2 Linear Transformations and Matrices

A standard approach in algebra is to study collections of sets with a common structure and the maps between them which preserve that structure. In linear algebra this means vector spaces and the maps which behave nicely with respect to their defining structure, namely addition and scalar multiplication. Otherwise said, *linear maps* should preserve *linear combinations*.

2.1 Linear Maps, Compositions and Isomorphisms

Definition 2.1. Let V and W be vector spaces over the same field $\mathbb{F}$. A function $\mathrm{T} : V \to W$ is *linear* if $\forall \mathbf{v}_1, \mathbf{v}_2 \in V,\ \lambda \in \mathbb{F}$

$$\left.\begin{array}{l} 1.\ \mathrm{T}(\mathbf{v}_1 + \mathbf{v}_2) = \mathrm{T}(\mathbf{v}_1) + \mathrm{T}(\mathbf{v}_2) \\ 2.\ \mathrm{T}(\lambda \mathbf{v}_1) = \lambda \mathrm{T}(\mathbf{v}_1) \end{array}\right\} \quad \text{equivalently} \ \ \mathrm{T}(\lambda \mathbf{v}_1 + \mathbf{v}_2) = \lambda \mathrm{T}(\mathbf{v}_1) + \mathrm{T}(\mathbf{v}_2)$$

The set of linear maps from V to W is denoted $\mathcal{L}(V, W)$. If $V = W$ we simply write $\mathcal{L}(V)$.

Warning! The definition looks very similar to that of a *subspace*. Make sure you know the difference!

You have already met many examples of linear maps in your mathematical career.

Examples 2.2.

1. For any V, W, the *zero function* $0 \in \mathcal{L}(V, W)$ maps everything to $\mathbf{0}_W \in W$, while the *identity function* $\mathrm{I} \in \mathcal{L}(V)$ leaves everything untouched:

$$\forall \mathbf{v} \in V, \quad 0(\mathbf{v}) = \mathbf{0}_W, \quad \mathrm{I}(\mathbf{v}) = \mathbf{v}$$

2. If $\mathbf{v} \in \mathbb{F}^n$ and $A \in M_{m\times n}(\mathbb{F})$, then *left-multiplication by A* is the linear map

$$\mathrm{L}_A : \mathbb{F}^n \to \mathbb{F}^m : \mathbf{v} \mapsto A\mathbf{v}$$

Verifying that this is linear is tedious: e.g., the i^{th} entry of the vector $A(\mathbf{x} + \mathbf{y})$ is

$$[A(\mathbf{x} + \mathbf{y})]_i = \sum_{j=1}^{n} a_{ij}(x_j + y_j) = \sum_{j=1}^{n} a_{ij}x_j + \sum_{j=1}^{n} a_{ij}y_j = [A\mathbf{x}]_i + [A\mathbf{y}]_i$$

precisely the i^{th} entry of the vector $A\mathbf{x} + A\mathbf{y}$. Scalar multiplication is similar.

3. Differentiation: Since $(\lambda f + g)' = \lambda f' + g'$, the function $\mathrm{T} : f \mapsto \frac{\mathrm{d}f}{\mathrm{d}x}$ is linear map defined on any vector space of differentiable functions.

4. Integration: $\mathrm{T} : C([a, b]) \to \mathbb{R} : f \mapsto \int_a^b f(x)\,\mathrm{d}x$ is linear, where $C([a, b])$ is the set of continuous functions $f : [a, b] \to \mathbb{R}$.

The following are easy to prove straight from Definition 2.1.

Lemma 2.3.

1. *Linear maps preserve zero:* $\mathrm{T}(\mathbf{0}_V) = \mathrm{T}(0\mathbf{v}) = 0\mathrm{T}(\mathbf{v}) = \mathbf{0}_W$.
2. *$\mathcal{L}(V, W)$ is a vector space whose identity is the zero function* $0 \in \mathcal{L}(V, W)$.

More important is the fact that linear maps are, as advertised, precisely those functions which preserve linear combinations.

Lemma 2.4. *A function* $\mathrm{T} : V \to W$ *is linear if and only if*

$$\forall \mathbf{v}_i \in V, a_i \in \mathbb{F}, \quad \mathrm{T}\left(\sum_{i=1}^{n} a_i \mathbf{v}_i\right) = \sum_{i=1}^{n} a_i \mathrm{T}(\mathbf{v}_i)$$

In particular, T *is defined by what it does to a basis: if* β *is a basis and we know all the values* $\mathrm{T}(\mathbf{v}_j)$ *for every* $\mathbf{v}_j \in \beta$*, then we know* T.

The proof is an exercise.

Example 2.5. The standard basis of $\mathbb{R}^3$ is $\beta = \{\mathbf{i}, \mathbf{j}, \mathbf{k}\}$. Suppose $\mathrm{T} \in \mathcal{L}(\mathbb{R}^3, \mathbb{R}^2)$ is such that

$$\mathrm{T}(\mathbf{i}) = \begin{pmatrix} 3 \\ 4 \end{pmatrix}, \quad \mathrm{T}(\mathbf{j}) = \begin{pmatrix} 1 \\ 9 \end{pmatrix}, \quad \mathrm{T}(\mathbf{k}) = \begin{pmatrix} -4 \\ 7 \end{pmatrix}$$

Since T is defined on a basis, we can easily compute the entire map:

$$\mathrm{T}\begin{pmatrix} x \\ y \\ z \end{pmatrix} = x\mathrm{T}(\mathbf{i}) + y\mathrm{T}(\mathbf{j}) + z\mathrm{T}(\mathbf{k}) = x\begin{pmatrix} 3 \\ 4 \end{pmatrix} + y\begin{pmatrix} 1 \\ 9 \end{pmatrix} + z\begin{pmatrix} -4 \\ 7 \end{pmatrix} = \begin{pmatrix} 3 & 1 & -4 \\ 4 & 9 & 7 \end{pmatrix}\begin{pmatrix} x \\ y \\ z \end{pmatrix}$$

It should be no surprise that the linear map T is in fact L_A where $A = \left(\begin{smallmatrix} 3 & 1 & -4 \\ 4 & 9 & 7 \end{smallmatrix}\right)$. We'll explore the relationship between linear maps, bases and matrices more fully in Section 2.3.

Compositions & Inverses

As functions, linear maps may be composed with each other, and might have inverses. Likely the only new information in the definition is notational.

Definition 2.6.
1. If $\mathrm{T} \in \mathcal{L}(V, W)$ and $\mathrm{U} \in \mathcal{L}(W, X)$, then the *composition* of T and U is defined by[a]

   $$\mathrm{UT} : V \to X : \mathbf{v} \mapsto \mathrm{U}(\mathrm{T}(\mathbf{v}))$$

   In the special case that $\mathrm{U} = \mathrm{T}$ (necessarily $V = W = X$), the composition is written T^2; similarly for higher powers $\mathrm{T}^3, \mathrm{T}^4$, etc.

2. $\mathrm{T} \in \mathcal{L}(V, W)$ is *invertible*, or an *isomorphism*, if has an *inverse*; a *function* $\mathrm{U} : W \to V$ for which

   $$\mathrm{TU} = \mathrm{I}_W \quad \text{and} \quad \mathrm{UT} = \mathrm{I}_V$$

   where $\mathrm{I}_W, \mathrm{I}_V$ are the identity maps on V, W respectively.

   Vector spaces V, W are *isomorphic* if there exists an isomorphism $\mathrm{T} \in \mathcal{L}(V, W)$.

[a] For brevity, we write UT instead of $\mathrm{U} \circ \mathrm{T}$.

Examples 2.7. 1. If $\mathrm{T} \in \mathcal{L}(\mathbb{R}^2, P_2(\mathbb{R}))$ is defined by $\mathrm{T}\begin{pmatrix} a \\ b \end{pmatrix} = a + bx$, then T is an isomorphism, with inverse $\mathrm{U} : a + bx \mapsto \begin{pmatrix} a \\ b \end{pmatrix}$.

2. (a) Let $A = \begin{pmatrix} 1 & 2 \\ -3 & 1 \end{pmatrix}$ and consider the linear map $\mathrm{L}_A \in \mathcal{L}(\mathbb{R}^2)$. Then

$$\mathrm{L}_A^2 \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 1 & 2 \\ -3 & 1 \end{pmatrix} \left(\begin{pmatrix} 1 & 2 \\ -3 & 1 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} \right) = \begin{pmatrix} 1 & 2 \\ -3 & 1 \end{pmatrix} \begin{pmatrix} x + 2y \\ -3x + y \end{pmatrix} = \begin{pmatrix} -5x + 4y \\ -6x - 5y \end{pmatrix}$$

Unsurprisingly, this is the linear map L_{A^2}: indeed $A^2 = \begin{pmatrix} -5 & 4 \\ -6 & -5 \end{pmatrix}$.

(b) Now consider $\mathrm{U} = \mathrm{L}_B$ where $B = \frac{1}{7}\begin{pmatrix} 1 & -2 \\ 3 & 1 \end{pmatrix} = A^{-1}$ is the inverse matrix. Plainly

$$\mathrm{TU}\begin{pmatrix} x \\ y \end{pmatrix} = A\left(B\begin{pmatrix} x \\ y \end{pmatrix}\right) = AB\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} x \\ y \end{pmatrix} \qquad \mathrm{UT}\begin{pmatrix} x \\ y \end{pmatrix} = (BA)\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} x \\ y \end{pmatrix}$$

whence U is an inverse of T, which is therefore an isomorphism.

3. Let $V = C^\infty(\mathbb{R})$ be the vector space of infinitely differentiable functions $f : \mathbb{R} \to \mathbb{R}$. Then $\mathrm{T}(f)(x) = f'(x) + f(x)$ and $\mathrm{U}(f)(x) = \int_0^1 f(x)\,\mathrm{d}x$ are linear maps $\mathrm{T}, \mathrm{U} \in \mathcal{L}(V)$. We compute:

$$\mathrm{UT}(f)(x) = \mathrm{U}\left(f'(x) + f(x)\right) = \int_0^1 f'(x) + f(x)\,\mathrm{d}x = f(1) - f(0) + \int_0^1 f(x)\,\mathrm{d}x$$

4. Define $\mathrm{T}, \mathrm{U} \in \mathcal{L}(P_1(\mathbb{R}))$ by $\mathrm{T}(f)(x) = f(x) + 3f'(x)$ and $\mathrm{U}(f)(x) = f(x) - 3f'(x)$. Then,

$$\mathrm{TU}(f)(x) = \mathrm{T}\big(f(x) - 3f'(x)\big) = f(x) - 3f'(x) + 3\big(f'(x) - 3f''(x)\big) = f(x)$$

since $f''(x) = 0$. We can similarly check that $\mathrm{UT} = \mathrm{I}_{P_1(\mathbb{R})}$ so that U is an inverse of T.

The examples suggest some simple results.

