Tutorial Questions

Question 1.

Let $X = \{a\}$ *and* $Y = \{b, c\}$.

a. List all possible functions from X *to* Y *(you may draw a diagram).*

b. List all functions from Y *to* X.

c. Compose a function from the first list with a function from the second list (in either order). What are all the different functions that arise in this way?

Solution. a. To specify a function $f\colon X \to Y$, we need to describe the image of each element of X. Since X has only one element a, any function $f\colon X \to Y$ is uniquely determined by $f(a) \in Y = \{b, c\}$. That is, there are exactly two functions f_1 and f_2 in Y^X, which are defined by $f_1(a) = b$and $f_2(a) = c$.

b. To define a function $g\colon Y \to X$ is equivalent to choose the value of $g(y) \in X$ for all $y \in Y$. Since $X = \{a\}$ has only one element, we are forced to choose $g(y) = a$ for all $y \in Y$. So there exists a unique function $g\colon Y \to X$, defined by $g(b) = g(c) = a$.

c. Let us give the details for $f_1 \circ g$:

$$\begin{array}{rcccl} f_1 \circ g\colon & Y & \xrightarrow{g} & X & \xrightarrow{f_1} Y \\ & b & \longmapsto & a & \longmapsto b \\ & c & \longmapsto & a & \longmapsto b. \end{array}$$

So $f_1 \circ g$ is the constant function b. That is: $(f_1 \circ g)(y) = b$ for all $y \in Y$. Similarly, $f_2 \circ g$ is the constant function c. Composing in the other direction we obtain that both $g \circ f_1$ and $g \circ f_2$ are the constant function a.

Observe that we have $g \circ f_1 = g \circ f_2 = \mathrm{Id}_X$ with $f_1 \neq f_2$, but $f_1 \circ g \neq f_2 \circ g$ and none of them is equal to Id_Y.

■

Question 2.

Let X, Y, Z *be sets and suppose that* $f\colon X \to Y$ *and* $g\colon Y \to Z$ *are functions.*

a. Show that if f *and* g *are injective then* $g \circ f$ *is injective.*

b. Show that if f *and* g *are surjective then* $g \circ f$ *is surjective.*

c. If $g \circ f$ *is injective then what can you say about* f *or* g*? What if* $g \circ f$ *is surjective?*

Solution. a. Suppose that both f and g are injective and let $a, b \in X$ be two elements of X such that $(g \circ f)(a) = (g \circ f)(b)$. We need to prove that in this case one necessary have $a = b$. From $(g \circ f)(a) = (g \circ f)(b)$ one obtain $g(f(a)) = g(f(b))$ and so $f(a) = f(b)$ by injectivity of g and finally $a = b$ by injectivity of f.

b. Suppose that both f and g are surjective and let $z \in Z$ be any element of Z. By surjectivity of g, there exists $y \in Y$ such that $g(y) = z$. By surjectivity of f, there exists $x \in X$ such that $f(x) = y$. But then $(g \circ f)(x) = g(y) = z$, proving surjectivity of $g \circ f$.

c. We will show that $g \circ f$ injective implies f injective, but g might not be injective; and that $g \circ f$ surjective implies g surjective, but f might not be surjective.

We start by providing a counterexample to the two incorrect implications. Let f_1 and g be the functions from the solution to Exercise 1.a. Then $g \circ f_1 = \mathrm{Id}_X$ is both injective and surjective. However, g is not injective as $g(b) = g(c)$ even if $b \neq c$, and f_1 is not surjective as $c \notin \mathrm{Im}(f_1)$.

Suppose now that $g \circ f$ is injective. Let $a, b \in X$ be any two elements with $f(a) = f(b)$. By applying g on both sides we have $(g \circ f)(a) = (g \circ f)(b)$ and hence $a = b$ by injectivity of $g \circ f$. We have just proved injectivity of f.

Finally, suppose that $g \circ f$ is surjective. Let $z \in Z$ be any element. Then, by surjectivity of $g \circ f$, there exists $x \in X$ such that $z = (g \circ f)(x) = g(f(x))$. Since $y := f(x)$ is in Y, we have $z \in \mathrm{Im}(g)$, proving surjectivity of g.

■

Question 3.

Recall that $\mathbf{F}$ *stands for* $\mathbf{R}$ *or* $\mathbf{C}$*. Let* m *be an integer. Prove that* $\mathbf{F}^m$ *satisfies the* distributive properties*, that is, for all* $a, b \in \mathbf{F}$ *and* $u, v \in \mathbf{F}^m$ *we have:*

1. $a(u + v) = au + av$;
2. $(a + b)v = av + bv$.

Solution. Let $u = [u_1, \dots, u_m]^\intercal$ and $v = [v_1, \dots, v_m]^\intercal$. Then

$$\begin{aligned} a(u+v) &= a\left(\begin{bmatrix} u_1 \\ \vdots \\ u_m \end{bmatrix} + \begin{bmatrix} u_1 \\ \vdots \\ u_m \end{bmatrix}\right) = a\begin{bmatrix} u_1 + v_1 \\ \vdots \\ u_m + v_m \end{bmatrix} \\ &= \begin{bmatrix} a(u_1 + v_1) \\ \vdots \\ a(u_m + v_m) \end{bmatrix} = \begin{bmatrix} au_1 + av_1 \\ \vdots \\ au_m + av_m \end{bmatrix} \\ &= \begin{bmatrix} au_1 \\ \vdots \\ au_m \end{bmatrix} + \begin{bmatrix} av_1 \\ \vdots \\ av_m \end{bmatrix} = au + av, \end{aligned}$$

where we used the distributivity property $a(u_1+v_1)=au_1+av_1$ in $\mathbf{F}$. ■

Recall that the **identity map** on a set X is the function $\mathrm{Id}_X\colon X\to X$ given by $\mathrm{Id}_X(x)=x$ for all $x\in X$.

Question 4.

Let $f\colon X\to Y$ and $g\colon Y\to X$ be functions. Suppose that $g\circ f=\mathrm{Id}_X$. Does it follow that $f\circ g=\mathrm{Id}_Y$?

Solution. No. See the solution of Question 1 for a counterexample. ■

Further Questions

Question 5.

Let X and Y be finite sets and $f\colon X\to Y$ be a function.

a. *Assume that f is injective. Find a function $g\colon Y\to X$ such that $g\circ f=\mathrm{Id}_X$.*

b. *Assume that f is surjective. Find a function $g\colon Y\to X$ such that $f\circ g=\mathrm{Id}_Y$.*

Solution. a. Suppose X non-empty and let x_0 be any element of X. Define $g\colon Y\to X$ by:

$$g(y):=\begin{cases}\text{the unique } x\in X \text{ such that } f(x)=y & \text{if } y\in\mathrm{Im}(f)\\ x_0 & \text{otherwise.}\end{cases}$$

It is then straightforward that $g\circ f=\mathrm{Id}_X$.

To go further

If $X=\varnothing$ there is a unique function from X to Y: the empty function. So if $X=\varnothing=Y$, the empty function is its own inverse. If Y is not empty, then there is no function from Y to $\varnothing$.

Observe that we did not use that X and Y are finite. That is, we actually proved a bit more than what was stated. Moreover, one can check that the g we constructed is surjective. So we have proved: if $f\colon X\to Y$ is an injective function between arbitrary sets, there exists a surjective function $g\colon Y\to X$ with $g\circ f=\mathrm{Id}_X$.

b. If f is surjective, then $Y=\mathrm{Im}(f)$. One can define $g\colon Y\to X$ by for each $y\in Y=\mathrm{Im}(f)$ choosing a $x\in X$ such that $f(x)=y$. One can check that such a g is injective and $f\circ g=\mathrm{Id}_Y$. So we have proved: if $f\colon X\to Y$ is an surjective function with Y finite, there exists an injective function $g\colon Y\to X$ with $f\circ g=\mathrm{Id}_Y$.

To go further

If Y is infinite, we need the axiom of choice (AC) to be able to make infinitely many simultaneous choices. So, if we assume (AC) we have: if $f\colon X \to Y$ is an surjective function between arbitrary sets, there exists an injective function $g\colon Y \to X$ with $f \circ g = \mathrm{Id}_Y$. Actually, this statement is equivalent to (AC). This is one of the reason to assume (AC): it implies that injective and surjective functions behave "dually".

■

The following questions are intended as revision for Year 1 Linear Algebra.

Question 6.

Solve the system of linear equations

$$\begin{aligned} x_1 + 3x_2 - 5x_3 &= 4 \\ x_1 + 4x_2 - 8x_3 &= 7 \\ -3x_1 - 7x_2 + 9x_3 &= -6. \end{aligned}$$

Solution. By doing the row transformations $r_2 \mapsto r_2 - r_1$ and $r_3 \mapsto r_3 + 3r_1$ we obtain the equivalent system

$$\begin{aligned} x_1 + 3x_2 - 5x_3 &= 4 \\ x_2 - 3x_3 &= 3 \\ 2x_2 - 6x_3 &= 6. \end{aligned}$$

Doing one last row transformation $r_3 \mapsto r_3 - 2r_2$ one have the, also equivalent, system

$$\begin{aligned} x_1 + 3x_2 - 5x_3 &= 4 \\ x_2 - 3x_3 &= 3 \\ 0 &= 0. \end{aligned}$$

We conclude that x_3 is free, $x_2 = 3 + 3x_3$ and $x_1 = 4 - 3x_2 + 5x_3 = -5 - 4x_3$. ■

Question 7.

Let $A = \begin{bmatrix} 2 & -2 & 1 \\ 3 & 5 & 0 \\ -1 & 3 & 4 \end{bmatrix}$. *Compute* $\det(A)$ *and* A^{-1}.

Solution. There are many ways to compute the determinant. We will compute it by developing the determinant according to the last column, giving:

$$\begin{aligned}\det(A) &= 1\cdot\det\begin{bmatrix}3 & 5\\ -1 & 3\end{bmatrix} + 0 + 4\cdot\det\begin{bmatrix}2 & -2\\ 3 & 5\end{bmatrix}\\ &= (9+5) + 4(10+6) = 78.\end{aligned}$$

They are also many methods to compute the inverse of A. We will show to of them. Firstly, we will use the formula

$$A^{-1} = \frac{1}{\det(A)}\operatorname{adj}(A),$$

where $\operatorname{adj}(A) = ((-1)^{i+j}\det(\mathrm{cofact}_{ij}))^{\intercal}_{ij}$. Recall that cofact_{ij} is the matrix obtained from A by erasing the i-row and j-column. Let us unravel this for our specific example. So we have

$$\operatorname{adj}(A)^{\intercal} = ((-1)^{i+j}\det(\mathrm{cofact}_{ij}))_{ij} = \begin{bmatrix}a & b & c\\ d & e & f\\ g & h & i\end{bmatrix},$$

where we need to compute the coefficients a to i. For example,

$$a = 1\cdot\det\begin{bmatrix}5 & 0\\ 3 & 4\end{bmatrix} = 20 \qquad b = -1\cdot\det\begin{bmatrix}3 & 0\\ -1 & 4\end{bmatrix} = -12.$$

By computing all the coefficients, one obtain

$$\operatorname{adj}(A)^{\intercal} = \begin{bmatrix}20 & -12 & 14\\ 11 & 9 & -4\\ -5 & 3 & 16\end{bmatrix},$$

And finally

$$A^{-1} = \frac{1}{78}\begin{bmatrix}20 & 11 & -5\\ -12 & 9 & 3\\ 14 & -4 & 16\end{bmatrix}.$$

While the above method always work, it is lengthy. Another way to compute the inverse is to use operations on rows. The idea is to start with $[A \mid \mathrm{Id}]$ and to do raw transformations to obtain $[\mathrm{Id} \mid B]$. The matrix B obtained with this process is A^{-1}. If we apply that to our concrete example, we obtain

$$A^{-1} = \begin{bmatrix}10/39 & 11/78 & -5/78\\ -2/13 & 3/26 & 1/26\\ 7/39 & -2/39 & 8/39\end{bmatrix} = \frac{1}{78}\begin{bmatrix}20 & 11 & -5\\ -12 & 9 & 3\\ 14 & -4 & 16\end{bmatrix},$$

which is identical to the answer obtained via the cofactors method. The details are shown on the next page.

$$\begin{tikzcd}
\left[\begin{array}{ccc|ccc} 2 & -2 & 1 & 1 & 0 & 0 \\ 3 & 5 & 0 & 0 & 1 & 0 \\ -1 & 3 & 4 & 0 & 0 & 1 \end{array}\right] \arrow[r, mapsto, "\substack{r_1 \mapsto \nicefrac{1}{2}\, r_1 \\ r_2 \mapsto r_2 - \nicefrac{3}{2}\, r_1 \\ r_3 \mapsto r_3 + \nicefrac{1}{2}\, r_1}"] & \left[\begin{array}{ccc|ccc} 1 & -1 & \nicefrac{1}{2} & \nicefrac{1}{2} & 0 & 0 \\ 0 & 8 & -\nicefrac{3}{2} & -\nicefrac{3}{2} & 1 & 0 \\ 0 & 2 & \nicefrac{9}{2} & \nicefrac{1}{2} & 0 & 1 \end{array}\right] \arrow[dl, mapsto, "\substack{r_2 \mapsto \nicefrac{1}{8}\, r_2 \\ r_3 \mapsto r_3 - \nicefrac{1}{4}\, r_2}"'] \\
\left[\begin{array}{ccc|ccc} 1 & -1 & \nicefrac{1}{2} & \nicefrac{1}{2} & 0 & 0 \\ 0 & 1 & -\nicefrac{3}{16} & -\nicefrac{3}{16} & \nicefrac{1}{8} & 0 \\ 0 & 0 & \nicefrac{39}{8} & \nicefrac{7}{8} & -\nicefrac{1}{4} & 1 \end{array}\right] \arrow[r, mapsto, "r_3 \mapsto \nicefrac{8}{39}\, r_3"] & \left[\begin{array}{ccc|ccc} 1 & -1 & \nicefrac{1}{2} & \nicefrac{1}{2} & 0 & 0 \\ 0 & 1 & -\nicefrac{3}{16} & -\nicefrac{3}{16} & \nicefrac{1}{8} & 0 \\ 0 & 0 & 1 & \nicefrac{7}{39} & -\nicefrac{2}{39} & \nicefrac{8}{39} \end{array}\right] \arrow[dl, mapsto, "\substack{r_1 \mapsto r_1 - \nicefrac{1}{2}\, r_3 \\ r_2 \mapsto r_2 + \nicefrac{3}{16}\, r_3}"'] \\
\left[\begin{array}{ccc|ccc} 1 & -1 & 0 & \nicefrac{16}{39} & \nicefrac{1}{39} & -\nicefrac{4}{39} \\ 0 & 1 & 0 & -\nicefrac{2}{13} & \nicefrac{3}{26} & \nicefrac{1}{26} \\ 0 & 0 & 1 & \nicefrac{7}{39} & -\nicefrac{2}{39} & \nicefrac{8}{39} \end{array}\right] \arrow[r, mapsto, "r_1 \mapsto r_1 + r_2"] & \left[\begin{array}{ccc|ccc} 1 & 0 & 0 & \nicefrac{10}{39} & \nicefrac{11}{78} & -\nicefrac{5}{78} \\ 0 & 1 & 0 & -\nicefrac{2}{13} & \nicefrac{3}{26} & \nicefrac{1}{26} \\ 0 & 0 & 1 & \nicefrac{7}{39} & -\nicefrac{2}{39} & \nicefrac{8}{39} \end{array}\right].
\end{tikzcd}$$

∎

Tutorial Questions

Question 1.

Let V be a vector space. Prove that $(-1)v = -v$ for all $v \in V$. (Before starting, think about what the minus sign on each side means.)

Solution. The left hand side of the equation is $-1_{\mathbf{F}} \cdot_{\mathbf{F}} v$. That is: the scalar multiplication of $-1 \in \mathbf{F}$ and $v \in V$, where -1 is the additive inverse of 1 in $\mathbf{F}$. The right hand side of the equation is $-_V v$, that is the additive inverse of v.

We have

$$v + (-1) \cdot v = 1 \cdot v + (-1) \cdot v = (1 + (-1))v = 0 \cdot v = 0.$$

We conclude that $(-1)v$ is an additive inverse of v and hence $(-1)v = -v$ by unicity of the additive inverse. ■

Question 2.

Determine if the following subsets of $\mathbf{R}^3$ are subspaces. Give reasons.

a. $U = \left\{ \begin{bmatrix} x \\ y \\ z \end{bmatrix} \in \mathbf{R}^3 \;\middle|\; 2x - 3y + z = 0 \right\}.$

b. $W = \left\{ \begin{bmatrix} x \\ y \\ z \end{bmatrix} \in \mathbf{R}^3 \;\middle|\; xyz = 0 \right\}.$

Solution. a. U is a subspace of $\mathbf{R}^3$. Firstly, $2 \cdot 0 - 3 \cdot 0 + 0 = 0$ so $[0, 0, 0]^\mathsf{T}$ is in U. Secondly, let $u_1 = [x_1, y_1, z_1]^\mathsf{T}$ and $u_2 = [x_2, y_2, z_2]^\mathsf{T}$ be two elements of U and let $\lambda \in \mathbf{R}$ be a scalar. By assumption, we have

$$2x_1 - 3y_1 + z_1 = 0 \qquad \text{and} \qquad 2x_2 - 3y_2 + z_2 = 0.$$

By multiplying the first equation by λ and adding the result to the second equation, we obtain

$$\begin{aligned} 0 &= \lambda(2x_1 - 3y_1 + z_1) + (2x_2 - 3y_2 + z_2) \\ &= 2(\lambda x_1 + x_2) - 3(\lambda y_1 + y_2) + (\lambda z_1 + z_2). \end{aligned}$$

Therefore, $\lambda u_1 + u_2$ is in U, finishing the proof that U is a subspace.

