Solutions & Notes of Homework 3

Junda Wu, Yang He

June 4, 2025

It is my experience that proofs involving matrices can be shortened by 50% if one throws the matrices out.

—*Emil Artin*

Attention: *Theorem 3.10.5 in the lecture notes* **is not correct**. *It should be modified as follows:*

Theorem 1 (linear transformation lemma) *Suppose that $v_1, \ldots, v_n$ is a basis of V and $w_1, \ldots, w_n \in W$. Then there is a unique linear transformation $T : V \to W$ such that $Tv_i = w_i$.*

In fact, the original theorem in lecture notes constructs an **invertible** *linear transformation. We formulate it as the following theorem (pay attention to the dimensions). Try to figure out why. You can also solve Exercise 3 in Section 3D in Linear Algebra Done Right.*

Theorem 2 *Suppose that $v_1, \ldots, v_n$ is a basis of V and $w_1, \ldots, w_n$ is a basis of W. Then the linear transformation $T : V \to W$ such that $Tv_i = w_i$ is invertible.*

Note 0.1: *The theorem numbered in the form x.x is sourced from Linear Algebra Done Right (fourth Edition).*

Problem 1

Suppose $\mathbf{v}_1, \ldots, \mathbf{v}_m$ is a list of vectors in V. Define $T \in \mathcal{L}(\mathbb{R}^m, V)$ by

$$T(\mathbf{x}) = x_1\mathbf{v}_1 + \cdots + x_m\mathbf{v}_m,$$

for $\mathbf{x} = \begin{bmatrix} x_1 \\ \vdots \\ x_m \end{bmatrix} \in \mathbb{R}^m$. (Injective or Surjective.)

(a) What property of T corresponds to $\mathbf{v}_1, \ldots, \mathbf{v}_m$ spanning V? Why?

(b) What property of T corresponds to $\mathbf{v}_1, \ldots, \mathbf{v}_m$ being linearly independent? Why?

Solution: (a) Surjective.

If T is surjective, then $\operatorname{range} T = V$, which implies for any $\mathbf{v} \in V$, there exists an $\mathbf{x} \in \mathbb{R}^m$ such that $T(\mathbf{x}) = \mathbf{v}$, thus every $\mathbf{v} \in V$ can be written as $\mathbf{v} = x_1\mathbf{v}_1 + \cdots + x_m\mathbf{v}_m$, i.e., a linear combination of $\mathbf{v}_1, \ldots, \mathbf{v}_m$. Therefore, $\mathbf{v}_1, \ldots, \mathbf{v}_m$ span V.

On the other hand, if $\mathbf{v}_1, \ldots, \mathbf{v}_m$ span V, then for any $\mathbf{v} \in V$, we have $\mathbf{v} = a_1\mathbf{v}_1 + \cdots + a_m\mathbf{v}_m$ ($a_1, \ldots, a_m \in \mathbb{R}$), which implies $T\left(\begin{bmatrix} a_1 & \cdots & a_m \end{bmatrix}^{\mathrm{T}}\right) = \mathbf{v}$. So for any $\mathbf{v} \in V$, there exists a certain vector $\mathbf{x} \in \mathbb{R}^m$ such that $T(\mathbf{x}) = \mathbf{v}$, i.e., T is surjective.

(b) Injective.

If T is injective, then $\operatorname{null} T = \{\mathbf{0}\}$, which implies $T(\mathbf{x}) = x_1\mathbf{v}_1 + \cdots + x_m\mathbf{v}_m = \mathbf{0}$ if and only if $\mathbf{x} = \mathbf{0}$, i.e., $x_1 = x_2 = \cdots = x_m = 0$. Therefore $\mathbf{v}_1, ..., \mathbf{v}_m$ is linearly independent.

On the other hand, if $\mathbf{v}_1, ..., \mathbf{v}_m$ is linearly independent, let $\mathbf{x} \in \operatorname{null} T$, $T(\mathbf{x}) = x_1\mathbf{v}_1 + \cdots + x_m\mathbf{v}_m = \mathbf{0}$. Because $\mathbf{v}_1, ..., \mathbf{v}_m$ is linearly independent, $x_1, ..., x_m$ can only be 0, and $\mathbf{x}$ must be $\mathbf{0}$. Thus, $\operatorname{null} T = \{\mathbf{0}\}$, which implies T is injective.

! **Note 1.1:** *We only use boldface letters in this Problems 1, 4 and 5 to distinguish between vectors in $\mathbb{F}^m$ and their components.*

Problem 2

(a) Suppose $T \in \mathcal{L}(V, W)$ is injective and $v_1, ..., v_n$ is linearly independent in V. Prove that $T(v_1), ..., T(v_n)$ is linearly independent in W.

(b) Suppose V is finite-dimensional and that $T \in \mathcal{L}(V, W)$. Prove that there exists a subspace U of V such that $U \cap \operatorname{null} T = \{0\}$ and $T(V) = T(U)$. Find a basis.

✔ **Proof:** (a) Since $v_1, ..., v_n$ is linearly independent, then for scalars $a_1, ..., a_n \in \mathbb{F}$,

$$a_1v_1 + \cdots + a_nv_n = 0 \iff a_1 = \cdots = a_n = 0.$$

Because T is injective, only 0 can be mapped into 0. Then

$$T(a_1v_1 + \cdots + a_nv_n) = 0 \iff a_1 = \cdots = a_n = 0.$$

From the linearity of T, we have $a_1T(v_1)+\cdots+a_nT(v_n) = T(a_1v_1+\cdots+a_nv_n)$. Therefore $T(v_1), ..., T(v_n)$ is linearly independent in W.

(b) Suppose that $\dim \operatorname{null} T = m$, $\dim V = n$, $m \leq n$. Let $u_1, ..., u_m$ be a basis of $\operatorname{null} T$, which can be extended into a basis of V: $u_1, ..., u_m, v_1, ..., v_{n-m}$. For any $v \in V$, there exist scalars $a_1, \ldots, a_m, b_1, \ldots, b_{n-m} \in \mathbb{F}$ such that

$$\begin{aligned} T(v) &= T(a_1u_1 + \cdots + a_mu_m + b_1v_1 + \cdots + b_{n-m}v_{n-m}) \\ &= \underbrace{T(a_1u_1 + \cdots + a_mu_m)}_{=0} + T(b_1v_1 + \cdots + b_{n-m}v_{n-m}) \\ &= T(b_1v_1 + \cdots + b_{n-m}v_{n-m}) \end{aligned}$$

Let $U = \operatorname{span}(v_1, ..., v_{n-m})$, then $T(V) \subset T(U)$. Note that every $u \in U$ can be expressed as $b_1v_1 + \cdots + b_{n-m}v_{n-m}$ ($b_1, \ldots, b_{n-m} \in \mathbb{F}$), so reverse the equality above and we show that $T(U) \subset T(V)$. It follows from 2.33 that $U \cap \operatorname{null} T = \{0\}$. The basis of U consists of the extended vectors $v_1, ..., v_{n-m}$. ■