Lemma 2.8. *A composition of linear maps is linear.*

Proof. This follows from the linearity of both T and U:

$$\begin{aligned} \mathrm{UT}(\lambda \mathbf{v}_1 + \mathbf{v}_2) &= \mathrm{U}\Big(\mathrm{T}(\lambda \mathbf{v}_1 + \mathbf{v}_2)\Big) = \mathrm{U}\Big(\lambda \mathrm{T}(\mathbf{v}_1) + \mathrm{T}(\mathbf{v}_2)\Big) && \text{(linearity of T)} \\ &= \lambda \mathrm{U}\big(\mathrm{T}(\mathbf{v}_1)\big) + \mathrm{U}\big(\mathrm{T}(\mathbf{v}_2)\big) && \text{(linearity of U)} \\ &= \lambda \mathrm{UT}(\mathbf{v}_1) + \mathrm{UT}(\mathbf{v}_2) \end{aligned}$$

■

Lemma 2.9. *Let* $\mathrm{T} \in \mathcal{L}(V, W)$ *be an isomorphism. Then:*

1. *The inverse is unique: we call this function* T^{-1}.
2. *The inverse is an isomorphism:* $\mathrm{T}^{-1} \in \mathcal{L}(W, V)$ *is linear and invertible with* $(\mathrm{T}^{-1})^{-1} = \mathrm{T}$.
3. *If* $\mathrm{S} \in \mathcal{L}(W, X)$ *is invertible, then* $\mathrm{ST} \in \mathcal{L}(V, X)$ *is invertible with* $(\mathrm{ST})^{-1} = \mathrm{T}^{-1}\mathrm{S}^{-1}$.

We leave the proof as an exercise. We'll return to invertibility later.

Exercises 2.1

1. Show explicitly that the following are linear maps:
   (a) $\mathrm{T} : \mathbb{R}^2 \mapsto \mathbb{R}^3 : \begin{pmatrix} x \\ y \end{pmatrix} \mapsto \left(\begin{smallmatrix} 3x \\ 2y-x \\ x \end{smallmatrix} \right)$
   (b) $\mathrm{T} : P_3(\mathbb{R}) \to P_3(\mathbb{R})$ defined by $\mathrm{T}(f)(x) = (x-1)f''(x) + 7f(x)$
   (c) $\mathrm{T} : M_n(\mathbb{F}) \to M_n(\mathbb{F}) : A \mapsto 3A - 2A^T$, where A^T is the transpose
2. Give a reason why the function $\mathrm{T} : \mathbb{R}^2 \to \mathbb{R}^2 : \begin{pmatrix} x \\ y \end{pmatrix} \mapsto \begin{pmatrix} xy \\ 3x - y \end{pmatrix}$ is non-linear.
3. Let $\mathrm{T}(f)(x) = f(x) + xf'(x)$ where $f(x) \in P_2(\mathbb{R})$.
   (a) Show that $\mathrm{T} \in \mathcal{L}(P_2(\mathbb{R}))$.
   (b) Compute the linear map $\mathrm{T}^2 \in \mathcal{L}(P_2(\mathbb{R}))$; that is, express $\mathrm{T}^2(f)(x)$ in terms of f and its derivatives.
4. Let $\mathrm{T}, \mathrm{U} \in \mathcal{L}(P_2(\mathbb{R}))$ be defined by
   $$\mathrm{T}(f)(x) = 2f(x) + f''(x), \qquad \mathrm{U}(f)(x) = \frac{1}{4}\big(2f(x) - f''(x)\big)$$
   Prove that $\mathrm{U} = \mathrm{T}^{-1}$.
5. Prove or disprove: $\mathrm{T} : \mathbb{R}^2 \to P_1(\mathbb{R}) : \left(\begin{smallmatrix} a \\ b \end{smallmatrix} \right) \mapsto ax - 2a + b$ is an isomorphism.
   (*Try to guess an inverse!*)
6. Prove, the 'equivalently' claim in Definition 2.1, that $\mathrm{T} : V \to W$ is linear if and only if
   $$\forall \mathbf{v}_1, \mathbf{v}_2 \in V,\ \lambda \in \mathbb{F},\ \mathrm{T}(\lambda \mathbf{v}_1 + \mathbf{v}_2) = \lambda \mathrm{T}(\mathbf{v}_1) + \mathrm{T}(\mathbf{v}_2)$$
7. Prove explicitly that if $\mathrm{T}_1, \mathrm{T}_2 \in \mathcal{L}(V, W)$, then $\mathrm{T}_1 + \mathrm{T}_2$ is also a linear map.
8. Prove Lemma 2.4.
9. Prove all three parts of Lemma 2.9.
10. With reference to Lemma 2.9, explain why 'Isomorphic' is an equivalence relation on any set of vector spaces.
11. Prove that a linear map $\mathrm{T} \in \mathcal{L}(V, W)$ is an isomorphism if and only if it is *bijective*, that is,
    (a) *Injective*: $\forall \mathbf{v}_1, \mathbf{v}_2 \in V, \mathrm{T}(\mathbf{v}_1) = \mathrm{T}(\mathbf{v}_2) \implies \mathbf{v}_1 = \mathbf{v}_2$.
    (b) *Surjective*: $\forall \mathbf{w} \in W, \exists \mathbf{v} \in V$ such that $\mathbf{w} = \mathrm{T}(\mathbf{v})$.

2.2 The Rank–Nullity Theorem

We define two sets which are crucial for understanding linear maps.

Definition 2.10. Suppose $\mathrm{T} \in \mathcal{L}(V, W)$. Its *range/image* and *nullspace/kernel* are the sets

$$\mathcal{R}(\mathrm{T}) := \{\mathrm{T}(\mathbf{v}) \in W : \mathbf{v} \in V\}, \qquad \mathcal{N}(\mathrm{T}) := \{\mathbf{v} \in V : \mathrm{T}(\mathbf{v}) = \mathbf{0}_W\}$$

Example 2.11. Let $\mathrm{T} \in \mathcal{L}(P_3(\mathbb{R}), P_2(\mathbb{R}))$ be 'differentiate:'

$$\mathrm{T}(a + bx + cx^2 + dx^3) = b + 2cx + 3dx^2$$

Plainly $\mathcal{N}(\mathrm{T}) = \{a : a \in \mathbb{R}\} \subseteq P_3(\mathbb{R})$ is the space of constants, and

$$\mathcal{R}(\mathrm{T}) = \mathrm{Span}\{0, 1, 2x, 3x^2\} = \mathrm{Span}\{1, 2x, 3x^2\} = P_2(\mathbb{R})$$

The example immediately suggests that the range and nullspace are not merely sub*sets*...

Lemma 2.12. *The nullspace and range of* $\mathrm{T} \in \mathcal{L}(V, W)$ *are subspaces of* V *and* W *respectively.*

Proof. Everything follows from the formula $\mathrm{T}(\lambda\mathbf{v}_1 + \mathbf{v}_2) = \lambda\mathrm{T}(\mathbf{v}_1) + \mathrm{T}(\mathbf{v}_2)$.
$\mathcal{N}(\mathrm{T})$ is non-empty since $\mathrm{T}(\mathbf{0}) = \mathbf{0}$. Moreover, if $\mathbf{v}_1, \mathbf{v}_2 \in \mathcal{N}(\mathrm{T})$, so is $\lambda\mathbf{v}_1 + \mathbf{v}_2$, whence $\mathcal{N}(\mathrm{T})$ is a subspace of V. The range is similar. ∎

Definition 2.13. The *rank* and *nullity* of a linear map $\mathrm{T} \in \mathcal{L}(V, W)$ are

$$\mathrm{rank}\,\mathrm{T} := \dim \mathcal{R}(\mathrm{T}) \qquad \mathrm{null}\,\mathrm{T} := \dim \mathcal{N}(\mathrm{T})$$

Examples 2.14.
1. Revisiting Example 2.11, we see that $\mathrm{rank}\,\mathrm{T} = 3$ and $\mathrm{null}\,\mathrm{T} = 1$.
2. Let $\mathrm{L}_A(\mathbf{v}) = A\mathbf{v}$ where $A \in M_{m\times n}(\mathbb{F})$. Applied to the standard basis β of $\mathbb{F}^n$, we see that

   $$\mathcal{R}(\mathrm{L}_A) = \mathrm{Span}\,\mathrm{L}_A(\beta) = \mathrm{Span}\{A\mathbf{e}_1, \ldots, A\mathbf{e}_n\}$$

   Since $A\mathbf{e}_j$ is the j^{th} column of A, we see that $\mathcal{R}(\mathrm{L}_A)$ *is the column space of* A.
   For example, if $A = \begin{pmatrix} 1 & 2 & 3 \\ 0 & 1 & 1 \\ 1 & 0 & 1 \\ 0 & 3 & 3 \end{pmatrix}$, then

   $$\mathcal{R}(L_A) = \mathrm{Span}\left\{\begin{pmatrix}1\\0\\1\\0\end{pmatrix}, \begin{pmatrix}2\\1\\0\\3\end{pmatrix}, \begin{pmatrix}3\\1\\3\\3\end{pmatrix}\right\} = \mathrm{Span}\left\{\begin{pmatrix}1\\0\\1\\0\end{pmatrix}, \begin{pmatrix}2\\1\\0\\3\end{pmatrix}\right\}$$

   since the third column is a linear combination of the others. We conclude that $\mathrm{rank}\,\mathrm{L}_A = 2$. To find the nullspace requires solving the system $A\mathbf{x} = \mathbf{0}$. After a few row operations,

   $$\mathcal{N}(\mathrm{L}_A) = \left\{\begin{pmatrix}x\\y\\z\end{pmatrix} : x + z = 0 = y + z\right\} = \mathrm{Span}\left\{\begin{pmatrix}1\\1\\-1\end{pmatrix}\right\} \implies \mathrm{null}\,\mathrm{L}_A = 1$$

These values in fact satisfy one of the most crucial relationships in linear algebra.

Theorem 2.15 (Rank–Nullity). *If* $\mathrm{T} \in \mathcal{L}(V, W)$*, then* $\operatorname{rank} \mathrm{T} + \operatorname{null} \mathrm{T} = \dim V$*.*

Examples 2.16. First revisit Examples 2.14.