One can show that U is in a plane containing the origin in $\mathbf{R}^3$.

b. W is not a subspace of $\mathbf{R}^3$. Indeed, both $[1, 1, 0]^\mathsf{T}$ and $[0, 0, 1]^\mathsf{T}$ belong to W, but there sum $[1, 1, 0]^\mathsf{T} + [0, 0, 1]^\mathsf{T} = [1, 1, 1]^\mathsf{T}$ is not in W.

Geometrically, W is the union of the three coordinates planes (the xy-plane, the xz-plane and the yz plane), see Figure 1.

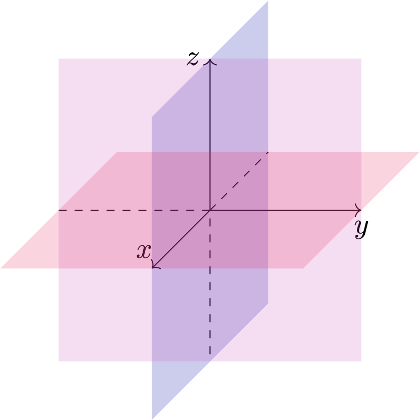


Figure 1: The union of the three coordinates planes.

Observe that the fact that U is a subspace of $\mathbf{F}^3$ but W is not remains true for an arbitrary field $\mathbf{F}$. ■

Question 3.

A sequence $(x_0, x_1, x_2, \dots) \in \mathbf{R}^{\mathbf{N}}$ *is* increasing *if* $x_m \leq x_{m+1}$ *for all* $m \in \mathbf{N}$. *Does the set* X *of all increasing sequences form a subspace of* $\mathbf{R}^{\mathbf{N}}$*? Give reasons.*

Solution. No, because X is not closed under scalar multiplication. For example $(0, 1, 1, \dots)$ is in X, but $-1 \cdot (0, 1, 1, \dots) = (0, -1, -1, \dots)$ is not. ■

Question 4.

Let V *be a vector space and* $U \subseteq V$ *be a subspace. Suppose that* $u \in U$ *and* $v \in V$. *Prove that* $v \in U$ *if and only if* $u + v \in U$.

Solution. If v is in U, then $u + v$ is also in U because U is a subspace.

If $u + v$ is in U, then $v = -1 \cdot u + (u + v)$ is also in U because U is a subspace. ■

Question 5.

Let U *be the xy-plane in* $\mathbf{R}^3$. *Find distinct subspaces* W_1, W_2 *of* $\mathbf{R}^3$ *such that* $U \oplus W_1 = U \oplus W_2 = \mathbf{R}^3$.

Solution. Let $W_1 := \{[0, 0, z]^\intercal \mid z \in \mathbf{R}\}$ and let $W_2 := \{[0, z, z]^\intercal \mid z \in \mathbf{R}\}$. Then $U \oplus W_1 = U \oplus W_2 = \mathbf{R}^3$.

For each W_i, we need to prove three things: 1. that it is a subspace of $\mathbf{R}^3$, 2. that $U + W_i = \mathbf{R}^3$, 3. and finally that $U + W_i$ is a direct sum. We will only check the last two properties, and only do it for W_2. The proof for W_1 is similar (and even simpler) and the proof that the W_i are subspaces of $\mathbf{R}^3$ is elementary.

Let $[x,y,z]^\mathsf{T}$ be any vector of $\mathbf{R}^3$. Then $[x,y,z]^\mathsf{T} = [x,y-z,0]^\mathsf{T} + [0,z,z]^\mathsf{T}$ is in $U+W_2$ and so $U+W_2=\mathbf{R}^3$. Finally, suppose that $0=u+w$ with $u\in U$ and $w\in W_2$. So $[0,0,0]^\mathsf{T} = [x,y,0]^\mathsf{T} + [0,z,z]^\mathsf{T}$ for some $x,y,z\in\mathbf{R}$. But this implies

$$\begin{cases} 0 = x \\ 0 = y+z \\ 0 = z \end{cases}$$

and we hence conclude that both u and w are the 0 vector. This shows that the sum $U+W_2$ is direct.

Observe that a similar proof shows that $U\oplus W=\mathbf{R}^3$ if W is any line containing the origin and not included in U.

To go further

The above statement remains true if we replace $\mathbf{R}$ by $\mathbf{F}$, in which case a *line* simply means a subspace of dimension 1. That is, a *line containing the origin* is any subspace of the form $\{[\lambda_1 z, \lambda_2 z, \lambda_3 z]^\mathsf{T} \mid z\in\mathbf{F}\}$ for some $\lambda_1,\lambda_2,\lambda_3$ with at least one $\lambda_i\neq 0$.

■

Further Questions

Question 6.

Is addition of subspaces commutative? That is, for subspaces U,W of V, is it true that $U+W=W+U$? Explain.

Solution. Addition of subspaces is commutative, because addition of vectors is. Indeed, a vector $v\in V$ belongs to $U+W$ if and only if there exists $u\in U$ and $w\in W$ such that $v=u+w$. By commutativity of *vectors addition*: $u+w=w+u$. So $v\in V$ belongs to $U+W$ if and only if there exists $u\in U$ and $w\in W$ such that $v=w+u$, that is if and only if v belongs to $W+U$.

Observe that we did not used the fact that U and W are subspaces. Therefore, addition of subsets (of a vector space) is commutative. ■

Question 7.

Does every subspace of V have an additive inverse*? That is, for each subspace U of V, does there exist a subspace W such that $U+W=\{0\}$?*

Solution. No. This is the case if and only if $U=\{0\}$, in which case $W=\{0\}$. Indeed, we always have $U\subseteq U+W$. So if $U+W=\{0\}$ for some W, then $U=\{0\}$. ■

Question 8.

Recall that a function $f\colon \mathbf{R} \to \mathbf{R}$ *is* even *if* $f(-x) = f(x)$ *for all* $x \in \mathbf{R}$*; and* odd *if* $f(-x) = -f(x)$ *for all* $x \in \mathbf{R}$*. Let* U_e *and* U_o *respectively denote the set of even and odd functions in* $\mathbf{R}^{\mathbf{R}}$*. Prove that* $\mathbf{R}^{\mathbf{R}} = U_e \oplus U_o$*.*

Solution. Once again, we need to prove 3 things: 1. that both U_e and U_o are subspaces of $\mathbf{R}^{\mathbf{R}}$, 2. that $U_e + U_o = \mathbf{R}^{\mathbf{R}}$, 3. and finally that $U_e + U_o$ is a direct sum.

Firstly, the constant function 0 is both even and odd. Moreover, if f and g are two even functions (respectively odd functions) and $\lambda \in \mathbf{R}$ a scalar, one easily verify that $\lambda f + g$ is still even (respectively odd). Indeed, this directly follows from $(\lambda f + g)(x) = \lambda f(x) + g(x)$. We conclude that both U_e and U_o are subspaces of $\mathbf{R}^{\mathbf{R}}$.

Then, let f be any function of $\mathbf{R}^{\mathbf{R}}$. Define two functions f_e and f_o by:

$$f_e(x) := \frac{f(x) + f(-x)}{2}, \qquad f_o(x) := \frac{f(x) - f(-x)}{2}.$$

On directly have $(f_e + f_o)(x) = f(x)$ for all $x \in \mathbf{R}$ and so $f_e + f_o = f$. We also have

$$f_e(-x) = \frac{f(-x) + f(x)}{2} = \frac{f(x) + f(-x)}{2} = f_e(x)$$

and

$$f_o(-x) = \frac{f(-x) - f(x)}{2} = -\frac{f(x) - f(-x)}{2} = -f_o(x).$$

So f_e is even and f_o is odd, showing that f is in $U_e + U_o$. This finishes the proof that $U_e + U_o = \mathbf{R}^{\mathbf{R}}$.

Finally, let f be in $U_e \cap U_o$. Then for every $x \in \mathbf{R}$, let $y = -x$. We have:

$$f(-y) = f(-y) = -f(-y).$$

But this implies that $f(x) = f(-y) = -f(-y) = -f(x)$ and so $f(x) = 0$. Therefore, f is the zero function, $U_e \cap U_o = \{0\}$ and the sum is direct. ■

Tutorial Questions

Question 1.

Let (u_1, u_2, u_3) *be a list of vectors of a vector space* V*.*

a. *Show that if* (u_1, u_2, u_3) *span* V *then so does the list* $(u_1 - u_2, u_2 - u_3, u_3)$*.*

b. *Show that if* (u_1, u_2, u_3) *is linearly independent then so is the list* $(u_1 - u_2, u_2 - u_3, u_3)$*.*

Solution. a. We will show that $\operatorname{span}(u_1, u_2, u_3) = \operatorname{span}(u_1 - u_2, u_2 - u_3, u_3)$, even they are not necessarily equal to V. The $\supseteq$ inclusion is clear, so let us prove $\operatorname{span}(u_1, u_2, u_3) \subseteq \operatorname{span}(u_1 - u_2, u_2 - u_3, u_3)$. Let $v \in \operatorname{span}(u_1, u_2, u_3)$. There exists scalars λ_1, λ_2 and λ_3 such that

$$\begin{aligned} v &= \lambda_1 u_1 + \lambda_2 u_2 + \lambda_3 u_3 \\ &= \lambda_1 (u_1 - u_2) + (\lambda_2 + \lambda 1)(u_2 - u_3) + (\lambda_3 + \lambda_2) u_3. \end{aligned}$$

Since all of λ_1, $\lambda_2 + \lambda 1$ and $\lambda_3 + \lambda_2$ are in $\mathbf{F}$, we have just shown that v belongs to $\operatorname{span}(u_1 - u_2, u_2 - u_3, u_3)$, which finishes the proof.

b. Suppose that (u_1, u_2, u_3) is linearly independent and let λ_1, λ_2 and λ_3 be scalars such that

$$0 = \lambda_1 (u_1 - u_2) + \lambda_2 (u_2 - u_3) + \lambda_3 u_3.$$

We have to show that this implies that all the λ_i are zeros. We have

$$0 = \lambda_1 u_1 + (\lambda_2 - \lambda_1) u_2 + (\lambda_3 - \lambda_2) u_3.$$

By linear independence of (u_1, u_2, u_3), all the coefficients λ_1, $\lambda_2 - \lambda_1$ and $\lambda_3 - \lambda_2$ are zeros. This directly implies that all the λ_i are 0, as required. ■

Question 2.

Let V *be a finite dimensional vector space and suppose that* $U \subseteq V$ *is a subspace. Prove that if* $\dim(U) = \dim(V)$ *then* $U = V$*.*

Solution. By assumption, $\dim(U) = \dim(V)$. So there exists a finite basis $\mathscr{B} = (u_1, \ldots, u_n)$ of U, where $n = \dim(V)$. Then $\mathscr{B}$ is a linearly independent family of vectors of U, hence also of vectors of V. Since it has $\dim(V)$ elements it is spanning in V. So $U = \operatorname{span}(\mathscr{B}) = V$. ■

Question 3.

Consider the subspace $U = \{p \in \mathcal{P}(\mathbf{R})_2 \mid p(1) = 0\}$ *of* $\mathcal{P}(\mathbf{R})_2$.

a. *Find a basis for* U. *(Hint: if a polynomial* p *satisfies* $p(a) = 0$ *then* $(x - a)$ *is a factor of* $p(x)$.*)*

b. *Extend the basis you found to a basis of* $\mathcal{P}(\mathbf{R})_2$.

c. *Find a subspace* W *of* $\mathcal{P}(\mathbf{R})_2$ *such that* $U \oplus W = \mathcal{P}(\mathbf{R})_2$.

Solution. We will show the result for a general field $\mathbf{F}$.

a. On one hand, we know that $\mathcal{P}(\mathbf{F})_2$ is of dimension 3 (with basis $(1, x, x^2)$). On the other hand, $U \neq \mathcal{P}(\mathbf{F})_2$ (for example, because $x \notin U$). By Question 2, this implies that U is of dimension at most 2. So if we find a linearly independent family of length 2 we are done. For example, one can take the family $(p_1 = x - 1, p_2 = x^2 - 1)$. We claim that this family is linearly independent. Indeed, if $0 = \lambda p_1 + \mu p_2$, then

$$0 = \mu x^2 + \lambda x + (-\lambda - \mu)$$

and therefore $\mu = 0 = \lambda$, which proves the claim.

b. We already have a length 2 linearly independent family. So for any $q \in \mathcal{P}(\mathbf{F})_2$ *not* in $\operatorname{span}(p_1, p_2) = U$, the family (p_1, p_2, q) is linearly independent by the linear dependence lemma. Since it is a length 3 linear independent family in a dimension 3 vector space, it is linearly independent.

For a concrete example, one can take $q = 1$ the constant polynomial 1. But $q = x$ or $q = x^2$ also work.

c. Let $q \notin U$ be any polynomial in $\mathcal{P}(\mathbf{F})_2$ but not in U. We claim that for $W := \operatorname{span}(q)$ we have $U \oplus W = \mathcal{P}(\mathbf{F})_2$. It is clear from above that $U + W = \mathcal{P}(\mathbf{F})_2$. It remains to prove that the sum is direct. Let p be in $U \cap W$. We need to show that $p = 0$. We have $p = \lambda q \in U$. If $p \neq 0$, then $q = \lambda^{-1} p \in U$; a contradiction.

■

Question 4.

Let U *and* W *be 3-dimensional subspaces of* $\mathbf{R}^5$. *What are the possible values of* $\dim(U \cap W)$*? Can* $U + W$ *be a direct sum?*

Solution. Once again, we will prove the statement for a general $\mathbf{F}$.

We have

$$\dim(U + W) = \dim(U) + \dim(W) - \dim(U + W) \leq \dim(\mathbf{R}^5) = 5.$$

In other words, $\dim(U \cap W) = \dim(U) + \dim(V) - \dim(U + W) \geq 3 + 3 - 5 = 1$. But we also have $\dim(U + W) \leq \dim(U) = 3$. So, $\dim(U + W)$ belongs to $\{1, 2, 3\}$. We now show that all this value can happens.

If $U = W_1$, for example $U = W_1 = \{[x_1, x_2, x_3, 0, 0]^\intercal \mid x_i \in \mathbf{F}\}$, then $\dim(U \cap W_1) = 3$.

For U as above and $W_2 = \{[0, x_2, x_3, x_4, 0]^\intercal \mid x_i \in \mathbf{F}\}$ the subspace $U \cap W_2 = \{[0, x_2, x_3, 0, 0]^\intercal \mid x_i \in \mathbf{F}\}$ is of dimension 2.

Finally, for U as above and $W_3 = \{[0, 0, x_3, x_4, x_5]^\intercal \mid x_i \in \mathbf{F}\}$ the subspace $U \cap W_3 = \{[0, 0, x_3, 0, 0]^\intercal \mid x_i \in \mathbf{F}\}$ is of dimension 1.

Since $\dim(U \cap W) \geq 1$, the sum $U + W$ is never direct. Indeed, the sum $U + W$ is direct if and only if $U \cap W = \{0\}$, if and only if $\dim(U \cap W) = 0$. ■

Question 5.

Let V be a vector space. Suppose that for each integer $m > 0$, there exists a list of m linearly independent vectors in V. Prove that V is infinite dimensional.

Solution. By Grassman's exchange lemma, if V has a linearly independent family of length m, then any spanning family has length at least m. In particular, any base of V has at least m elements, for all $m \in \mathbf{N}$. We conclude that a basis cannot be finite. ■

Further Questions

Question 6.

Prove that $\mathbf{R}^{\mathbf{N}}$ is infinite dimensional. (More generally, if S is an infinite set then $(\mathbf{F}^S)_0$ and $\mathbf{F}^S$ are infinite dimensional.)

Solution. By Question 5, it is enough to exhibit arbitrary long linearly independent lists. For any $i \in \mathbf{N}_{\geq 1}$, define $e_i = (0, \dots, 0, 1, 0, \dots)$ to be list with a 1 in coordinate i and 0 everywhere else. We claim that for any m, the family $(e_1, \dots, e_m)$ is linearly independent. Indeed, let λ_i be scalars such that

$$0 = \sum_{i=1}^{m} \lambda_i e_i.$$

Then

$$(0, 0, \dots) = (\lambda_1, \dots, \lambda_m, 0, \dots)$$

forcing all the λ_i to be 0.

A similar proof shows that $(\mathbf{F}^S)_0$ (and thus also $\mathbf{F}^S$) are infinite dimensional as soon as S is infinite. Indeed, for $s \in S$ one can define a map $\chi_S \colon S \to \mathbf{F}$ by

$$\chi_s(t) = \begin{cases} 1 & \text{if} t = s \\ 0 & \text{otherwise.} \end{cases}$$

We now claim that for any finite subset $A \subseteq S$, the family $(\chi_a)_{a\in A}$ is linearly independent. Indeed, let $(\lambda_a)_{a\in A}$ be a family of scalars such that

$$0 = \sum_{a\in A} \lambda_a \chi_a.$$

Then, for $a \in A$ one can evaluate both sides of the above equality on a to obtain $0 = \lambda_a \cdot 1$, and thus $\lambda_a = 0$.

To go further

It is in fact possible to show that for any S, finite or infinite, we have $\dim((\mathbf{F}^S)_0) = \#S$. Recall that for finites S, one have $(\mathbf{F}^S)_0 = \mathbf{F}^S$.

$\blacksquare$

Question 7.

Explain why all the subspaces of $\mathbf{R}^3$ *are precisely* $\{0\}$, *all straight lines through the origin, all planes through the origin, and* $\mathbf{R}^3$.

Solution. First of all, let us remind that a line in $\mathbf{R}^3$ is a subset of the form

$$L_{u,v} = \{\lambda u + v \mid \lambda \in \mathbf{R}\}$$

for some vectors v and $u \neq 0$. A line $L_{u,v}$ goes through the origin if and only if $v = 0$. Let us fix $u \in \mathbf{R}^3$. One easily check that $0 \in L_{u,0}$ and that $L_{u,0}$ is closed under addition and scalar multiplication. In consequence $L_{u,0}$ is a subspace of $\mathbf{R}^3$, of dimension 1 ((u) is a basis).

