}}(V, W)$. Then,*

$$\dim V = \text{nul } f + \text{rank } f.$$

Proof: By Corollary 1.5.19 we know that there is a subspace $U \subset V$ such that $V = \ker f \oplus U$. Let $\mathcal{B} = (b_1, ..., b_r)$ be an ordered basis for U. Then, we claim that $\mathcal{C} = (f(b_1), ..., f(b_r))$ is an ordered basis of im$f$.

First, it is easy to see that the set $\{f(b_1), ..., f(b_r)\} \subset W$ is a subset of im$f$. If $v \in \text{im} f$, then $v = z + u$, where $z \in \ker f$, $u \in U$ (since $V = \ker f \oplus U$). Moreover, if $u = \lambda_1 b_1 + ... + \lambda_r b_r$ then

$$f(v) = f(z+u) = f(z) + f(u) = 0_W + f(\lambda_1 b_1 + ... + \lambda_r b_r) = \lambda_1 f(b_1) + ... + \lambda_r f(b_r) \in \text{span}_{\mathbb{K}} \mathcal{C}.$$

Hence, since $\text{im} f = \{f(v) \in W \mid v \in V\}$ then we must have $\text{span}_{\mathbb{K}} \mathcal{C} = \text{im} f$.

It remains to show that $\{f(b_1), ..., f(b_r)\}$ is linearly independent: indeed, suppose we have a linear relation

$$\lambda_1 f(b_1) + ... + \lambda_r f(b_r) = 0_W.$$

Then, since f is linear, this implies that $\lambda_1 b_1 + ... + \lambda_r b_r \in \ker f$ and $\lambda_1 b_1 + ... + \lambda_r b_r \in U$ (because $\mathcal{B}$ is a basis of U). Hence,

$$\lambda_1 b_1 + ... + \lambda_r b_r \in \ker f \cap U = \{0_V\},$$

so that

$$\lambda_1 b_1 + ... + \lambda_r b_r = 0_V.$$

Now, as $\mathcal{B}$ is linearly independent then

$$\lambda_1 = \lambda_2 = \cdots = \lambda_r = 0.$$

Hence, $\mathcal{C}$ is linearly independent and therefore a basis of im$f$.

Now, using Corollary 1.5.19, we see that

$$\dim V = \text{nul } f + r = \text{nul } f + \text{rank } f,$$

by the previous discussion. $\square$

Lemma 1.7.14. *Let A be an $m \times n$ matrix. Then, the rank of A is equal to the maximal number of linearly independent columns of A.*

Proof: Let us write

$$A = [a_1 \; a_2 \; \cdots \; a_n],$$

so that the i^{th} column of A is the vector $a_i \in \mathbb{K}^m$.

Consider the linear morphism $T_A \in \text{Hom}_{\mathbb{K}}(\mathbb{K}^n, \mathbb{K}^m)$. Then, we have defined rank A = rank T_A = $\dim \text{im} T_A$. Then, since

$$T_A\left(\begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}\right) = x_1 a_1 + ... + x_n a_n,$$

we see that

$$(*) \qquad \text{span}_{\mathbb{K}}\{a_1, \dots, a_n\} = \text{im}\, T_A.$$

Suppose that $A \neq 0_{m,n}$. Thus, one of the columns of A is nonzero. Suppose that $a_i \neq 0_{\mathbb{K}^m}$. Then, $\{a_i\}$ is a linearly independent set and can be extended to a basis of $\text{im}\, T_A$ using vectors from $\{a_1, \dots, a_n\}$, by $(*)$. Hence, $\text{rank}\, A = \dim \text{im}\, T_A$ is equal to the number of columns of A that form a basis of $\text{im}\, T_A$. Moreover, by Proposition 1.5.15, every linearly independent set in $\text{im}\, T_A$ has size no greater than $\text{rank}\, A$. In particular, every linearly independent subset of the columns of A has size no greater than $\text{rank}\, A$ while there does exist some subset having size exactly $\text{rank}\, A$.

If $A = 0_{m,n}$ then $T_A \in \text{Hom}_{\mathbb{K}}(V, W)$ is the zero morphism and $\text{rank}\, T_A = \dim\{0_W\} = 0$. The result follows. $\square$

The proof that we have just given for the Rank Theorem implies the following result.