1. If $\mathrm{T} = \frac{\mathrm{d}}{\mathrm{d}x} \in \mathcal{L}(P_3(\mathbb{R}), P_2(\mathbb{R}))$, then, in accordance with the rank–nullity theorem,

$$\operatorname{rank} \mathrm{T} + \operatorname{null} \mathrm{T} = 3 + 1 = 4 = \dim P_3(\mathbb{R})$$

2. If $A = \left(\begin{smallmatrix} 1 & 2 & 3 \\ 0 & 1 & 1 \\ 1 & 0 & 1 \\ 0 & 3 & 3 \end{smallmatrix}\right)$ so that $\mathrm{L}_A \in \mathcal{L}(\mathbb{R}^3, \mathbb{R}^4)$, then

$$\operatorname{rank} \mathrm{L}_A + \operatorname{null} \mathrm{L}_A = 2 + 1 = 3 = \dim \mathbb{R}^3$$

3. Let $\mathrm{T} \in \mathcal{L}(M_2(\mathbb{R}))$ be defined by $\mathrm{T}(A) = A + A^T$ where A^T is the transpose, that is,

$$\mathrm{T}\begin{pmatrix} a & b \\ c & d \end{pmatrix} = \begin{pmatrix} 2a & b+c \\ b+c & 2d \end{pmatrix}$$

   It should be clear that

$$\mathcal{R}(\mathrm{T}) = \left\{ \left(\begin{smallmatrix} p & q \\ q & r \end{smallmatrix}\right) : p, q, r \in \mathbb{R} \right\} \quad \mathcal{N}(\mathrm{T}) = \left\{ \left(\begin{smallmatrix} 0 & s \\ -s & 0 \end{smallmatrix}\right) : s \in \mathbb{R} \right\}$$

   are, respectively, the subspaces of symmetric and skew-symmetric matrices. Plainly $\operatorname{rank} \mathrm{T} + \operatorname{null} \mathrm{T} = 3 + 1 = 4 = \dim M_2(\mathbb{R})$.

4. Let $\mathrm{T} \in \mathcal{L}(P(\mathbb{R}))$ be defined by $\mathrm{T}(f)(x) = f(x) + f(-x)$. It is easy to check that this is linear. Moreover,

$$f \in \mathcal{N}(\mathrm{T}) \iff f(-x) = -f(x) \iff f \text{ is odd } \iff f \in \operatorname{Span}\{x, x^3, x^5, x^7, \cdots\}$$

   We may also check that the range consists of all even polynomials $\mathcal{R}(\mathrm{T}) = \operatorname{Span}\{1, x^2, x^4, \ldots\}$. The rank–nullity theorem holds even for this infinite-dimensional example, though it isn't very instructive.[1]

Before proving the rank–nullity theorem, we consider what a linear map does to a basis.

Lemma 2.17. *If* β *is a basis of* V*, then* $\mathrm{T}(\beta) := \{\mathrm{T}(\mathbf{v}) : \mathbf{v} \in \beta\}$ *is a spanning set for* $\mathcal{R}(\mathrm{T})$*.*

Proof. Let $\mathrm{T}(\mathbf{v}) \in \mathcal{R}(\mathrm{T})$. Then $\exists \mathbf{v}_i \in \beta$ such that $\mathbf{v} = a_1\mathbf{v}_1 + \cdots + a_n\mathbf{v}_n$. By Lemma 2.4,

$$\mathrm{T}(\mathbf{v}) = \mathrm{T}\left(\sum_{i=1}^n a_i \mathbf{v}_i\right) = \sum_{i=1}^n a_i \mathrm{T}(\mathbf{v}_i) \in \operatorname{Span}(\mathrm{T}(\beta))$$

Thus $\mathcal{R}(\mathrm{T}) \leq \operatorname{Span} \mathrm{T}(\beta)$.

Conversely, the right hand side of Lemma 2.4 is a general element of $\operatorname{Span} \mathrm{T}(\beta)$ which is certainly in the range of T. Thus $\operatorname{Span} \mathrm{T}(\beta) \leq \mathcal{R}(\mathrm{T})$, and the subspaces are equal. ∎

[1] If you're happy with addition of infinite cardinals, the rank–nullity theorem reads $\aleph_0 + \aleph_0 = \aleph_0$.

Proof of Rank–Nullity Theorem. Suppose ν is a basis of $\mathcal{N}(\mathrm{T})$. By the Extension Theorem, we may extend to a basis $\beta = \nu \cup \rho$ of V, where $\nu \cap \rho = \emptyset$. The proof now rests on two claims:

1. $\mathrm{T}(\rho)$ is a basis of $\mathcal{R}(\mathrm{T})$. First we verify the spanning property: by Lemma 2.17,

$$\mathcal{R}(\mathrm{T}) = \operatorname{Span}\mathrm{T}(\beta) = \operatorname{Span}\{\mathrm{T}(\mathbf{n}), \mathrm{T}(\mathbf{r}) : \mathbf{n} \in \nu, \mathbf{r} \in \rho\} = \operatorname{Span}\{\mathrm{T}(\mathbf{r}) : \mathbf{r} \in \rho\} = \operatorname{Span}\mathrm{T}(\rho)$$

   For linear independence, suppose $\mathbf{r}_1, \ldots, \mathbf{r}_r \in \rho$ and compute:

$$\begin{aligned} \mathbf{0}_W = \sum_{i=1}^{r} a_i \mathrm{T}(\mathbf{r}_i) = \mathrm{T}\left(\sum_{i=1}^{r} a_i \mathbf{r}_i\right) &\implies \sum_{i=1}^{r} a_i \mathbf{r}_i \in \mathcal{N}(\mathrm{T}) \cap \operatorname{Span}\rho = \{\mathbf{0}_V\} && (*) \\ &\implies a_1 = \cdots = a_n = 0 && (\dagger) \end{aligned}$$

   We conclude that $\mathrm{T}(\rho)$ is a linearly independent spanning set of $\mathcal{R}(\mathrm{T})$.

2. $|\rho| = |\mathrm{T}(\rho)|$. We leave this to Exercise 2.1.3.

In conclusion:

$$\dim V = |\beta| = |\nu| + |\rho| \overset{(2)}{=} \operatorname{null}\mathrm{T} + |\mathrm{T}(\rho)| \overset{(1)}{=} \operatorname{null}\mathrm{T} + \operatorname{rank}\mathrm{T}$$

■

Note that the proof works when V is infinite-dimensional though the result is not so useful.

Injective & Surjective Linear Maps: Isomorphisms Revisited

It turns out that injectivity and surjectivity may be checked by considering the rank and nullity.

Theorem 2.18. *Suppose* $\mathrm{T} \in \mathcal{L}(V, W)$.

1. T *is injective* $\iff \mathcal{N}(\mathrm{T}) = \{\mathbf{0}\} \iff \operatorname{null}\mathrm{T} = 0$.
2. T *surjective* $\iff \mathcal{R}(\mathrm{T}) = W \implies \operatorname{rank}\mathrm{T} = \dim W$.
   Additionally, if W is finite-dimensional, then $\operatorname{rank}\mathrm{T} = \dim W \implies \mathrm{T}$ *surjective.*
3. *If* $\dim V = \dim W$ *is finite, then*

$$\mathrm{T} \text{ is injective} \iff \operatorname{null}\mathrm{T} = 0 \iff \operatorname{rank}\mathrm{T} = \dim V \iff \mathrm{T} \text{ is surjective}$$

   In view of Exercise 2.1.11, if any one of these conditions holds then T *is an isomorphism*

Proof. 1. Injectivity means $\mathrm{T}(\mathbf{v}_1) = \mathrm{T}(\mathbf{v}_2) \implies \mathbf{v}_1 = \mathbf{v}_2$. The result follows follows quickly from,

$$\mathrm{T}(\mathbf{v}_1) = \mathrm{T}(\mathbf{v}_2) \overset{\text{linearity}}{\iff} \mathrm{T}(\mathbf{v}_1 - \mathbf{v}_2) = \mathbf{0} \iff \mathbf{v}_1 - \mathbf{v}_2 \in \mathcal{N}(\mathrm{T})$$

2. $\mathcal{R}(\mathrm{T}) = W$ is the definition of surjectivity: certainly this implies $\operatorname{rank}\mathrm{T} = \dim W$. In finite dimensions we have the converse:

$$\mathcal{R}(\mathrm{T}) \leq W \text{ and } \dim\mathcal{R}(\mathrm{T}) = \operatorname{rank}\mathrm{T} = \dim W \implies \mathcal{R}(\mathrm{T}) = W$$

3. This follows immediately from the rank–nullity theorem and parts 1 & 2.

■

Examples 2.19. 1. Continuing a previous example, $\mathrm{T} = \frac{\mathrm{d}}{\mathrm{d}x} \in \mathcal{L}(P_3(\mathbb{R}), P_2(\mathbb{R}))$ is surjective ($\operatorname{rank} \mathrm{T} = 3 = \dim P_2(\mathbb{R})$) but not injective ($\operatorname{null} \mathrm{T} = 1 \neq 0$). The non-injectivity of T corresponds to the famous '$+C$' from calculus:

$$\frac{\mathrm{d}}{\mathrm{d}x}(p(x) + C) = \frac{\mathrm{d}}{\mathrm{d}x} p(x)$$

2. Consider $\mathrm{L}_A \in \mathcal{L}(\mathbb{R}^2, \mathbb{R}^3)$ where $A = \left(\begin{smallmatrix} 1 & 3 \\ 0 & 1 \\ 1 & 1 \end{smallmatrix}\right)$ and compute the range and nullspace:

$\mathcal{R}(\mathrm{L}_A) = \operatorname{Span}\left\{ \left(\begin{smallmatrix} 1 \\ 0 \\ 1 \end{smallmatrix}\right), \left(\begin{smallmatrix} 3 \\ 1 \\ 1 \end{smallmatrix}\right) \right\} \implies \operatorname{rank} \mathrm{L}_A = 2 \neq \dim \mathbb{R}^3 \implies \mathrm{L}_A$ is not surjective.

$\mathcal{N}(\mathrm{L}_A) = \{\mathbf{0}\} \implies \mathrm{L}_A$ is injective.

3. Let $\mathrm{T} \in \mathcal{L}(P_2(\mathbb{R}))$ be defined by $\mathrm{T}(p)(x) = p'(x) + (x^2 - 1)\int_0^1 p(t)\, \mathrm{d}t$. Observe that

$$\begin{aligned} \mathrm{T}(a + bx + cx^2) &= b + 2cx + (x^2 - 1)\left(a + \frac{1}{2}b + \frac{1}{3}c\right) \\ &= \frac{1}{2}b - a - \frac{1}{3}c + 2cx + \left(a + \frac{1}{2}b + \frac{1}{3}c\right)x^2 \\ &= 0 \iff c = 0 = \frac{1}{2}b - a = a + \frac{1}{2}b \iff a = b = c = 0 \end{aligned}$$

Since $\mathcal{N}(\mathrm{T}) = \{0\}$, we see that T is bijective and thus an isomorphism.

Corollary 2.20. *Suppose that V, W are vector spaces over the same field.*

1. *If $\mathrm{T} \in \mathcal{L}(V, W)$ is an isomorphism and β is a basis of V, then $\mathrm{T}(\beta)$ is a basis of W.*
2. *V and W are isomorphic if and only if $\dim V = \dim W$.*

The proof is little mostly a special case of the rank–nullity theorem: indeed you should re-read the proof to convince yourself that the restriction $\mathrm{T}_{\operatorname{Span}\rho} : \operatorname{Span}\rho \to \mathcal{R}(\mathrm{T})$ is indeed an isomorphism!

