4
Coordinates

Every finite-dimensional vector space is isomorphic to a coordinate space. Choosing such isomorphisms for the source and target of a linear map allows one to identify the map with a matrix. This chapter explores these matrix representations.

4.0 Change of Basis

HOW ARE COORDINATE VECTORS RELATIVE TO DIFFERENT BASES RELATED? Coordinate vectors provide a map from a finite-dimensional vector space into the coordinate space of the same dimension.

4.0.0 Proposition. *Let $\mathcal{B} := (\boldsymbol{v}_1, \boldsymbol{v}_2, \dots, \boldsymbol{v}_n)$ be an ordered basis for the $\mathbb{K}$-vector space V. The map $\boldsymbol{w} \mapsto (\boldsymbol{w})_{\mathcal{B}}$, sending a vector $\boldsymbol{w}$ in V to its coordinate vector $(w)_{\mathcal{B}}$ in $\mathbb{K}^n$, is an invertible linear map.*

In particular, the vector space V is isomorphic to $\mathbb{K}^n$.

Proof. Fix two vectors $\boldsymbol{w}$ and $\boldsymbol{w}'$ in V. Since the vectors $\boldsymbol{v}_1, \boldsymbol{v}_2, \dots, \boldsymbol{v}_n$ span V, there exists scalars $c_1, c_2, \dots, c_n$ and scalars $c'_1, c'_2, \dots, c'_n$ such that $\boldsymbol{w} = c_1\, \boldsymbol{v}_1 + c_2\, \boldsymbol{v}_2 + \cdots + c_n\, \boldsymbol{v}_n$ and $\boldsymbol{w}' = c'_1\, \boldsymbol{v}_1 + c'_2\, \boldsymbol{v}_2 + \cdots + c'_n\, \boldsymbol{v}_n$, so we have $(\boldsymbol{w})_{\mathcal{B}} = \begin{bmatrix} c_1 & c_2 & \cdots & c_n \end{bmatrix}^{\mathsf{T}}$ and $(\boldsymbol{w}')_{\mathcal{B}} = \begin{bmatrix} c'_1 & c'_2 & \cdots & c'_n \end{bmatrix}^{\mathsf{T}}$. It follows that, for all scalars d and d' in $\mathbb{K}$, we have

$$d\, \boldsymbol{w} + d'\, \boldsymbol{w}' = (dc_1 + d'c'_1)\, \boldsymbol{v}_1 + (dc_2 + d'c'_2)\, \boldsymbol{v}_2 + \cdots + (dc_n + d'c'_n)\, \boldsymbol{v}_n$$

and

$$(d\, \boldsymbol{w} + d'\, \boldsymbol{w}')_{\mathcal{B}} = \begin{bmatrix} dc_1 + d'c'_1 \\ dc_2 + d'c'_2 \\ \vdots \\ dc_n + d'c'_n \end{bmatrix} = d \begin{bmatrix} c_1 \\ c_2 \\ \vdots \\ c_n \end{bmatrix} + d' \begin{bmatrix} c'_1 \\ c'_2 \\ \vdots \\ c'_n \end{bmatrix} = d(\boldsymbol{w})_{\mathcal{B}} + d'(\boldsymbol{w}')_{\mathcal{B}}\,,$$

proving linearity.

For any scalars $b_1, b_2, \dots, b_n$ in $\mathbb{K}$, the coordinate vector of the vector $b_1\, \boldsymbol{v}_1 + b_2\, \boldsymbol{v}_2 + \cdots + b_n\, \boldsymbol{v}_n$ in V is $\begin{bmatrix} b_1 & b_2 & \cdots & b_n \end{bmatrix}^{\mathsf{T}}$, so the linear map $\boldsymbol{w} \mapsto (\boldsymbol{w})_{\mathcal{B}}$ is surjective. Since each vector $\boldsymbol{u}$ in V is a unique linear combination of a basis [2.3.0], the linear map $\boldsymbol{w} \mapsto (\boldsymbol{w})_{\mathcal{B}}$ is injective. Thus, the characterization of invertibility [3.2.5] shows that the map $\boldsymbol{w} \mapsto (w)_{\mathcal{B}}$ is invertible. □

4.0.1 Problem. Let $f_1(x) := 1 - \cos(x)$, $f_2(x) := 1 - 3\cos(x) + \sin(x)$, and $f_3(x) := 1 - \cos(x) + \sin(x)$. Show that the functions f_1, f_2, f_3 are a basis for the $\mathbb{R}$-vector space of trigonometric polynomials having degree at most 1.

Solution. The space of trigonometric polynomials [2.0.7] of degree at most 1 has $\mathcal{T} := \big(1, \cos(x), \sin(x)\big)$ is its canonical ordered basis. Since this vector space has dimension 3, it suffices to show that the functions f_1, f_2, f_3 are linearly independent. Because the map sending a trigonometric polynomial to its coordinate vector relative to $\mathcal{T}$ is linear, the functions f_1, f_2, f_3 are linearly independent if and only if the vectors $(f_1)_{\mathcal{T}}$, $(f_2)_{\mathcal{T}}$, $(f_3)_{\mathcal{T}}$ are linearly independent. Since $(f_1)_{\mathcal{T}} = \begin{bmatrix} 1 & -1 & 0 \end{bmatrix}^{\mathsf{T}}$, $(f_2)_{\mathcal{T}} = \begin{bmatrix} 1 & -3 & 1 \end{bmatrix}^{\mathsf{T}}$, $(f_1)_{\mathcal{T}} = \begin{bmatrix} 1 & -1 & 1 \end{bmatrix}^{\mathsf{T}}$, and

