MATRIX ANALYSIS

A Quick Guide

Huan Bui

Colby College
Physics & Statistics
Class of 2021

February 22, 2019

Preface

Greetings,

Matrix Analysis: A Quick Guide to is compiled based on my MA353: Matrix Analysis notes with professor Leo Livshits. The sections are based on a number of resources: *Linear Algebra Done Right* by Axler, *A Second Course in Linear Algebra* by Horn and Garcia, *Matrices and Linear Transformations* by Cullen, *Matrices: Methods and Applications* by Barnett, *Problems and Theorems in Linear Algebra* by Prasolov, *Matrix Operations* by Richard Bronson, and professor Leo Livshits' own textbook (in the making). Prerequisites: some prior exposure to a first course in linear algebra.

The development of this text will come in layers. The first layer, one that I am working on during the course of S'19 MA353, will be an overview of the key topics listed in the table of contents. As the semester progresses, I will be constantly updating the existing notes, as well as adding prof. Livshits' problems and my solutions to the problems. The second layer will come after the course is over, when concepts will have hopefully "come together."

I will decide how much narrative I should put into the text as text is developed over the semester. I'm thinking that I will only add detailed explanations wherever I find fit or necessary for my own studies. I will most likely keep the text as condensed as I can.

Enjoy!

Contents

1 List of Special Matrices & Their Properties

1. **Hermitian/Self-adjoint**: $H = H^\dagger$. A Hermitian matrix is matrix that is equal to its own conjugate transpose:

$$H \text{ is Hermitian} \iff H_{ij} = \bar{H}_{ji}$$

Properties 1.1.

(a) H is Hermitian $\iff \langle w, Hv \rangle = \langle Hw, v \rangle$, where $\langle , \rangle$ denotes the inner product.

(b) H is Hermitian $\iff \langle v, Hv \rangle \in \mathbb{R}$.

(c) H is Hermitian $\iff$ it is *unitarily diagonalizable* with *real eigenvalues.*

Unitary: $U^*U = UU^* = I = U^\dagger U = UU^\dagger$. The real analogue of a unitary matrix is an orthogonal matrix. The following list contain the properties of U:

(a) U preserves the inner product:

$$\langle Ux, Uy \rangle = \langle x, y \rangle.$$

(b) U is normal: it commutes with $U^* = U^\dagger$.

(c) U is diagonalizable:

$$U = VDV^*,$$

where D is diagonal and unitary, and V is unitary.

(d) $|\det(U)| = 1$ (hence the real analogue to U is an orthogonal matrix)

(e) Its eigenspaces are orthogonal.

(f) U can be written as

$$U = e^{iH},$$

where H is a Hermitian matrix.

(g) Any square matrix with unit Euclidean norm is the average of two unitary matrices.

2. **Idempotent:** M idempotent $\iff M^2 = M$.

(a) Singularity: its number of independent rows (and columns) is less than its number of rows (and columns).

(b) When an idempotent matrix is subtracted from the identity matrix, the result is also idempotent.

"Proof".

$$[I - M][I - M] = I - M - M + M^2 = I - M - M + M = I - M.$$

$\square$

(c) M is idempotent $\iff \forall n \in \mathbb{N}$, $A^n = A$.

(d) Eigenvalues: an idempotent matrix is always diagonalizable and its eigenvalues are either 0 or 1. (think "projection")

(e) Trace: the trace of an idempotent matrix equals the rank of the matrix and thus is always an integer. So

$$\operatorname{tr}(A) = \dim(\operatorname{Im} A).$$

3. **Nilpotent:** a nilpotent matrix is a square matrix N such that

$$N^k = 0$$

for some positive integer k. The smallest such k is sometimes called the **index** of N.

The following statements are equivalent:

(a) N is nilpotent.

(b) The minimal polynomial for N is x^k for some positive integer $k \leq n$.

(c) The characteristic polynomial for N is x^n.

(d) The only complex eigenvalue for N is 0.

(e) $\operatorname{tr} N^k = 0$ for all $k > 0$.

Properties 1.2.

(a) The degree of an $n \times n$ nilpotent matrix is always less than or equal to n.

(b) $\det N = \operatorname{tr}(N) = 0$.

(c) Nilpotent matrices are not invertible.

(d) The only nilpotent diagonalizable matrix is the zero matrix.