Proof. 1. This is a special case of part 1 of the proof of the rank–nullity theorem: we have $\mathcal{N}(\mathrm{T}) = \{\mathbf{0}\}$, whence $\nu = \emptyset$ and $\rho = \beta$.

2. ($\Rightarrow$) This is part 2 of the same proof: $\dim V = |\beta| = |\mathrm{T}(\beta)| = \dim W$.

($\Leftarrow$) Let $\dim V = \dim W$ and choose any bases β, γ of V, W. These have the same cardinality, whence $\exists f : \beta \to \gamma$ a bijection. By Lemma 2.4, f defines a unique linear map $\mathrm{T} \in \mathcal{L}(V, W)$: if $\mathbf{v}_1, \ldots, \mathbf{v}_n \in \beta$, define

$$\mathrm{T}\left(\sum_{i=1}^{n} a_i \mathbf{v}_i\right) := \sum_{i=1}^{n} a_i f(\mathbf{v}_i)$$

It is straightforward, if tedious, to check that T is an isomorphism. ∎

Examples 2.21. 1. $\mathbb{R}^6, P_5(\mathbb{R}), M_{2\times 3}(\mathbb{R}), M_{3\times 2}(\mathbb{R})$ are isomorphic since they all have dimension 6 over the same field $\mathbb{R}$. Explicit isomorphisms can be found by lining up the standard bases: e.g.

$$\mathrm{T}: M_{2\times 3}(\mathbb{R}) \to P_5(\mathbb{R}) \text{ by } \mathrm{T}\begin{pmatrix} a & b & c \\ d & e & f \end{pmatrix} = a + bx + cx^2 + dx^3 + ex^4 + fx^5$$

2. Be careful with base fields!

   (a) $P_5(\mathbb{R})$ is not isomorphic to $\mathbb{C}^6$ since the (implied) base fields are different $\mathbb{R} \neq \mathbb{C}$.

   (b) Viewing $\mathbb{C}^6$ as a vector space over $\mathbb{R}$, we still fail to have isomorphicity since

   $$\dim_{\mathbb{R}} P_5(\mathbb{R}) = 6 \neq 12 = \dim_{\mathbb{R}} \mathbb{C}^6$$

   (c) As vector spaces over $\mathbb{R}$ the spaces $P_5(\mathbb{R})$ and $\mathbb{C}^3$ are isomorphic since both have *real* dimension 6. A suitable isomorphism is

   $$\mathrm{T}(a + bx + cx^2 + dx^3 + ex^5 + fx^5) = \begin{pmatrix} a+ib \\ c+id \\ e+if \end{pmatrix}$$

Exercises 2.2 1. For each function $\mathrm{T}: V \to W$: prove that T is linear, compute $\mathcal{N}(\mathrm{T})$ and $\mathcal{R}(\mathrm{T})$, and the rank and nullity, verify the Rank–Nullity theorem, and determine whether the function is injective or surjective.

   (a) $\mathrm{T}: \mathbb{R}^2 \to \mathbb{R}^3 : \begin{pmatrix} x \\ y \end{pmatrix} \mapsto \begin{pmatrix} x+y \\ 0 \\ 2x-y \end{pmatrix}$

   (b) $\mathrm{T}: M_n(\mathbb{F}) \to \mathbb{F} : A \mapsto \operatorname{tr} A$, where $\operatorname{tr} A = \sum_{i=1}^{n} A_{ii}$ is the *trace* of A.

   (c) $\mathrm{T}: P_4(\mathbb{R}) \to P_4(\mathbb{R}) : f \mapsto g$ where $g(x) = f(x) - (x^2+1)f''(x)$.

2. For each linear map, find the range and nullspace and compute the rank and nullity.

   (a) $\mathrm{T} = \mathrm{L}_A \in \mathcal{L}(\mathbb{C}^4)$ where $A = \begin{pmatrix} 1 & i & 1 & -i \\ 0 & 1 & 0 & -1 \\ i & 0 & i & 0 \\ i & -1 & i & 1 \end{pmatrix}$

   (b) $\mathrm{T} \in \mathcal{L}(V)$ where $\mathrm{T}(f)(x) = f''(x) - 4f(x)$ and $V = \operatorname{Span}\{e^{2x}, e^{-2x}, xe^{2x}, xe^{-2x}\}$.

3. Verify the following claims made during the proof of the rank–nullity theorem.

   (a) $(*)$ $\mathcal{N}(\mathrm{T}) \cap \operatorname{Span} \rho = \{\mathbf{0}_V\}$.

   (b) $(\dagger)$ $\sum_{i=1}^{r} a_i \mathbf{r}_i = \mathbf{0}_V \implies a_1 = \cdots = a_n = 0$.

   (c) $|\rho| = |\mathrm{T}(\rho)|$: do this by proving that $\mathrm{T}_\rho : \rho \to \mathrm{T}(\rho)$ is a bijection.

4. Suppose that $\dim V > \dim W$. Prove that there are no injective functions $\mathrm{T} \in \mathcal{L}(V, W)$.

5. Let $\mathrm{T} = \frac{\mathrm{d}}{\mathrm{d}x} \in \mathcal{L}(P(\mathbb{R}))$. Is T injective? Surjective? Why is this not a problem for Theorem 2.18?

6. Which of the following pairs are isomorphic? If yes, state an explicit isomorphism.

   (a) $\mathbb{F}^3$ and $P_3(\mathbb{F})$ (b) $\mathbb{F}^4$ and $P_3(\mathbb{F})$ (c) $M_2(\mathbb{R})$ and $P_3(\mathbb{R})$

   (d) $V = \{A \in M_2(\mathbb{R}) : \operatorname{tr} A = 0\}$ and $\mathbb{R}^4$ (e) $V = \{A \in M_2(\mathbb{R}) : \operatorname{tr} A = 0\}$ and $\mathbb{R}^3$

7. Let $\mathrm{T} \in \mathcal{L}(V, W)$ be an isomorphism and U a subspace of V. Prove that $\mathrm{T}(U) := \{\mathrm{T}(\mathbf{u}) : \mathbf{u} \in U\}$ is a subspace of W and that $\dim \mathrm{T}(U) = \dim U$.

8. (Hard) We prove the *first isomorphism theorem* for vector spaces and see its relation to the rank–nullity theorem. This should be familiar if you've studied group theory: you will need to recall the exercise on cosets from the first chapter. Throughout U is a subspace of V.

   (a) Let ${}^{V}\!/_{U} = \{\mathbf{v} + U : \mathbf{v} \in V\}$ be the set of cosets of U in V. Prove that the *canonical map*

   $$\gamma : V \to {}^{V}\!/_{U} : \mathbf{v} \mapsto \mathbf{v} + U$$

   is linear and has nullspace U.
   (*Thus every subspace of V is the nullspace of some linear map γ with* $\operatorname{dom} \gamma = V$)

   (b) Let $\mathrm{T} \in \mathcal{L}(V, W)$. Prove that

   $$\mathbf{v}_1 + \mathcal{N}(\mathrm{T}) = \mathbf{v}_2 + \mathcal{N}(\mathrm{T}) \iff \mathrm{T}(\mathbf{v}_1) = \mathrm{T}(\mathbf{v}_2)$$

   (c) Prove that the following is a well-defined isomorphism of vector spaces:

   $$\mu : {}^{V}\!/_{\mathcal{N}(\mathrm{T})} \to \mathcal{R}(\mathrm{T}) : \mathbf{v} + \mathcal{N}(\mathrm{T}) \mapsto \mathrm{T}(\mathbf{v})$$

   (d) By extending a basis of U to V, show that for any subspace $U \leq V$ we have

   $$\dim {}^{V}\!/_{U} + \dim U = \dim V$$

   Hence conclude the rank–nullity theorem.

2.3 The Matrix Representation of a Linear Map

Recall that if β is a basis of a n-dimensional vector space V over $\mathbb{F}$, then any vector $\mathbf{v} \in V$ has a unique co-ordinate representation $[\mathbf{v}]_\beta \in \mathbb{F}^n$. The same thing can be done for a linear map, resulting in a tight relationship between linear maps, bases and matrices.

Example 2.22. The linear map on $\mathbb{R}^2$ defined by $\mathrm{T}\left(\begin{smallmatrix} x \\ y \end{smallmatrix}\right) = \left(\begin{smallmatrix} y \\ x \end{smallmatrix}\right)$ is plainly left-multiplication by the matrix $A = \left(\begin{smallmatrix} 0 & 1 \\ 1 & 0 \end{smallmatrix}\right)$. Otherwise said, $\mathrm{T} = \mathrm{L}_A$. Observe that the *columns* of A are the result of applying T to the standard basis $\epsilon = \{\mathbf{i}, \mathbf{j}\}$:

$$\mathrm{T}(\mathbf{i}) = \begin{pmatrix} 0 \\ 1 \end{pmatrix} \qquad \mathrm{T}(\mathbf{j}) = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$$

As the next result shows, given any linear map between finite dimensional spaces, choosing bases yields a representation of the map in terms of matrix multiplication.

Theorem 2.23 (Matrix representations). *Suppose that $\beta = \{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ and $\gamma = \{\mathbf{w}_1, \ldots, \mathbf{w}_m\}$ are bases of V and W respectively.*

1. *If $\mathrm{T} \in \mathcal{L}(V, W)$ then the matrix*

$$A = \Big([\mathrm{T}(\mathbf{v}_1)]_\gamma \ \cdots \ [\mathrm{T}(\mathbf{v}_n)]_\gamma \Big) \in M_{m\times n}(\mathbb{F}) \tag{$\dagger$}$$

   with j^{th} column $[\mathrm{T}(\mathbf{v}_j)]_\gamma$ is the unique matrix satisfying

$$\forall \mathbf{v} \in V, \quad [\mathrm{T}(\mathbf{v})]_\gamma = A[\mathbf{v}]_\beta \tag{$*$}$$

2. *Given $A \in M_{m\times n}(\mathbb{F})$, there is a unique linear map T satisfying $(*)$.*

Proof. 1. Suppose A is defined by $(\dagger)$ with ij^{th} entry $a_{ij} = \Big[[\mathrm{T}(\mathbf{v}_j)]_\gamma\Big]_i$, let $\mathbf{v} = b_1\mathbf{v}_1 + \cdots + b_n\mathbf{v}_n \in V$ be given, and compute: the column vector $A[\mathbf{v}]_\beta \in \mathbb{F}^m$ has i^{th} row

$$\begin{aligned} \Big(A[\mathbf{v}]_\beta\Big)_i &= \sum_{j=1}^n a_{ij}b_j = \sum_{j=1}^n \Big([\mathrm{T}(\mathbf{v}_j)]_\gamma\Big)_i b_j = \sum_{j=1}^n \Big(b_j[\mathrm{T}(\mathbf{v}_j)]_\gamma\Big)_i = \left(\sum_{j=1}^n [b_j\mathrm{T}(\mathbf{v}_j)]_\gamma\right)_i \\ &= \left(\left[\mathrm{T}\left(\sum_{j=1}^n b_j\mathbf{v}_j\right)\right]_\gamma\right)_i = \Big([\mathrm{T}(\mathbf{v})]_\gamma\Big)_i && \text{(linearity/Lemma 2.4)} \end{aligned}$$

   A therefore satisfies $(*)$. Conversely, if $\epsilon = \{\mathbf{e}_1, \ldots, \mathbf{e}_n\}$ is the standard basis of $\mathbb{F}^n$ and A satisfies $(*)$, then

$$[\mathrm{T}(\mathbf{v}_j)]_\gamma = A[\mathbf{v}_j]_\beta = A\mathbf{e}_j$$

   is the j^{th} column of A: this proves uniqueness.