$$\begin{bmatrix} 1 & 1 & 1 \\ -1 & -3 & -1 \\ 0 & 1 & 1 \end{bmatrix} \xrightarrow[\sim]{\substack{\vec{r}_1 \mapsto \vec{r}_1 - \vec{r}_3 \\ \vec{r}_2 \mapsto \vec{r}_2 + 3\vec{r}_3}} \begin{bmatrix} 1 & 0 & 0 \\ -1 & 0 & 2 \\ 0 & 1 & 1 \end{bmatrix} \xrightarrow[\sim]{\vec{r}_2 \mapsto \vec{r}_2 + \vec{r}_1} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 2 \\ 0 & 1 & 1 \end{bmatrix} \xrightarrow[\sim]{\substack{\vec{r}_2 \mapsto -0.5\vec{r}_2 \\ \vec{r}_3 \mapsto \vec{r}_3 - 0.5\vec{r}_2}} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix} \xrightarrow[\sim]{\substack{\vec{r}_2 \mapsto \vec{r}_3 \\ \vec{r}_3 \mapsto \vec{r}_1}} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix},$$

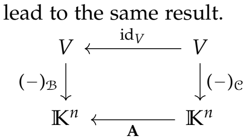
we deduce that $(f_1)_{\mathcal{T}}$, $(f_2)_{\mathcal{T}}$, $(f_3)_{\mathcal{T}}$ are linearly independent. □

4.0.2 Theorem (Change of basis)**.** *Fix ordered bases $\mathcal{B} := (\boldsymbol{v}_1, \boldsymbol{v}_2 \dots, \boldsymbol{v}_n)$ and $\mathcal{C} := (\boldsymbol{w}_1, \boldsymbol{w}_2, \dots, \boldsymbol{w}_n)$ for a $\mathbb{K}$-vector space V. The matrix*

$$\mathbf{A} := \Big[(\boldsymbol{w}_1)_{\mathcal{B}} \;\; (\boldsymbol{w}_2)_{\mathcal{B}} \;\; \cdots \;\; (\boldsymbol{w}_n)_{\mathcal{B}} \Big],$$

whose k-th column is the coordinate vector of $\boldsymbol{w}_k$ relative to $\mathcal{B}$, is invertible. Moreover, for any vector $\boldsymbol{u}$ in V, we have $(\boldsymbol{u})_{\mathcal{B}} = \mathbf{A}\,(\boldsymbol{u})_{\mathcal{C}}$.

This theorem says that all directed paths from V to $\mathbb{K}^n$ in the diagram below lead to the same result.



Proof. For all $1 \leqslant j \leqslant n$ and all $1 \leqslant k \leqslant n$, let the scalar $a_{j,k}$ be the (j,k)-entry in the matrix $\mathbf{A}$. The definition of the matrix $\mathbf{A}$ implies that $\boldsymbol{w}_k = a_{1,k}\,\boldsymbol{v}_1 + a_{2,k}\,\boldsymbol{v}_2 + \cdots + a_{n,k}\,\boldsymbol{v}_n$ for all $1 \leqslant k \leqslant n$. For any vector $\boldsymbol{u} := c_1\,\boldsymbol{w}_1 + c_2\,\boldsymbol{w}_2 + \cdots + c_n\,\boldsymbol{w}_n$ where $c_1, c_2, \dots, c_n$ are scalars in $\mathbb{K}$, we have

$$\boldsymbol{u} = \sum_{k=1}^{n} c_k\,\boldsymbol{w}_k = \sum_{k=1}^{n} c_k \left(\sum_{j=1}^{n} a_{j,k}\,\boldsymbol{v}_j \right) = \sum_{j=1}^{n} \left(\sum_{k=1}^{n} a_{j,k}\,c_k \right) \boldsymbol{v}_j$$

Since $\mathcal{B}$ is a basis for the vector space V, the vector $\boldsymbol{u}$ is a unique linear combination [2.3.0] of the vectors $\boldsymbol{v}_1, \boldsymbol{v}_2, \dots, \boldsymbol{v}_n$. It follows that

$$(\boldsymbol{u})_{\mathcal{B}} = \begin{bmatrix} a_{1,1}c_1 + a_{1,2}c_2 + \cdots + a_{1,n}c_n \\ a_{2,1}c_1 + a_{2,2}c_2 + \cdots + a_{2,n}c_n \\ \vdots \\ a_{n,1}c_1 + a_{n,2}c_2 + \cdots + a_{n,n}c_n \end{bmatrix} = \begin{bmatrix} a_{1,1} & a_{1,2} & \cdots & a_{1,n} \\ a_{2,1} & a_{2,2} & \cdots & a_{2,n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n,1} & a_{n,2} & \cdots & a_{n,n} \end{bmatrix} \begin{bmatrix} c_1 \\ c_2 \\ \vdots \\ c_n \end{bmatrix} = \mathbf{A}\,(\boldsymbol{u})_{\mathcal{C}}\,.$$

It remains to show that the matrix $\mathbf{A}$ is invertible. Consider a vector $\begin{bmatrix} c_1 & c_2 & \cdots & c_n \end{bmatrix}^{\mathsf{T}}$ in $\operatorname{Ker}(\mathbf{A})$. Setting $\boldsymbol{u} := c_1\,\boldsymbol{w}_1 + c_2\,\boldsymbol{w}_2 + \cdots + c_n\,\boldsymbol{w}_n$, it follows that $\mathbf{0} = \mathbf{A}(\boldsymbol{u})_{\mathcal{C}} = (\boldsymbol{u})_{\mathcal{B}}$. Hence, we deduce that

$$\boldsymbol{u} = 0\,\boldsymbol{v}_1 + 0\,\boldsymbol{v}_2 + \cdots + 0\,\boldsymbol{v}_n = \mathbf{0}\,,$$

which implies that $(\boldsymbol{u})_{\mathcal{C}} = \mathbf{0}$ and $c_1 = c_2 = \cdots = c_n = 0$. Since $\operatorname{Ker}(\mathbf{A}) = \{\mathbf{0}\}$, combining the characterizations of injectivity [3.1.4] and invertibility [3.2.5] shows that $\mathbf{A}$ is invertible. □

Knowing the change of basis matrix sometimes allows one to avoid solving a linear system via row reduction.