Similarly, a plane in $\mathbf{R}^3$ is a subset of the form

$$P_{u_1,u_2,v} = \{\lambda_1 u_1 + \lambda_2 u_2 + v \mid \lambda_1, \lambda_2 \in \mathbf{R}\}$$

for some vectors v and non-collinear (that is linearly independent) u_1 and u_2. Once again, a plane $P_{u_1,u_2,v}$ goes through the origin if and only if $v = 0$ and $P_{u_1,u_2,0}$ is a subspace of $\mathbf{R}^3$, of dimension 2 ((u_1, u_2) is a basis).

Finally, it is clear that both $\{0\}$ and $\mathbf{R}^3$ are subspaces of $\mathbf{R}^3$ (of respective dimension 0 and 3). We will now prove that any subspace U of $\mathbf{R}^3$ is of the form $\{0\}$, $L_{u,0}$ (for some non-zero u), $P_{u_1,u_2,0}$ (for some linearly independent u_1 and u_2) or $\mathbf{R}^3$. To do this, we will do a case-by-case analysis depending on $\dim(U) \in \{0, \dots, 3\}$.

If $\dim(U) = 0$, then $U = \{0\}$.

If $\dim(U) = 3 = \dim(\mathbf{R}^3)$, then $U = \mathbf{R}^3$ by Question 2.

Suppose now that $\dim(U = 1)$ and take any $u \neq 0$ in U. Then $\dim(\operatorname{span}(u)) = 1 = \dim(U)$ and thus $U = \operatorname{span}(u) = \{\lambda u \mid \lambda \in \mathbf{R}\} = L_{u,0}$ is a line through the origin. We also have the converse: any line through the origin is a subspace of dimension 1.

Suppose finally that $\dim(U) = 2$. Take any $u_1 \neq 0$ in U. Since $\dim(\operatorname{span}(u)) = 1 < 2 = \dim(U)$, there exists $u_2 \in U \setminus \operatorname{span}(u_1)$. By the linear dependence lemma, the family (u_1, u_2) is linearly independent and hence a basis of U. So $U = \mathrm{span}(u_1, u_2) = \{\lambda_1 u_1 + \lambda_2 u_2 \mid \lambda_i \in \mathbf{R}\} = P_{u_1,u_2,0}$ is a plane containing the origin. We also have the converse: any plane containing the origin is a subspace of dimension 2. ■

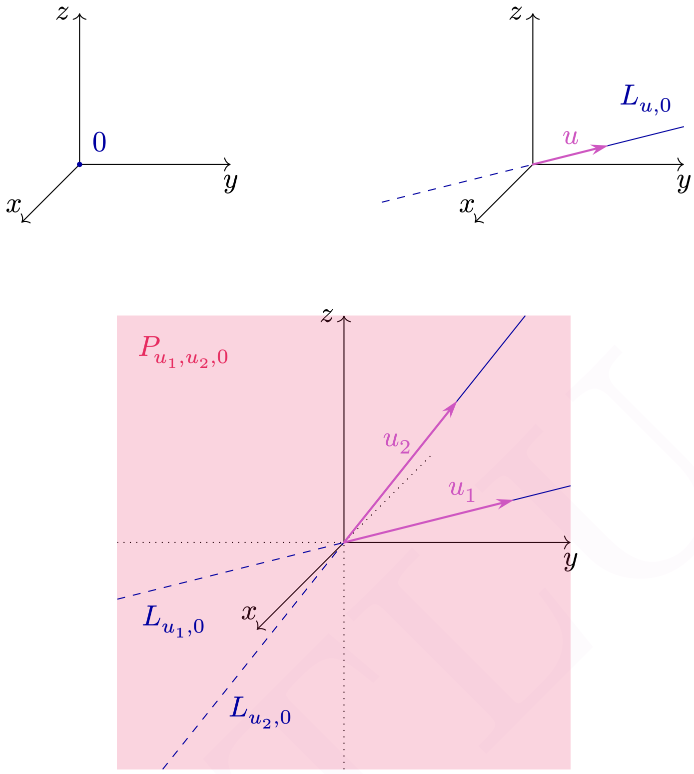


Figure 2: Subspaces of dimension 0 (the origin), 1 (line through the origin) and 2 (plane through the origin) of $\mathbf{R}^3$. The last kind of subspace is $\mathbf{R}^3$ itself.

Tutorial Questions

Question 1.

Determine whether the following functions are linear maps.

a. $T_1\colon \mathbf{R}^2 \to \mathbf{R}$ *given by* $T_1(\left[\begin{smallmatrix}x\\y\end{smallmatrix}\right]) = xy$.

b. $T_2\colon \mathbf{C} \to \mathbf{C}$ *given by* $T_2(z) = \bar{z}$, *where* $\mathbf{C}$ *is regarded as a real vector space* ($\overline{a+bi} = a - bi$ *is the complex conjugation).*

c. $T_2\colon \mathbf{C} \to \mathbf{C}$ *given by* $T_2(z) = \bar{z}$, *where* $\mathbf{C}$ *is regarded as a complex vector space.*

d. $T_3\colon \mathcal{P}(\mathbf{R}) \to \mathbf{R}$ *given by* $T_3(p) = p(2)$.

Solution. a. T_1 is not linear. Indeed, $T_1\left(\left[\begin{smallmatrix}1\\0\end{smallmatrix}\right] + \left[\begin{smallmatrix}0\\1\end{smallmatrix}\right]\right) = T_1\left(\left[\begin{smallmatrix}1\\1\end{smallmatrix}\right]\right) = 1\cdot 1 = 1$ while $T_1\left(\left[\begin{smallmatrix}1\\0\end{smallmatrix}\right]\right) + T_1\left(\left[\begin{smallmatrix}0\\1\end{smallmatrix}\right]\right) = 1\cdot 0 + 0\cdot 1 = 0 \neq 1$.

To go further

We have actually showed for any $\mathbf{F}$ the map $T_1\colon \mathbf{F}^2 \to \mathbf{F}$, $\left[\begin{smallmatrix}x\\y\end{smallmatrix}\right] \mapsto xy$ is not linear.

b. T_2 is $\mathbf{R}$-linear. Indeed, for $z, w \in \mathbf{C}$ and $\lambda \in \mathbf{R}$ we have $T_2(\lambda z + w) = \overline{\lambda z + w} = \overline{\lambda}\overline{z} + \overline{w} = \lambda\overline{z} + \overline{w} = \lambda T_2(z) + T_2(w)$. The above equalities can be verified explicitly by writing $z = a + bi$ and $w = c + di$.

c. T_2 is not $\mathbf{C}$-linear. Indeed, $iT_2(1) = i\overline{1} = i$, but $T_2(i1) = \overline{i} = -i \neq i$.

d. T_3 is linear. Indeed, let p and q by two polynomials in $\mathcal{P}(\mathbf{R})$ and $\lambda \in \mathbf{R}$. Then $T_3(\lambda p + q) = (\lambda p + q)(2) = \lambda p(2) + q(2) = \lambda T(p) + T(q)$.

To go further

The same proof shows that for any field $\mathbf{F}$, and any element $a \in \mathbf{F}$, the map $\mathcal{P}(\mathbf{F}) \to \mathbf{F}, p \mapsto p(a)$ is linear. Even more generally, if S is any set then the map $\mathbf{F}^S \to \mathbf{F}, f \mapsto f(a)$ is linear. This map is called the **evaluation at** a.

■

For a linear map $T \in \mathcal{L}(U, V)$, we write $\ker(T) = \mathrm{null}(T)$ for its kernel (also called nullspace): $\ker(T) = \{u \in U \mid T(u) = 0\}$. We write $\mathrm{Im}(T) = \mathrm{range}(T)$ for the image of T (also called range): $\mathrm{Im}(T) = \{v \in V \mid \exists u \in U : T(u) = v\}$.

Question 2.

Let U, V, and W be vector spaces, and suppose $T \in \mathcal{L}(U,V)$ and $S \in \mathcal{L}(V,W)$ are linear maps.

a. *Prove that* $\ker(T)$ *is a subspace of* $\ker(ST)$.

b. *Prove that* $\operatorname{Im}(ST)$ *is a subspace of* $\operatorname{Im}(S)$.

Solution. We have

$$\begin{tikzcd} U \arrow[r, "T"] \arrow[rr, bend right, "ST"'] & V \arrow[r, "S"] & W. \end{tikzcd}$$

a. Both T and ST have domain U, so both $\ker(T)$ and $\ker(ST)$ are subspaces of U. It is thus enough to prove that $\ker(T)$ is a *subset* of $\ker(ST)$. Let u be any element of $\ker(T)$, so $T(u) = 0$. Then $(ST)(u) = S(T(u)) = S(0) = 0$ and hence u is in $\ker(ST)$ as desired.

b. Both S and ST have codomain W, so both $\operatorname{Im}(S)$ and $\operatorname{Im}(ST)$ are subspaces of W. It is thus enough to prove that $\operatorname{Im}(ST)$ is a *subset* of $\operatorname{Im}(S)$. Let w be any element in $\operatorname{Im}(ST)$. So there exists $u \in U$ with $(ST)u = w$. In particular, there exists $v = T(u)$ in V (the domain of S) such that $S(v) = S(T(u)) = (ST)(u) = w$ and hence w is in $\operatorname{Im}(S)$ as desired.

Remark 0.1.

The inclusion $\ker(T) \subseteq \ker(ST)$ might be interpreted as ST *is less injective than* T. We recover the particular case: if ST is injective then so is T.
The inclusion $\operatorname{Im}(ST) \subseteq \operatorname{Im}(S)$ might be interpreted as ST *is less surjective than* S. We recover the particular case: if ST is surjective then so is S.

■

Question 3.

Let $T \in \mathcal{L}(V,W)$ be an injective linear map. Suppose that $(v_1, \dots, v_k)$ is a linearly independent list of vectors in V. Prove that $(Tv_1, \dots, Tv_k)$ is a linearly independent list in W.

Solution. Let $\lambda_1, \dots, \lambda_k$ be scalars such that

$$0 = \sum_{j=1}^{k} \lambda_j T v_j.$$

We need to prove that all the λ_j are 0. By linearity of T we have $0 = T\left(\sum_{j=1}^{k} \lambda_j v_j\right)$. By injectivity of T one obtains

$$0 = \sum_{j=1}^{k} \lambda_j v_j.$$

By linear independence of the $(v_1, \dots, v_k)$ we conclude that all the λ_j are 0. ■

Question 4.

Let V, W be vector spaces such that W is finite-dimensional. Let $T \in \mathcal{L}(V, W)$ be a linear map.

a. *Prove that T is surjective if and only if there exists a linear map $S \in \mathcal{L}(W, V)$ such that $TS = \mathrm{Id}_W$. (Hint: Define S on any basis of W then extend linearly.)*

b. *Prove that T is injective if and only if there exists a linear map $S \in \mathcal{L}(W, V)$ such that $ST = \mathrm{Id}_V$. (Hint: Similar to the above, but choose a basis of W carefully.)*

Solution. First of all, we know that for functions $f\colon X \to Y$ and $g\colon Y \to Z$ we have:

- if $g \circ f$ is injective then f is injective;
- if $g \circ f$ is surjective then g is surjective.

Since Id_X is bijective, this directly shows:

a. if $ST = \mathrm{Id}_V$ then T is injective;

b. if $TS = \mathrm{Id}_W$ then T is surjective.

Observe that we don't need to suppose that W is finite dimensional, or that S is a linear map (S a function is enough).

For the other direction, we will use bases to translate our problem for linear maps on vector spaces to a problem about functions on sets. We can do this, because a linear map $T\colon V \to W$ is uniquely determined by what it does on a given basis $\mathscr{B}$ of V.

a. Let $n = \dim W$. Suppose that T is surjective and let $\mathscr{C} = (w_1, \dots, w_n)$ be any basis of W. For any $w_j \in \mathscr{C}$, choose a preimage $v_j \in V$ so $T(v_j) = w_j$. There exists at least one such v_j by surjectivity. We define $S\colon \mathscr{C} \to V$ by $S(w_j) = v_j$ and extend it to a linear map $S \in \mathcal{L}(W, V)$. Then for any w_j we have $TSw_j = w_j$. Since TS is the identity on the basis $\mathscr{C}$, it is the identity everywhere: $TS = \mathrm{Id}_W$.

b. Suppose that T is injective and let $\mathscr{D} = (w_1, \dots, w_k)$ be a basis of $\mathrm{Im}(T)$. Since T is injective, for every $j \in \{1, \dots, k\}$ there exists a unique $v_j \in V$ with $Tv_j = w_j$. We claim that $\mathscr{A} = (v_1, \dots, v_k)$ is a basis of V.

   First, we prove that $\mathscr{A}$ is a spanning family for V. Let v be any vector in V. Then $Tv \in \mathrm{Im}(T) = \mathrm{span}(\mathscr{D})$ so there exists λ_j, $j \in \{1, \dots, k\}$ such that

$$T(v) = \sum_{j=1}^{k} \lambda_j w_j = \sum_{j=1}^{k} \lambda_j T v_j = T\Big(\sum_{j=1}^{k} \lambda_j v_j\Big).$$

By injectivity, $v = \sum_{j=1}^{k} \lambda_j v_j$ is in $\operatorname{span}(\mathscr{A})$.

We now prove that $\mathscr{A}$ is linearly independent. Suppose that

$$0 = \sum_{j=1}^{k} \lambda_j v_j$$

for some λ_j. Applying T on both sides, one obtain $0 = \sum_{j=1}^{k} \lambda_j w_j$ and conclude that all the λ_j are 0, proving linear independence of $\mathscr{A}$.

Now, since $\mathscr{D}$ is a basis of $\operatorname{Im}(T) \subseteq W$, we can extend it to a basis $\mathscr{C} = (w_1, \dots, w_k, w_{k+1}, \dots, w_m)$ of W. We define S on $\mathscr{C}$ by:

$$S(w_j) = \begin{cases} v_j & \text{if } j \in \{1, \dots, k\} \\ 0 & \text{otherwise} \end{cases}$$

and then extend it to a linear map $S \in \mathcal{L}(W, V)$. One have $ST(v_j) = v_j$ for all $j \in \{1, \dots, k\}$ and so $ST = \mathrm{Id}_V$.

Infinite dimensional vector spaces

The statements remain true for a general, not necessarily finite dimensional, vector space W if we assume (AC). We use (AC) in two different places. Firstly, for the existence of a basis of W, and secondly when T is injective and we *choose* a preimage of w_j.

■

Question 5.

Rephrase the following problems in terms of linear maps between vector spaces. (Do not solve the problems.)

a. *Find all smooth functions* $f\colon \mathbf{R} \to \mathbf{R}$ *that satisfy* $f'' - 3f' + 2f = 0$.

b. *Find all real sequences* $(x_0, x_1, \dots)$ *that satisfy* $x_{n+2} = x_{n+1} + x_n + 1$ *for all* $n \in \mathbf{N}$.

Solution. a. The space of smooth functions is $\mathcal{C}^\infty(\mathbf{R}) = \{f\colon \mathbf{R} \to \mathbf{R} \mid f \text{ smooth}\} = \{f\colon \mathbf{R} \to \mathbf{R} \mid \forall n : f^{(n)} \text{ exists}\}$. Let $D\colon \mathcal{C}^\infty(\mathbf{R}) \to \mathcal{C}^\infty(\mathbf{R}), f \mapsto f'$ be the differentiation operator. This is a linear map. Define $T := D^2 - 3D + 2\,\mathrm{Id}$, where $D^2 = D \circ D$. This is also a linear map from $\mathcal{C}^\infty(\mathbf{R})$ to itself and $T(f) = f'' - 3f' + 2f$. We have $f'' - 3f' + 2f = 0$ if and only if $f \in \ker(T)$.

b. The space of real sequences is $\mathbf{R}^{\mathbf{N}} = \{(x_0, x_1, x_2, \dots) \mid \forall j : x_j \in \mathbf{R}\}$. Let $S\colon \mathbf{R}^{\mathbf{N}} \to \mathbf{R}^{\mathbf{N}}$ be the backward shift: $S(x_0, x_1, \dots) = (x_1, x_2, \dots)$. Finally, let $c = (1, 1, \dots)$ be the constant sequence 1. We want to find all sequences $x \in \mathbf{R}^{\mathbf{N}}$ such that

$S^2(x) = S(x) + x + c$. So if we define $T = S^2 - S - \mathrm{Id}$, a sequence x satisfies the questions if and only if $T(x) = c$. ∎

Further Questions

Question 6.

Let U, V, and W be vector spaces, and suppose $S_1, S_2 \in \mathcal{L}(U,V)$ and $T \in \mathcal{L}(V,W)$ are linear maps. Prove that the following distributive property holds: $T(S_1 + S_2) = TS_1 + TS_2$.

Solution. This is similar to the proof of $(T_1 + T_2)S = T_1S + T_2S$ viewed in class. Let u be any vector in U. Then

$$\begin{aligned}(T(S_1+S_2))u &= T((S_1+S_2)u)\\ &= T(S_1u + S_2u)\\ &= T(S_1u) + T(S_2u)\\ &= (TS_1)u + (TS_2)u = (TS_1 + TS_2)u\end{aligned}$$

and thus $T(S_1 + S_2) = TS_1 + TS_2$. ∎

Question 7.

Let $T \in \mathcal{L}(U,V)$ and $S \in \mathcal{L}(V,W)$ be linear maps. Show that $ST = 0 \in \mathcal{L}(U,W)$ (the zero map) if and only if $\mathrm{Im}(T)$ is a subspace of $\ker S$.