2. The co-ordinate representation $[\mathrm{T}(\mathbf{v})]_\gamma$ is unique and so $(*)$ uniquely defines T. ∎

Definition 2.24. The matrix defined in (†) is the *matrix representation of* T *with respect to* β *and* γ: we write $A = [\mathrm{T}]_\beta^\gamma$. In the simplest case when $V = W$ and $\beta = \gamma$, we write $[\mathrm{T}]_\beta$.

The Theorem can be summarized by commutative diagrams: both options for travelling from V to $\mathbb{F}^m$ produce the same result.

$$\begin{tikzcd} V \arrow[r, "\mathrm{T}"] \arrow[d, "{[\ ]_\beta}"'] & W \arrow[d, "{[\ ]_\gamma}"] \\ \mathbb{F}^n \arrow[r, "A={[\mathrm{T}]_\beta^\gamma}"] & \mathbb{F}^m \end{tikzcd} \qquad \begin{tikzcd} \mathbf{v} \arrow[r, mapsto] \arrow[d, mapsto] & \mathrm{T}(\mathbf{v}) \arrow[d, mapsto] \\ {[\mathbf{v}]_\beta} \arrow[r, mapsto] & {[\mathrm{T}(\mathbf{v})]_\gamma = [\mathrm{T}]_\beta^\gamma[\mathbf{v}]_\beta} \end{tikzcd}$$

The big take-away is this:

$$\boxed{\text{Linear Map} \underset{\text{Bases}}{\overset{\text{Choose}}{\leftrightsquigarrow}} \text{Matrix Multiplication}}$$

More precisely, given bases of finite dimensional vector spaces, any linear map between them is equivalent to multiplication by a unique matrix.

Examples 2.25. 1. Recall the course introduction and the linear map T defined by 'rotate clockwise by 30° around the origin in $\mathbb{R}^2$.' With respect to the standard basis $\epsilon = \{\mathbf{i}, \mathbf{j}\}$, the matrix of T is

$$[\mathrm{T}]_\epsilon = \big([\mathrm{T}(\mathbf{i})]_\epsilon \quad \mathrm{T}[(\mathbf{j})]_\epsilon\big) = \begin{pmatrix} \frac{\sqrt{3}}{2} & \frac{1}{2} \\ -\frac{1}{2} & \frac{\sqrt{3}}{2} \end{pmatrix}$$

To rotate, say $\mathbf{v} = 2\mathbf{i} - 4\mathbf{j}$, we would compute

$$[\mathrm{T}(\mathbf{v})]_\epsilon = [\mathrm{T}]_\epsilon[\mathbf{v}]_\epsilon = \begin{pmatrix} \frac{\sqrt{3}}{2} & \frac{1}{2} \\ -\frac{1}{2} & \frac{\sqrt{3}}{2} \end{pmatrix} \begin{pmatrix} 2 \\ -4 \end{pmatrix} = \begin{pmatrix} \sqrt{3} - 2 \\ -1 - 2\sqrt{3} \end{pmatrix}$$

whence $\mathrm{T}(\mathbf{v}) = (\sqrt{3} - 2)\mathbf{i} + (-1 - 2\sqrt{3})\mathbf{j}$.

2. Recall Example 2.14.1. Let the standard bases of $P_3(\mathbb{R})$ and $P_2(\mathbb{R})$ be $\beta = \{1, x, x^2, x^3\}$ and $\gamma = \{1, x, x^2\}$. The matrix of $\mathrm{T} = \frac{\mathrm{d}}{\mathrm{d}x} \in \mathcal{L}(P_3(\mathbb{R}), P_2(\mathbb{R}))$ is then

$$[\mathrm{T}]_\beta^\gamma = \Big([\mathrm{T}(1)]_\gamma \ [\mathrm{T}(x)]_\gamma \ [\mathrm{T}(x^2)]_\gamma \ [\mathrm{T}(x^3)]_\gamma\Big) = \Big([0]_\gamma \ [1]_\gamma \ [2x]_\gamma \ [3x^2]_\gamma\Big) = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 3 \end{pmatrix}$$

For instance, compare the calculations

$$[\mathrm{T}]_\gamma^\beta \begin{pmatrix} 2 \\ 5 \\ 3 \\ 1 \end{pmatrix} = \begin{pmatrix} 5 \\ 6 \\ 3 \end{pmatrix} \leftrightsquigarrow \frac{\mathrm{d}}{\mathrm{d}x}(2 + 5x + 3x^2 + x^3) = 5 + 6x + 3x^2$$

3. Linear maps often look nice with respect to a sensible basis. For example, let $T \in \mathcal{L}(\mathbb{R}^2)$ is defined by 'reflect in the line $y = 2x$.'

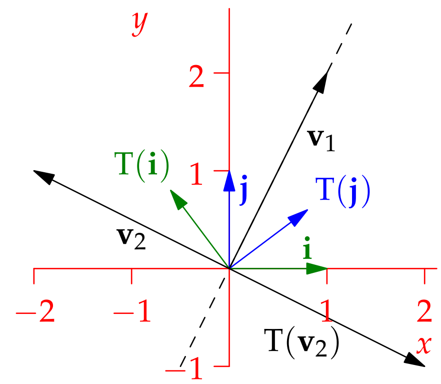
   The vectors $\mathbf{v}_1 = \binom{1}{2}$ and $\mathbf{v}_2 = \binom{-2}{1}$ satisfy

   $$T(\mathbf{v}_1) = \mathbf{v}_1 \qquad T(\mathbf{v}_2) = -\mathbf{v}_2$$

   Clearly $\beta = \{\mathbf{v}_1, \mathbf{v}_2\}$ is a basis of $\mathbb{R}^2$, with respect to which

   $$[T]_\beta = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$$

   This is much easier than finding the matrix with respect to the standard basis (exercise)

   $$[T]_\epsilon = \frac{1}{5}\begin{pmatrix} -3 & 4 \\ 4 & 3 \end{pmatrix}$$

   We will revisit this idea later: the vectors $\{\mathbf{v}_1, \mathbf{v}_2\}$ are *eigenvectors* for T; the matrix of a linear map with respect to an eigenbasis is always *diagonal*, with the *eigenvalues* down the diagonal.

4. Here is another example that simplifies nicely in terms of eigenvectors. Consider the linear map $T = L_A \in \mathcal{L}(\mathbb{R}^3)$ where

   $$A = \frac{1}{5}\begin{pmatrix} 4 & 0 & 2 \\ 0 & 5 & 0 \\ 2 & 0 & 1 \end{pmatrix}$$

   Plainly $A = [T]_\epsilon$ is the matrix of T with respect to the standard basis $\epsilon = \{\mathbf{i}, \mathbf{j}, \mathbf{k}\}$. Now consider

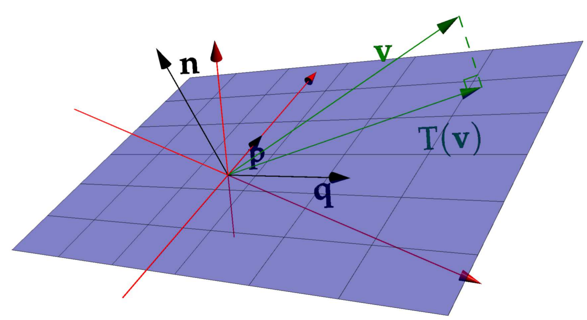
   $$\beta = \{\mathbf{n}, \mathbf{p}, \mathbf{q}\} = \left\{ \begin{pmatrix} -1 \\ 0 \\ 2 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 2 \\ 0 \\ 1 \end{pmatrix} \right\}$$

   It is easy to verify that β is linearly independent and thus a basis of $\mathbb{R}^3$. Moreover,

   $$T(\mathbf{n}) = \mathbf{0}, \qquad T(\mathbf{p}) = \mathbf{p}, \qquad T(\mathbf{q}) = \mathbf{q}$$

   from which the matrix is very simple

   $$[T]_\beta = \Big([T(\mathbf{n})]_\beta \ [T(\mathbf{p})]_\beta \ [T(\mathbf{q})]_\beta \Big)$$
   $$= \Big([\mathbf{0}]_\beta \ [\mathbf{p}]_\beta \ [\mathbf{q}]_\beta \Big) = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

   Since $\mathbf{p}$ and $\mathbf{q}$ are perpendicular ($\mathbf{p} \cdot \mathbf{n} = 0 = \mathbf{q} \cdot \mathbf{n}$) to $\mathbf{n}$, the matrix makes the physical interpretation of the linear map clear: T is the *orthogonal projection* onto the subspace $\text{Span}\{\mathbf{p}, \mathbf{q}\}$. It should also be from the map's interpretation as a projection that $\mathcal{N}(T) = \text{Span}\{\mathbf{n}\}$ and $\mathcal{R}(T) = \text{Span}\{\mathbf{p}, \mathbf{q}\}$.

Composition and Matrix Multiplication

It seems reasonable to expect that composition of linear maps corresponds to matrix multiplication. We merely have to be (very) careful with bases! Before seeing this, we engage in a little book-keeping.