4.0.3 Problem. Let $\mathcal{M} := (1, t, t^2)$ and $\mathcal{B} := \big(1, t-1, (t-1)^2\big)$ be ordered bases for the $\mathbb{Q}$-vector space $\mathbb{Q}[t]_{\leqslant 2}$. Given the polynomial $f := a_0 + a_1\,(t-1) + a_2\,(t-1)^2$, find rational scalars b_0, b_1, b_2 such that $f = b_0 + b_1\,t + b_2\,t^2$.

Solution. Since we have $(1)_{\mathcal{M}} = \begin{bmatrix}1 & 0 & 0\end{bmatrix}^{\mathsf{T}}$, $(t-1)_{\mathcal{M}} = \begin{bmatrix}-1 & 1 & 0\end{bmatrix}^{\mathsf{T}}$, and $((t-1)^2)_{\mathcal{M}} = \begin{bmatrix}1 & -2 & 1\end{bmatrix}^{\mathsf{T}}$, change of basis [4.0.2] gives

$$\begin{bmatrix}b_0\\ b_1\\ b_2\end{bmatrix} = (f)_{\mathcal{M}} = \mathbf{A}\,(f)_{\mathcal{B}} = \begin{bmatrix}1 & -1 & 1\\ 0 & 1 & -2\\ 0 & 0 & 1\end{bmatrix}\begin{bmatrix}a_0\\ a_1\\ a_2\end{bmatrix} = \begin{bmatrix}a_0 - a_1 + a_2\\ a_1 - 2a_2\\ a_2\end{bmatrix}. \qquad \square$$

Exercises

4.0.4 Problem. Determine which of the following statements are true. If a statement is false, then provide a counterexample.

i. A change of basis matrix is always a square matrix.
ii. A change of basis matrix is always invertible.
iii. A change of basis matrix is always equal to its own inverse.
iv. When the two chosen bases on a vector space are equal, the change of basis matrix is the identity matrix.
v. The zero matrix can never be a change of basis matrix.

4.0.5 Problem. Consider the four polynomials $f_0(t) := 1$, $f_1(t) := t$, $f_2(t) := t(t-1)$ and $f_3(t) := t(t-1)(t-2)$.

i. Show that $\mathcal{B} := (f_0, f_1, f_2, f_3)$ is an ordered basis for $\mathbb{Q}[t]_{\leqslant 3}$.
ii. Suppose that we have the equation

$$a_0 + a_1\,t + a_2\,t^2 + a_3\,t^3 = b_0\,f_0(t) + b_1\,f_1(t) + b_2\,f_2(t) + b_3\,f_3(t)$$

where $a_0, a_1, \dots, a_3, b_0, b_1, \dots, b_3 \in \mathbb{Q}$. If $\vec{\boldsymbol{a}} = \begin{bmatrix}a_0 & a_1 & a_2 & a_3\end{bmatrix}^{\mathsf{T}}$ and $\vec{\boldsymbol{b}} = \begin{bmatrix}b_0 & b_1 & b_2 & b_3\end{bmatrix}^{\mathsf{T}}$, then find matrices $\mathbf{M}$ and $\mathbf{N}$ such that $\mathbf{M}\,\vec{\boldsymbol{a}} = \vec{\boldsymbol{b}}$ and $\mathbf{N}\,\vec{\boldsymbol{b}} = \vec{\boldsymbol{a}}$.
iii. Find the coordinates of t^2 and t^3 with respect to $\mathcal{B}$.

4.1 Matrix of a Linear Map

How are linear maps and matrices related? Let $T\colon V \to W$ be a linear map. Choose $\mathcal{B} := (\boldsymbol{v}_1, \boldsymbol{v}_2, \dots, \boldsymbol{v}_n)$ and $\mathcal{C} := (\boldsymbol{w}_1, \boldsymbol{w}_2, \dots, \boldsymbol{w}_m)$ to be ordered bases for the $\mathbb{K}$-vector spaces V and W respectively. For all $1 \leqslant k \leqslant n$, the vector $T[\boldsymbol{v}_k]$ lying in W is a unique linear combination [2.3.0] of the basis vectors $\boldsymbol{w}_1, \boldsymbol{w}_2, \dots, \boldsymbol{w}_m$. Hence, there exists scalars $a_{1,k}, a_{2,k}, \dots, a_{m,k}$ in $\mathbb{K}$ such that

$$T[\boldsymbol{v}_k] = a_{1,k}\,\boldsymbol{w}_1 + a_{2,k}\,\boldsymbol{w}_2 + \cdots + a_{m,k}\,\boldsymbol{w}_m\,.$$

Since a linear map is determined by its values on a basis [3.0.7], the collection of scalars $a_{j,k}$ determines the map T. More formally, we make the following definition.