Solution. "⇒" Suppose $ST = 0$. Let v be any element in $\mathrm{Im}(T)$. By assumption, there exists $u \in U$ such that $Tu = v$. But then, $Sv = S(Tu) = (ST)u = 0u = 0$ and thus v is in $\ker(S)$ as desired.

"⇐" Suppose that $\mathrm{Im}(T) \subseteq \ker(S)$. Then for any $u \in U$ we have $(ST)(u) = S(Tu) = 0$, showing that ST is the 0 map. ∎

Tutorial Questions

Question 1.

Let $D\colon \mathcal{P}(\mathbf{R}) \to \mathcal{P}(\mathbf{R})$ *be the differentiation operator. Is* D *invertible?*

Solution. The operator D is not injective and hence not invertible. Indeed, both the constant polynomial 0 and the constant polynomial 1 are sent onto 0. ■

Question 2.

Let U *be the subspace of* $\mathcal{C}^\infty(\mathbf{R})$ *spanned by the basis* $\mathscr{B}$ *given by* $(\mathrm{e}^x, x\mathrm{e}^x, x^2\mathrm{e}^x)$. *Let* $D \in \mathcal{L}(U)$ *be the differentiation operator. Find the matrix* $[D]^{\mathscr{B}}_{\mathscr{B}}$.

Solution. We compute each column of $[D]^{\mathscr{B}}_{\mathscr{B}}$.

Firstly, $D(\mathrm{e}^x) = \mathrm{e}^x = 1\cdot\mathrm{e}^x + 0\cdot x\mathrm{e}^x + 0\cdot x^2\mathrm{e}^x$ so $[D(\mathrm{e}^x)]_{\mathscr{B}} = \begin{bmatrix}1\\0\\0\end{bmatrix}$. Then, $D(x\mathrm{e}^x) = \mathrm{e}^x + x\mathrm{e}^x$ and therefore $[D(x\mathrm{e}^x)]_{\mathscr{B}} = \begin{bmatrix}1\\1\\0\end{bmatrix}$. Finally, $D(x^2\mathrm{e}^x) = 2x\mathrm{e}^x + x^2\mathrm{e}^x$ and therefore $[D(x\mathrm{e}^x)]_{\mathscr{B}} = \begin{bmatrix}0\\2\\1\end{bmatrix}$.

Altogether, we have

$$[D]^{\mathscr{B}}_{\mathscr{B}} = \begin{bmatrix}1 & 1 & 0\\0 & 1 & 2\\0 & 0 & 1\end{bmatrix}.$$

■

Question 3.

Let $T \in \mathcal{L}(V,W)$ *be a linear map and suppose* $S_1, S_2 \in \mathcal{L}(W,V)$ *satisfy* $S_1T = \mathrm{Id}_V$ *and* $TS_2 = \mathrm{Id}_W$. *Prove that* T *is invertible. What is the inverse of* T*?*

Solution. We have

$$S_1 = S_1\,\mathrm{Id}_W = S_1TS_2 = Id_VS_2 = S_2.$$

Therefore $S_1 = S_2$ is the inverse of T, which is therefore invertible. ■

Question 4.

Let $T \in \mathcal{L}(V,W)$ *be a surjective linear map. Suppose that* U *is a subspace of* V *such that* $\ker(T) \oplus U = V$. *Prove that the restriction* $T|_U\colon U \to W$ *is an isomorphism (recall that the restriction is defined by* $T|_U(u) = Tu$ *for all* $u \in U$*).*

Solution. The map $T|_U$ is the restriction of a linear map, and thus linear. It remains to show that it is injective and surjective.

We first prove injectivity. We both have $\ker(T|_U) \subseteq U$ and $\ker(T|_U) \subseteq \ker(T)$. We conclude $\ker(T|_U) \subseteq U \cap \ker(T)$. But since the sum $U \oplus \ker(T)$ is direct, we have $U \cap \ker(T) = \{0\}$, proving that $T|_U$ is injective.

We now prove surjectivity. Let $w \in W$ be any vector. By surjectivity of T, there exists $v \in V$ such that $T(v) = w$. Since $V = \ker(T) \oplus U$, there exist (unique) $u_1 \in \ker(T)$ and $u_2 \in U$ such that $v = u_1 + u_2$. So $w = T(v) = T(u_1 + u_2) = T(u_1) + T(u_2) = 0 + T(u_2) = T|_U(u_2)$, proving surjectivity of $T|_U$. ■

Question 5.

Let $x_1, \dots, x_{n+1} \in \mathbf{R}$ be distinct. Define a linear map $T\colon \mathcal{P}(\mathbf{R})_n \to \mathbf{R}^{n+1}$ by $T(p) := [p(x_1), \dots, p(x_{n+1})]^\mathsf{T}$.

a. *Prove that T is injective. (Hint: By the division algorithm, any non-zero degree n polynomial has at most n distinct roots.)*

b. *Show that for any $y_1, \dots, y_{n+1} \in \mathbf{R}$, there exists a unique polynomial p of degree at most n such that $p(x_i) = y_i$ for all $1 \leq i \leq n+1$. (Hint: Show that T is an isomorphism.)*

Solution. a. Let p be any polynomial in $\ker(T)$. So $p(x_1) = \dots = p(x_{n+1}) = 0$. That is, p is a polynomial of degree at most n with (at least) $n+1$ distinct roots. We conclude that p is the zero polynomial and so that $\ker(T) = \{0\}$.

b. Both $\mathcal{P}(\mathbf{R})_n$ and $\mathbf{R}^{n+1}$ have dimension $n+1$. Therefore T being injective is also surjective, and hence bijective. In particular, for any $y_1, \dots, y_{n+1} \in \mathbf{R}$ there exists a unique polynomial $p \in \mathcal{P}(\mathbf{R})_n$ such that $T(p) = [y_1, \dots, y_{n+1}]^\mathsf{T}$.

■

Further Questions

Question 6.

[Continuing from Question 5, if you want to solve it explicitly] Let $(e_1, \dots, e_{n+1})$ be the standard basis for $\mathbf{R}^{n+1}$ and $p_i = T^{-1}(e_i) \in \mathcal{P}(\mathbf{R})_n$. In other words, p_i is the unique polynomial of degree at most n satisfying $p_i(x_i) = 1$ and $p_i(x_j) = 0$ for $i \neq j$.

a. *Find an explicit formula for $p_i(x)$. (Hint: a non-zero polynomial p satisfies $p(a) = 0$ if and only if $(x - a)$ is a factor of p.)*

b. *Write down $T^{-1}[y_1, \dots, y_{n+1}]^\mathsf{T}$ in terms of the p_i's and y_i's. (This is known as Lagrange's Interpolation Formula.)*

Solution. a. We first construct polynomials that satisfy $q_i(x_j) = 0$ if $i \neq j$, but might not have the correct value for $q_i(x_i) = 0$. This is quite easy.

Let $q_1(x) := (x - x_2) \cdots (x - x_{n+1})$ and define similarly $q_2 := (x - x_1)(x - x_3) \cdots (x - x_{n+1})$ and $q_3, \dots, q_n$ up to $q_{n+1}(x) := (x - x_1)(x - x_2) \cdots (x - x_{n+1} + 1)$. In other words,

$$q_i = \prod_{\substack{1 \leq k \leq n+1 \\ k \neq i}} (x - x_k).$$

Then for any $1 \leq j \leq n+1$ one have

$$q_i(x_j) = \begin{cases} 0 & \text{if } i \neq j, \\ \prod_{\substack{1 \leq k \leq n+1 \\ k \neq i}} (x_i - x_k) \neq 0 & \text{if } i = j. \end{cases}$$

It is now easy to construct the p_i, by normalising their value at x_i:

$$p_i = \frac{\prod_{\substack{1 \leq k \leq n+1 \\ k \neq i}} (x - x_k)}{\prod_{\substack{1 \leq k \leq n+1 \\ k \neq i}} (x_i - x_k)} = \prod_{\substack{1 \leq k \leq n+1 \\ k \neq i}} \frac{(x - x_k)}{(x_i - x_k)}.$$

b. We have $[y_1, \dots, y_{n+1}]^\intercal = y_1 e_1 + \cdots y_{n+1} e_{n+1}$. By linearity

$$T^{-1}[y_1, \dots, y_{n+1}]^\intercal = \sum_{i=1}^{n+1} y_i T^{-1}(e_i) = \sum_{i=1}^{n+1} y_i p_i = \sum_{i=1}^{n+1} y_i \prod_{\substack{1 \leq k \leq n+1 \\ k \neq i}} \frac{(x - x_k)}{(x_i - x_k)}.$$

■

Question 7.

Prove that V is isomorphic to $\mathcal{L}(\mathbf{F}, V)$. (Hint: consider the function $F \colon \mathcal{L}(\mathbf{F}, V) \to V$ defined by $F(T) = T(1)$. This function is called evaluation at 1.)

Solution. We need to prove that F is a bijective linear map.

We first prove injectivity. Let T be an element of the kernel of F. That is: $T \colon \mathbf{F} \to V$ is a linear map such that $T(1) = 0$. But then, for every $\lambda \in \mathbf{F}$ we have $T(\lambda) = T(\lambda \cdot 1) = \lambda T(1) = \lambda \cdot 0 = 0$. This proves that $\ker(F) = \{0\}$ and so F is injective.

For surjectivity, let v be any vector in V. Since 1 is a basis of $\mathbf{F}$, there exists a unique linear map $T \colon \mathbf{F} \to V$ satisfying $T(1) = v$. We have $F(T) = v$ as desired.

Alternatively, one can prove bijectivity of F by finding an inverse. Let $G \colon V \to \mathcal{L}(\mathbf{F}, V)$ be the map defined by $G(v) \colon \mathbf{F} \to V, \lambda \mapsto \lambda.v$. That is, each v is send by G onto the map $G(v) = T$ from the surjectivity proof. One easily verify that $(G \circ F)(T) = T$ and $(F \circ G)(v) = v$, proving that G is the inverse of F.

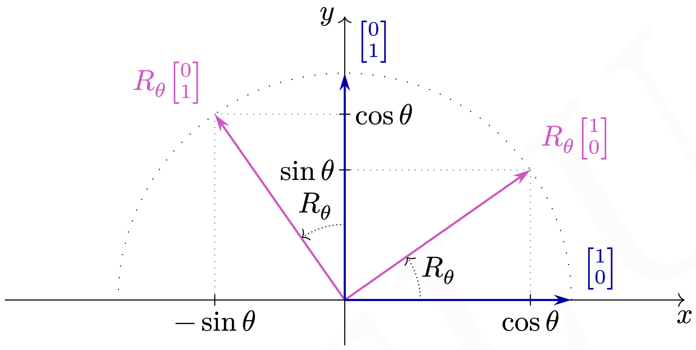
Finally, we show that F is linear. Let $S, T \mathcal{L}(\mathbf{F}, V)$ be two linear maps and let $\lambda \in \mathbf{F}$ be a scalar. Then $F(\lambda S + T) = (\lambda S + T)(1) = \lambda \cdot S(1) + T(1) = \lambda F(S) + F(T)$ as desired. ■

Tutorial Questions

Question 1.

Let $R_\theta \colon \mathbf{R}^2 \to \mathbf{R}^2$ *be a rotation about the origin through an angle of* θ. *Find the matrix* $[R_\theta]^{\mathscr{E}}_{\mathscr{E}}$, *where* $\mathscr{E}$ *is the standard basis for* $\mathbf{R}^2$. *Use this to prove that* $R_\theta \circ R_\phi = R_{\theta+\phi}$. (Hint: use the trigonometric formulas $\cos(a+b) = \cos(a)\cos(b) - \sin(a)\sin(b)$ and $\sin(a+b) = \sin(a)\cos(b) + \cos(a)\sin(b)$.)

Solution. We have $\mathscr{E} = \left(\begin{bmatrix}1\\0\end{bmatrix}, \begin{bmatrix}0\\1\end{bmatrix}\right)$. Simple trigonometry allows us to compute the coordinates of $R_\theta\begin{bmatrix}1\\0\end{bmatrix}$ and $R_\theta\begin{bmatrix}0\\1\end{bmatrix}$.



So we have

$$R_\theta\begin{bmatrix}1\\0\end{bmatrix} = \begin{bmatrix}\cos\theta\\ \sin\theta\end{bmatrix} \qquad \text{and} \qquad R_\theta\begin{bmatrix}0\\1\end{bmatrix} = \begin{bmatrix}-\sin\theta\\ \cos\theta\end{bmatrix},$$

which implies

$$[R_\theta]^{\mathscr{E}}_{\mathscr{E}} = \begin{bmatrix}\cos\theta & -\sin\theta\\ \sin\theta & \cos\theta\end{bmatrix}.$$

We then compute

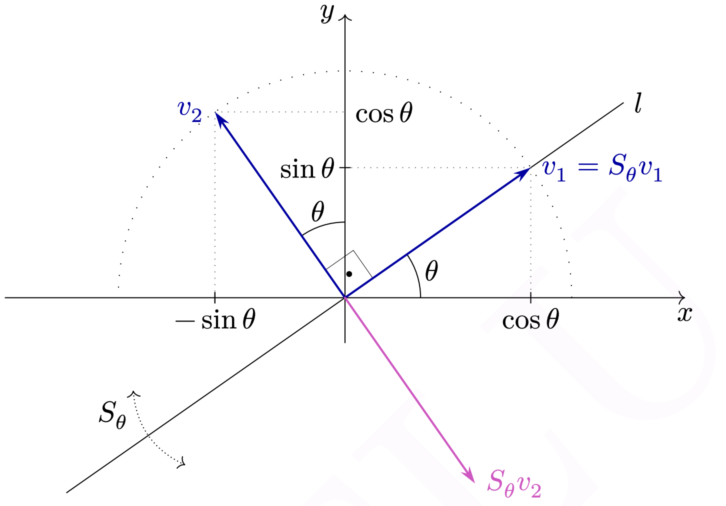
$$\begin{aligned}
[R_\theta \circ R_\phi]^{\mathscr{E}}_{\mathscr{E}} &= [R_\theta]^{\mathscr{E}}_{\mathscr{E}}[R_\phi]^{\mathscr{E}}_{\mathscr{E}} \\
&= \begin{bmatrix}\cos\theta & -\sin\theta\\ \sin\theta & \cos\theta\end{bmatrix}\begin{bmatrix}\cos\phi & -\sin\phi\\ \sin\phi & \cos\phi\end{bmatrix} \\
&= \begin{bmatrix}\cos\theta\cos\phi - \sin\theta\sin\phi & -\cos\theta\sin\phi - \sin\theta\cos\phi\\ \sin\theta\cos\phi + \cos\theta\sin\phi & -\sin\theta\sin\phi + \cos\theta\cos\phi\end{bmatrix} \\
&= \begin{bmatrix}\cos(\theta+\phi) & -\sin(\theta+\phi)\\ \sin(\theta+\phi) & \cos(\theta+\phi)\end{bmatrix} = [R_{\theta+\phi}]^{\mathscr{E}}_{\mathscr{E}}.
\end{aligned}$$

We conclude $R_\theta \circ R_\phi = R_{\theta+\phi}$ as desired. ■

Question 2.

Let $S_\theta\colon \mathbf{R}^2 \to \mathbf{R}^2$ be a reflection across the line through the origin which makes an angle of θ with the positive x-axis. Let $\mathscr{B}$ be the basis of $\mathbf{R}^2$ given by $v_1 = \left[\begin{smallmatrix}\cos\theta\\ \sin\theta\end{smallmatrix}\right]$ and $v_2 = \left[\begin{smallmatrix}-\sin\theta\\ \cos\theta\end{smallmatrix}\right]$. Find $[S_\theta]_{\mathscr{B}}^{\mathscr{B}}$. Then find $[S_\theta]_{\mathscr{E}}^{\mathscr{E}}$ by applying the change of basis formula, where $\mathscr{E}$ is the standard basis of $\mathbf{R}^2$.

Solution.



Since S_θ is a reflection along $l = \operatorname{span}(v_1)$ and $v_2 \perp v_1$ we have $S_\theta(v_1) = v_1$ and $S_\theta(v_2) = v_2$. Therefore, $[S_\theta(v_1)]_{\mathscr{B}} = \left[\begin{smallmatrix}1\\0\end{smallmatrix}\right]$ and $[S_\theta(v_2)]_{\mathscr{B}} = \left[\begin{smallmatrix}0\\-1\end{smallmatrix}\right]$. Altogether we obtain

$$[S_\theta]_{\mathscr{B}}^{\mathscr{B}} = \begin{bmatrix}1 & 0\\ 0 & -1\end{bmatrix}.$$

The change of basis formula gives us:

$$[S_\theta]_{\mathscr{E}}^{\mathscr{E}} = [\mathrm{Id}]_{\mathscr{B}}^{\mathscr{E}}[R_\theta]_{\mathscr{B}}^{\mathscr{B}}[\mathrm{Id}]_{\mathscr{E}}^{\mathscr{B}}.$$

So we need to compute $[\mathrm{Id}]_{\mathscr{B}}^{\mathscr{E}}$ and $[\mathrm{Id}]_{\mathscr{E}}^{\mathscr{B}} = [\mathrm{Id}]_{\mathscr{B}}^{\mathscr{E}\,-1}$. We have $\mathrm{Id}\, v_1 = \cos\theta e_1 + \sin\theta e_2$ and $\mathrm{Id}\, v_2 = -\sin\theta e_1 + \cos\theta e_2$. This gives us

$$[\mathrm{Id}]_{\mathscr{B}}^{\mathscr{E}} = \begin{bmatrix}\cos\theta & -\sin\theta\\ \sin\theta & \cos\theta\end{bmatrix} \qquad \text{and} \qquad [\mathrm{Id}]_{\mathscr{E}}^{\mathscr{B}} = ([\mathrm{Id}]_{\mathscr{B}}^{\mathscr{E}})^{-1} = \begin{bmatrix}\cos\theta & \sin\theta\\ -\sin\theta & \cos\theta\end{bmatrix}.$$

Finally, one have

$$\begin{aligned}
[S_\theta]_{\mathscr{E}}^{\mathscr{E}} &= \begin{bmatrix}\cos\theta & -\sin\theta\\ \sin\theta & \cos\theta\end{bmatrix}\begin{bmatrix}1 & 0\\ 0 & -1\end{bmatrix}\begin{bmatrix}\cos\theta & \sin\theta\\ -\sin\theta & \cos\theta\end{bmatrix}\\
&= \begin{bmatrix}\cos\theta & -\sin\theta\\ \sin\theta & \cos\theta\end{bmatrix}\begin{bmatrix}\cos\theta & \sin\theta\\ \sin\theta & -\cos\theta\end{bmatrix}\\
&= \begin{bmatrix}\cos^2\theta - \sin^2\theta & 2\cos\theta\sin\theta\\ 2\cos\theta\sin\theta & \sin^2\theta - \cos^2\theta\end{bmatrix} = \begin{bmatrix}\cos 2\theta & \sin 2\theta\\ \sin 2\theta & -\cos 2\theta\end{bmatrix}.
\end{aligned}$$

■

Question 3.