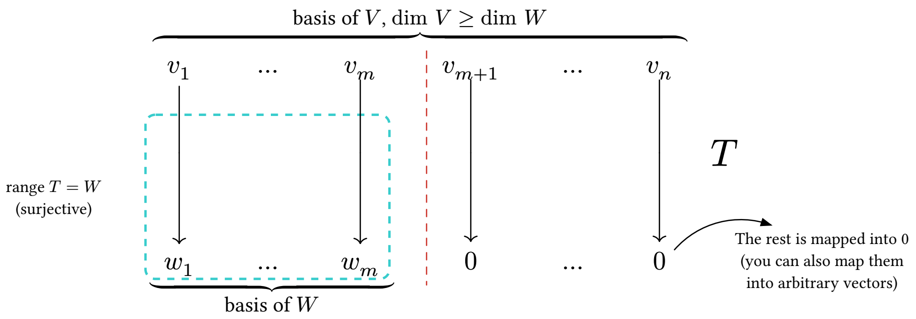
Problem 3

(a) Suppose V and W are both finite-dimensional. Prove that there exists an surjective linear transformation from V onto W if and only if $\dim V \geq \dim W$.

(b) Suppose V and W are finite-dimensional and that U is a subspace of V. Prove that there exists $T \in \mathcal{L}(V, W)$ such that $\operatorname{null} T = U$ if and only if $\dim U \geq \dim V - \dim W$.

Proof: (a) $\implies$: There exists a surjective linear map $T : V \to W$. Then from fundamental theorem of linear maps (3.21), we have $\dim V \geq \dim \operatorname{range} T = \dim W$, completing the first part of the proof.

$\impliedby$:



Suppose $v_1, ..., v_n$ is a basis of V and $w_1, ..., w_m$ is a basis of $W (n \geq m)$. Then from linear map lemma (3.4), we can construct a linear map $T : V \to W$ such that

$$T(v_1) = w_1, ..., T(v_m) = w_m, T(v_{m+1}) = 0, ..., T(v_n) = 0.$$

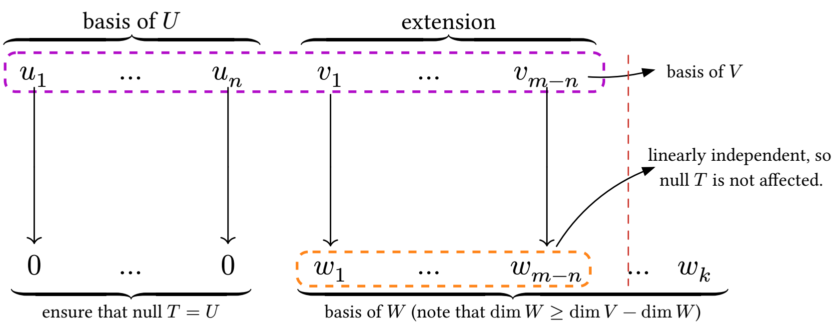
For every $w \in W$, we have

$$\begin{aligned}
w &= a_1 w_1 + \cdots + a_m w_m \\
&= a_1 T(v_1) + \cdots + a_m T(v_m) \\
&= a_1 T(v_1) + \cdots + a_m T(v_m) + \underbrace{a_{m+1} T(v_{m+1}) + \cdots + a_n T(v_n)}_{=0} \\
&= T(a_1 v_1 + \cdots + a_m v_m) \in \operatorname{range} T,
\end{aligned}$$

where $a_1, \ldots, a_n \in \mathbb{F}$. Thus $W \subset \operatorname{range} T$. Clearly we have $\operatorname{range} T \subset W$. Then $\operatorname{range} T = W$, T is surjective, completing the proof.

(b) $\implies$: From 3.21, $\dim U = \dim \operatorname{null} T = \dim V - \dim \operatorname{range} T \geq \dim V - \dim W$, completing the first part of the proof.

$\impliedby$:



Let $\dim U = n, \dim V = m, \dim W = k$. Let $u_1, ..., u_n$ be a basis of U, and let $w_1, \ldots, w_k$ be a basis of W. Because U is a subspace of V, $u_1, ..., u_n$ can be extended to a basis of V: $u_1, ..., u_n, v_1, ..., v_{m-n}$. From 3.4, define $T \in \mathcal{L}(V, W)$ such that

$$T(u_1) = \cdots = T(u_n) = 0, T(v_1) = w_1, ..., T(v_{m-n}) = w_{m-n}.$$

We can easily verify that $U \subset \operatorname{null} T$, next we verify $\operatorname{null} T \subset U$. Let $T(u) = 0$, where $u = k_1u_1 + \cdots + k_nu_n + l_1v_1 + \cdots + l_{m-n}v_{m-n} (k_1, \ldots, k_n, l_1, \ldots, l_{m-n} \in \mathbb{F})$. Then

$$\begin{aligned}
T(u) &= T(k_1u_1 + \cdots + k_nu_n + l_1v_1 + \cdots + l_{m-n}v_{m-n}) \\
&= k_1T(u_1) + \cdots + k_nT(u_n) + l_1T(v_1) + \cdots + l_{m-n}T(v_{m-n}) \\
&= l_1w_1 + \cdots + l_{m-n}w_{m-n} = 0.
\end{aligned}$$

Because $w_1, ..., w_{m-n}$ is linearly independent (note that $k \geq m - n$), we have $l_1 = \cdots = l_{m-n} = 0$. Hence $u = k_1u_1 + \cdots + k_nu_n \in U$, and thus $\operatorname{null} T = U$. This completes the proof. ■

Problem 4

Find the standard matrices of the following linear transformations.

(a) $T \in \mathcal{L}(\mathbb{R}^2, \mathbb{R}^3)$ with $T(x) = \begin{bmatrix} 2x_1 - x_2 \\ -x_1 + x_2 \\ x_2 \end{bmatrix}$ for $x = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$.

(b) $T \in \mathcal{L}(\mathbb{R}^2, \mathbb{R}^2)$ is a vertical shear transformatin that leaves $\mathbf{e}_2$ unchanged and maps $\mathbf{e}_1$ into $2\mathbf{e}_2 + \mathbf{e}_1$.

(c) $T \in \mathcal{L}(\mathbb{R}^2, \mathbb{R}^2)$ first performs a horizontal shear transformatin that leaves $\mathbf{e}_1$ unchanged and maps $\mathbf{e}_2$ into $-2\mathbf{e}_1 + \mathbf{e}_2$ and then reflects points through the line $x_2 = -x_1$.