4.1.0 Definition. The ***matrix*** of a linear map $T\colon V \to W$ relative to the ordered bases $\mathcal{B}$ and $\mathcal{C}$ for the $\mathbb{K}$-vector spaces V and W is

$$(T)_{\mathcal{C}}^{\mathcal{B}} := \begin{bmatrix} a_{1,1} & a_{1,2} & \cdots & a_{1,n} \\ a_{2,1} & a_{2,2} & \cdots & a_{2,n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m,1} & a_{m,2} & \cdots & a_{m,n} \end{bmatrix} \in \mathbb{K}^{m \times n} .$$

When $\boldsymbol{v}_k$ denotes the k-th vector in the ordered basis $\mathcal{B}$, the k-th column of the matrix $(T)_{\mathcal{C}}^{\mathcal{B}}$ is the coordinate vector of image $T[\boldsymbol{v}_k]$ relative to ordered basis $\mathcal{C}$.

Our notation highlights the relation between the coordinate vectors relative to an ordered basis. Given a linear map $T\colon V \to W$, an ordered basis $\mathcal{B}$ for V, and an order basis $\mathcal{C}$ for W, we have $(T)_{\mathcal{C}}^{\mathcal{B}}\,(\boldsymbol{v})_{\mathcal{B}} = (T[\boldsymbol{v}])_{\mathcal{C}}$ for all $\boldsymbol{v} \in V$. This is equivalent to saying that all directed paths from V to $\mathbb{K}^m$ in the diagram below lead to the same result.

$$\begin{array}{ccc} V & \xleftarrow{\;T\;} & W \\ {\scriptstyle (-)_{\mathcal{C}}}\downarrow & & \downarrow{\scriptstyle (-)_{\mathcal{B}}} \\ \mathbb{K}^n & \xleftarrow[(T)_{\mathcal{C}}^{\mathcal{B}}]{} & \mathbb{K}^m \end{array}$$

4.1.1 Remark. The change of basis matrix [4.0.2] is matrix of the identity map relative to two ordered bases; $\mathbf{A} = (\mathrm{id}_V)_{\mathcal{B}}^{\mathcal{C}}$.

4.1.2 Problem. Let $\mathcal{M} := (1, t, t^2, \dots, t^n)$ be the monomial basis for the $\mathbb{K}$-vector space $\mathbb{K}[t]_{\leqslant n}$. For the linear operator $T\colon \mathbb{K}[t]_{\leqslant n} \to \mathbb{K}[t]_{\leqslant n}$ is defined by $T[t^k] := (t+1)^k$, compute the matrix $(T)_{\mathcal{M}}^{\mathcal{M}}$.

Solution. The Binomial Theorem [2.3.4] gives

$$T[t^k] = (t+1)^k = \sum_{j=0}^{k} \binom{k}{j}\, t^j ,$$

so we obtain

$$(T)_{\mathcal{M}}^{\mathcal{M}} = \begin{bmatrix} \binom{0}{0} & \binom{1}{0} & \binom{2}{0} & \binom{3}{0} & \binom{4}{0} & \cdots & \binom{n}{1} \\ \binom{0}{1} & \binom{1}{1} & \binom{2}{1} & \binom{3}{1} & \binom{4}{1} & \cdots & \binom{n}{1} \\ \binom{0}{2} & \binom{1}{2} & \binom{2}{2} & \binom{3}{2} & \binom{4}{2} & \cdots & \binom{n}{2} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ \binom{0}{n} & \binom{1}{n} & \binom{2}{n} & \binom{3}{n} & \binom{4}{n} & \cdots & \binom{n}{n} \end{bmatrix} = \begin{bmatrix} 1 & 1 & 1 & 1 & 1 & \cdots & 1 \\ 0 & 1 & 2 & 3 & 4 & \cdots & n \\ 0 & 0 & 1 & 3 & 6 & \cdots & \frac{n(n-1)}{2} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & 0 & 0 & \cdots & 1 \end{bmatrix} . \square$$

4.1.3 Problem. Consider linear operator $T\colon \mathbb{K}[t]_{\leqslant 2} \to \mathbb{K}[t]_{\leqslant 2}$ defined, for all polynomials f in $\mathbb{K}[t]_{\leqslant 2}$, by $T[f] := f'' + 2f' + f$. Find the matrix of T with respect to the monomial basis $\mathcal{M} := (1, t, t^2)$. Using this matrix, solve the equation $T[f] = 1 + t + t^2$, and compute $\operatorname{Ker}(T)$.

Solution. Since $T[1] = 1$, $T[t] = 2 + t$, and $T[t^2] = 2 + 4t + t^2$, we have

$$(T)_{\mathcal{M}}^{\mathcal{M}} = \begin{bmatrix} 1 & 2 & 2 \\ 0 & 1 & 4 \\ 0 & 0 & 1 \end{bmatrix} .$$

To solve the equation $T[f] = 1 + t + t^2$, we consider the matrix equation $(T)_{\mathcal{M}}^{\mathcal{M}}\,(f)_{\mathcal{M}} = \begin{bmatrix} 1 & 1 & 1 \end{bmatrix}^{\mathsf{T}}$. Elementary row operations give

$$\begin{bmatrix} 1 & 2 & 2 & 1 \\ 0 & 1 & 4 & 1 \\ 0 & 0 & 1 & 1 \end{bmatrix} \xrightarrow[\sim]{\substack{r_1 \mapsto r_1 - 2r_3 \\ r_2 \mapsto r_2 - 4r_3}} \begin{bmatrix} 1 & 2 & 0 & -1 \\ 0 & 1 & 0 & -3 \\ 0 & 0 & 1 & 1 \end{bmatrix} \xrightarrow[\sim]{r_1 \mapsto r_1 - 2r_2} \begin{bmatrix} 1 & 0 & 0 & 5 \\ 0 & 1 & 0 & -3 \\ 0 & 0 & 1 & 1 \end{bmatrix} ,$$

so we deduce that $T[5 - 3t + t^2] = 1 + t + t^2$. Since the reduced row echelon form of $(T)_{\mathcal{M}}^{\mathcal{M}}$ is the identity matrix, the matrix $(T)_{\mathcal{M}}^{\mathcal{M}}$ is invertible. Thus, the operator T is invertible and $\operatorname{Ker}(T) = \{0\}$. □