Definition 2.26. The *identity matrix* $I_n \in M_n(\mathbb{F})$ has ij^{th} entry the *Kronecker delta symbol*

$$(I_n)_{ij} = \delta_{ij} = \begin{cases} 1 & \text{if } i = j \\ 0 & \text{if } i \neq j \end{cases}$$

Lemma 2.27. *Let V be an n-dimensional vector space with basis β and let $\mathrm{T} \in \mathcal{L}(V)$. Then*

$$[\mathrm{T}]_\beta = I_n \iff \mathrm{T} = \mathrm{I} \text{ is the identity map on } V$$

Proof. $(\Leftarrow)$ If $\mathrm{T} = \mathrm{I}$, then $\mathrm{T}(\mathbf{v}_i) = \mathbf{v}_i$ for each $\mathbf{v}_i \in \beta$. Plainly $[\mathrm{T}]_\beta = I_n$.
$(\Rightarrow)$ By the uniqueness of the matrix representation, $\mathrm{T} = \mathrm{I}$ is the only linear map with matrix I_n. ∎

Theorem 2.28. *Suppose $\mathrm{T} \in \mathcal{L}(V, W)$ and $\mathrm{U} \in \mathcal{L}(W, X)$ are linear maps and that V, W, X are finite-dimensional with bases β, γ, δ respectively. Then*

$$[\mathrm{UT}]_\beta^\delta = [\mathrm{U}]_\gamma^\delta [\mathrm{T}]_\beta^\gamma$$

In the common situation where $V = W = X$ and $\beta = \gamma = \delta$, this reduces to $[\mathrm{UT}]_\beta = [\mathrm{U}]_\beta [\mathrm{T}]_\beta$

Proof. Label the bases $\beta = \{\mathbf{v}_1, \ldots, \mathbf{v}_l\}$, $\gamma = \{\mathbf{w}_1, \ldots, \mathbf{w}_m\}$ and $\delta = \{\mathbf{x}_1, \ldots, \mathbf{x}_n\}$, and the matrices $A = [\mathrm{T}]_\beta^\gamma$, $B = [\mathrm{U}]_\gamma^\delta$ and $C = [\mathrm{UT}]_\beta^\delta$. Observe first

$$[\mathrm{T}(\mathbf{v}_k)]_\gamma = \begin{pmatrix} A_{1k} \\ \vdots \\ A_{mk} \end{pmatrix} \implies \mathrm{T}(\mathbf{v}_k) = \sum_{j=1}^m \mathbf{w}_j A_{jk} \qquad (k^{\text{th}} \text{ column of } A)$$

$$[\mathrm{U}(\mathbf{w}_j)]_\delta = \begin{pmatrix} B_{1j} \\ \vdots \\ B_{nj} \end{pmatrix} \implies \mathrm{U}(\mathbf{w}_j) = \sum_{i=1}^n \mathbf{x}_i B_{ij} \qquad (j^{\text{th}} \text{ column of } B)$$

$$[\mathrm{UT}(\mathbf{v}_k)]_\delta = \begin{pmatrix} C_{1k} \\ \vdots \\ C_{nk} \end{pmatrix} \implies \mathrm{UT}(\mathbf{v}_k) = \sum_{i=1}^n \mathbf{x}_i C_{ik} \qquad (k^{\text{th}} \text{ column of } C)$$

Now put it together:

$$\sum_{i=1}^n \mathbf{x}_i C_{ik} = \mathrm{UT}(\mathbf{v}_k) = \mathrm{U}\left(\sum_{j=1}^m \mathbf{w}_j A_{jk}\right) = \sum_{j=1}^m \mathrm{U}(\mathbf{w}_j) A_{jk} = \sum_{j=1}^m \sum_{i=1}^n \mathbf{x}_i B_{ij} A_{jk} = \sum_{i=1}^n \mathbf{x}_i \left(\sum_{j=1}^m B_{ij} A_{jk}\right)$$

Since $\delta = \{\mathbf{x}_1, \ldots, \mathbf{x}_n\}$ is a basis we conclude that

$$C_{ik} = \sum_{j=1}^m B_{ij} A_{jk}$$

Otherwise said, $C = BA$. ∎

Taking the special case where $U = T^{-1}$ and $\delta = \beta$, we instantly conclude:

Corollary 2.29. *Suppose* $T \in \mathcal{L}(V, W)$ *is a map between* n*-dimensional spaces* V, W *with bases* β, γ*. Then* T *is an isomorphism if and only if its matrix* $[T]_\beta^\gamma \in M_n(\mathbb{F})$ *is invertible. Moreover*

$$[T^{-1}]_\gamma^\beta = \left([T]_\beta^\gamma\right)^{-1}$$

Examples 2.30. 1. Recall (Example 2.25.1) that the matrix of 'rotate clockwise by 30°' with respect to the standard basis $\epsilon = \{\mathbf{i}, \mathbf{j}\}$ of $\mathbb{R}^2$ is

$$[T]_\epsilon = \begin{pmatrix} \frac{\sqrt{3}}{2} & \frac{1}{2} \\ -\frac{1}{2} & \frac{\sqrt{3}}{2} \end{pmatrix}$$

It follows that T^2 (rotate clockwise by 60°) and T^3 (90°) have matrices

$$[T^2]_\epsilon = [T]_\epsilon [T]_\epsilon = \begin{pmatrix} \frac{\sqrt{3}}{2} & \frac{1}{2} \\ -\frac{1}{2} & \frac{\sqrt{3}}{2} \end{pmatrix}^2 = \begin{pmatrix} \frac{1}{2} & \frac{\sqrt{3}}{2} \\ -\frac{\sqrt{3}}{2} & \frac{1}{2} \end{pmatrix} \quad \text{and} \quad [T^3]_\epsilon = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$$

Moreover, the inverse of T (namely 'rotate 30° *counter-clockwise*') has matrix

$$[T^{-1}]_\epsilon = [T]_\epsilon^{-1} = \begin{pmatrix} \frac{\sqrt{3}}{2} & \frac{1}{2} \\ -\frac{1}{2} & \frac{\sqrt{3}}{2} \end{pmatrix}^{-1} = \begin{pmatrix} \frac{\sqrt{3}}{2} & -\frac{1}{2} \\ \frac{1}{2} & \frac{\sqrt{3}}{2} \end{pmatrix}$$

2. Recall Example 2.25.4, and suppose $T \in \mathcal{L}(\mathbb{R}^3)$ is projection onto the plane perpendicular to $\mathbf{n} = -\mathbf{i} + 2\mathbf{k}$. Also let U be rotation by 60° clockwise around the $\mathbf{k}$-axis when viewed from above. With respect to the standard basis $\epsilon = \{\mathbf{i}, \mathbf{j}, \mathbf{k}\}$, and following the previous example,

$$[UT]_\beta = [U]_\beta [T]_\beta = \begin{pmatrix} \frac{1}{2} & \frac{\sqrt{3}}{2} & 0 \\ -\frac{\sqrt{3}}{2} & \frac{1}{2} & 0 \\ 0 & 0 & 1 \end{pmatrix} \cdot \begin{pmatrix} \frac{4}{5} & 0 & \frac{2}{5} \\ 0 & 1 & 0 \\ \frac{2}{5} & 0 & \frac{1}{5} \end{pmatrix} = \frac{1}{10} \begin{pmatrix} 4 & 5\sqrt{3} & 2 \\ -4\sqrt{3} & 5 & -2\sqrt{3} \\ 4 & 0 & 2 \end{pmatrix}$$

3. Let $\beta = \{e^{-x}\cos 2x, e^{-x}\sin 2x\}$, $V = \operatorname{Span}\beta$ and consider $T = \frac{\mathrm{d}}{\mathrm{d}x} \in \mathcal{L}(V)$. By computing

$$T(e^{-x}\cos 2x) = -e^{-x}\cos 2x - 2e^{-x}\sin 2x, \qquad T(e^{-x}\sin 2x) = 2e^{-x}\cos 2x - e^{-x}\sin 2x$$

we see that the matrix of T with respect to β, and its inverse are

$$[T]_\beta = \begin{pmatrix} -1 & 2 \\ -2 & -1 \end{pmatrix} \implies [T^{-1}]_\beta = [T]_\beta^{-1} = \frac{1}{5}\begin{pmatrix} -1 & -2 \\ 2 & -1 \end{pmatrix}$$

Since T^{-1} computes *anti-derivatives*, the upshot is that we can do this using linear algebra!

$$\int a e^{-x}\cos 2x + b e^{-x}\sin 2x \,\mathrm{d}x = -\frac{a+2b}{5}e^{-x}\cos 2x + \frac{2a-b}{5}e^{-x}\sin 2x$$

Of course if you *really* prefer integration by parts...

A final bit of book-keeping: co-ordinate isomorphisms and matrices

Two colloquialisms are sometimes uttered in an attempt to summarize or simplify linear algebra.

1. Every vector space is really $\mathbb{F}^n$ in disguise.
2. Every linear map is really matrix multiplication in disguise.

While strictly incorrect, we can make these statements precise, at least under the imposition of two caveats: *when dimensions are finite* and *after choosing bases.*

Corollary 2.31. 1. *Suppose β is a basis of V and that $\dim_{\mathbb{F}} V = n$. Then the co-ordinate representation ϕ_β is an isomorphism*

$$\phi_\beta : V \to \mathbb{F}^n : \mathbf{v} \mapsto [\mathbf{v}]_\beta$$

2. *Additionally, suppose γ is a basis of W and that $\dim_{\mathbb{F}} W = m$. Then the vector space of linear maps $\mathcal{L}(V, W)$ is isomorphic to the space of matrices $M_{m\times n}(\mathbb{F})$ via the isomorphism*

$$\Phi_\beta^\gamma : \mathcal{L}(V, W) \to M_{m\times n}(\mathbb{F}) : \mathrm{T} \mapsto [\mathrm{T}]_\beta^\gamma$$

This is really just the finite-dimensional version of Corollary 2.20, but with explicit isomorphisms.

Proof. The co-ordinate representation ϕ_β is linear: if $\{\mathbf{e}_1, \dots, \mathbf{e}_n\} \subseteq \mathbb{F}^n$ is the standard basis,

$$\phi_\beta(a_1\mathbf{v}_1 + \cdots + a_n\mathbf{v}_n) = \begin{pmatrix} a_1 \\ \vdots \\ a_n \end{pmatrix} = a_1\mathbf{e}_1 + \cdots + a_n\mathbf{e}_n = a_1\phi_\beta(\mathbf{v}_1) + \cdots + a_n\phi_\beta(\mathbf{v}_n)$$

Since $\mathcal{N}(\phi_\beta) = \{\mathbf{0}_V\}$, it is also plainly injective and thus an isomorphism.
Part 2 is similar. ∎

A peculiar difficulty with this discussion is that it can be hard to disentangle a *matrix* $A \in M_{m\times n}(\mathbb{F})$ from its associated *linear map* $\mathrm{L}_A : \mathbb{F}^n \to \mathbb{F}^m$. For reference, we summarize everything here.

Corollary 2.32. *Let $A, B \in M_{m\times n}(\mathbb{F})$.*

1. *If β, γ are the standard bases, then $[\mathrm{L}_A]_\beta^\gamma = A$*
2. $\mathrm{L}_A = \mathrm{L}_B \iff A = B$
3. $\mathrm{L}_{A+B} = \mathrm{L}_A + \mathrm{L}_B$ *and* $\mathrm{L}_{\lambda A} = \lambda \mathrm{L}_A$ *for all* $\lambda \in \mathbb{F}$
4. *If $\mathrm{T} \in \mathcal{L}(\mathbb{F}^n, \mathbb{F}^m)$, then there is a unique $C \in M_{m\times n}(\mathbb{F})$ such that $\mathrm{T} = \mathrm{L}_C$*
5. *If $E \in M_{n\times p}(\mathbb{F})$, then $\mathrm{L}_{AE} = \mathrm{L}_A\mathrm{L}_E$*
6. *If $m = n$, then $\mathrm{L}_{I_n} = \mathrm{I}$*

Everything should be straightforward to prove given our previous results.