✔ **Solution:** (a) $T(\mathbf{e}_1) = \begin{bmatrix} 2 \\ -1 \\ 0 \end{bmatrix}$, $T(\mathbf{e}_2) = \begin{bmatrix} -1 \\ 1 \\ 1 \end{bmatrix}$, so the standard matrix is $\begin{bmatrix} 2 & -1 \\ -1 & 1 \\ 0 & 1 \end{bmatrix}$.

(b) $[\mathbf{e}_1 \quad \mathbf{e}_2] \begin{bmatrix} 1 & 0 \\ 2 & 1 \end{bmatrix} = [2\mathbf{e}_2 + \mathbf{e}_1 \quad \mathbf{e}_2]$, so the standard matrix is $\begin{bmatrix} 1 & 0 \\ 2 & 1 \end{bmatrix}$.

(c) Denote the vertical shear transformation as T_1 and the reflection as T_2. $T_1(\mathbf{e}_1) = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$, $T_1(\mathbf{e}_2) = \begin{bmatrix} -2 \\ 1 \end{bmatrix}$; $T(\mathbf{e}_1) = T_2(T_1(\mathbf{e}_1)) = \begin{bmatrix} 0 \\ -1 \end{bmatrix}$, $T(\mathbf{e}_2) = T_2(T_1(\mathbf{e}_2)) = \begin{bmatrix} -1 \\ 2 \end{bmatrix}$. So the standard matrix is $\begin{bmatrix} 0 & -1 \\ -1 & 2 \end{bmatrix}$.

Problem 5

Let

$$\mathbf{b}_1 = \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}, \mathbf{b}_2 = \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}, \mathbf{b}_3 = \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix}$$

and let L be the linear transformation from $\mathbb{R}^2$ into $\mathbb{R}^3$ define by

$$L(\mathbf{x}) = (x_1 - x_2)\mathbf{b}_1 + x_2\mathbf{b}_2 + (x_1 + x_2)\mathbf{b}_3,$$

find the matrix A representing L with respect to the ordered bases $\{\mathbf{e}_1, \mathbf{e}_2\}$ and $\{\mathbf{b}_1, \mathbf{b}_2, \mathbf{b}_3\}$.

✔ **Solution:** Rearrange the equation in the problem and we have

$$L(\mathbf{x}) = [\mathbf{b}_1 \quad \mathbf{b}_2 \quad \mathbf{b}_3]\begin{bmatrix} x_1 - x_2 \\ x_2 \\ x_1 + x_2 \end{bmatrix} \implies L(\mathbf{e}_1) = [\mathbf{b}_1 \quad \mathbf{b}_2 \quad \mathbf{b}_3]\begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}, L(\mathbf{e}_2) = [\mathbf{b}_1 \quad \mathbf{b}_2 \quad \mathbf{b}_3]\begin{bmatrix} -1 \\ 1 \\ 1 \end{bmatrix},$$

hence $A = \begin{bmatrix} [L(\mathbf{e}_1)]_{\{\mathbf{b}_1,\mathbf{b}_2,\mathbf{b}_3\}} & L(\mathbf{e}_2)_{\{\mathbf{b}_1,\mathbf{b}_2,\mathbf{b}_3\}} \end{bmatrix} = \begin{bmatrix} 1 & -1 \\ 0 & 1 \\ 1 & 1 \end{bmatrix}$.

Problem 6

Let D be the differentiation operator on $\mathbb{P}_2(\mathbb{R})$. Find the matrix B representing D with respect to $[1, 2x, x^2]$, the matrix A representing D with respect to $[2, 4x, 4x^2 - 4]$, and the nonsingular matrix S such that $B = S^{-1}AS$.

✔ **Solution:** $(x^n)' = nx^{n-1}$, $(x^2)' = 2x, x' = 1, 1' = 0$, so

$$\begin{bmatrix} D(1) & D(2x) & D(x^2) \end{bmatrix} = \begin{bmatrix} 0 & 2 & 2x \end{bmatrix} = \begin{bmatrix} 1 & 2x & x^2 \end{bmatrix}\begin{bmatrix} 0 & 2 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix} \implies B = \begin{bmatrix} 0 & 2 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix},$$

$$\begin{bmatrix} D(2) & D(4x) & D(4x^2-4) \end{bmatrix} = \begin{bmatrix} 0 & 4 & 8x \end{bmatrix} = \begin{bmatrix} 2 & 4x & 4x^2-4 \end{bmatrix}\begin{bmatrix} 0 & 2 & 0 \\ 0 & 0 & 2 \\ 0 & 0 & 0 \end{bmatrix} \implies A = \begin{bmatrix} 0 & 2 & 0 \\ 0 & 0 & 2 \\ 0 & 0 & 0 \end{bmatrix}.$$

Since $\begin{bmatrix} 1, 2x, x^2 \end{bmatrix} = \begin{bmatrix} 2, 4x, 4x^2 - 4 \end{bmatrix}\begin{bmatrix} \frac{1}{2} & 0 & \frac{1}{2} \\ 0 & \frac{1}{2} & 0 \\ 0 & 0 & \frac{1}{4} \end{bmatrix}$, we have $S = \begin{bmatrix} \frac{1}{2} & 0 & \frac{1}{2} \\ 0 & \frac{1}{2} & 0 \\ 0 & 0 & \frac{1}{4} \end{bmatrix}$.

Problem 7

Suppose V and W are finite-dimensional and $T \in \mathcal{L}(V, W)$. Prove that $\dim \operatorname{range} T = 1$ if and only if there exist a basis of V and a basis of W such that with respect to these bases, all entries of the matrix representation $\mathcal{M}(T)$ equal 1.

! **Note 7.1:** *The key is to construct a basis.*

! **Note 7.2:** *Note that for linear maps we often use the product notation Tv as well as the usual function notation $T(v)$. We will use the former notation in the following proof and proof of Problem 8(c) and Problem 9.*

✔ **Proof:** $\Longleftarrow$: Let $v_1, \dots, v_n$ be a basis of V and $w_1, \dots, w_m$ be a basis of W, which satisfy the condition. Then we have

$$\begin{bmatrix} Tv_1 & Tv_2 & \cdots & Tv_n \end{bmatrix} = \begin{bmatrix} w_1 & w_2 & \cdots & w_m \end{bmatrix}\begin{bmatrix} 1 & 1 & \cdots & 1 \\ \vdots & \vdots & & \vdots \\ 1 & 1 & \cdots & 1 \end{bmatrix},$$

thus

$$Tv_1 = Tv_2 = \cdots = Tv_n = w_1 + \cdots + w_m.$$

Denote $w_1+\cdots+w_m$ as w, thus $T(v)=T(k_1v_1+\cdots+k_nv_n)=(k_1+\cdots+k_n)w$. Hence $\operatorname{range} T=\operatorname{span}(w)$. Given that w is nonzero, $\dim\operatorname{range} T=1$, completing the first part of the proof.