4.1.4 Proposition. *Let $\mathcal{B} := (\boldsymbol{v}_1, \boldsymbol{v}_2, \dots, \boldsymbol{v}_n)$ and $\mathcal{C} := (\boldsymbol{w}_1, \boldsymbol{w}_2 \dots, \boldsymbol{w}_m)$ be ordered bases for the $\mathbb{K}$-vector spaces V and W respectively. The map from $\operatorname{Hom}(V, W)$ to $\mathbb{K}^{m \times n}$ defined by $T \mapsto (T)_{\mathcal{C}}^{\mathcal{B}}$ is an invertible linear map, so the $\mathbb{K}$-vector space $\operatorname{Hom}(V, W)$ is isomorphic to $\mathbb{K}^{m \times n}$.*

Proof. Consider two linear maps $T \colon V \to W$ and $S \colon V \to W$. For any scalars c and d in $\mathbb{K}$, we have $(c\,T + d\,S)[\boldsymbol{v}_k] = c\,T[\boldsymbol{v}_k] + d\,T[\boldsymbol{v}_k]$ for all $1 \leqslant k \leqslant n$, because $\operatorname{Hom}(V, W)$ has pointwise operations. For all $1 \leqslant j \leqslant m$ and all $1 \leqslant k \leqslant n$, let the scalars $a_{j,k}$ and $b_{j,k}$, denote the (j, k)-entries in the matrices $(T)_{\mathcal{C}}^{\mathcal{B}}$ and $(S)_{\mathcal{C}}^{\mathcal{B}}$ respectively. We obtain

$$(c\,T + d\,S)_{\mathcal{C}}^{\mathcal{B}} = \left[c\,a_{j,k} + d\,b_{j,k}\right] = c\,\left[a_{j,k}\right] + d\,\left[b_{j,k}\right] = c\,(T)_{\mathcal{C}}^{\mathcal{B}} + d\,(S)_{\mathcal{C}}^{\mathcal{B}}\,,$$

because $\mathbb{K}^{m \times n}$ is equipped with entrywise operations. Therefore, the map $T \mapsto (T)_{\mathcal{C}}^{\mathcal{B}}$ is linear.

As a consequence of the characterization of invertibility [3.2.5], it suffices to prove that this map is bijective.

- Suppose that the map $T \colon V \to W$ belongs to the kernel. It follows that $(T)_{\mathcal{C}}^{\mathcal{B}} = \mathbf{0}$ and $T[\boldsymbol{v}_k] = \mathbf{0}$ for all $1 \leqslant k \leqslant n$. Since $(\boldsymbol{v}_1, \boldsymbol{v}_2, \dots, \boldsymbol{v}_n)$ is a basis of V and a linear map is determined by its values on a basis [3.0.7], we see that $T = 0$. Hence, the characterization of injectivity [3.1.4] shows that the map $T \mapsto (T)_{\mathcal{C}}^{\mathcal{B}}$ is injective.
- Given an $(m \times n)$-matrix $\mathbf{A}$ whose (j, k)-entry is the scalar $a_{j,k}$ for all $1 \leqslant j \leqslant m$ and all $1 \leqslant k \leqslant n$, consider the linear map $S \colon V \to W$ defined by $S[\boldsymbol{v}_k] = a_{1,k}\,\boldsymbol{w}_1 + a_{2,k}\,\boldsymbol{w}_2 + \cdots + a_{m,k}\,\boldsymbol{w}_m$; see [3.0.7]. Since we have $(S)_{\mathcal{C}}^{\mathcal{B}} = \mathbf{A}$, the map $T \mapsto (T)_{\mathcal{C}}^{\mathcal{B}}$ is surjective. □

4.1.5 Corollary. *For any two finite-dimensional $\mathbb{K}$-vectors spaces V and W, we have $\dim \operatorname{Hom}(V, W) = \dim(V)\,\dim(W)$.*

Proof. By choosing ordered bases $\mathcal{B}$ and $\mathcal{C}$ for V and W respectively, Proposition 4.1.4 gives the invertible linear map $T \mapsto (T)_{\mathcal{C}}^{\mathcal{B}}$. This invertible linear map sends any basis of $\operatorname{Hom}(V, W)$ to a basis for $\mathbb{K}^{m \times n}$. We conclude that $\dim \operatorname{Hom}(V, W) = \dim(V)\,\dim(W)$. □

Exercises

4.1.6 Problem. Determine which of the following statements are true. If a statement is false, then provide a counterexample. Assume that V and W are finite-dimensional vector spaces with order basis $\mathcal{B}$ and $\mathcal{C}$ respectively. Let $T \colon V \to W$ and $S \colon V \to W$ be linear maps.

i. When $m = \dim(V)$ and $n = \dim(W)$, the matrix $(T)_{\mathcal{B}}^{\mathcal{C}}$ is had m rows and n columns.

ii. The matrix $(T)_{\mathcal{B}}^{\mathcal{C}}$ is always invertible.
iii. The matrix $(\mathrm{id}_V)_{\mathcal{B}}^{\mathcal{B}}$ is always identity matrix.
iv. We have $(T)_{\mathcal{B}}^{\mathcal{C}} = (S)_{\mathcal{B}}^{\mathcal{C}}$ if and only if $T = S$.
v. For all scalars b and c, we have $(b\,T + c\,S)_{\mathcal{B}}^{\mathcal{C}} = b\,(T)_{\mathcal{B}}^{\mathcal{C}} + c\,(S)_{\mathcal{B}}^{\mathcal{C}}$.
vi. The vector space $\mathrm{Hom}(V, W)$ is always equal to the vector space $\mathrm{Hom}(W, V)$.
vii. The vector space $\mathrm{Hom}(V, W)$ is always isomorphic to the vector space $\mathrm{Hom}(W, V)$.