Exercises 2.3 1. Let $\mathrm{T} = \mathrm{L}_A \in \mathcal{L}(\mathbb{R}^2)$ be left-multiplication by $A = \left(\begin{smallmatrix} 1 & -2 \\ 3 & -1 \end{smallmatrix}\right)$.

   (a) Find the matrix $[\mathrm{T}]_\beta$ with respect to the basis $\beta = \{\left(\begin{smallmatrix} 1 \\ 1 \end{smallmatrix}\right), \left(\begin{smallmatrix} 2 \\ 3 \end{smallmatrix}\right)\}$.

   (b) Compute $\mathrm{T}(3\mathbf{i} + 4\mathbf{j})$ in two different ways, and make sure your answers agree!

2. Consider the linear map $\mathrm{T} = \frac{\mathrm{d}}{\mathrm{d}x} \in \mathcal{L}(P_3(\mathbb{R}), P_2(\mathbb{R}))$.

   (a) Compute $[\mathrm{T}]_\beta^\gamma$ with respect to the bases $\beta = \{1, x, x^2, x^3\}$ and $\gamma = \{1 - x, 1 + x, x^2 - 1\}$.

   (b) Verify that multiplication by $[\mathrm{T}]_\beta^\gamma$ correctly computes the derivative of the polynomial $p(x) = 2 + 5x + 3x^2 + x^3$.

   (c) Let $\mathrm{U} \in \mathcal{L}(P_2(\mathbb{R}), P_1(\mathbb{R}))$ also be 'differentiate,' so that $\mathrm{UT} = \frac{\mathrm{d}^2}{\mathrm{d}x^2}$ is the second-derivative. Compute the matrices of T, U and check that $[\mathrm{UT}]_\beta^\delta = [\mathrm{U}]_\gamma^\delta [\mathrm{T}]_\beta^\gamma$ when:

      i. β, γ and δ are the standard bases of $P_3(\mathbb{R})$, $P_2(\mathbb{R})$ and $P_1(\mathbb{R})$ respectively.

      ii. $\beta = \{1, x, x^2, x^3\}$, $\gamma = \{1 - x, 1 + x, x^2 - 1\}$ and $\delta = \{1, x\}$

3. Define $\mathrm{T} \in \mathcal{L}(P_3(\mathbb{R}), P_4(\mathbb{R}))$ by $\mathrm{T}(g)(x) = (x-1)g(x)$, let $f(x) = x + 2x^2 - 3x^3$ and suppose β, γ are the standard bases.

   (a) Compute $[f]_\beta$ and $[(x-1)f]_\gamma$.

   (b) Compute the matrix $[\mathrm{T}]_\beta^\gamma$ and check explicitly that $[(x-1)f]_\gamma = [\mathrm{T}]_\beta^\gamma [f]_\beta$.

4. Let $\mathrm{T} : P_3(\mathbb{R}) \to P_3(\mathbb{R})$ be the function defined by

$$\mathrm{T}(f)(x) = 2\int_0^1 f(t)\,\mathrm{d}t - 3\frac{\mathrm{d}}{\mathrm{d}x} f(x)$$

   (a) Give a short argument to justify the fact that T is linear.

   (b) Compute the matrix $[\mathrm{T}]_\beta$ of T with respect to the standard basis β of $P_3(\mathbb{R})$.

   (c) Find an explicit expression for the linear map $\mathrm{T}^2 \in \mathcal{L}(P_3(\mathbb{R}))$; that is, express $\mathrm{T}^2(f)(x)$ in terms of the integral and derivatives of $f(x)$.

   (d) Compute $[\mathrm{T}^2]_\beta$ using part (a), and check that it equals $[\mathrm{T}]_\beta^2$.

5. (*Recall Example 2.25.3*) Let $\mathrm{T} \in \mathcal{L}(\mathbb{R}^2)$ be the linear map 'reflect across the line $y = 2x$.' With respect to the standard basis, show that its matrix is $[\mathrm{T}]_\epsilon = \frac{1}{5}\left(\begin{smallmatrix} -3 & 4 \\ 4 & 3 \end{smallmatrix}\right)$.

6. Find a basis of $\mathbb{R}^2$ with respect to which the linear map 'reflect across the line $x + 3y = 0$' has a diagonal matrix. Now find the matrix of this map with respect to the standard basis.

7. Let B be a fixed invertible $n \times n$ matrix. Prove that the following map is an isomorphism:

$$\Psi : M_n(\mathbb{F}) \to M_n(\mathbb{F}) : A \mapsto B^{-1}AB$$

8. Compute the integral $\int (2x - 3x^2)e^{3x}\,\mathrm{d}x$ *without* using integration by parts.

   (*Hint: Let $\beta = \{e^{3x}, xe^{3x}, x^2e^{3x}\}$ and invert the matrix of $\frac{\mathrm{d}}{\mathrm{d}x}$ with respect to β…*)

9. Give explicit proofs of Corollary 2.32 parts 1 & 5.

10. In the context of Corollary 2.31; suppose $\dim_{\mathbb{F}} V = n$ and that $\mathrm{T} : V \to \mathbb{F}^n$ is an isomorphism. Prove that $\mathrm{T} = \phi_\beta$ for some basis β.

2.4 The Change of Co-ordinate Matrix

Suppose V is finite-dimensional over $\mathbb{F}$ with distinct bases β, ϵ. We know from Corollary 2.31 that the co-ordinate maps are isomorphisms $V \to \mathbb{F}^n$:

$$\phi_\beta(\mathbf{v}) = [\mathbf{v}]_\beta, \qquad \phi_\epsilon(\mathbf{v}) = [\mathbf{v}]_\epsilon$$

Since inverses and compositions of isomorphisms are also isomorphisms, it follows that

$$\phi_\epsilon \circ \phi_\beta^{-1} : \mathbb{F}^n \to \mathbb{F}^n : [\mathbf{v}]_\beta \mapsto [\mathbf{v}]_\epsilon$$

is an isomorphism. Corollaries 2.32 and 2.31 force this isomorphism to be left-multiplication by an invertible matrix:

$$\exists Q_\beta^\epsilon \in M_n(\mathbb{F}) \text{ such that } \forall \mathbf{v} \in V,\ [\mathbf{v}]_\epsilon = \phi_\epsilon \circ \phi_\beta^{-1}(\mathbf{v}) = Q_\beta^\epsilon [\mathbf{v}]_\beta \tag{$*$}$$

Definition 2.33. Q_β^ϵ is the *change of co-ordinate matrix* from β to ϵ.

Indeed $(*)$ makes it obvious how to compute: if $\beta = \{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ then $[\mathbf{v}_j]_\beta = \mathbf{e}_j$ is the j^{th} standard basis (column) vector of $\mathbb{F}^n$, and so...

Lemma 2.34. *The change of co-ordinate matrix from β to ϵ is the matrix of the identity linear map with respect to these bases: if $\beta = \{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$, then,*

$$Q_\beta^\epsilon = [\mathrm{I}_V]_\beta^\epsilon = \begin{pmatrix} [\mathbf{v}_1]_\epsilon & \cdots & [\mathbf{v}_n]_\epsilon \end{pmatrix}$$

It follows immediately that $Q_\epsilon^\beta = (Q_\beta^\epsilon)^{-1}$ and that $Q_\epsilon^\gamma Q_\beta^\epsilon = Q_\beta^\gamma$.

We are prioritizing ϵ in the above notation because, in many situations, one of the bases is a standard basis. In such a case, one can simply state Q_β^ϵ and invert to obtain Q_ϵ^β, as the next example illustrates.

Example 2.35. Consider the basis $\beta = \{1 - 3x, 2 + 5x\}$ of $P_1(\mathbb{R})$ and let $\epsilon = \{1, x\}$ be the standard basis. Then

$$Q_\beta^\epsilon = [\mathrm{I}_{P_1(\mathbb{R})}]_\beta^\epsilon = \begin{pmatrix} [1-3x]_\epsilon & [2+5x]_\epsilon \end{pmatrix} = \begin{pmatrix} 1 & 2 \\ -3 & 5 \end{pmatrix} \implies Q_\epsilon^\beta = \begin{pmatrix} 1 & 2 \\ -3 & 5 \end{pmatrix}^{-1} = \frac{1}{11}\begin{pmatrix} 5 & -2 \\ 3 & 1 \end{pmatrix}$$

To write, for instance $p(x) = 3 - x$ in terms of β, compute

$$[p]_\beta = Q_\epsilon^\beta [p]_\epsilon = \frac{1}{11}\begin{pmatrix} 5 & -2 \\ 3 & 1 \end{pmatrix}\begin{pmatrix} 3 \\ -1 \end{pmatrix} = \frac{1}{11}\begin{pmatrix} 17 \\ 8 \end{pmatrix}$$
$$\implies 3 - x = \frac{17}{11}(1 - 3x) + \frac{8}{11}(2 + 5x)$$

While this approach doesn't save any time for a single calculation, it is much more efficient when one needs to convert many vectors to another basis. It is important to remember that the change of co-ordinate matrix merely tells you how the *co-ordinates* of a vector $\mathbf{v} \in V$ change when a basis changes: nothing happens to $\mathbf{v}$ itself!

Example 2.36. Here is a 3-dimensional example. Consider the basis $\beta = \{1+x, 2-x^2, 4-x^2\}$ of $P_2(\mathbb{R})$ and let $\epsilon = \{1, x, x^2\}$ be the standard basis. Then

$$Q_\beta^\epsilon = \Big([1+x]_\epsilon \; [2-x^2]_\epsilon \; [4-x^2]_\epsilon\Big) = \begin{pmatrix} 1 & 2 & 4 \\ 1 & 0 & 0 \\ 0 & -1 & -1 \end{pmatrix} \implies Q_\epsilon^\beta = \frac{1}{2}\begin{pmatrix} 0 & 2 & 0 \\ -1 & 1 & -4 \\ 1 & -1 & 2 \end{pmatrix}$$

To check that this makes sense, we check the co-ordinate representation of, say, $p(x) = 2 + 3x + 4x^2$ with respect to β:

$$[p]_\beta = Q_\epsilon^\beta [p]_\epsilon = \frac{1}{2}\begin{pmatrix} 0 & 2 & 0 \\ -1 & 1 & -4 \\ 1 & -1 & 2 \end{pmatrix}\begin{pmatrix} 2 \\ 3 \\ 4 \end{pmatrix} = \begin{pmatrix} 3 \\ -\frac{15}{2} \\ \frac{7}{2} \end{pmatrix}$$

$$p(x) = 3(1+x) - \frac{15}{2}(2-x^2) + \frac{7}{2}(4-x^2)$$

which is easily verified to by multiplying out. Of course, all this is predicated on being willing to invert a 3×3 matrix!

This process can be combined with matrix representations of linear maps.