Consider a basis $\mathscr{B}$ *of* $\mathbf{R}^2$ *given by* $v_1 = \left[\begin{smallmatrix}3\\4\end{smallmatrix}\right]$ *and* $v_2 = \left[\begin{smallmatrix}-4\\3\end{smallmatrix}\right]$*. Let* T *be a transformation that stretches in the* v_1*-direction by a factor of* 2*, and in the* v_2*-direction by a factor of* 3*. Compute* $[T]^{\mathscr{E}}_{\mathscr{E}}$*.*

Solution. We have $Tv_1 = 2v_1$ and $Tv_2 = 3v_2$, which gives $[T]^{\mathscr{B}}_{\mathscr{B}} = \begin{bmatrix}2 & 0\\0 & 3\end{bmatrix}$. The change of basis matrices are given by

$$[\mathrm{Id}]^{\mathscr{E}}_{\mathscr{B}} = \begin{bmatrix}3 & -4\\4 & 3\end{bmatrix} \qquad \text{and} \qquad [\mathrm{Id}]^{\mathscr{B}}_{\mathscr{E}} = ([\mathrm{Id}]^{\mathscr{E}}_{\mathscr{B}})^{-1} = \frac{1}{25}\begin{bmatrix}3 & 4\\-4 & 3\end{bmatrix}.$$

Finally,

$$[T]^{\mathscr{E}}_{\mathscr{E}} = \begin{bmatrix}3 & -4\\4 & 3\end{bmatrix}\begin{bmatrix}2 & 0\\0 & 3\end{bmatrix}\frac{1}{25}\begin{bmatrix}3 & 4\\-4 & 3\end{bmatrix} = \frac{1}{25}\begin{bmatrix}3 & -4\\4 & 3\end{bmatrix}\begin{bmatrix}6 & 8\\-12 & 9\end{bmatrix} = \frac{1}{25}\begin{bmatrix}66 & -12\\-12 & 59\end{bmatrix}.$$

■

Question 4.

Let U *be the subspace of* $\mathscr{C}^\infty(\mathbf{R})$ *spanned by the basis* $\mathscr{B}$ *given by* $(\mathrm{e}^x, x\mathrm{e}^x, x^2\mathrm{e}^x)$*. Let* $D \in \mathscr{L}(U)$ *be the differentiation operator. Recall from Tutorial 5 that*

$$[D]^{\mathscr{B}}_{\mathscr{B}} = \begin{bmatrix}1 & 1 & 0\\0 & 1 & 2\\0 & 0 & 1\end{bmatrix}.$$

Is D *invertible? If so, compute* $[D^{-1}]^{\mathscr{B}}_{\mathscr{B}}$ *and use this to find the unique function* $f \in U$ *satisfying* $f' = x^2\mathrm{e}^x$*.*

Solution. Since $\det([D]^{\mathscr{B}}_{\mathscr{B}}) = 1$, the matrix $[D]^{\mathscr{B}}_{\mathscr{B}}$ is invertible and so D is invertible. We compute

$$\left[\underbrace{\begin{matrix}1 & 1 & 0\\0 & 1 & 2\\0 & 0 & 1\end{matrix}}_{[D]^{\mathscr{B}}_{\mathscr{B}}}\,\middle|\,\underbrace{\begin{matrix}1 & 0 & 0\\0 & 1 & 0\\0 & 0 & 1\end{matrix}}_{\mathrm{Id}}\right] \xrightarrow[r_2 \mapsto r_2 - 2r_3]{r_1 \mapsto r_1 - r_2 + 2r_3} \left[\underbrace{\begin{matrix}1 & 0 & 0\\0 & 1 & 0\\0 & 0 & 1\end{matrix}}_{\mathrm{Id}}\,\middle|\,\underbrace{\begin{matrix}1 & -1 & 2\\0 & 1 & -2\\0 & 0 & 1\end{matrix}}_{([D]^{\mathscr{B}}_{\mathscr{B}})^{-1}}\right].$$

So

$$([D]^{\mathscr{B}}_{\mathscr{B}})^{-1} = \begin{bmatrix}1 & -1 & 2\\0 & 1 & -2\\0 & 0 & 1\end{bmatrix}$$

and f is given by

$$[f]_{\mathscr{B}} = ([D]_{\mathscr{B}}^{\mathscr{B}})^{-1}[x^2\mathrm{e}^x]_{\mathscr{B}} = \begin{bmatrix} 1 & -1 & 2 \\ 0 & 1 & -2 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} = \begin{bmatrix} 2 \\ -2 \\ 1 \end{bmatrix}.$$

In other words, $f = 2\mathrm{e}^x - 2x\mathrm{e}^x + x^2\mathrm{e}^x$. ■

Further Questions

Question 5.

[Continuing from Question 2] Compute $[S_\theta \circ S_\phi]_{\mathscr{E}}^{\mathscr{E}}$. Explain what the transformation $S_\theta \circ S_\phi$ does geometrically.

Solution.

$$\begin{aligned}
[S_\theta \circ S_\phi]_{\mathscr{E}}^{\mathscr{E}} &= [S_\theta]_{\mathscr{E}}^{\mathscr{E}}[S_\phi]_{\mathscr{E}}^{\mathscr{E}} \\
&= \begin{bmatrix} \cos 2\theta & \sin 2\theta \\ \sin 2\theta & -\cos 2\theta \end{bmatrix} \begin{bmatrix} \cos 2\phi & \sin 2\phi \\ \sin 2\phi & -\cos 2\phi \end{bmatrix} \\
&= \begin{bmatrix} \cos 2\theta \cos 2\phi + \sin 2\theta \sin 2\phi & \cos 2\theta \sin 2\phi - \sin 2\theta \cos 2\phi \\ \sin 2\theta \cos 2\phi - \cos 2\theta \sin 2\phi & \sin 2\theta \sin 2\phi + \cos 2\theta \cos 2\phi \end{bmatrix} \\
&= \begin{bmatrix} \cos(2\theta - 2\phi) & -\sin(2\theta - 2\phi) \\ \sin(2\theta - 2\phi) & \cos(2\theta - 2\phi) \end{bmatrix} = [R_{2\theta-2\phi}]_{\mathscr{E}}^{\mathscr{E}}.
\end{aligned}$$

We conclude that the linear map $S_\theta \circ S_\phi$ is the rotation of angle $2\theta - 2\phi$. ■

Question 6.

Let B and D be invertible $n \times n$ matrices over $\mathbf{F}$. Let $\mathscr{B} = (b_1, \dots, b_n)$ and $\mathscr{D} = (d_1, \dots, d_n)$ be bases of $\mathbf{F}^n$, where b_j and d_j are respectively the j^{th} column of B and D. Explain how you can compute $[\mathrm{Id}_{\mathbf{F}^n}]_{\mathscr{B}}^{\mathscr{D}}$ from B and D.

Solution. Let $\mathscr{E}$ be the standard basis of $\mathbf{F}^n$. Since the b_js and d_js are column vectors in $\mathbf{F}^n$ we have $b_j = [b_j]_{\mathscr{E}}$ and $d_j = [d_i]_{\mathscr{E}}$. So

$$B = [b_1\ b_2\ \dots b_n] = [[b_1]_{\mathscr{E}}\ [b_2]_{\mathscr{E}}\ \dots\ [b_n]_{\mathscr{E}}] = [\mathrm{Id}]_{\mathscr{B}}^{\mathscr{E}} \qquad \text{and} \qquad D = [\mathrm{Id}]_{\mathscr{D}}^{\mathscr{E}}.$$

Since D is invertible, we obtain

$$[\mathrm{Id}_{\mathbf{F}^n}]_{\mathscr{B}}^{\mathscr{D}} = [\mathrm{Id}]_{\mathscr{E}}^{\mathscr{D}}[\mathrm{Id}]_{\mathscr{B}}^{\mathscr{E}} = D^{-1}B.$$

■

Tutorial Questions

Question 1.

Let $P \in \mathcal{L}(V)$ be a projection, i.e. $P^2 = P$. Let $b \in V$. Show that there exists a solution $x \in V$ to $Px = b$ if and only if $Pb = b$.

Solution. If $Pb = b$, then $x = b$ is a solution to $Px = b$.

Suppose that x is a solution to $b = Px$. Then applying P to both sides we have $Pb = P^2 x = Px = b$. ■

Question 2.

Suppose that $P \in \mathcal{L}(V)$ is a projection. Prove that $\operatorname{Id} - P \in \mathcal{L}(V)$ is also a projection (it is sometimes called the complementary projection *to P). What is the relationship between the image and kernel of P and $\operatorname{Id} - P$?*

Solution. The identity function $\operatorname{Id} = \operatorname{Id}_V$ is in $\mathcal{L}(V)$. So P is in $\mathcal{L}(V)$ if and only if $\operatorname{Id} - P$ is in $\mathcal{L}(V)$, which proves that $\operatorname{Id} - P$ is linear. We have

$$(\operatorname{Id} - P)^2 = \operatorname{Id}^2 - \operatorname{Id} P - P \operatorname{Id} + P^2 = \operatorname{Id} - 2P + P = \operatorname{Id} - P,$$

showing that $\operatorname{Id} - P$ is a projection.

A vector v is in $\ker(P)$ if and only if $P(v) = 0$, if and only if $(\operatorname{Id} - P)v = v$, if and only if (by Question 1) v is in the image of $\operatorname{Id} - P$. We conclude that $\ker(P) = \operatorname{Im}(\operatorname{Id} - P)$.

Since $\operatorname{Id} - P$ is a projection, we have $\ker(\operatorname{Id} - P) = \operatorname{Im}(\operatorname{Id} - (\operatorname{Id} - P)) = \operatorname{Im}(P)$. ■

Question 3.

Show that the matrix

$$A = \frac{1}{3}\begin{bmatrix} 2 & 1 & 1 \\ 1 & 2 & -1 \\ 1 & -1 & 2 \end{bmatrix}$$

represents a projection map on $\mathbf{F}^3$. Use this to determine whether the equations $Ax = b$ and $Ax = c$ are consistent, where $b = [1, 1, 0]^\mathsf{T}$ and $c = [1, -1, -1]^\mathsf{T}$. Write b and c in the form $u + u'$ where $u \in \operatorname{col} A$ and $u' \in \operatorname{null} A$.

Solution. Recall that a $m \times n$ matrix A, represents the linear map $\mathrm{L}_A \in \mathcal{L}(\mathbf{F}^n, \mathbf{F}^m)$ defined by $\mathrm{L}_A(v) = Av$. The linear map L_A is a projection if and only if $A^2 = A$.

We need to show that $A^2 = A$. A simple computation gives us

$$A^2 = \frac{1}{9}\begin{bmatrix} 6 & 3 & 3 \\ 3 & 6 & -3 \\ 3 & -3 & 6 \end{bmatrix} = A.$$

We verify

$$Ab = \frac{1}{3}\begin{bmatrix}3\\3\\0\end{bmatrix} = \begin{bmatrix}1\\1\\0\end{bmatrix} = b \qquad \text{and} \qquad Ac = \frac{1}{3}\begin{bmatrix}0\\0\\0\end{bmatrix} \neq c.$$

By Question 1, $Ax = b$ is consistent, but $Ax = c$ is not.

We have $\operatorname{col}(A) = \operatorname{Im}(\mathrm{L}_A)$ and $\operatorname{null} A = \ker(\mathrm{L}_A)$. It immediately follows that $b = b + 0$ and $c = 0 + c$ as b belongs to $\operatorname{Im}(\mathrm{L}_A)$ and c belongs to $\ker(\mathrm{L}_A)$. ■

Question 4.

Let $\mathbf{F}$ *be either* $\mathbf{R}$ *or* $\mathbf{C}$ *and let* V *be an* $\mathbf{F}$*-vector space. Suppose that* $P \in \mathcal{L}(V)$ *is an operator. Let* $T = 2P - \mathrm{Id} \in \mathcal{L}(V)$*.*

a. *Prove that* P *is a projection if and only if* $T^2 = \mathrm{Id}$*.*

b. *Suppose that* P *is an orthogonal projection to a line* L *through the origin in* $\mathbf{R}^2$*. Give a geometric description of the operator* T*.*

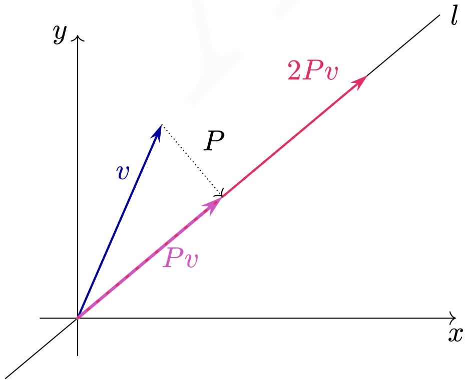
Solution. a. We have $T^2 = (2P - \mathrm{Id})^2 = 4P^2 - 4P + \mathrm{Id} = \mathrm{Id} + 4(P^2 - P)$. Therefore, $T^2 = \mathrm{Id}$ if and only if $P^2 = P$.

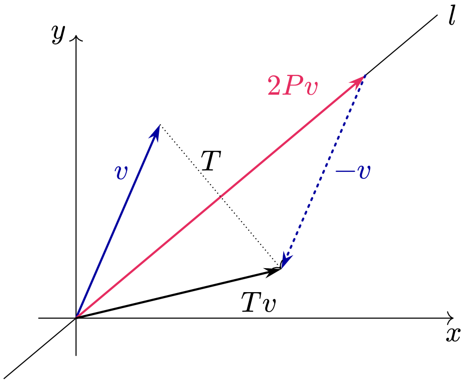
To go further

In any field $\mathbf{F}$, if P is a projection, then $(2P - \mathrm{Id})^2 = \mathrm{Id}$. The converse is not true in general. Indeed, it might happens that $4 = 0$ in $\mathbf{F}$. This is the case in $\mathbf{F}_2 = \{0, 1\}$ for example. If $4 = 0$ in $\mathbf{F}$ then one can show that $2 = 0$ and thus $-1 = 1$. This implies $T = \mathrm{Id}$ is independent of P.

For a concrete example, let $V = (\mathbf{F}_2)^2$. This is a $\mathbf{F}_2$-vector space with 4 elements: $\left[\begin{smallmatrix}0\\0\end{smallmatrix}\right]$, $\left[\begin{smallmatrix}0\\1\end{smallmatrix}\right]$, $\left[\begin{smallmatrix}1\\0\end{smallmatrix}\right]$ and $\left[\begin{smallmatrix}1\\1\end{smallmatrix}\right]$. Then the map $S\left[\begin{smallmatrix}x\\y\end{smallmatrix}\right] = \left[\begin{smallmatrix}y\\x\end{smallmatrix}\right]$ is a linear map but not a projection as $S^2 = \mathrm{Id} \neq S$. However, $(2S - \mathrm{Id})\left[\begin{smallmatrix}x\\y\end{smallmatrix}\right] = \left[\begin{smallmatrix}2y-x\\2x-y\end{smallmatrix}\right] = \left[\begin{smallmatrix}-x\\-y\end{smallmatrix}\right] = \left[\begin{smallmatrix}x\\y\end{smallmatrix}\right]$.

b. T is the orthogonal reflection across L as demonstrated in the following pictures.





■

Question 5.

Let $T \in \mathcal{L}(V, W)$ *and* $S \in \mathcal{L}(W, V)$*. Given* $b \in W$*, consider the linear problem* $Tx = b$*. Prove the following:*

a. *If* $TS = \mathrm{Id}_W$ *then* $x = Sb$ *is a solution to* $Tx = b$ *(in particular, the problem is consistent for all* $b \in W$*).*

b. *If* $ST = \mathrm{Id}_V$ *then* $Tx = b$ *is consistent if and only if* $TSb = b$*; in which case there is a unique solution* $x = Sb$*.*

Solution. a. We have $T(Sb) = (TS)b = \mathrm{Id}\, b = b$.

b. If $b = TSb = T(Sb)$, the system is consistent.

If the system is consistent, then $Tx = b$ for some x. Applying S on both sides we obtain $Sb = STx = x$. This shows both that Sb is a solution and also that it is the unique solution.

■

Further Questions

Question 6.

Let $D \in \mathcal{L}(\mathcal{P}(\mathbf{R}))$ *be the differentiation operator. Find a right-inverse for* D*. Does* D *have a left-inverse?*

Solution. Intuitively, we want the right inverse to be the integration map $\int p(x)\,\mathrm{d}x$. Integration is defined up to an additive constant only. Since we want our map to be linear we probably want to impose $T0 = 0$ and more generally that the constant coefficient of Tp is 0 for all polynomial.