4.1.7 Problem. Consider the following three complex (2×2)-matrices:

$$\mathbf{X} := \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}, \qquad \mathbf{H} := \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}, \qquad \mathbf{Y} := \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}.$$

Problem 3.1.8 shows that $\mathcal{B} := (\mathbf{X}, \mathbf{H}, \mathbf{Y})$ is an ordered basis for the linear subspace $\mathfrak{sl}(2, \mathbb{C})$ of traceless complex (2×2)-matrices. For any fixed complex (2×2)-matrix $\mathbf{A}$, let $\mathrm{ad}_{\mathbf{A}} \colon \mathfrak{sl}(2, \mathbb{C}) \to \mathbb{C}^{2 \times 2}$ be the function defined by $\mathrm{ad}_{\mathbf{A}}(\mathbf{B}) := \mathbf{A}\,\mathbf{B} - \mathbf{B}\,\mathbf{A}$.

i. Show that $\mathrm{ad}_{\mathbf{A}}$ is a linear map.
ii. Show that the image of $\mathrm{ad}_{\mathbf{A}}$ is contained in $\mathfrak{sl}(2, \mathbb{C})$.
iii. Determine the matrices $(\mathrm{ad}_{\mathbf{X}})_{\mathcal{B}}^{\mathcal{B}}$, $(\mathrm{ad}_{\mathbf{H}})_{\mathcal{B}}^{\mathcal{B}}$, and $(\mathrm{ad}_{\mathbf{Y}})_{\mathcal{B}}^{\mathcal{B}}$.

4.1.8 Problem. Let $J \colon \mathbb{R}[t]_{\leqslant 2} \to \mathbb{R}[t]_{\leqslant 2}$ be the linear operator defined, for all polynomials p in $\mathbb{R}[t]_{\leqslant 2}$, by

$$\big(J[p]\big)(t) := \frac{1}{2} \int_{-1}^{1} (3 + 6st - 15s^2t^2)\, p(s)\, ds\,.$$

i. Let $\mathcal{M} := (1, t, t^2)$ denote the monomial basis for $\mathbb{R}[t]_{\leqslant 2}$. Compute the matrix $(J)_{\mathcal{B}}^{\mathcal{B}}$.
ii. Find bases for $\mathrm{Ker}(J)$ and $\mathrm{Im}(J)$.
iii. Show that J^{-1} exists and find an expression for $J^{-1}[a + b\,t + c\,t^2]$.
iv. Find polynomial p in $\mathbb{R}[t]_{\leqslant 2}$ such that $J[p] = (1 + t)^2$.
v. Find polynomial q in $\mathbb{R}[t]_{\leqslant 2}$ such that $J^2[q] = t^2$.

4.2 Similar Matrices

IS THERE AN EQUIVALENCE RELATION ON LINEAR OPERATORS? We start with a new characterization of an invertible matrix.

4.2.0 Proposition. *A matrix is invertible if and only if it is the matrix of the identity map relative to some pair of ordered bases.*

Proof. Let n be a nonnegative integer.

$\Rightarrow$: Suppose that the $(n \times n)$-matrix $\mathbf{A}$ is invertible. For all $1 \leqslant j \leqslant n$ and all $1 \leqslant k \leqslant n$, let the scalar $a_{j,k}$ be the (j, k)-entry in $\mathbf{A}$. The characterization of invertible matrices shows that the column vectors $\boldsymbol{a}_1, \boldsymbol{a}_2, \dots, \boldsymbol{a}_n$ in the matrix $\mathbf{A}$ are a basis for $\mathbb{K}^n$. For the standard basis $\mathcal{E} := (\boldsymbol{e}_1, \boldsymbol{e}_2, \dots, \boldsymbol{e}_n)$ and for all $1 \leqslant k \leqslant n$, we have $\boldsymbol{a}_k = a_{1,k}\,\boldsymbol{e}_1 + a_{2,k}\,\boldsymbol{e}_2 + \cdots + a_{n,k}\,\boldsymbol{e}_n$, so $\mathbf{A} = (\mathrm{id}_{\mathbb{K}^n})_{\mathcal{A}}^{\mathcal{E}}$.

$\Leftarrow$: For any two ordered bases $\mathcal{B}$ and $\mathcal{C}$ on an n-dimensional vector space V, the change of basis theorem [4.0.2] establishes that the $(\mathrm{id}_V)^{\mathcal{B}}_{\mathcal{C}}$ is invertible. □

4.2.1 Definition. For any two square matrices $\mathbf{A}$ and $\mathbf{B}$, we say that $\mathbf{A}$ is ***similar*** to $\mathbf{B}$ if there is an invertible matrix $\mathbf{P}$ such that $\mathbf{P}^{-1}\,\mathbf{A}\,\mathbf{P} = \mathbf{B}$ and we write $\mathbf{A} \approx \mathbf{B}$.

4.2.2 Problem. Use the matrix $\mathbf{P} := \begin{bmatrix} 1 & -1 \\ 1 & 0 \end{bmatrix}$, to demonstrate that the matrix $\mathbf{A} := \begin{bmatrix} 3 & -1 \\ 0 & 2 \end{bmatrix}$ is similar to the matrix $\mathbf{B} := \begin{bmatrix} 2 & 0 \\ 0 & 3 \end{bmatrix}$.