2 List of Operations

1. **Conjugate transpose** is what its name suggests.

2. **Classical adjoint/Adjugate/adjunct** of a square matrix is the transpose of its cofactor matrix.

3 List of Algorithms

4 Complex Numbers

4.1 A different point of view

We often think of complex numbers as

$$a + ib$$

where $a, b \in \mathbb{R}$ and $i = \sqrt{-1}$. While there is nothing "bad" about this way of thinking - in fact thinking of complex numbers as $a+ib$ allows us to very quickly and intuitively do arithmetics operations on them - a "matrix representation" of complex numbers can give us some insights on "what we actually do" when we perform complex arithmetics.

Let us think of

$$\begin{pmatrix} a & -b \\ b & a \end{pmatrix}$$

as a different representation of the same object - the same complex number "$a + ib$." Note that it does not make sense to say the matrix representation **equals** the complex number itself. But we shall see that a lot of the properties of complex numbers are carried into this matrix representation under interesting matricial properties.

$$\boxed{\begin{pmatrix} a & -b \\ b & a \end{pmatrix} \sim a + ib}$$

First, let us break the the matrix down:

$$a + ib = a \times 1 + i \times b \sim \begin{pmatrix} a & -b \\ b & a \end{pmatrix} = a \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} + b \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} = aI + b\mathcal{I}.$$

Right away, we can make some "mental connections" between the representations:

$$I \sim 1$$
$$\mathcal{I} \sim i.$$

Now, we know that complex number multiplications commute:

$$(a + ib)(c + id) = (c + id)(a + ib).$$

Matrix multiplications are not commutative. So, we might wonder whether commutativity holds under the this new representation of complex numbers. Well, the answer is yes. We can readily verify that

$$(aI + b\mathcal{I})(cI + b\mathcal{I}) = (cI + b\mathcal{I})(aI + b\mathcal{I}).$$

How about additions? Let's check:

$$(a+ib)+(c+id)=(a+c)+i(b+d)\sim\begin{pmatrix}a+c & -(b+d)\\ (b+d) & a+c\end{pmatrix}=\begin{pmatrix}a & -b\\ b & a\end{pmatrix}+\begin{pmatrix}c & -d\\ d & c\end{pmatrix}.$$

Ah! Additions work. So, the new representation of complex numbers seems to be working flawlessly. However, we have yet to gain any interesting insights into the connections between the representations. To do that, we have to look into changing the form of the matrix. First, let's see what conjugation does:

$$(a+ib)^*=a-ib\sim\begin{pmatrix}a & b\\ -b & a\end{pmatrix}=\begin{pmatrix}a & -b\\ b & a\end{pmatrix}^\top$$

Ah, so conjugation to a complex number in the traditional representation is the same as transposition in the matrix representations. What about the amplitude square? Let us call

$$M=\begin{pmatrix}a & -b\\ b & a\end{pmatrix}.$$

We have

$$(a+ib)(a-ib)\sim\begin{pmatrix}a & -b\\ b & a\end{pmatrix}\begin{pmatrix}a & b\\ -b & a\end{pmatrix}=MM^\top=(a^2+b^2)I=\det(M)I$$

Interesting. But observe that if $\det(M)\neq 0$

$$\frac{1}{\det(M)}MM^\top=I.$$

This tells us that

$$M^\top=M^{-1},$$

where M^{-1} is the inverse of M, and, not surprisingly, it corresponds to the reciprocal to the complex number $a+ib$. We can readily show that

$$M^{-1}\sim(a+ib)^{-1}=\frac{1}{a^2+b^2}(a-ib).$$

Remember that we can also think of a complex number as a column vector:

$$c+id\sim\begin{pmatrix}c\\ d\end{pmatrix}.$$

Let us look back at complex number multiplication under matrix representation:

$$(a+ib)(c+id)=(ac-bd)+i(bc+ad)\sim\begin{pmatrix}a & -b\\ b & a\end{pmatrix}\begin{pmatrix}c\\ d\end{pmatrix}=\begin{pmatrix}ac-bd\\ bc+ad\end{pmatrix}.$$