Formally, let $T\colon \mathcal{P}(\mathbf{R}) \to \mathcal{P}(\mathbf{R})$ be the unique linear map such that $T(x^n) = \frac{1}{n+1}x^{n+1}$. We have $T(a_0 + \cdots + a_n x^n) = a_0 x + \cdots + \frac{1}{n+1}x^{n+1}$. One easily verify that we have $DT(p) = p$ for any polynomial.

The map D has no left-inverse, as the equality $SD = \mathrm{Id}$ would implies that D is injective, which is absurd. ■

Question 7.

Suppose $T \in \mathcal{L}(V, W)$ *and* $S \in \mathcal{L}(W, V)$ *satisfy* $TST = T$ *and* $STS = S$*. Show that* TS *and* ST *are both projections. What are their images/kernel? (Hint: The answers can be chosen from* $\ker(S)$, $\mathrm{Im}(S)$, $\ker(T)$, $\mathrm{Im}(T)$*.)*

Solution. We always have $\ker(T) \subseteq \ker(ST)$ and $\operatorname{Im}(ST) \subseteq \operatorname{Im}(S)$ (Tutorial 4, Question 2).

Suppose that $TST = T$. Then $(ST)^2 = STST = S(TST) = ST$ and so ST is a projection. A vector v is in $\ker(ST)$ if and only if $STv = 0$. This implies $0 = TSTv = Tv$. We conclude that $\ker(ST) \subseteq \ker(T)$ as soon as $TST = T$.

Suppose now that $STS = S$. Then $(ST)^2 = STST = (STS)T = ST$ is a projection. Finally, let $v \in \operatorname{Im}(S)$. Then there exists w such that $v = Sw$. But then $STv = STSw = Sw = v$ and therefore $\operatorname{Im}(S) = \operatorname{Im}(ST)$.

The proofs for TS are similar. ■

Tutorial Questions

Question 1.

Let V be an inner product space. Suppose that $u, v \in V$ are vectors satisfying $\langle u,u\rangle = 4$, $\langle v,v\rangle = 6$ and $\langle u,v\rangle = -2$. What is $\|3u-2v\|$?

Solution. We have

$$\begin{aligned}\|3u-2v\|^2 &= \langle 3u-2v, 3u-2v\rangle = 9\langle u,u\rangle - 6\langle u,v\rangle - 6\langle v,u\rangle + 4\langle v,v\rangle \\ &= 9\cdot 4 - 6\cdot(-2) - 6\cdot\overline{-2} + 4\cdot 6 = 84.\end{aligned}$$

So $\|3u-2v\| = \sqrt{84} = 2\sqrt{21}$. ■

Question 2.

Consider the vector space $\mathcal{P}(\mathbf{R})$ equipped with the inner product $\langle p,q\rangle = \int_0^1 p(x)q(x)\,\mathrm{d}x$. Write the polynomial x as a sum $p+q$, where p is a scalar multiple of 1 and q is orthogonal to 1.

Solution. Recall that if V is an inner product space and u and v are two vectors with $v \neq 0$ we have $u = \frac{\langle u,v\rangle}{\langle v,v\rangle}v + w$ with w orthogonal to v. In our case, we have $p = \frac{\langle x,1\rangle}{\langle 1,1\rangle}1$ and $q = x - p$. To find p we compute

$$\langle x,1\rangle = \int_0^1 x\,\mathrm{d}x = \left[\frac{1}{2}x^2\right]_0^1 = \frac{1}{2} \qquad \text{and} \qquad \langle 1,1\rangle = \int_0^1 1\,\mathrm{d}x = [x]_0^1 = 1.$$

So $p = \frac{1}{2}$ and $q = x - \frac{1}{2}$. ■

Question 3.

Prove the converse of the Pythagorean Theorem:

> If u and v are vectors in a real inner product space V satisfying $\|u\|^2 + \|v\|^2 = \|u+v\|^2$ then u and v must be orthogonal.

Explain why the assumption that V is a real inner product space is necessary.

Solution. We have

$$\begin{aligned}\|u+v\|^2 &= \langle u+v, u+v\rangle = \langle u,u\rangle + \langle u,v\rangle + \langle v,u\rangle + \langle v,v\rangle \\ &= \|u\|^2 + \|v\|^2 + \langle u,v\rangle + \langle v,u\rangle.\end{aligned}$$

Therefore, the equality $\|u\|^2 + \|v\|^2 = \|u+v\|^2$ implies (in fact is equivalent to) $0 = \langle u,v\rangle + \langle v,u\rangle = \langle u,v\rangle + \overline{\langle u,v\rangle}$. If V is a real inner product space we conclude $0 = \langle u,v\rangle$. In other words, u and v are orthogonal.

The above proof does not work for complex inner product spaces. But what about the conclusion? It also fails for any complex inner product space $V \neq \{0\}$. Indeed, let $v \neq 0$ be a non-zero vector and let $u = \mathrm{i}v$. Then u and v are not orthogonal since $\langle u, v\rangle = \mathrm{i} \cdot \langle v, v\rangle = \neq 0$. However, we have $\langle u, v\rangle + \langle v, u\rangle = \mathrm{i}\langle v, v\rangle + \bar{\mathrm{i}}\langle v, v\rangle = 0$ and thus $\|u\|^2 + \|v\|^2 = \|u + v\|^2$. ■

Question 4.

Recall that the **mean** *of n real numbers $x_1, \dots, x_n$ is equal to $\frac{1}{n}\sum_{j=1}^n x_j$. Prove that the square of the mean of $x_1, \dots, x_n$ is less than or equal to the mean of $x_1^2, \dots, x_n^2$.*

Solution. We want to prove that

$$\left(\frac{1}{n}\sum_{j=1}^n x_n\right)^2 \leq \frac{1}{n}\sum_{j=1}^n x_n^2,$$

or equivalently that

$$\left(\sum_{j=1}^n x_n\right)^2 \leq n\sum_{j=1}^n x_n^2.$$

In order to do that, we will use the Cauchy-Schwarz inequality: $\left|\langle u, v\rangle\right|^2 \leq \|u\|^2\|v\|^2$ for $u = [x_1, \dots, x_n]^\mathsf{T}$ and $v = [1, \dots, 1]^\mathsf{T}$ in $\mathbf{R}^n$ and the standard dot product and the corresponding norm $\|v\| = \sqrt{v \bullet v}$. One easily check

$$\begin{aligned}(u \bullet v)^2 &= (x_1 + \dots + x_n)^2,\\ \|u\|^2 &= x_1^2 + \dots + x_n^2,\\ \|v\|^2 &= n\end{aligned}$$

and thus the desired formula follows. ■

Question 5.

[From Calculus] Let $\gamma\colon \mathbf{R} \to \mathbf{R}^n$ be a **smooth path**, *i.e., $\gamma(t) = [\gamma_1(t), \dots, \gamma_n(t)]^\mathsf{T}$ where each $\gamma_i\colon \mathbf{R} \to \mathbf{R}$ is a smooth function. Its derivative is given by $\gamma'(t) = [\gamma_1'(t), \dots, \gamma_n'(t)]^\mathsf{T}$. Prove that if γ and δ are smooth paths in $\mathbf{R}^n$, then for any t we have $(\gamma(t) \bullet \delta(t))' = (\gamma'(t) \bullet \delta(t)) + (\gamma(t) \bullet \delta'(t))$ for the standard dot product on $\mathbf{R}^n$.*

Solution. For a fixed t, the dot product is $\gamma(t) \bullet \delta(t) = \sum_{i=1}^n \gamma_i(t)\delta_i(t)$. So the function $t \mapsto (\gamma(t) \bullet \delta(t))$ is a smooth function from $\mathbf{R}$ to $\mathbf{R}$ and

$$\begin{aligned}(\gamma(t) \bullet \delta(t))' &= \left(\sum_{i=1}^n \gamma_i(t)\delta_i(t)\right)' = \sum_{i=1}^n (\gamma_i(t)\delta_i(t))'\\ &= \sum_{i=1}^n (\gamma_i'(t)\delta_i(t) + \gamma_i(t)\delta_i'(t)) = \sum_{i=1}^n \gamma_i'(t)\delta_i(t) + \sum_{i=1}^n \gamma_i(t)\delta_i'(t)\\ &= (\gamma'(t) \bullet \delta(t)) + (\gamma(t) \bullet \delta'(t)).\end{aligned}$$

■

Further Questions

Question 6.

[Continuing from Question 5] Suppose that $\gamma\colon \mathbf{R} \to \mathbf{R}^n$ *is a smooth path of constant speed, i.e.* $\|\gamma'(t)\|$ *is constant. Prove that* $\gamma'(t)$ *and* $\gamma''(t)$ *are orthogonal for all* t*. (This shows that the velocity and acceleration along a constant speed path are always orthogonal.)*

Solution. Since $\|\gamma'(t)\|^2 = \gamma'(t) \bullet \gamma'(t)$ is constant, we have

$$0 = (\gamma'(t) \bullet \gamma'(t))' = (\gamma''(t) \bullet \gamma'(t)) + (\gamma'(t) \bullet \gamma''(t)) = 2\gamma''(t) \bullet \gamma'(t)$$

since $\gamma''(t) \bullet \gamma'(t)$ is a real number. Therefore $0 = \gamma''(t) \bullet \gamma'(t)$ for every $t \in \mathbf{R}$. ■

Question 7.

[From Statistics] Suppose that we have some data given as n *pairs of real numbers* $(x_1, y_1), \dots, (x_n, y_n)$*. The* ***sample correlation coefficient*** *is given by the following formula:*

$$r = \frac{\sum_{i=1}^{n}(x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^{n}(x_i - \bar{x})^2}\sqrt{\sum_{i=1}^{n}(y_i - \bar{y})^2}},$$

where $\bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_n$ *and* $\bar{y} = \frac{1}{n}\sum_{i=1}^{n} y_n$*. Prove that* $-1 \leq r \leq 1$*. What can you say about the value of* r *if:*

a. *A fixed constant is added to all* x *values (or all* y *values)?*

b. *We scale each* x *value (or each* y *value) by a fixed positive constant?*

Solution. Let $x = [x_1, \dots, x_n]^\intercal$, $y = [y_1, \dots, y_n]^\intercal$ and $u = [1, \dots, 1]^\intercal$ be vectors in $\mathbf{R}^n$. Then

$$\bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i = \frac{x \bullet u}{u \bullet u} \quad \text{and} \quad \bar{y} = \frac{1}{n}\sum_{i=1}^{n} y_i = \frac{y \bullet u}{u \bullet u}.$$

Now, let

$$\tilde{x} = x - \frac{x \bullet u}{u \bullet u}u = [x_1 - \bar{x}, \dots, x_n - \bar{x}]^\intercal \quad \text{and} \quad \tilde{y} = y - \frac{y \bullet u}{u \bullet u}u = [y_1 - \bar{y}, \dots, y_n - \bar{y}]^\intercal.$$

We therefore have

$$r = \frac{\tilde{x} \bullet \tilde{y}}{\|\tilde{x}\|\|\tilde{y}\|}.$$

By the Cauchy-Schwarz inequality we have $|\tilde{x} \bullet \tilde{y}| \leq \|\tilde{x}\|\|\tilde{y}\|$, that is: $-1 \leq r \leq 1$.

a. Adding a fixed constant λ to all x_i is the same as replacing x by $x + \lambda u$. But then $\overline{x + \lambda u} = \bar{x} + \lambda$ and $\tilde{x}$ is replaced by

$$\widetilde{x + \lambda u} = [(x_1 + \lambda) - (\bar{x} + \lambda), \dots, (x_n + \lambda) - (\bar{x} + \lambda)]^\mathsf{T} = \tilde{x}.$$

So r remains the same.

b. Scaling each value by a positive λ is the same as replacing x by $\lambda x = [\lambda x_1, \dots, \lambda x_n]^\mathsf{T}$. We have $\overline{\lambda x} = \lambda \bar{x}$ and thus $\widetilde{\lambda x} = \lambda \tilde{x}$. Therefore r becomes

$$r = \frac{\lambda \tilde{x} \bullet \tilde{y}}{\|\lambda \tilde{x}\| \|\tilde{y}\|} = \frac{\lambda \tilde{x} \bullet \tilde{y}}{|\lambda| \|\tilde{x}\| \|\tilde{y}\|} = \frac{\tilde{x} \bullet \tilde{y}}{\|\tilde{x}\| \|\tilde{y}\|}.$$

So r remains the same (observe that it is crucial that λ is positive).

■

Tutorial Questions

For this tutorial, the field $\mathbf{F}$ is either $\mathbf{R}$ or $\mathbf{C}$.

Question 1.

Consider the vector space $\mathcal{P}(\mathbf{R})_2$ with the inner product given by $\langle p, q\rangle = \int_0^1 p(x)q(x)\,\mathrm{d}x$. Apply the Gram-Schmidt procedure to $(x, 1, x^2)$ to obtain an orthonormal basis for $\mathcal{P}(\mathbf{R})$. (Note that the order matters.)

Solution. Let $v_1 = x$, $v_2 = 1$ and $v_3 = x^2$. Then $u_1 = v_1 = x$. We compute

$$\langle u_1, u_1\rangle = \int_0^1 x^2\,\mathrm{d}x = \frac{1}{3} \qquad \text{and} \qquad \langle v_2, u_1\rangle = \int_0^1 x\,\mathrm{d}x = \frac{1}{2}.$$

This gives us

$$u_2 = v_2 - \frac{\langle v_2, u_1\rangle}{\langle u_1, u_1\rangle}u_1 = 1 - \frac{1/2}{1/3}x = 1 - \frac{3}{2}x.$$

We now compute

$$\langle v_3, u_1\rangle = \int_0^1 x^3\,\mathrm{d}x = \frac{1}{4}, \qquad \langle v_3, u_2\rangle = \int_0^1 x^2 - \frac{3}{2}x^3\,\mathrm{d}x = \frac{1}{3} - \frac{3}{2}\frac{1}{4} = -\frac{1}{24},$$
$$\langle u_2, u_2\rangle = \int_0^1 1 - 3x + \frac{9}{4}x^2\,\mathrm{d}x = 1 - 3\frac{1}{2} + \frac{9}{4}\frac{1}{3} = \frac{1}{4}.$$

We then have

$$u_3 = v_3 - \frac{\langle v_3, u_1\rangle}{\langle u_1, u_1\rangle}u_1 - \frac{\langle v_3, u_2\rangle}{\langle u_2, u_2\rangle}u_2 = x^2 - \frac{1/4}{1/3}x - \frac{-1/24}{1/4}\Big(1 - \frac{3}{2}x\Big) = x^2 - x + \frac{1}{6}.$$

We compute a last inner product

$$\langle u_3, u_3\rangle = \int_0^1 x^4 - 2x^3 + \frac{4}{3}x^2 - \frac{1}{3}x + \frac{1}{36}\,\mathrm{d}x = \frac{1}{5} - 2\frac{1}{4} + \frac{4}{3}\frac{1}{3} - \frac{1}{3}\frac{1}{2} + \frac{1}{36} = \frac{1}{180}.$$

Finally, we normalize:

$$e_1 = \frac{u_1}{\|u\|_1} = \sqrt{3}x, \qquad e_2 = \frac{u_2}{\|u\|_2} = 2 - 3x, \qquad e_3 = \frac{u_3}{\|u\|_3} = 6\sqrt{5}\Big(x^2 - x + \frac{1}{6}\Big).$$

■

Question 2.

The ***conjugate transpose*** *of an $m \times n$ matrix A is the $n \times m$ matrix $A^* = \overline{A^\intercal} = \overline{A}^\intercal$.*

a. *Show that $u \bullet v = v^* u$ for vectors u and v in $\mathbf{F}^n$.*

b. *Prove that* $(\operatorname{col} A)^\perp = \operatorname{null} A^*$ *for the standard dot product (where* $\operatorname{col} A$ *is the column space of* A*).*

Solution. a. Let $u = [u_1, \dots, u_n]^\top$ and $v = [v_1, \dots, v_n]^\top$ be any two vectors in $\mathbf{F}^n$. We have

$$u \bullet v = \sum_{j=1}^{n} u_j \bar{v}_j = [\bar{v}_1 \ \dots \ \bar{v}_n] \begin{bmatrix} u_1 \\ \vdots \\ u_n \end{bmatrix} = v^* u.$$

b. Recall that $\operatorname{col}(A) = \{Av \mid v \in \mathbf{F}^n\}$. Now, let u be any vector in $\mathbf{F}^m$. We have

$$\begin{aligned} u \in \operatorname{col}(A)^\perp &\iff \forall v \in \mathbf{F}^n : u \bullet Av = 0 \\ &\iff \forall v \in \mathbf{F}^n : 0 = (Av)^* u = v^* A^* u = A^* u \bullet v \\ &\iff A^* u = 0 \\ &\iff u \in \operatorname{null}(A^*). \end{aligned}$$

■

Question 3.

Let V *be an inner product space and suppose that* $P \in \mathcal{L}(V)$ *is a projection onto a subspace* U*. Prove that* P *is an orthogonal projection if and only if* $\|Pv\| \leq \|v\|$ *for all* $v \in V$*.* (Hint: Show that $\ker P$ must be orthogonal to U.)

Solution. "⇒" Suppose that P is an orthogonal projection, so $P = P_{U,U^\perp}$. Then for any v in V the vectors $Pv \in U$ and $v - Pv \in U^\perp$ are orthogonal. By the Pythagorean theorem, $\|v\|^2 = \|Pv\|^2 + \|v - Pv\|^2$. Since $\|v - Pv\|^2$ is non-negative, we conclude $\|Pv\|^2 \leq \|v\|^2$ and hence $\|Pv\|^2 \leq \|v\|^2$.

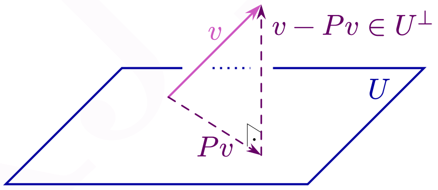


Figure 3: An orthogonal projection of v onto U and the decomposition $v = Pv + (v - Pv)$ with $v - Pv \in U^\perp$.