Solution. Since $\mathbf{P}^{-1} = \begin{bmatrix} 1 & -1 \\ 1 & 0 \end{bmatrix}^{-1} = \begin{bmatrix} 0 & 1 \\ -1 & 1 \end{bmatrix}$, we have

$$\mathbf{P}^{-1}\,\mathbf{A}\,\mathbf{P} = \begin{bmatrix} 0 & 1 \\ -1 & 1 \end{bmatrix} \begin{bmatrix} 3 & -1 \\ 0 & 2 \end{bmatrix} \begin{bmatrix} 1 & -1 \\ 1 & 0 \end{bmatrix} = \begin{bmatrix} 0 & 2 \\ -3 & 3 \end{bmatrix} \begin{bmatrix} 1 & -1 \\ 1 & 0 \end{bmatrix} = \begin{bmatrix} 2 & 0 \\ 0 & 3 \end{bmatrix} = \mathbf{B}\,,$$

which establishes that $\mathbf{A} \approx \mathbf{B}$. □

4.2.3 Lemma (Similarity is an equivalence relation)**.** *For any three square matrices* $\mathbf{A}$, $\mathbf{B}$, *and* $\mathbf{C}$ *having the same number of columns, we have the following properties.*

(reflexivity)	*The matrix* $\mathbf{A}$ *is similar to itself;* $\mathbf{A} \approx \mathbf{A}$*;*
(symmetry)	*When* $\mathbf{A} \approx \mathbf{B}$, *we have* $\mathbf{B} \approx \mathbf{A}$*;*
(transitivity)	*When* $\mathbf{A} \approx \mathbf{B}$ *and* $\mathbf{B} \approx \mathbf{C}$, *we have* $\mathbf{A} \approx \mathbf{C}$.

Proof. As $\mathbf{A} = \mathbf{I}\,\mathbf{A} = \mathbf{I}^{-1}\,\mathbf{A}\,\mathbf{I}$, we see that $\mathbf{A} \approx \mathbf{A}$. The relation $\mathbf{A} \approx \mathbf{B}$ means that there exists an invertible $\mathbf{P}$ such that $\mathbf{P}^{-1}\,\mathbf{A}\,\mathbf{P} = \mathbf{B}$. Hence, we have $\mathbf{A} = \mathbf{P}\,\mathbf{B}\,\mathbf{P}^{-1}$. Setting $\mathbf{Q} := \mathbf{P}^{-1}$ yields $\mathbf{A} = \mathbf{Q}^{-1}\,\mathbf{B}\,\mathbf{Q}$, so we deduce that $\mathbf{B} \approx \mathbf{A}$. The relations $\mathbf{A} \approx \mathbf{B}$ and $\mathbf{B} \approx \mathbf{C}$ imply that there exists invertible matrices $\mathbf{P}$ and $\mathbf{Q}$ such that $\mathbf{P}^{-1}\,\mathbf{A}\,\mathbf{P} = \mathbf{B}$ and $\mathbf{Q}^{-1}\,\mathbf{B}\,\mathbf{Q} = \mathbf{C}$. It follows that

$$(\mathbf{P}\,\mathbf{Q})^{-1}\,\mathbf{A}\,(\mathbf{P}\,\mathbf{Q}) = \mathbf{Q}^{-1}\,(\mathbf{P}^{-1}\,\mathbf{A}\,\mathbf{P})\,\mathbf{Q} = \mathbf{Q}^{-1}\,\mathbf{B}\,\mathbf{Q} = \mathbf{C}\,,$$

so we conclude that $\mathbf{A} \approx \mathbf{C}$. □

4.2.4 Proposition (Properties of similar matrices)**.** *For any two similar matrices* $\mathbf{A}$ *and* $\mathbf{B}$, *we have the following.*

i. $\det(\mathbf{A}) = \det(\mathbf{B})$;
ii. *The matrix* $\mathbf{A}$ *is invertible if and only if* $\mathbf{B}$ *is invertible;*

Proof. As the matrices $\mathbf{A}$ and $\mathbf{B}$ are similar, there exists an invertible matrix $\mathbf{P}$ such that $\mathbf{P}^{-1}\,\mathbf{A}\,\mathbf{P} = \mathbf{B}$.

i. The characterization of the determinant proves that $\det(\mathbf{I}) = 1$, so the multiplicativity of determinants and the commutativity of scalar multiplication give

$$\begin{aligned} \det(\mathbf{B}) &= \det(\mathbf{P}^{-1}\,\mathbf{A}\,\mathbf{P}) = \det(\mathbf{P}^{-1})\det(\mathbf{A})\det(\mathbf{P}) \\ &= \det(\mathbf{A})\det(\mathbf{P}^{-1}\,\mathbf{P}) = \det(\mathbf{A})\det(\mathbf{I}) = \det(\mathbf{A})\,. \end{aligned}$$

ii. Since a matrix is invertible if and only if its determinant is nonzero, the assertion follows immediately from part *i*. □

The converse of this proposition is false.

4.2.5 Problem. Prove that the matrices $\mathbf{I}$ and $\mathbf{A} := \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}$ have the same determinant, but are not similar.