$\Longrightarrow$: From fundamental theorem of linear transformation (3.21), $\dim V=\dim\operatorname{null} T+\dim\operatorname{range} T=\dim\operatorname{null} T+1$. Suppose that $v_1,...,v_{n-1}$ is a basis of $\operatorname{null} T$. Extend it to a basis of V: $v_1,\ldots,v_n$, and from 3.21 we know that $T(v_n)$ is a basis of $\operatorname{range} T$. Denote it as w_1. Extend the basis to form a basis in W: $w_1,...,w_m$. Then

$$\begin{bmatrix} Tv_1 & Tv_2 & \cdots & Tv_n \end{bmatrix}=\begin{bmatrix} w_1 & w_2 & \cdots & w_m \end{bmatrix}\begin{bmatrix} 0 & \cdots & 0 & 1 \\ 0 & \cdots & 0 & 0 \\ \vdots & \ddots & \vdots & \vdots \\ 0 & \cdots & \cdots & 0 \end{bmatrix}_{m\times n}. \tag{1}$$

Proof 1: Choose a new basis of V:

$$v_1+v_n,\cdots,v_{n-1}+v_n,v_n, \tag{2}$$

so that every basis vector is mapped into the same value w_m. To verify that it is indeed a basis, let $a_1,...,a_n\in\mathbb{F}$, then for the equation

$$\begin{aligned} a_1(v_1+v_n)+\cdots+a_{n-1}(v_{n-1}+v_n)+a_nv_n&=0,\\ \Longrightarrow\ a_1v_1+\cdots+a_{n-1}v_{n-1}+\left(\sum_{i=1}^{n}a_i\right)v_n&=0, \end{aligned}$$

and we can first find that $a_1=0$, and then $a_2=0,a_3=0,\ldots,a_n=0$. A list of linearly independent vectors that has appropriate length is a basis.

Then we choose a new basis of W:

$$w_1-w_2-\cdots-w_m,...,w_{m-1},w_m, \tag{3}$$

so that $w_1=(w_1-w_2-\cdots-w_m)+\sum_{i=2}^{m}w_i$, i.e., the sum of all the new basis vectors. To verify that it is indeed a basis, let $a_1,...,a_m\in\mathbb{F}$, then for the equation

$$\begin{aligned} a_1(w_1-w_2-\cdots-w_m)+a_2w_2+\cdots+a_mw_m&=0,\\ \Longrightarrow\ a_1w_1+(a_2-a_1)w_2+\cdots+(a_m-a_1)w_m&=0, \end{aligned}$$

and we can first find that $a_1=0$, and then $a_2=a_1=0,\ldots,a_m=a_1=0$. A list of linearly independent vectors that has appropriate length is a basis.

Therefore, with (2) and (3),

$$T(v_1+v_n)=\cdots=T(v_{n-1}+v_n)=T(v_n)=w_1=(w_1-w_2-\cdots-w_m)+\sum_{i=2}^{m}w_i,$$

and the matrix representation with respect to these two bases is the matrix whose entries all equal 1. This completes the proof.

Proof 2: *(The proof below shows the construction process using language of matrices. Compare it with the proof above.)*

First we perform elementary column manipulation to (1) for a proper basis of V. Multiply $\begin{bmatrix} 1 & & & \\ 1 & 1 & & \\ 1 & 1 & \ddots & \\ 1 & 1 & \cdots & 1 \end{bmatrix}_{n\times n}$ on the right of both sides and the left hand side becomes

$$\begin{bmatrix} T\left(\sum_{i=1}^{n} v_i\right) & T\left(\sum_{i=2}^{n} v_i\right) & \cdots & T(v_n) \end{bmatrix}, \tag{4}$$

while the right hand side becomes

$$\begin{bmatrix} w_1 & w_2 & \cdots & w_m \end{bmatrix} \begin{bmatrix} 1 & 1 & \cdots & 1 \\ 0 & 0 & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & 0 \end{bmatrix}_{m\times n}.$$

Next, perform elementary row manipulation in the equation above to find a proper basis of W and to transform the matrix that have 1's in the top row into a matrix that has 1 everywhere:

$$\begin{aligned} & \begin{bmatrix} w_1 & w_2 & \cdots & w_m \end{bmatrix} \begin{bmatrix} 1 & 1 & \cdots & 1 \\ 0 & 0 & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & 0 \end{bmatrix}_{m\times n} \\ = & \begin{bmatrix} w_1 & w_2 & \cdots & w_m \end{bmatrix} \begin{bmatrix} 1 & & & & \\ -1 & 1 & & & \\ & \ddots & \ddots & & \\ & & -1 & 1 & \\ & & & -1 & 1 \end{bmatrix}_{m\times m} \begin{bmatrix} 1 & 0 & \cdots & 0 \\ 1 & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 1 & 1 & \cdots & 1 \end{bmatrix}_{m\times m} \begin{bmatrix} 1 & 1 & \cdots & 1 \\ 0 & 0 & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & 0 \end{bmatrix}_{m\times n} \\ = & \begin{bmatrix} w_1 - w_2 & w_2 - w_3 & \cdots & w_{m-1} - w_m & w_m \end{bmatrix} \begin{bmatrix} 1 & 1 & \cdots & 1 \\ 1 & 1 & \cdots & 1 \\ \vdots & \vdots & & \vdots \\ 1 & 1 & \cdots & 1 \end{bmatrix}_{m\times n}. \end{aligned} \tag{5}$$

Finally, we verify that the new lists of vectors in the procedure above are basis of V and W respectively. First, consider the vectors $\sum_{i=1}^{n} v_i, \sum_{i=2}^{n} v_i, ..., v_n$ is a basis of V. Let $a_1, ..., a_n \in \mathbb{F}$, then for the equation

$$\begin{aligned} & a_1 \left(\sum_{i=1}^{n} v_i\right) + a_2 \left(\sum_{i=2}^{n} v_i\right) + \cdots + a_n v_n = 0, \\ \implies & a_1 v_1 + (a_1 + a_2) v_2 + (a_1 + a_2 + a_3) v_3 + \cdots + \left(\sum_{i=1}^{n} a_i\right) v_n = 0, \end{aligned}$$

and we can first find that $a_1 = 0$, and then $a_2 = 0, a_3 = 0, \ldots, a_n = 0$. A list of linearly independent vectors that has appropriate length is a basis.

Then we verify that $w_1 - w_2, w_2 - w_3, ..., w_{m-1} - w_m, w_m$ is a basis of W. Let $a_1, ..., a_m \in \mathbb{F}$, then for the equation

$$
\begin{aligned}
& a_1(w_1 - w_2) + a_2(w_2 - w_3) + \cdots + a_{m-1}(w_{m-1} - w_m) + a_m w_m = 0, \\
\Longrightarrow\ & a_1 w_1 + (a_2 - a_1) w_2 + (a_3 - a_2) w_3 + \cdots + (a_m - a_{m-1}) w_m = 0,
\end{aligned}
$$

and thus $a_1 = 0, a_2 = a_1 = 0, \ldots, a_m = a_{m-1} = 0$. A list of linearly independent vectors that has appropriate length is a basis.