"⇐" We will give a proof assuming that V is finite dimensional. For the general case, see the "Infinite dimensional vector spaces" box below.

We claim:

If $\|Pv\| \leq \|v\|$ *for all* $v \in V$*, then* $\ker(P)^\perp$ *is contained in* $\operatorname{Im}(P)$*.*

Before proving this claim, let us see how it implies that P is an orthogonal projection (i.e. that $\operatorname{Im}(P)^\perp = \ker(P)$ if V is finite dimensional. Since V is finite dimensional, so is $\ker(P)$.

Therefore $V = \mathrm{Im}(P) \oplus \ker(P) = \ker(P) \oplus \ker(P)^\perp$. Everything being finite dimensional, we conclude $\dim(V) = \dim(\mathrm{Im}\, P) + \dim(\ker P) = \dim(\ker P) + \dim(\ker(P)^\perp)$. So $\dim(\ker(P)^\perp) = \dim(\mathrm{Im}\, P)$, which by the claim implies $\ker(P)^\perp = \mathrm{Im}(P)$.

Let us now prove the claim. Let v be any vector in $\ker(P)^\perp$. We have $Pv = v + (Pv - v)$ with $v \in \ker(P)^\perp$. One easily check $P(Pv - v)) = P^2v - Pv = 0$ so $Pv - v$ is in $\ker(P)$ and hence orthogonal to v. By the Pythagorean Theorem we have $\|Pv\|^2 = \|v\|^2 + \|Pv - v\|^2$. By assumption, we also have $\|Pv\|^2 \leq \|v\|^2$. Altogether, this implies that $\|Pv - v\|^2 = 0$ and hence $v = Pv$ is in the image of P as desired.

Infinite dimensional vector spaces

The above proof works only for V finite dimensional. The result is still true for V infinite dimensional, but requires a different proof.
Since P is a projection onto U, we have $U = \mathrm{Im}(P)$ and $V = \mathrm{Im}(P) \oplus \ker(P)$. We claim:

> *If $\|Pv\| \leq \|v\|$ for all $v \in V$, then $\ker(P)$ is contained in $\mathrm{Im}(P)^\perp$ (and hence P is the orthogonal projection onto $\mathrm{Im}(P)$).*

Let us prove this claim. Let $u \in \mathrm{Im}(P)$ and $w \in \ker(P)$. If $w = 0$, then it is orthogonal to u. We can hence suppose that $w \neq 0$. We want to show that u and w are orthogonal. The trick is to consider the orthogonal projection of u onto $\mathrm{span}(w)$. More precisely, we will evaluate the norm of $v := u - \frac{\langle u, w\rangle}{\langle w, w\rangle} w$. Since u is in the image of P and w in the kernel, we have $P(v) = u$. So by assumption, $\|u\| \leq \|v\|$. We hence have

$$\begin{aligned}
0 \leq \|v\|^2 - \|u\|^2 &= \left\| u - \frac{\langle u, w\rangle}{\|w\|^2} w \right\|^2 - \|u\|^2 \\
&= \left\| \frac{\langle u, w\rangle}{\|w\|^2} w \right\|^2 - \left\langle u, \frac{\langle u, w\rangle}{\|w\|^2} w \right\rangle - \left\langle \frac{\langle u, w\rangle}{\|w\|^2} w, u \right\rangle \\
&= \frac{|\langle u, w\rangle|^2}{\|w\|^2} - \frac{\overline{\langle u, w\rangle}}{\|w\|^2} \langle u, w\rangle - \frac{\langle u, w\rangle}{\|w\|^2} \langle w, u\rangle \\
&= -\frac{|\langle u, w\rangle|^2}{\|w\|^2}.
\end{aligned}$$

We conclude that $\frac{|\langle u, w\rangle|^2}{\|w\|^2} \leq 0$ and thus $\langle u, w\rangle = 0$ as desired.

■

Question 4.

Prove that a $m \times m$ matrix A represents an orthogonal projection (with respect to the standard dot product) if and only if $A^2 = A = A^*$. (Hint: For the only if direction, show that A^* represents a projection and has the same image and kernel as A.

Solution. "$\Leftarrow$" Since $A = A^2$, it represents a projection L_A and $\mathbf{F}^n = \operatorname{Im}(\mathrm{L}_A) \oplus \ker(\mathrm{L}_A) = \operatorname{col}(A) \oplus \operatorname{null}(A)$. By Question 2 $(\operatorname{col} A)^\perp = \operatorname{null}(A^*) = \operatorname{null}(A)$. So A represents an orthogonal projection since $\operatorname{col}(A)$ and $\operatorname{null}(A)$ are orthogonal.

"$\Rightarrow$" Suppose that L_A is an orthogonal projection, so $A^2 = A$ and $\mathbf{F}^n = \operatorname{col}(A) \oplus \operatorname{null}(A)$ with $\operatorname{col}(A)^\perp = \operatorname{null}(A)$. We have

$$(A^*)^2 = (\overline{A^\intercal})^2 = \overline{(A^\intercal)^2} = \overline{(A^2)^\intercal} = \overline{A^\intercal} = A^*,$$

so A^* represents a projection and $\mathbf{F}^n = \operatorname{col}(A^*) \oplus \operatorname{null}(A^*)$. By Question 2 $(\operatorname{col} A)^\perp = \operatorname{null}(A^*)$, and applying this to A^* and $(A^*)^* = A$ we also have $(\operatorname{col} A^*)^\perp = \operatorname{null}(A)$, or in other words $\operatorname{col} A^* = (\operatorname{null} A)^\perp$. Since A is an orthogonal projection, we have $\operatorname{null}(A^*) = (\operatorname{col} A)^\perp = \operatorname{null}(A)$ and $\operatorname{col}(A^*) = (\operatorname{null} A)^\perp = \operatorname{col}(A)$. That is, A^* represents the orthogonal projection onto $\operatorname{col}(A)$ with direction parallel to $\operatorname{null}(A)$. By unicity of such a projection, we conclude $A^* = A$. ■

Further Questions

Question 5.

Let A and B be matrices satisfying $ABA = A$, $BAB = B$, $(AB)^* = AB$ and $(BA)^* = BA$. Prove that AB and BA are both orthogonal projection matrices. What are their respective images? (Hint: see Tutorial 7 Question 7.)

Solution. By Tutorial 7 Question 7, AB represents a projection onto $\operatorname{col}(AB)$ along direction $\operatorname{null}(AB)$, $\operatorname{col}(AB) = \operatorname{col}(A)$ and $\operatorname{null}(B) = \operatorname{null}(AB)$. Altogether, AB represents a projection onto $\operatorname{col}(A)$ along direction $\operatorname{null}(B)$. Finally, $(AB)^* = AB$ implies (Question 4) that the projection represented by AB is orthogonal. ■

Question 6.

Let $T \in \mathcal{L}(V, W)$ be a linear map between finite-dimensional inner product spaces.

- *a. Show that for each $w \in \operatorname{Im}(T)$ there exists a unique $v \in \ker(T)^\perp$ such that $Tv = w$.*
- *b. Prove that there is a unique linear map $S \in \mathcal{L}(W, V)$ satisfying $TST = T$, $STS = S$, and TS, ST are orthogonal projections to $\operatorname{Im}(T)$, $\ker(T)^\perp$ respectively. (This S is called the* ***pseudo-inverse*** *of T.)*

Solution. a. Since w is in the image of T, there exists $u \in V$ such that $Tu = w$. Since V is finite dimensional, we have $V = \ker(T) \oplus \ker(T)^{\perp}$. Therefore, there exists (unique) $z \in \ker(T)$ and $v \in \ker(T)^{\perp}$ such that $u = z + v$. We have $w = Tu = Tz + Tv = 0 + Tv = Tv$. This finishes the proof of existence. Now, suppose that v' also satisfies $Tv' = w$ and $v' \in \ker(T)^{\perp}$. Then $T(v - v') = 0$ implies that $v - v'$ belongs to $\ker(T) \cap \ker(T)^{\perp} = \{0\}$. In other words, we necessarily have $v' = v$, showing unicity.

b. By finite dimension, we have $W = \operatorname{Im}(T) \oplus \operatorname{Im}(T)^{\perp}$. Use this to define a function

$$\begin{aligned} S\colon W = \operatorname{Im}(T) \oplus \operatorname{Im}(T)^{\perp} &\longrightarrow V \\ w + z &\longmapsto \text{the unique } v \in \ker(T)^{\perp} \text{ such that } Tv = w. \end{aligned}$$

The existence and unicity of v makes this is a well-defined function. We now need to show it is linear. So let $w_1 + z_1$ and $w_2 + z_2$ be two vectors in $W = \operatorname{Im}(T) \oplus \operatorname{Im}(T)^{\perp}$ and let λ be a scalar. Let v_1 and v_2 be the unique $v_i \in \ker(T)^{\perp}$ such that $Tv_i = w_i$. Then $\lambda v_1 + v_2$ is in $\ker(T)^{\perp}$ and $T(\lambda v_1 + v_2) = \lambda w_1 + w_2$. It follows that $S(\lambda(w_1 + z_1) + (w_2 + z_2)) = \lambda v_1 + v_2 = \lambda S(w_1 + z_1) + S(w_2 + z_2)$. This shows the linearity of S.

For any $w + z$ in $W = \operatorname{Im}(T) \oplus \operatorname{Im}(T)^{\perp}$ we have $STS(w + z) = ST(v) = S(w) = S(w + z)$. For any $v \in V$, one has $TSTv = TSw = Tv$. It follows (Tutorial 7 Question 7) that TS is a projection onto $\operatorname{Im}(T)$ along direction $\ker(S)$, while ST is a projection onto $\operatorname{Im}(S)$ along direction $\ker(T)$.

For orthogonality, we need to show that $\ker(S)^{\perp} = \operatorname{Im}(T)$ and $\ker(T)^{\perp} = \operatorname{Im}(S)$. Let z be a vector in $\operatorname{Im}(T)^{\perp}$, therefore we have $z = 0 + z \in Im(T) \oplus Im(T)^{\perp}$. Then $S(z)$ is the unique w in $\ker T^{\perp}$ such that $T(w) = 0$ and therefore $w \in \ker T \cap \ker T^{\perp}$ and $S(z) = w = 0$. This proves $\operatorname{Im}(T)^{\perp} \subseteq \ker(S)$.

Now, let $w + z$ be any element in $\ker(S)$. So $0 = S(w + z) = S(w)$ satisfies $T0 = w$. So $w + z = 0 + z$ is in $\operatorname{Im}(T)^{\perp}$. This proves the second inclusion $\ker(S) \subseteq \operatorname{Im}(T)^{\perp}$.

We have $\operatorname{Im}(S) \subseteq \ker(T)^{\perp}$ by definition of S. Now, for v in $\ker(T)^{\perp}$ we have Tv in $\operatorname{Im}(T)$ and hence $S(Tv) = v$, so v belongs to $\operatorname{Im}(S)$. This proves the inclusion $\ker(T)^{\perp} \subseteq \operatorname{Im}(S)$.

We proved that there exists an S with the desired properties. To finish the proof, we need to show that such an S is unique. Suppose that Q also satisfies the same properties. Then both TS and TQ are orthogonal projections to $\operatorname{Im}(T)$ along direction $\operatorname{Im}(T)^{\perp}$. By unicity of orthogonal projections, $TS = TQ$. One prove similarly $ST = QT$. Finally, $S = STS = STQ = QTQ = Q$.

■

Tutorial Questions

Question 1.

Consider the matrix $A = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}$.

a. *Find the eigenvalues of* A *in* $\mathbf{C}$.

b. *For what values of* θ *does* A *have real eigenvalues?*

Solution. a. First we compute the characteristic polynomial:

$$\begin{aligned} \chi_A(t) &= \det \begin{bmatrix} t-\cos\theta & \sin\theta \\ -\sin\theta & t-\cos\theta \end{bmatrix} \\ &= (t-\cos\theta)^2 + \sin^2\theta = t^2 - 2t\cos\theta + \cos^2\theta + \sin^2\theta \\ &= t^2 - 2t\cos\theta + 1 = (t-\lambda_1)(t-\lambda_2). \end{aligned}$$

Using Vieta's formula, we conclude:

$$\lambda_\pm = \frac{-(-2\cos\theta) \pm \sqrt{(2\cos\theta)^2 - 4\cdot 1\cdot 1}}{2\cdot 1} = \cos\theta \pm \sin\theta \mathrm{i}.$$

b. A has real eigenvalues if and only if $\sin\theta = 0$ if and only if $\theta = k\pi$ for some $k \in \mathbf{Z}$.

■

Question 2.

Let V *be a vector space and let* $P = P_{U,W} \in \mathcal{L}(V)$ *be a projection to a subspace* U *along a complementary subspace* W.

a. *What are the possible eigenvalues of* P*?*

b. *What are the eigenspaces of* P*?*

Solution. a. By assumption $V = \operatorname{Im}(P) \oplus \ker(P)$. Then $Pv = \lambda v$ and $P^2 v = P\lambda v = \lambda P v = \lambda^2 v$. Since P is a projection, $P^2 = P$ and so $Pv = \lambda^2 v$. We conclude that $\lambda^2 v = \lambda v$ and since $v \neq 0$ that $\lambda^2 = \lambda$. Therefore $\lambda \in \{0,1\}$.

b. The possible eigenvalues of P are 0 and 1. We have $\mathrm{E}(0,P) = \ker(P - 0\,\mathrm{Id}) = \ker(P)$ and $\mathrm{E}(1,P) = \ker(P - 1\,\mathrm{Id}) = \operatorname{Im}(P)$.

■

Question 3.

Suppose that $T \in \mathcal{L}(\mathbf{F}^4)$ *is an operator such that* $\dim \mathrm{E}(3,T) = 3$. *Prove that at least one of* $T - 2\,\mathrm{Id}$ *or* $T - 4\,\mathrm{Id}$ *is invertible.*

Solution. Since $\mathbf{F}^4$ is finite dimensional, the operator $T - \lambda \operatorname{Id}$ is not invertible if and only if λ is an eigenvalue of T, if and only if $\dim \mathrm{E}(\lambda, T) \geq 1$. We also have that $4 = \dim \mathbf{F}^4 \geq \sum_{j=1}^{k} \mathrm{E}(\lambda_j, T)$ where the sum is taken over all the distinct eigenvalues. We already know that 3 is an eigenvalue, let us say $\lambda_1 = 3$, so

Proposition 5.1.11

$$4 = \dim \mathbf{F}^4 \geq 3 + \sum_{j=2}^{k} \mathrm{E}(\lambda_j, T).$$

We conclude that T has at most one other eigenvalue $\lambda \neq 3$ and thus $T - \lambda \operatorname{Id}$ is *not* invertible for at most one $\lambda \neq 3$. In particular, at least one of $T - 2\operatorname{Id}$ or $T - 4\operatorname{Id}$ is invertible. ■

Question 4.

Define an operator $T \in \mathcal{L}(\mathcal{P}(\mathbf{R})_3)$ by $Tp(x) = xp'(x)$ for all $x \in \mathbf{R}$. Find all eigenvalues and eigenvectors of T.

Solution. Trying to solve $\lambda p(x) = xp'(x)$ to find an eigenvector $p \neq 0$ is a priori difficult. So we will instead compute the characteristic polynomial of T. Let $\mathscr{B} = (1, x, x^2, x^3)$ be the standard basis of $\mathcal{P}(\mathbf{R})_3$. We have $T(1) = 0$, $T(x) = x$, $Tx^2 = 2x^2$ and $Tx^3 = 3x^3$. So

$$[T]_{\mathscr{B}}^{\mathscr{B}} = \begin{bmatrix} 0 & & & \\ & 1 & & \\ & & 2 & \\ & & & 3 \end{bmatrix}.$$

Since this is a diagonal matrix, we conclude that the eigenvalues of T are 0, 1, 2 and 3.

Eigenvalue $\lambda = 0$ We have $\operatorname{null}([T]_{\mathscr{B}}^{\mathscr{B}} - 0\operatorname{Id}) = \operatorname{null}([T]_{\mathscr{B}}^{\mathscr{B}}) = \{[a, 0, 0, 0]^{\mathsf{T}} \mid a \in \mathbf{R}\}$. So the $\mathrm{E}(0, T) = \ker(T - 0\operatorname{Id}) = \{a \mid a \in \mathbf{R}\}$ (the constant polynomials). The eigenvectors are $\mathrm{E}(0, T) \setminus \{0\}$.

Eigenvalue $\lambda = 1$ We have

$$\operatorname{null}([T]_{\mathscr{B}}^{\mathscr{B}} - 1\operatorname{Id}) = \operatorname{null} \begin{bmatrix} -1 & & & \\ & 0 & & \\ & & 1 & \\ & & & 2 \end{bmatrix} = \left\{ \begin{bmatrix} 0 \\ b \\ 0 \\ 0 \end{bmatrix} \middle| \, b \in \mathbf{R} \right\}.$$

So the $\mathrm{E}(1, T) = \ker(T - 1\operatorname{Id}) = \{bx \mid b \in \mathbf{R}\}$. The eigenvectors are $\mathrm{E}(1, T) \setminus \{0\}$.

Eigenvalues $\lambda = 2$ **and** $\lambda = 3$ The same method as above gives $\mathrm{E}(2, T) = \ker(T - 1\operatorname{Id}) = \{cx^2 \mid c \in \mathbf{R}\}$ and $\mathrm{E}(3, T) = \ker(T - 1\operatorname{Id}) = \{dx^3 \mid d \in \mathbf{R}\}$. ■

Question 5.

Let $V = \mathcal{C}^\infty(\mathbf{R})$ be the space of smooth real functions and let $D \in \mathcal{L}(V)$ be the differentiation operator. Show that every $\lambda \in \mathbf{R}$ is an eigenvalue for D, and find a corresponding eigenvector.