Theorem 1.7.15. *Let $f \in \text{Hom}_{\mathbb{K}}(V, W)$ be a $\mathbb{K}$-linear morphism and denote $r = \text{rank}\, f$. Then, there exists ordered bases $\mathcal{B} \subset V, \mathcal{C} \subset W$ such that*

$$[f]_{\mathcal{B}}^{\mathcal{C}} = \begin{bmatrix} I_r & 0_{r,n-r} \\ 0_{m-r,r} & 0_{m-r,n-r} \end{bmatrix},$$

where $n = \dim V$, $m = \dim W$ and $0_{i,j} \in Mat_{i,j}(\mathbb{K})$ is the zero matrix.

Proof: Consider an ordered basis $\mathcal{B}_1 = (b_1, \dots, b_{n-r})$ of $\ker f$ and extend to an ordered basis

$$\mathcal{B} = (b_1, \dots, b_{n-r}, b_{n-r+1}, \dots, b_n)$$

of V. Then, as in the proof of the Rank Theorem, we see that $(f(b_{n-r+1}), \dots, f(b_n))$ is an ordered basis of $\text{im} f$. Extend this to an ordered basis

$$\mathcal{C} = (f(b_{n-r+1}), \dots, f(b_n), c_1, \dots, c_{m-r}),$$

of W. Then,

$$[f]_{\mathcal{B}}^{\mathcal{C}} = \begin{bmatrix} I_r & 0_{r,n-r} \\ 0_{m-r,r} & 0_{m-r,n-r} \end{bmatrix}.$$

$\square$

Corollary 1.7.16. *Let $A \in Mat_{m,n}(\mathbb{K})$ such that $\text{rank}\, A = r$. Then, there exists $P \in \text{GL}_n(\mathbb{K}), Q \in \text{GL}_m(\mathbb{K})$ such that*

$$Q^{-1}AP = \begin{bmatrix} I_r & 0_{r,n-r} \\ 0_{m-r,r} & 0_{m-r,n-r} \end{bmatrix}.$$

Corollary 1.7.17. *Let $A, B \in Mat_{m,n}(\mathbb{K})$. Then, A, B are the matrices of the same linear map with respect to different bases if and only if they have the same rank.*

Proof: By the previous Corollary we can find $Q_1, Q_2 \in \text{GL}_m(\mathbb{K}), P_1, P_2 \in \text{GL}_n(\mathbb{K})$ such that

$$Q_1^{-1}AP_1 = \begin{bmatrix} I_r & 0_{r,n-r} \\ 0_{m-r,r} & 0_{m-r,n-r} \end{bmatrix} = Q_2^{-1}BP_2.$$

Then, we have

$$Q_2Q_1^{-1}AP_1P_2^{-1} = B.$$

Recall that $Q_2Q_1^{-1} = (Q_1Q_2^{-1})^{-1}$. Then, as $Q_1Q_2^{-1}$ and $P_1P_2^{-1}$ are invertible matrices (products of invertible matrices are invertible) their different sets of columns are linearly independent and therefore form ordered bases $\mathcal{C} \subset \mathbb{K}^m$ and $\mathcal{B} \subset \mathbb{K}^n$. Then, if we consider the linear map T_A, the above equation says that

$$P_{\mathcal{C} \leftarrow \mathcal{S}^{(m)}}[T_A]_{\mathcal{S}^{(n)}}^{\mathcal{S}^{(m)}} P_{\mathcal{S}^{(n)} \leftarrow \mathcal{B}} = B,$$

so that Proposition 1.7.6 implies that $[T_A]_{\mathcal{B}}^{\mathcal{C}} = B$. $\square$

Remark 1.7.18.
1. The rank of a matrix A is usually <u>defined</u> to be the maximum number of linearly independent columns of A. However, we have shown that our definition is equivalent to this definition.

2. Theorem 1.7.15 is just a restatement in terms of linear morphisms of a fact that you might have come across before: every $m \times n$ matrix can be row-reduced to reduced echelon form using row operations. Moreover, if we allow 'column operations', then any $m \times n$ matrix can be row/column-reduced to a matrix of the form appearing in Theorem 1.7.15.

   This requires the use of *elementary (row-operation) matrices* and we will investigate this result during discussion.

3. Corollary 1.7.17 allows us to provide a classification of $m \times n$ matrices based on their rank: namely, we can say that A and B are *equivalent* if there exists $Q \in \mathrm{GL}_m(\mathbb{K})$, $P \in \mathrm{GL}_n(\mathbb{K})$ such that

$$B = Q^{-1}AP.$$

   Then, this notion of equivalence defines an *equivalence relation* on $Mat_{m,n}(\mathbb{K})$. Hence, we can partition $Mat_{m,n}(\mathbb{K})$ into dictinct equivalence classes. Corollary 1.7.17 says that the equivalence classes can be labelled by the rank of the matrices that each class contains.

1.8 Dual Spaces (non-examinable)