Multiplication actually works in this "mixed" way of representing complex numbers as well. Now, observe that what we just did was performing a linear transformation on a vector in $\mathbb{R}^2$. It is always interesting to look at the geometrical interpretation of this transformation. To do this, let us call N the "normalized" version of M:

$$N = \frac{1}{\sqrt{a^2+b^2}} \begin{pmatrix} a & -b \\ b & a \end{pmatrix}.$$

We immediately recognize that N is an orthogonal matrix. This means N is an orthogonal transformation (length preserving). Now, it is reasonable to define

$$\cos\theta = \frac{a}{\sqrt{a^2+b^2}}$$
$$\sin\theta = \frac{b}{\sqrt{a^2+b^2}}.$$

We can write N as

$$N = \begin{pmatrix} \cos\theta & -\sin\theta \sin\theta & \cos\theta \end{pmatrix},$$

which is a physicists' favorite matrix: the rotation by θ. So, let us write M in terms of N:

$$M = \begin{pmatrix} a & -b \\ b & a \end{pmatrix} = \sqrt{a^2+b^2} N = \sqrt{a^2+b^2} \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}.$$

We can interpret M as a rotation by θ, followed by a scaling by $\sqrt{(}a^2+b^2)$. But what $\sqrt{a^2+b^2}$ exactly is just the "length" or the "amplitude" of the complex number $a+ib$, if we think of it as an arrow in a plane.

4.2 Relevant properties and definitions

1. The *modulus* of $z = a+ib$ is the "amplitude" of z, denoted by $|z| = \sqrt{a^2+b^2} = z\bar{z}$.
2. The modulus is *multiplicative*, i.e.
$$|wz| = |w||z|.$$
3. Triangle inequality:
$$|z+w| \leq |z| + |w|.$$
We can readily show this geometrically, or algebraically.
4. The *argument* of $z = a+ib$ is θ, where
$$\theta = \begin{cases} \tan^{-1}\left(\frac{b}{a}\right), \text{ if } a > 0 \\ \frac{\pi}{2} + k2\pi, k \in \mathbb{R} \text{ if } a = 0, b > 0 \\ -\frac{\pi}{2} + k2\pi, k \in \mathbb{R} \text{ if } a = 0, b < 0 \\ \text{Undefined if } a = b = 0. \end{cases}$$

5. The *conjugate* of $a+ib$ is $a-ib$. Conjugation is *additive* and *multiplicative*, i.e.

$$\overline{z+w} = \bar{z} + \bar{w}$$
$$\overline{wz} = \bar{w}\bar{z}.$$

Note that we can also show the multiplicative property with the matrix representation as well:

$$\overline{wz} \sim (WZ)^\top = Z^\top W^\top \sim \bar{z}\bar{w} = \bar{w}\bar{z}.$$

6. Euler's identity, generalized to de Moivre's formula:

$$z^n = r^n e^{in\theta}.$$

5 Vector Spaces & Linear Functions

5.1 Review of Linear Spaces and Subspaces

Properties 5.1. of linear spaces:

1. Commutativity and associativity of addition
2. Existence of an additively neutral element (null element). Zero multiples of elements give the null element: $0 \cdot V = \mathbf{0}$
3. Every element has an (unique) additively antipodal element
4. Scalar multiplication distributes over addition
5. Multiplicative identity:
6. $ab \cdot V = a \cdot (bV)$
7. $(a + b)V = aV + bV$

W is a subspace of V if

1. $S \subseteq V$
2. S is non-empty
3. S is closed under addition and scalar multiplication

Properties 5.2. that are interesting/important/maybe-not-so-obvious:

1. If S is a subspace of V and $S \neq V$ then S is a proper subspace of V.
2. If X in a subspace of Y and Y is a subspace of Z, then X is a subspace of W.
3. Non-trivial linear (non-singleton) spaces are infinite.

5.2 Review of Linear Maps

Consider linear spaces V and W and elements $v \in V$ and $w \in W$ and scalars $\alpha, \beta \in \mathbb{R}$, a function $F : V \to W$ is a linear map if

$$F[\alpha v + \beta w] = \alpha F[v] + \beta F[w].$$

Properties 5.3.

1. $F[\mathbf{0}_V] = \mathbf{0}_W$
2. $G[w] = (\alpha \cdot F)[w] = \alpha \cdot F[w]$
3. Given $F : V \to W$ and $G : V \to W$, $H[v] = F[v] + G[v] = (F + G)[v]$ is call the sum of the functions F and G.