Therefore, with (4) and (5), the matrix representation with respect to these two bases is the matrix whose entries all equal 1. This completes the proof. ■

Attention: *You should always start at a random basis and construct the basis you need step-by-step. You can't just assume that $Tv_i = c_i w (c_i \neq 0)$, because you impose additional conditions to v_i.*

Note 7.3: *When I solved this problem for the first time (in spring semester of 2022), I didn't use the "extend" technique and just assumed that the basis of* $\operatorname{range} T$ *is* w_1 *and* $Tv_1 = k_1 w_1, Tv_2 = k_2 w_1, ..., Tv_n = k_n w_1$ *and* $k_1, ..., k_s \neq 0, k_{s+1} = ... = k_n = 0$ *(*$s \leq n$*, without loss of generality). Then*

$$
\begin{bmatrix} Tv_1 & Tv_2 & \cdots & Tv_n \end{bmatrix} = \begin{bmatrix} w_1 & w_2 & \cdots & w_m \end{bmatrix} \begin{bmatrix} k_1 & \cdots & k_s & 0 & \cdots & 0 \\ 0 & 0 & \cdots & & \cdots & 0 \\ \vdots & \vdots & & & & \vdots \\ 0 & 0 & \cdots & & \cdots & 0 \end{bmatrix}_{m \times n}.
$$

I then multiplied $\begin{bmatrix} \frac{1}{k_1} &&&&&& \\ & \frac{1}{k_2} &&&&& \\ && \ddots &&&& \\ &&& \frac{1}{k_s} & \frac{1}{k_s} & \cdots & \frac{1}{k_s} \\ &&&& 1 && \\ &&&&& \ddots & \\ &&&&&& 1 \end{bmatrix}$ *on the right side and get*

$$
\begin{aligned}
& \begin{bmatrix} T(\frac{1}{k_1} v_1) & T(\frac{1}{k_2} v_2) & \cdots & T(\frac{1}{k_s} v_s) & T(\frac{1}{k_s} v_s + v_{s+1}) & \cdots & T(\frac{1}{k_s} v_s + v_n) \end{bmatrix} \\
= & \begin{bmatrix} w_1 & w_2 & \cdots & w_m \end{bmatrix} \begin{bmatrix} 1 & 1 & \cdots & 1 \\ 0 & 0 & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & 0 \end{bmatrix}_{m \times n}.
\end{aligned}
$$

Compare this with the current proof, you can see that the "extend" technique provides clearer insight into the "structure" of the linear transformation.

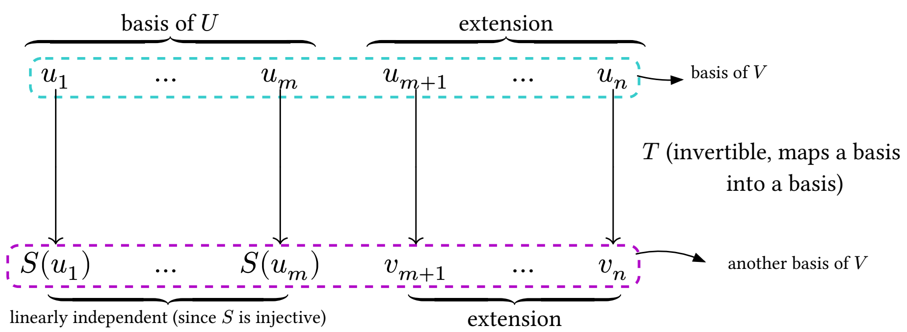
Problem 8

(a) Suppose V is finite-dimensional, U is a subspace of V, and $S \in \mathcal{L}(U, V)$. Prove there exists an invertible operator $T \in \mathcal{L}(V)$ such that $T(u) = S(u)$ for every $u \in U$ if and only if S is injective.
(b) Suppose V, W are finite-dimensional and $T_1, T_2 \in \mathcal{L}(V, W)$. Prove that $\operatorname{null} T_1 = \operatorname{null} T_2$ if and only if there exists an invertible operator $S \in \mathcal{L}(W)$ such that $T_1 = ST_2$.
(c) Suppose V, W are finite-dimensional and $T_1, T_2 \in \mathcal{L}(V, W)$. Prove that $\dim \operatorname{null} T_1 = \dim \operatorname{null} T_2$ if and only if there exists invertible operators $S \in \mathcal{L}(W)$ and $R \in \mathcal{L}(V)$ such that $T_1 = ST_2R$.

Note 8.1: *The key is to construct a linear operator/transformation through linear map(transformation) lemma.*

Proof: (a) First we assume there exists an invertible operator $T \in \mathcal{L}(V)$. Then T is injective. Then $T(u) = 0$ if and only if $u = 0$. Because $T(u) = S(u)$, we have $S(u) = 0$ if and only if $u = 0$. Thus $\operatorname{null} S = \{0\}$, and S is injective, completing the first part of the proof.

Next we assume that S in injective.



Let $u_1, ..., u_m$ be a basis of U. Since U is a subspace of V, $u_1, ..., u_m$ can be extended to a basis of V: $u_1, ..., u_m, u_{m+1}, ..., u_n$. Define $T \in \mathcal{L}(V)$ such that

$$T(u_1) = S(u_1), ..., T(u_m) = S(u_m),$$

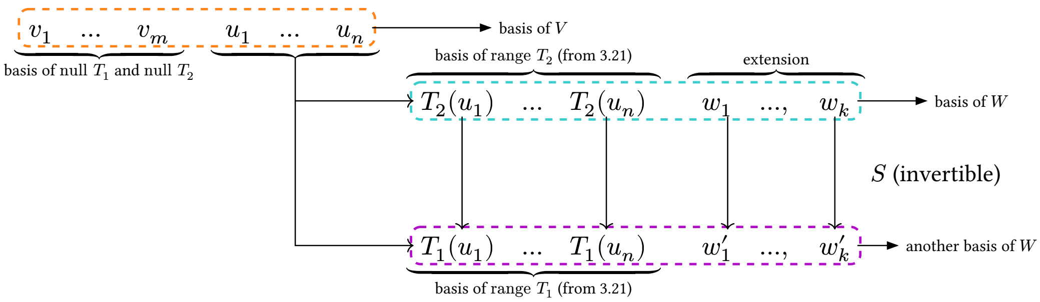
then $T(u) = S(u)$ for every $u \in U$. Since S is injective, $S(u_1), ..., S(u_m)$ is linearly independent (refer to Problem 2(a)), so it can be extended to another basis of V: $S(u_1), ..., S(u_m), v_{m+1}, ..., v_n$. Let

$$T(u_{m+1}) = v_{m+1}, ..., T(u_n) = v_n.$$

Thus T is well-defined. From **Theorem 2** (turn back to the first page of this document), T is invertible, completing the proof.