Solution. Although we have $\det(\mathbf{I}) = 1 = \det(\mathbf{A})$, we also have $\mathbf{I} \not\approx \mathbf{A}$ because $\mathbf{P}^{-1}\,\mathbf{I}\,\mathbf{P} = \mathbf{P}^{-1}\,\mathbf{P} = \mathbf{I} \neq \mathbf{A}$ for any invertible matrix $\mathbf{P}$. □

4.2.6 Lemma (Multiplicative property)**.** *Let $\mathcal{A}$, $\mathcal{B}$, and $\mathcal{C}$ be ordered bases for the $\mathbb{K}$-vector spaces U, V, and W respectively. For all linear maps $S\colon U \to V$ and $T\colon V \to W$, we have $(T\,S)_{\mathcal{C}}^{\mathcal{A}} = (T)_{\mathcal{C}}^{\mathcal{B}}\,(S)_{\mathcal{B}}^{\mathcal{A}}$.*

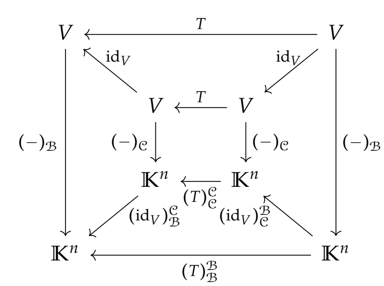
Proof. Suppose that $\mathcal{A} := (\boldsymbol{u}_1, \boldsymbol{u}_2, \dots, \boldsymbol{u}_n)$, $\mathcal{B} := (\boldsymbol{v}_1, \boldsymbol{v}_2, \dots, \boldsymbol{v}_m)$ and $\mathcal{C} := (\boldsymbol{w}_1, \boldsymbol{w}_2 \dots, \boldsymbol{w}_\ell)$. For all $1 \leqslant i \leqslant \ell$, all $1 \leqslant j \leqslant m$ and all $1 \leqslant k \leqslant n$, let the scalar $a_{j,k}$ denote the (j,k)-entry in the matrix $(S)_{\mathcal{B}}^{\mathcal{A}}$ and let the scalar $b_{i,j}$ denote the (i,j)-entry in the matrix $(T)_{\mathcal{C}}^{\mathcal{B}}$. It follows that

$$\begin{aligned} T\big[S[\boldsymbol{u}_k]\big] = T\left[\sum_{j=1}^{m} a_{j,k}\,\boldsymbol{v}_j\right] &= \sum_{j=1}^{m} a_{j,k}\,T[\boldsymbol{v}_j] \\ &= \sum_{j=1}^{m} a_{j,k}\left(\sum_{i=1}^{\ell} b_{i,j}\,\boldsymbol{w}_i\right) = \sum_{i=1}^{\ell}\left(\sum_{j=1}^{m} b_{i,j}\,a_{j,k}\right)\boldsymbol{w}_i\,. \end{aligned}$$

Hence, (i,k)-entry in $(n \times \ell)$-matrix the $(T\,S)_{\mathcal{C}}^{\mathcal{A}}$ equals $\sum_{j=1}^{m} b_{i,j}\,a_{j,k}$ which is $(T)_{\mathcal{C}}^{\mathcal{B}}\,(S)_{\mathcal{B}}^{\mathcal{A}}$ by the definition of matrix multiplication. □

4.2.7 Proposition. *Let $\mathcal{B}$ and $\mathcal{C}$ be ordered bases for a finite-dimensional vector space V. For any linear operator $T\colon V \to V$, we have*

$$(T)_{\mathcal{C}}^{\mathcal{C}} = (\mathrm{id}_V)_{\mathcal{C}}^{\mathcal{B}}\,(T)_{\mathcal{B}}^{\mathcal{B}}(\mathrm{id}_V)_{\mathcal{B}}^{\mathcal{C}} \qquad \text{and} \qquad \big((\mathrm{id}_V)_{\mathcal{C}}^{\mathcal{B}}\big)^{-1} = (\mathrm{id}_V)_{\mathcal{B}}^{\mathcal{C}}\,.$$



In other words, similar matrices represent the same linear operator relative to different ordered bases.

Proof. The multiplicative property for the matrices associated to linear maps shows that

$$(\mathrm{id}_V)_{\mathcal{C}}^{\mathcal{B}}\,(T)_{\mathcal{B}}^{\mathcal{B}}\,(\mathrm{id}_V)_{\mathcal{B}}^{\mathcal{C}} = (\mathrm{id}_V\;T)_{\mathcal{C}}^{\mathcal{B}}\,(\mathrm{id}_V)_{\mathcal{B}}^{\mathcal{C}} = (T\;\mathrm{id}_V)_{\mathcal{C}}^{\mathcal{C}} = (T)_{\mathcal{C}}^{\mathcal{C}}\,,$$

$(\mathrm{id}_V)_{\mathcal{C}}^{\mathcal{B}}\,(\mathrm{id}_V)_{\mathcal{B}}^{\mathcal{C}} = (\mathrm{id}_V)_{\mathcal{C}}^{\mathcal{C}} = \mathbf{I}$, and $(\mathrm{id}_V)_{\mathcal{B}}^{\mathcal{C}}\,(\mathrm{id}_V)_{\mathcal{C}}^{\mathcal{B}} = (\mathrm{id}_V)_{\mathcal{B}}^{\mathcal{B}} = \mathbf{I}$. □

Exercises

4.2.8 Problem. Determine which of the following statements are true. If a statement is false, then provide a counterexample. Assume that U, V and W are finite-dimensional vector spaces with order basis $\mathcal{A}$, $\mathcal{B}$, and $\mathcal{C}$ respectively. Let $S\colon U \to V$ and $T\colon V \to W$ be linear maps.

i. Every invertible matrix is the change of basis matrix from some pair of ordered bases.
ii. We always have $(T^{-1})_{\mathcal{C}}^{\mathcal{B}}\,(T)_{\mathcal{B}}^{\mathcal{C}} = I$.
iii. We always have $(S^{-1})_{\mathcal{A}}^{\mathcal{B}} = \big((S)_{\mathcal{B}}^{\mathcal{A}}\big)^{-1}$.
iv. The matrix $(\mathrm{id}_V)_{\mathcal{B}}^{\mathcal{B}}$ is always identity matrix.
v. Similar matrices represent the same linear operator relative to different ordered bases.