Solution. For every $\lambda \in \mathbf{R}$ we have $De^{\lambda x} = \lambda e^{\lambda x}$. So $e^{\lambda x}$ is an eigenvector with eigenvalue λ. ■

Further Questions

Question 6.

Le $\lambda_1, \dots, \lambda_k$ *be distinct real numbers. Show that the list* $(e^{\lambda_1 x}, \dots, e^{\lambda_k x})$ *is linearly independent in* $\mathcal{C}^\infty(\mathbf{R})$*.* Hint: Consider the differentiation operator restricted to the subpsace $U = \operatorname{span}(e^{\lambda_1 x}, \dots, e^{\lambda_k x})$ of $\mathcal{C}^\infty(\mathbf{R})$.

Solution. Since $De^{\lambda_j x} = \lambda_j e^{\lambda_j x}$, the subspace U is D invariant. Therefore, $D|_U$ is an operator on U. The λ_j for $j \in \{1, \dots, k\}$ are all eigenvalues of $D|_U$. This implies that U has dimension at least k. Therefore, U has dimension exactly k and $(e^{\lambda_1 x}, \dots, e^{\lambda_k x})$ is a basis for U (and hence linearly independent). ■

Question 7.

Let V *be a finite dimensional vector space and suppose that* $T \in \mathcal{L}(V)$ *has* $\dim(V)$ *distinct eigenvalues. Let* $S \in \mathcal{L}(V)$ *be an operator with the same eigenvectors as* T*, but not necessarily the same eigenvalues. Prove that* $TS = ST$*.*

Solution. Let $m := \dim(V)$ and let $\lambda_1, \dots, \lambda_m$ be the eigenvalues of T and $v_1, \dots, v_m$ be the corresponding eigenvectors. Since the v_j correspond to distinct eigenvalues, they are linearly independent and therefore $\mathscr{B} = (v_1, \dots, v_m)$ is a basis of V.

By assumption, there exists $\mu_1, \dots, \mu_m$ such that $Sv_j = \mu_j v_j$. Now, $TSv_j = T\mu_j v_j = \lambda_j \mu_j v_j$ while $STv_j = S\lambda_j v_j = \mu_j \lambda_j v_j = \lambda_j \mu_j v_j = TSv_j$. Since ST and TS agree on a basis, they are equal. ■

Tutorial Questions

Question 1.

Consider the operator map $T \in \mathcal{L}(\mathbf{C}^3)$ given by $T(x) = Ax$ where

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

Find the eigenvalues and the generalised eigenspaces of T.

Solution. We first compute the characteristic polynomial. Since A is a triangular matrix,$\chi_T(t) = (t-2)^2(t-3)$ and the eigenvalues are $\lambda_1 = 2$ (of algebraic multiplicity 2) and $\lambda_2 = 3$ (of algebraic multiplicity 1).

For $\lambda_1 = 2$

$$2\operatorname{Id} - A = \begin{bmatrix} 0 & -1 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & -1 \end{bmatrix} \quad \text{and} \quad (2\operatorname{Id} - A)^2 = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

We conclude

$$\mathrm{G}(2, T) = \operatorname{null}((2\operatorname{Id} - A)^2) = \left\{ \begin{bmatrix} z_1 \\ z_2 \\ 0 \end{bmatrix} \middle| \, z_1, z_2 \in \mathbf{C} \right\}.$$

For $\lambda_2 = 3$

$$3\operatorname{Id} - A = \begin{bmatrix} 1 & -1 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix}.$$

We conclude

$$\mathrm{G}(3, T) = \operatorname{null}(3\operatorname{Id} - A) = \left\{ \begin{bmatrix} 0 \\ 0 \\ z_3 \end{bmatrix} \middle| \, z_3 \in \mathbf{C} \right\}.$$

■

Question 2.

Let T be an operator on $\mathbf{C}^5$ whose eigenvalues are 3, 4, and 5. Prove that $(T - 3\operatorname{Id})^3(T - 4\operatorname{Id})^3(T - 5\operatorname{Id})^3 = 0$.

Solution. Since T is a complex operator with eigenvalues 3, 4 and 5, its characteristic polynomial is $\chi_T(t) = (t-3)^{\mathrm{m_{alg}}(3,T)}(t-4)^{\mathrm{m_{alg}}(4,T)}(t-5)^{\mathrm{m_{alg}}(3,T)}$ with all the $\mathrm{m_{alg}}(\lambda_j, T) \geq 1$ and $5 = \mathrm{m_{alg}}(3,T) + \mathrm{m_{alg}}(4,T) + \mathrm{m_{alg}}(5,T)$. We conclude that for $\lambda \in \{3,4,5\}$ we have $\mathrm{m_{alg}}(\lambda, T) \leq 3$.

The minimal polynomial $\mathrm{M}_T(t)$ divides $\chi_T(t)$ which itself divides $(t-3)^3(t-4)^3(t-5)^3$ by the above. Since $\mathrm{M}_T(T) = 0$ we conclude $(T - 3\,\mathrm{Id})^3(T - 4\,\mathrm{Id})^3(T - 5\,\mathrm{Id})^3 = 0$. ■

Question 3.

Let V be an m-dimensional complex vector space. Suppose that P is a projection onto an l-dimensional subspace of V. Show that $\chi_P(t) = (t-1)^l t^{m-l}$.

Solution. We know (see Tutorial 10, Question 2) that the only possible eigenvalues of P are 0 and 1 and that $\mathrm{E}(1, P) = \mathrm{Im}(P)$ while $\mathrm{E}(0, P) = \ker(P)$. Moreover, we have $V = \mathrm{G}(1, P) \oplus \mathrm{G}(0, P) = \mathrm{Im}\, P \oplus \ker P = \mathrm{E}(1, P) \oplus \mathrm{E}(0, P)$ (we use that P is a projection for the second equality). Therefore, we must have $\mathrm{E}(1, P) = \mathrm{G}(1, P)$ and $\mathrm{E}(0, P) = \mathrm{G}(0, P)$.

We conclude that $\mathrm{m}_{\mathrm{alg}}(1, P) = \dim \mathrm{G}(1, P) = \dim \mathrm{E}(1, P) = l$, while $\mathrm{m}_{\mathrm{alg}}(0, P) = \dim \ker(P) = \dim V - \dim \mathrm{Im}\, P = m - l$. This gives $\chi_P(t) = (t-1)^l t^{m-l}$ as desired. ■

Question 4.

Let $T \in \mathcal{L}(V)$ be an operator on a finite-dimensional complex vector space V. Suppose that the characteristic and minimal polynomials of T are respectively $\chi_T(z) = (z-2)(z+1)^4(z-3)^3$ and $\mathrm{M}_T(z) = (z-2)(z+1)^3(z-3)$. Write down a Jordan form of T.

Solution. From the characteristic polynomial, the eigenvalues of T are 2, -1 and 3.

The sum of the sizes of the Jordan blocks with eigenvalue 2 is $1 = \mathrm{m}_{\mathrm{alg}}(2)$. So for $\lambda = 2$ we only have one block of size 1.

For $\lambda = -1$, the sum of the sizes of the Jordan blocks is $\mathrm{m}_{\mathrm{alg}}(-1) = 4$, while the size of the biggest block is 3 (the exponent of $(z+1)$ in the minimal polynomial). This means that we have one block of size 3 and one of size 1.

Finally, for $\lambda = 3$, the sum of the sizes of the Jordan blocks is $\mathrm{m}_{\mathrm{alg}}(3) = 3$, while the size of the biggest block is 1 (the exponent of $(z-3)$ in the minimal polynomial). This means that we have 3 blocks of size 1.

Altogether, this gives us

$$\begin{bmatrix}
\boxed{2} & & & & & & & \\
& -1 & 1 & & & & & \\
& & -1 & 1 & & & & \\
& & & -1 & & & & \\
& & & & \boxed{-1} & & & \\
& & & & & \boxed{3} & & \\
& & & & & & \boxed{3} & \\
& & & & & & & \boxed{3}
\end{bmatrix}$$

■

Question 5.

Consider the matrix

$$A = \begin{bmatrix} 6 & 6 & -15 \\ 1 & 5 & -5 \\ 1 & 2 & -2 \end{bmatrix}.$$

a. *Find the characteristic polynomial and minimal polynomial of A*

b. *What is the Jordan form for A?*

c. *(Optional) Find a Jordan basis for A.*

Solution. a. We start by computing the characteristic polynomial:

$$\begin{aligned}\chi_A(t) &= \det \begin{bmatrix} t-6 & -6 & 15 \\ -1 & t-5 & 5 \\ -1 & -2 & t+2 \end{bmatrix} \\ &= (t-6)\det \begin{bmatrix} t-5 & 5 \\ -2 & t+2 \end{bmatrix} + 1\det \begin{bmatrix} -6 & 15 \\ -2 & t+2 \end{bmatrix} - 1\det \begin{bmatrix} -6 & 15 \\ t-5 & 5 \end{bmatrix} \\ &= (t-6)((t-5)(t+2)+10) + (-6(t+2)+30) - (-30-15(t-5)) \\ &= t^3 - 9t^2 + 27t - 27 = (t-3)^3.\end{aligned}$$

So A has only one eigenvalue, 3, of algebraic multiplicity 3. The geometric multiplicity of 3 is therefore 1, 2 or 3. We test

$$A - 3\,\mathrm{Id} = \begin{bmatrix} 3 & 6 & -15 \\ 1 & 2 & -5 \\ 1 & 2 & -5 \end{bmatrix} \neq 0,$$

and

$$(A - 3\,\mathrm{Id})^2 = \begin{bmatrix} 3 & 6 & -15 \\ 1 & 2 & -5 \\ 1 & 2 & -5 \end{bmatrix}^2 = 0.$$

We conclude that $\mathrm{m_{geo}}(3, A) = 2$ and $\mathrm{M}_A(t) = (t-3)^2$.

b. The Jordan form for A has one Jordan block of size $\mathrm{m_{geo}}(3, A) = 2$, and therefore one more block of size 1. That is, the Jordan form is

$$J = \begin{bmatrix} 3 & 1 & \\ & 3 & \\ & & 3 \end{bmatrix}.$$

c. To start, we need to find a vector v_1 in $\operatorname{null}(A - 3\operatorname{Id})^2 = \mathbf{F}^3$ but not in $\operatorname{null}(A - 3\operatorname{Id})$. We can for example take $v_1 = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$. Then we compute

$$v_2 := (A - 3\operatorname{Id})v_1 = \begin{bmatrix} 3 \\ 1 \\ 1 \end{bmatrix} \in \operatorname{null}(A - 3\operatorname{Id}).$$

Finally, we need to complete v_2 to a basis of $\operatorname{null}(A - 3\operatorname{Id})$. That is, we need to find a vector v_3 in $\operatorname{null}(A - 3\operatorname{Id}) = \left\{ \begin{bmatrix} x \\ y \\ z \end{bmatrix} \,\middle|\, x + 2y - 5z = 0 \right\}$ but not in $\operatorname{span} v_2$. For example, we can take $v_3 = \begin{bmatrix} 5 \\ 0 \\ 1 \end{bmatrix}$. A Jordan basis is then given by (v_1, v_2, v_3).

Remark 0.2.

Let P be the matrix $P = [v_1\ v_2\ v_3]$. We can easily verify that

$$AP = [Av_1\ Av_2\ Av_3] = [3v_1\ v_1 + 3v_2\ 3v_3] = [v_1\ v_2\ v_3] \begin{bmatrix} 3 & 1 & \\ & 3 & \\ & & 3 \end{bmatrix} = PJ.$$

That is, $P^{-1}AP = J$.

■

Further Questions

Question 6.

Let $N \in \mathcal{L}(V)$ be a nilpotent operator.

a. *Prove that N can only have 0 as an eigenvalue.*

b. *Show that $\operatorname{Id}_V - N$ is invertible.*

Solution. a. Let $v \neq 0$ be an eigenvector: $Nv = \lambda v$. Since N is nilpotent, there exists k with $N^k = 0$ and so $0 = N^k v = \lambda^k v$. Since $v \neq 0$ we necessarily have $\lambda^k = 0$, which implies that $\lambda = 0$.

b. Let k be an integer such that $N^k = 0$, then the inverse of $\operatorname{Id} - N$ is $\operatorname{Id} + N + \cdots + N^{k-1}$. This can easily be directly verified. The intuition to find it comes from the power series $\sum_{i=0}^{\infty} x^i = \frac{1}{1-x}$ for $|x| < 1$.

■

Question 7.

Let V be a finite-dimensional complex vector space. Suppose that $T \in \mathcal{L}(V)$ is an operator with only 0 as an eigenvalue.

a. Show that T is nilpotent.

b. What is the characteristic polynomial of T?

Solution. a. The minimal polynomial of T is $\mathrm{M}_T(t) = t^k$ for some $k \in \{1, \dots, \dim(V)\}$. Then $0_V = \mathrm{M}_T(T) = T^k$ and so T is nilpotent.

b. Since V is a complex vector space, its characteristic polynomial is

$$\chi_T(t) = (t - \lambda_1) \cdots (t - \lambda_{\dim(V)}) = t^{\dim(V)}.$$

■

Tutorial Questions

Question 1.

Let $V \neq \{0\}$ be a finite dimensional complex vector space, and let $T \in \mathcal{L}(V)$ be an operator. Prove that there does not exist a direct sum decomposition of V into two nonzero subspaces invariant under T if and only if $\mathrm{M}_T(z) = (z-\lambda)^{\dim V}$ for some $\lambda \in \mathbf{C}$.

Solution. Firstly, remind that the exponent of $z - \lambda$ in $\mathrm{M}_T(z)$ is the size of the biggest block of eigenvalue λ in a Jordan normal form of T. Therefore, $\mathrm{M}_T(z) = (z-\lambda)^{\dim V}$ if and only if the Jordan normal form of T has only one Jordan block (of eigenvalue λ).

Suppose that $V = V_1 \oplus V_2$ with V_1 and V_2 nonzero and T-invariant. For $j \in \{1, 2\}$, let $T_i := T|_{V_i}$ and let $\mathscr{B}_i$ be a Jordan basis for T_i (so $\mathscr{B}_i$ is a basis of V_i). Then $\mathscr{B} := \mathscr{B}_1 \cup \mathscr{B}_2$ is a basis of V for which

$$[T]_{\mathscr{B}}^{\mathscr{B}} = \begin{bmatrix} [T_1]_{\mathscr{B}_1}^{\mathscr{B}_1} & 0 \\ 0 & [T_2]_{\mathscr{B}_2}^{\mathscr{B}_2} \end{bmatrix}.$$

So $[T]_{\mathscr{B}}^{\mathscr{B}}$ is a Jordan normal form for T, with at least two blocks.

For the other direction, suppose that the Jordan normal form of T has at least two blocks and let $\mathscr{B}$ be a corresponding Jordan basis. Let $\mathscr{B}_1$ be the first vectors of $\mathscr{B}$ corresponding to the first block and let $\mathscr{B}_2 := \mathscr{B} \setminus \mathscr{B}_1$. So we have

$$[T]_{\mathscr{B}}^{\mathscr{B}} = \begin{array}{c} \begin{bmatrix} \boxed{J_1} & 0 \\ 0 & J_2 \end{bmatrix} \\ \underbrace{\hphantom{J_1}}_{\mathscr{B}_1}\underbrace{\hphantom{J_2}}_{\mathscr{B}_2} \end{array}.$$

where J_1 is a Jordan block and J_2 is composed of one or more Jordan blocks. Let $V_i := \mathrm{span}(\mathscr{B}_i)$. By construction, $V = V_1 \oplus V_2$. Moreover, the dimension of $V_i \neq \{0\}$ is equal to the size of J_i and hence at least 1. Finally, the fact that $[T]_{\mathscr{B}}^{\mathscr{B}}$ is diagonal by blocks, with diagonal blocks J_1 and J_2 is equivalent to the fact that V_1 and V_2 are T invariant. ■

Question 2.

Let $T \in \mathcal{L}(V)$ be an operator. Suppose that $\lambda \in \mathbf{F}$ is not an eigenvalue of T. Show that $\mathrm{G}(\lambda, T) = \{0\}$.

Solution. ■

Further Questions

Question 3.

Let $T \in \mathcal{L}(V)$ be an operator and $v \in V$ be a vector such that $T^{m-1}v \neq 0$ but $T^k v = 0$ for some $k \geq 1$. Prove that $v, Tv, \ldots, T^{k-1}v$ are linearly independent.

Solution. ■

Question 4.

Let $T \in \mathcal{L}(V)$.

a. *Show that $V = \mathrm{Im}(T^0) \supseteq \mathrm{Im}(T^1) \supseteq \cdots \supseteq \mathrm{Im}(T^k) \supseteq \mathrm{Im}((T^{k+1}) \cdots$.*

b. *Suppose that $\mathrm{Im}(T^k) = \mathrm{Im}(T^{k+1})$ for some $k \geq 0$. Show that $\mathrm{Im}(T^l) = \mathrm{Im}(T^l)$ for all $l > k$.*

c. *Suppose that $\dim V = m$. Prove that $\mathrm{Im}(T^m) = \mathrm{Im}(T^{m+1}) = \mathrm{Im}(T^{m+2}) = \cdots$.*

Solution. ■

Question 5.

Let $q(z)$ be a monic complex polynomial. Suppose that $p(z)$ is a monic polynomial multiple of $q(z)$ with the same roots as $q(z)$ (but not necessarily with the same multiplicities). Show that there exists an operator $T \in \mathcal{L}(\mathbf{C}^{\deg(p)})$, such that $\chi_T(z) = p(z)$ and $\mathrm{M}_T(z) = q(z)$.

Solution. ■

Question 6.

Let T be an invertible operator on an m-dimensional complex vector space V. Prove that $\chi_{T^{-1}}(z) = \frac{1}{\chi_T(0)} z^m \chi_T(\frac{1}{z})$.

Solution. ■