(b) First we assume that there exists an invertible operator $S \in \mathcal{L}(W)$ such that $T_1 = ST_2$. Then for any vector $v \in \operatorname{null} T_2$, $T_2(v) = 0$. Then $S(T_2(v)) = 0 = T_1(v)$, thus $\operatorname{null} T_2 \subset \operatorname{null} T_1$. For any vector $v \in \operatorname{null} T_1$, $T_1(v) = S(T_2(v)) = 0$. Since S is injective, $T_2(v) = 0$. Thus $\operatorname{null} T_1 \subset \operatorname{null} T_2$. Hence $\operatorname{null} T_1 = \operatorname{null} T_2$, completing the first part of the proof.

Next we assume that $\operatorname{null} T_1 = \operatorname{null} T_2$.



Denote $\operatorname{null} T_1 = \operatorname{null} T_2 = U$. Let $u_1, ..., u_m$ be a basis of U. Because U is a subspace of V, we can extend $u_1, ..., u_m$ to a basis of V: $u_1, ..., u_m, v_1, ..., v_n$. T_1, T_2 can be written as follows:

$$T_1(u_1) = \cdots = T_1(u_m) = 0, T_1(v_1) = w_1, ..., T_1(v_n) = w_n,$$

$$T_2(u_1) = \cdots = T_2(u_m) = 0, T_2(v_1) = w_1', ..., T_2(v_n) = w_n',$$

where $w_1, ..., w_n$ and $w_1', ..., w_n'$ are both lists of linearly independent vectors in W (from the proof of 3.21). Extend each list to a basis of W:

$$w_1, ..., w_n, x_1, ..., x_k \quad \text{and} \quad w_1', ..., w_n', y_1, ..., y_k,$$

and we can define $S \in \mathcal{L}(W)$ such that

$$S(w_1') = w_1, ..., S(w_n') = w_n, S(y_1) = x_1, ..., S(y_k) = x_k.$$

From **Theorem 2** (turn back to the first page of this document), S is invertible.

For any $v \in V$, there exist scalars $a_1, ..., a_m, b_1, ..., b_n$ such that

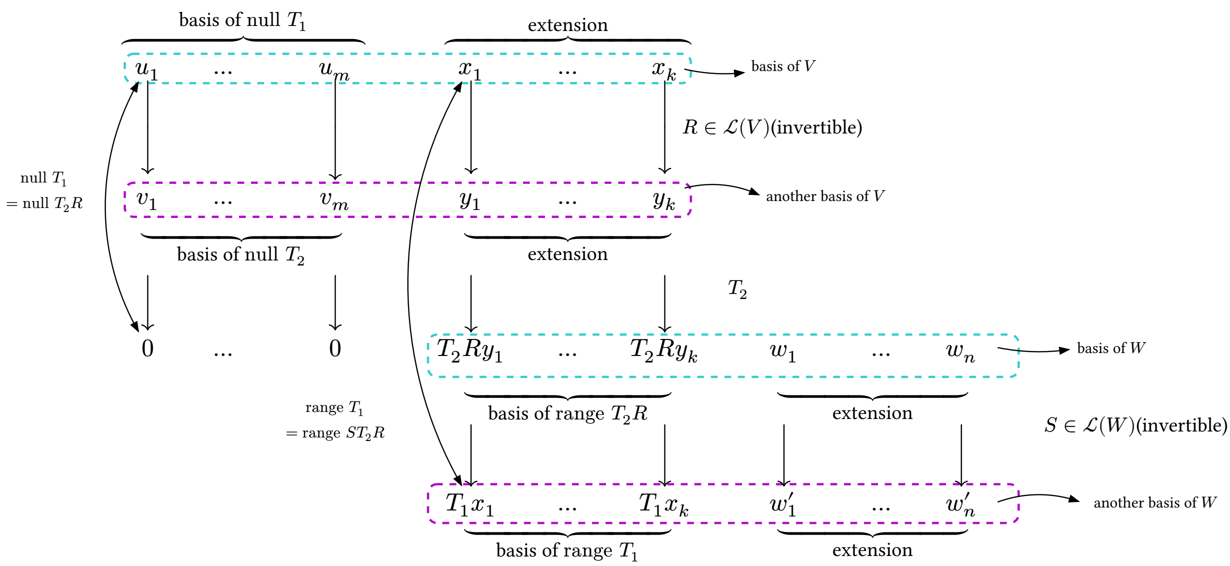
$$\begin{aligned}
ST_2(v) &= S(T_2(a_1u_1 + \cdots + a_mu_m + b_1v_1 + \cdots + b_nv_n)) \\
&= S(b_1w_1' + \cdots + b_nw_n') \\
&= b_1w_1 + \cdots + b_nw_n \\
&= b_1T_1(v_1) + \cdots + b_nT_1(v_n) \\
&= a_1T_1(u_1) + \cdots + a_mT_1(u_m) + b_1T_1(v_1) + \cdots + b_nT_1(v_n) = T_1(v),
\end{aligned}$$

hence $ST_2 = T_1$, completing the proof.

! **Note 8.2:** *To verify that ST_2 and T_1 are equal, you can also show that they take a basis to the same value.*

Solution Cont. (c) First we assume that if there exists invertible operators $S \in \mathcal{L}(W)$ and $R \in \mathcal{L}(V)$ such that $T_1 = ST_2R$. Suppose that $v_1, \ldots, v_m$ is a basis of $\operatorname{null} T_1$. Since $T_1 = ST_2R$, we have $ST_2Rv_1 = T_1v_1 = 0, \ldots, ST_2Rv_m = T_1v_m = 0$. Since S is invertible, it is injective, hence $T_2Rv_1 = \cdots = T_2Rv_m = 0$. Since $v_1, \ldots, v_m$ is linearly independent and R is injective, from Problem 2(a) we have $Rv_1, \ldots, Rv_m$ is a linearly independent list of vectors in $\operatorname{null} T_2$. The dimension of a vector space must be greater than the number of vectors in a linearly independent list. Therefore we have $\dim \operatorname{null} T_1 \leq \dim \operatorname{null} T_2$.

Note that $T_2 = S^{-1}T_1R^{-1}$, repeat the procedure above and we show that $\dim \operatorname{null} T_2 \leq \dim \operatorname{null} T_1$. Therefore, $\dim \operatorname{null} T_1 = \dim \operatorname{null} T_2$, completing the proof in one direction.