Theorem 2.37. *Let* $\mathrm{T} \in \mathcal{L}(V)$ *where* V *has finite bases* ϵ *and* β*. Then the matrices of* T *satisfy*

$$[\mathrm{T}]_\beta = Q_\epsilon^\beta [\mathrm{T}]_\epsilon Q_\beta^\epsilon = Q_\epsilon^\beta [\mathrm{T}]_\epsilon (Q_\epsilon^\beta)^{-1}$$

Proof. Simply apply the right hand side to the representation of any vector $\mathbf{v} \in V$ with respect to β:

$$Q_\epsilon^\beta [\mathrm{T}]_\epsilon Q_\beta^\epsilon [\mathbf{v}]_\beta = Q_\epsilon^\beta [\mathrm{T}]_\epsilon [\mathbf{v}]_\epsilon = Q_\epsilon^\beta [\mathrm{T}(\mathbf{v})]_\epsilon = [\mathrm{T}(\mathbf{v})]_\beta = [\mathrm{T}]_\beta [\mathbf{v}]_\beta$$

∎

The matrices of a linear map with respect to different bases are therefore *similar/conjugate.*

Example 2.38. We revisit Example 2.25.3 in this language. Let $\mathrm{T} : \mathbb{R}^2 \to \mathbb{R}^2$ be reflection in the line $y = 2x$, let $\epsilon = \{\mathbf{i}, \mathbf{j}\}$ be the standard basis and $\beta = \{\mathbf{v}_1, \mathbf{v}_2\} = \{\mathbf{i} + 2\mathbf{j}, -2\mathbf{i} + \mathbf{j}\}$ be chosen to point parallel/perpendicular to the line of reflection. Since $\mathrm{T}(\mathbf{v}_1) = \mathbf{v}_1$ and $\mathrm{T}(\mathbf{v}_2) = -\mathbf{v}_2$ we saw that

$$[\mathrm{T}]_\beta = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$$

The change of co-ordinate matrices are then

$$Q_\beta^\epsilon = [\mathrm{I}_{\mathbb{R}^2}]_\beta^\epsilon = ([\mathbf{v}_1]_\epsilon \; [\mathbf{v}_2]_\epsilon) = \begin{pmatrix} 1 & -2 \\ 2 & 1 \end{pmatrix}, \qquad Q_\epsilon^\beta = (Q_\beta^\epsilon)^{-1} = \frac{1}{5}\begin{pmatrix} 1 & 2 \\ -2 & 1 \end{pmatrix}$$

The matrix of T with respect to the standard basis ϵ is therefore

$$[\mathrm{T}]_\epsilon = Q_\beta^\epsilon [\mathrm{T}]_\beta Q_\epsilon^\beta = \begin{pmatrix} 1 & -2 \\ 2 & 1 \end{pmatrix}\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \cdot \frac{1}{5}\begin{pmatrix} 1 & 2 \\ -2 & 1 \end{pmatrix} = \frac{1}{5}\begin{pmatrix} -3 & 4 \\ 4 & 3 \end{pmatrix}$$

as we recovered earlier.

Change of basis in general (non-examinable)

The discussion generalizes to linear maps $\mathrm{T} \in \mathcal{L}(V, W)$ where we change bases of *both spaces.*

Theorem 2.39. *Suppose $\mathrm{T} \in \mathcal{L}(V, W)$, that V has bases β, ϵ, and W has bases γ, δ, where $\dim V = n$ and $\dim W = m$. Then*

$$[\mathrm{T}]_\beta^\gamma = Q_\delta^\gamma [\mathrm{T}]_\epsilon^\delta Q_\beta^\epsilon$$

where $Q_\beta^\epsilon \in M_n(\mathbb{F})$ and $Q_\delta^\gamma \in M_m(\mathbb{F})$ are change of co-ordinate matrices. The relationships between these objects is summarized in the picture:

$$\begin{tikzcd}
 & {[\mathbf{v}]_\epsilon \in \mathbb{F}^n} \arrow[r, "{[\mathrm{T}]_\epsilon^\delta}"'] & {[\mathrm{T}(\mathbf{v})]_\delta \in \mathbb{F}^m} \arrow[dd, dashed, bend right, "Q_\delta^\gamma"'] & \\
\mathbf{v} \in V \arrow[ur, "\phi_\epsilon"] \arrow[dr, "\phi_\beta"'] \arrow[rrr, "\mathrm{T}"'] & & & \mathrm{T}(\mathbf{v}) \in W \arrow[ul, "\phi_\delta"'] \arrow[dl, "\phi_\gamma"] \\
 & {[\mathbf{v}]_\beta \in \mathbb{F}^n} \arrow[uu, dashed, bend right, "Q_\beta^\epsilon"'] \arrow[r, "{[\mathrm{T}]_\beta^\gamma}"'] & {[\mathrm{T}(\mathbf{v})]_\gamma \in \mathbb{F}^m} &
\end{tikzcd}$$

Example 2.40. With respect to the standard bases $\epsilon = \{1, x, x^2, x^3\}$ and $\delta = \{1, x, x^2\}$, the derivative operator $\mathrm{T} = \frac{\mathrm{d}}{\mathrm{d}x} \in \mathcal{L}(P_3(\mathbb{R}), P_2(\mathbb{R}))$ has matrix

$$[\mathrm{T}]_\epsilon^\delta = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 3 \end{pmatrix}$$

Consider new bases $\beta = \{1 + x, 1 - x, 2x + x^2, x^3 - 1\}$ and $\gamma = \{1 - x, 2 + x^2, x\}$ of $P_3(\mathbb{R})$ and $P_2(\mathbb{R})$ respectively. The matrix of T with respect to β and γ is then

$$[\mathrm{T}]_\beta^\gamma = (Q_\gamma^\delta)^{-1} [\mathrm{T}]_\epsilon^\delta Q_\beta^\epsilon = \begin{pmatrix} 1 & 2 & 0 \\ -1 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix}^{-1} \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 3 \end{pmatrix} \begin{pmatrix} 1 & 1 & 0 & -1 \\ 1 & -1 & 2 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} = \begin{pmatrix} 1 & -1 & 2 & -6 \\ 0 & 0 & 0 & 3 \\ 1 & -1 & 4 & -6 \end{pmatrix}$$

We can check this on an example: written with respect to β, let

$$p(x) = 3(1 + x) + 2(1 - x) - 4(2x + x^2) + 5(x^3 - 1)$$

$$\implies [p']_\gamma = [\mathrm{T}]_\beta^\gamma [p]_\beta = \begin{pmatrix} 1 & -1 & 2 & -6 \\ 0 & 0 & 0 & 3 \\ 1 & -1 & 4 & -6 \end{pmatrix} \begin{pmatrix} 3 \\ 2 \\ -4 \\ 5 \end{pmatrix} = \begin{pmatrix} -37 \\ 15 \\ -45 \end{pmatrix}$$

which comports with

$$p'(x) = 3 - 2 - 4(2 + 2x) + 15x^2 = -37(1 - x) + 15(2 + x^2) - 45x$$

Exercises 2.4 1. Let $\epsilon = \{\mathbf{i}, \mathbf{j}\}$ be the standard basis of $\mathbb{R}^2$, and consider two further bases

$$\beta = \{-\mathbf{i} + 2\mathbf{j},\ 2\mathbf{i} - \mathbf{j}\}, \qquad \gamma = \{2\mathbf{i} + 5\mathbf{j},\ -\mathbf{i} - 3\mathbf{j}\}$$

Find the change of co-ordinate matrices Q_ϵ^β, Q_ϵ^γ and Q_β^γ.

2. Let $\epsilon = \{1, x, x^2\}$ and $\beta = \{1 - x, x + x^2, x^2 - 1\}$. Find the change of co-ordinate matrix Q_ϵ^β for $P_2(\mathbb{F})$. Check your answer by finding constants a, b, c such that

$$2 + 7x - 4x^2 = a(1 - x) + b(x + x^2) + c(x^2 - 1)$$

3. For each matrix A and basis β, find $[\mathrm{L}_A]_\beta$ and an invertible matrix Q such that $[\mathrm{L}_A]_\beta = Q^{-1}AQ$.

   (a) $A = \begin{pmatrix} 1 & 2 \\ 2 & 1 \end{pmatrix}$ and $\beta = \left\{ \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ -1 \end{pmatrix} \right\}$

   (b) $A = \begin{pmatrix} 13 & 1 & 4 \\ 1 & 13 & 4 \\ 4 & 4 & 10 \end{pmatrix}$ and $\beta = \left\{ \begin{pmatrix} 1 \\ 1 \\ -2 \end{pmatrix}, \begin{pmatrix} 1 \\ -1 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} \right\}$

4. Recall the *trace* of a $n \times n$ matrix: $\operatorname{tr} C = \sum_{j=1}^{n} c_{jj}$.

   (a) Prove that $\operatorname{tr} AB = \operatorname{tr} BA$, provided both AB and BA are square.

   (b) Prove that if A and B are similar matrices ($B = Q^{-1}AQ$ for some Q), then $\operatorname{tr} A = \operatorname{tr} B$.
   (*The matrices of a linear map with respect to any two bases therefore always have the same trace*)

5. Suppose $\beta = \{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ is a basis of $\mathbb{F}^n$ and that $\epsilon = \{\mathbf{e}_1, \ldots, \mathbf{e}_n\}$ is the standard basis. Prove that $Q_\epsilon^\beta \mathbf{v}_k = \mathbf{e}_k$ for each k.

   (*In this context, Q_ϵ^β is sometimes called a* change of basis *matrix, though this only makes sense in* $\mathbb{F}^n$)

6. Let R reflect in the line through the origin making angle θ with the positive x-axis in $\mathbb{R}^2$.

   (a) As in Example 2.38, use a change of co-ordinate matrix to find the matrix of R with respect to the standard basis $\epsilon = \{\mathbf{i}, \mathbf{j}\}$.

   (b) If the line of reflection has gradient m, state $[\mathrm{R}]_\epsilon$ in terms of m. When m is a rational number, what does this have to do with Pythagorean triples?

7. Let c and s be constants and consider the change of co-ordinates

$$\begin{cases} x = cu + sv \\ y = -su + cv \end{cases} \tag{$*$}$$

   That is, if $\mathbf{x} = [\mathbf{x}]_\epsilon = \binom{x}{y}$ is viewed with respect to $\epsilon = \{\mathbf{i}, \mathbf{j}\}$, then $[\mathbf{x}]_\beta = \binom{u}{v}$ with respect to some new basis β.

   (a) Find β.

   (b) The curve

$$7x^2 - 6\sqrt{3}xy + 13y^2 = 16$$

   represents a conic in the plane. Assume that $c = \cos\theta$ and $s = \sin\theta$ for some unknown angle θ. Substitute, using $(*)$, for u and v in order to find a value of $\theta \in [0, 90°)$ for which the conic has no uv-term.

   Use your understanding of the basis β and the resulting change of co-ordinates to sketch the original conic.