To prove the implication in the other direction, let $u_1, \ldots, u_m$ be a basis of $\operatorname{null} T_1$ and $v_1, \ldots, v_m$ be a basis of $\operatorname{null} T_2$. Extend these bases to bases of V:

$$u_1, \ldots, u_m, x_1, \ldots, x_k \quad \text{and} \quad v_1, \ldots, v_m, y_1, \ldots, y_k.$$

Define $R \in \mathcal{L}(V)$ as

$$Ru_1 = v_1, \ldots, Ru_m = v_m, Rx_1 = y_1, \ldots, Rx_k = y_k.$$

From **Theorem 2** (turn back to the first page of this document), R is invertible.

For $u \in \operatorname{null} T_1$, $T_2Ru = T_2(R(a_1u_1 + \cdots + a_mu_m)) = T_2(a_1v_1 + \cdots + a_mv_m) = 0$ ($a_1, \ldots, a_m$ are scalars), hence $\operatorname{null} T_1 \subset \operatorname{null} T_2R$; conversely, for $u \in \operatorname{null} T_2R$, there exist scalars $a_1, \ldots, a_m, b_1, \ldots, b_k$ such that

$$\begin{aligned} T_2Ru &= T_2R(a_1u_1 + \cdots + a_mu_m + b_1x_1 + \cdots + b_kx_k) \\ &= \underbrace{T_2(a_1v_1 + \cdots + a_mv_m)}_{=0} + T_2(b_1y_1 + \cdots + b_ky_k) \\ &= b_1T_2y_1 + \cdots + b_kT_2y_k \end{aligned}$$

From the proof of 3.21, we know that $T_2y_1, \ldots, T_2y_k$ is linearly independent. Since $T_2Ru = 0$, we have $b_1 = \cdots = b_k = 0$. Thus $u = a_1u_1 + \cdots + a_mu_m \in \operatorname{null} T_1$, hence $\operatorname{null} T_2R \subset \operatorname{null} T_1$. Therefore, $\operatorname{null} T_2R = \operatorname{null} T_1$, and from (b) we know that there exists an invertible $S \in \mathcal{L}(W)$ such that $T_1 = ST_2R$. ■

Problem 9

Suppose $V_1, \ldots, V_m$ and $W_1, \ldots, W_m$ are vector spaces.
(a) Prove that $\mathcal{L}(V_1 \times \cdots \times V_m, W)$ and $\mathcal{L}(V_1, W) \times \cdots \times \mathcal{L}(V_m, W)$ are isomorphic vector spaces.
(b) Prove that $\mathcal{L}(V, W_1 \times \cdots \times W_m)$ and $\mathcal{L}(V, W_1) \times \cdots \times \mathcal{L}(V, W_m)$ are isomorphic vector spaces.

Attention: *$V_1, \ldots, V_m$ and $W_1, \ldots, W_m$ are not necessarily finite-dimensional, so you can't utilize the dimension test!*

✔ **Proof:** (a) Define a map

$$\Phi : \mathcal{L}(V_1, W) \times \cdots \times \mathcal{L}(V_m, W) \to \mathcal{L}(V_1 \times \cdots \times V_m, W)$$

by

$$(\Phi(T_1, \ldots, T_m))(v_1, \ldots, v_m) = T_1 v_1 + \cdots + T_m v_m$$

for any $T_k \in \mathcal{L}(V_k, W)$ and any $(v_1, \ldots, v_m) \in V_1 \times \cdots \times V_m$, where $k \in \{1, \ldots, m\}$.

It is straightforward to verify that $\Phi(T_1, \ldots, T_m)$ is indeed a linear map for $(T_1, \ldots, T_m)$ and that Φ itself is linear.

To prove that these two involved vector spaces are isomorphic vector spaces, we need to prove that Φ is an isomorphism.

(a) Proof 1:

For any $T \in \mathcal{L}(V_1 \times \cdots \times V_m, W)$ and any $k \in \{1, \ldots, m\}$, there exists $T_k \in \mathcal{L}(V_k, W)$ (it's easy to verify that T_k is a linear map) such that

$$T_k u = T(0, \ldots, u, \ldots, 0),$$

with $u \in V_k$ in the k-th position and the rest 0.

Hence, for any $(v_1, \ldots, v_m) \in V_1 \times \cdots \times V_m$, we have

$$\begin{aligned} T(v_1, \ldots, v_m) &= T(v_1, 0, \ldots, 0) + \cdots + T(0, \ldots, v_m) \\ &= T_1 v_1 + \cdots + T_m v_m \\ &= (\Phi(T_1, \ldots, T_m))(v_1, \ldots, v_m) \end{aligned}$$

That is, $T = \Phi(T_1, \ldots, T_m)$. So Φ is surjective.

Now suppose $(T_1, \ldots, T_m) \in \mathcal{L}(V_1, W) \times \cdots \times \mathcal{L}(V_m, W)$ and $\Phi(T_1, \ldots, T_m) = 0$. Suppose for some $k \in \{1, \ldots, m\}$, $T_k \neq 0$, then there exists $u_k \in V_k$ such that $T_k u_k \neq 0$, and

$$(\Phi(T_1, \ldots, T_m))(0, \ldots, u_k, \ldots, 0) = T_k u_k \neq 0,$$

with $u_k \in V_k$ in the k-th position and the rest 0.

There would be a contradiction. So for all $k \in \{1, \ldots, m\}$, we must have $T_k = 0$, hence $(T_1, \ldots, T_m) = 0$. So Φ is injective.

Hence, Φ is an isomorphism, which implies the involved two vector spaces are isomorphic, as desired.

(a) Proof 2:

For $k \in \{1, \ldots, m\}$, define $\tau_k : V_k \to V_1 \times \cdots \times V_m$ by $\tau_k(v) = (0, \ldots, v, \ldots, 0)$, where the v is in the k^{th} position; it is straightforward to check that each τ_k is a linear map. Define a map

$$\Psi : \mathcal{L}(V_1 \times \cdots \times V_m, W) \to \mathcal{L}(V_1, W) \times \cdots \times \mathcal{L}(V_m, W),$$

where $\Psi(T)$ is given by $(T \circ \tau_1, \ldots, T \circ \tau_m)$. The linearity of each $T \circ \tau_k$ follows from the linearity of T and the linearity of τ_k. Let $(T_1, \ldots, T_m) \in \mathcal{L}(V_1, W) \times \cdots \times \mathcal{L}(V_m, W)$ be given and observe that

$$\Psi(\Phi(T_1, \ldots, T_m)) = (\Phi(T_1, \ldots, T_m) \circ \tau_1, \ldots, \Phi(T_1, \ldots, T_m) \circ \tau_m).$$

For any $k \in \{1, \dots, m\}$ and $v \in V_k$ we have

$$\begin{aligned}[\Phi(T_1, \dots, T_m)](\tau_k(v)) &= [\Phi(T_1, \dots, T_m)](0, \dots, v, \dots, 0) \\ &= T_1(0) + \cdots + T_k v + \cdots + T_m(0) \\ &= T_k v.\end{aligned}$$

Thus $\Phi(T_1, \dots, T_m) \circ \tau_k = T_k$ and it follows that $\Psi(\Phi(T_1, \dots, T_m)) = (T_1, \dots, T_m)$, i.e. $\Psi \circ \Phi$ is the identity map on $\mathcal{L}(V_1, W) \times \cdots \times \mathcal{L}(V_m, W)$. Now let $T \in \mathcal{L}(V_1 \times \cdots \times V_m, W)$ be given and observe that

$$\begin{aligned}[\Phi(\Psi(T))](v_1, \dots, v_m) &= [\Phi(T \circ \tau_1, \dots, T \circ \tau_m)](v_1, \dots, v_m) \\ &= (T \circ \tau_1)(v_1) + \cdots + (T \circ \tau_m)(v_m) \\ &= T(v_1, 0, \dots, 0) + \cdots + T(0, \dots, v_m) \\ &= T(v_1, \dots, v_m).\end{aligned}$$

Thus $\Phi(\Psi(T)) = T$, i.e. $\Phi \circ \Psi$ is the identity map on $\mathcal{L}(V_1 \times \cdots \times V_m, W)$. As the proof of 3.63 shows, it now follows that Ψ is a linear map. Thus Φ is an isomorphism from $\mathcal{L}(V_1, W) \times \cdots \times \mathcal{L}(V_m, W)$ to $\mathcal{L}(V_1 \times \cdots \times V_m, W)$ and its inverse is Ψ. Hence, These two vector spaces are isomorphic, as desired.

(b) Define a map

$$\Phi : \mathcal{L}(V, W_1) \times \cdots \times \mathcal{L}(V, W_m) \to \mathcal{L}(V, W_1 \times \cdots \times W_m)$$

by

$$\Phi(T_1, \dots, T_m)(v) = (T_1 v, \dots, T_m v)$$

for any $(T_1, \dots, T_m) \in \mathcal{L}(V, W_1) \times \cdots \times \mathcal{L}(V, W_m)$ and $v \in V$. It is straightforward to verify that $\Phi(T_1, \dots, T_m)$ is indeed a linear map, and that Φ itself is linear.

For each $k \in \{1, \dots, m\}$, define $p_k : W_1 \times \cdots \times W_m \to W_k$ by $p_k(w_1, \dots, w_m) = w_k$; it is straightforward to check that each p_k is a linear map.

(b) Proof 1:

Let $T \in \mathcal{L}(V, W_1 \times \cdots \times W_m)$ be given. Then $T_k \in \mathcal{L}(V, W_k)$, and for any $v \in V$,

$$\Phi(T_1, \dots, T_m)(v) = (T_1 v, \dots, T_m v) = (p_1(Tv), \dots, p_m(Tv)) = T(v).$$

So $T = \Phi(T_1, \dots, T_m)$, and hence Φ is surjective.

Now suppose $(T_1, \dots, T_m) \in \mathcal{L}(V, W_1) \times \cdots \times \mathcal{L}(V, W_m)$ satisfies $\Phi(T_1, \dots, T_m) = 0$, i.e.,
$\Phi(T_1, \dots, T_m)(v) = (T_1 v, \dots, T_m v) = 0$ for all $v \in V$. Then for each k, $T_k v = 0$ for all $v \in V$, so $T_k = 0$. Hence $(T_1, \dots, T_m) = 0$, and Φ is injective.

Therefore, Φ is a linear isomorphism from $\mathcal{L}(V, W_1) \times \cdots \times \mathcal{L}(V, W_m)$ to $\mathcal{L}(V, W_1 \times \cdots \times W_m)$, and the two vector spaces are isomorphic, as desired.

(b) Proof 2:

Define a map

$$\Psi : \mathcal{L}(V, W_1 \times \cdots \times W_m) \to \mathcal{L}(V, W_1) \times \cdots \times \mathcal{L}(V, W_m),$$

where $\Psi(T)$ is given by $(p_1 \circ T, \ldots, p_m \circ T)$. The linearity of each $p_k \circ T$ is given by the linearity of p_k and the linearity of T. Let $(T_1, \ldots, T_m) \in \mathcal{L}(V, W_1) \times \cdots \times \mathcal{L}(V, W_m)$ be given and observe that

$$\Psi(\Phi(T_1, \ldots, T_m)) = (p_1 \circ \Phi(T_1, \ldots, T_m), \ldots, p_m \circ \Phi(T_1, \ldots, T_m)).$$

For any $k \in \{1, \ldots, m\}$ and $v \in V$ we have

$$p_k((\Phi(T_1, \ldots, T_m))(v)) = p_k(T_1 v, \ldots, T_m v) = T_k v.$$

Thus $p_k \circ \Phi(T_1, \ldots, T_m) = T_k$ and it follows that $\Psi(\Phi(T_1, \ldots, T_m)) = (T_1, \ldots, T_m)$, i.e., $\Psi \circ \Phi$ is the identity map on $\mathcal{L}(V, W_1) \times \cdots \times \mathcal{L}(V, W_m)$. Now let $T \in \mathcal{L}(V, W_1 \times \cdots \times W_m)$ be given and observe that

$$[\Phi(\Psi(T))](v) = [\Phi(p_1 \circ T, \ldots, p_m \circ T)](v) = (p_1(Tv), \ldots, p_m(Tv)) = Tv.$$

Thus $\Phi(\Psi(T)) = T$, i.e. $\Phi \circ \Psi$ is the identity map on $\mathcal{L}(V, W_1 \times \cdots \times W_m)$. As the proof of 3.63 shows, it now follows that Ψ is a linear map. Thus Φ is an isomorphism from $\mathcal{L}(V, W_1) \times \cdots \times \mathcal{L}(V, W_m)$ to $\mathcal{L}(V, W_1 \times \cdots \times W_m)$, and its inverse is Ψ. Hence, These two vector spaces are isomorphic, as desired. ■

Problem 10

Suppose that v, x are vectors in V and that U, W are subspaces of V such that $v + U = x + W$. Prove that $U = W$.

✔ **Proof:** For any $u \in U$, there exists $w \in W$ such that $v + u = x + w$, i.e.,

$$u = w + x - v. \tag{6}$$

Note that for $0 \in V$ (which is also the additive identity in U and W), there exists $w_1 \in W$ such that $v + 0 = x + w_1$, which implies $x - v = -w_1 \in W$. Thus from (6) we have $u \in W$, which further implies that $U \subset W$.

Similarly we can show that $W \subset U$. Therefore, $U = W$. ■

Acknowledgement

Thanks YinMo19 for helping me draw the illustrations in Problem 3 and 8.