4. Linear combinations of linear maps are linear.
5. Compositions of linear maps are linear.
6. Compositions distributes over linear combinations of linear maps.
7. Inverses of linear functions (if they exist) are linear.
8. Inverse of a bijective linear function is a bijective linear function

5.3 Review of Kernels and Images

Definition 5.1. Let $F : V \to W$ be given. The kernel of F is defined as

$$\ker(F) = \{v \in V | F[v] = \mathbf{0}_W\}.$$

Properties 5.4. Let $F : V \to W$ a linear map be given. Also, consider a linear map G such that $F \circ G$ is defined

1. F is null $\iff \ker(F) = V \iff \mathrm{Im}(F) = \mathbf{0}_W$
2. $\ker(F)$ is a subspace of V
3. $\mathrm{Im}(F)$ is a subspace of W
4. F is injective $\iff \ker(F) = \mathbf{0}_V$
5. $\ker(F) \subseteq \ker(F \circ G)$.
6. F injective $\implies \ker(F) = \ker(F \circ G)$

Definition 5.2. Let $F : V \to W$ be given. The image of F is defined as

$$\mathrm{Im}(F) = \{w \in W | \exists v \in V, F[v] = w\}.$$

5.4 Atrices

Definition 5.3. Atrix functions: Let $V_1, \ldots, V_m \in \mathbf{V}$ be given. Consider $f : \mathbb{R}^m \to \mathbf{V}$ be defined by

$$f \begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_m \end{pmatrix} = \sum_{i=1}^{m} a_i V_i.$$

We denote f by

$$\begin{pmatrix} V_1 & V_2 & \ldots & V_m \end{pmatrix}.$$

We refer to the V_i's as the **columns** of f, even though there doesn't have to be any columns. Basically, f is simply a function that takes in an ordered list of coefficients and returns a linear combination of V_i with the respective coefficients. A matrix is a special atrix. Not every atrix is a matrix.

Properties 5.5. of atrices

1. The V_i's - the columns of an atrix - are the images of the standard basis tuples.

2. $\mathbf{e}_j \in \ker(V_1 \ldots V_m) \iff V_j = \mathbf{0}_V$, where $\mathbf{e}_j$ denotes a standard basis tuple with a 1 at the j^{th} position. To put in words, a kernel of an atrix contains a standard basis if and only if one of its columns in a null element.

3. $\text{Im}(f) \equiv \text{Im}(V_1 \ldots V_m) = \text{span}(V_1 \ldots V_m)$

4. B is a null atrix $\iff \ker(B) = \mathbb{R}^m \iff \text{Im}(B) = \mathbf{0}_V \iff V_j = \mathbf{0}_V \forall j = 1, 2, \ldots, m.$

5. f is a linear function $\mathbb{R}^m \to V \iff f$ is an atrix function $\mathbb{R}^m \to V$.

6. An atrix A is bijective/invertible, then its inverse A^{-1} is a linear function, but is an atrix only if A is a matrix.

7. Linear combinations of atrices are atrices:

$$\begin{aligned}\alpha \cdot \begin{pmatrix} V_1 & \cdots & V_m \end{pmatrix} + \beta \cdot \begin{pmatrix} W_1 & \ldots & W_m \end{pmatrix} \\ = \begin{pmatrix} \alpha V_1 + \beta W_1 & \ldots & \alpha V_m + \beta W_m \end{pmatrix}\end{aligned}$$

8. Compositions of two atrices are NOT defined unless the atrix going first is a matrix. Consider $F : \mathbb{R}^m \to W$ and $G : \mathbb{R}^n \to T$. $F \circ G$ is only defined if $T = \mathbb{R}^m$. This make G an $n \times m$ matrix. It follows that the atrix $F \circ G$ has the form

$$\begin{pmatrix} f(g_1) & f(g_2) & \ldots & f(g_m) \end{pmatrix}.$$

9. Consider $F : \mathbb{R}^m \to V$ and $G : \mathbb{R}^k \to V$. $\text{Im}(F) \subseteq \text{Im}(G) \iff F = G \circ C$, with $C \in \mathbb{M}_{k \times m}$, i.e. C is an $k \times m$ matrix.

10. Consider an atrix $A : \mathbb{R}^m \to V$. $A = (V_1 \ldots V_m)$. A is NOT injective.
$\iff \exists$ a non-trivial linear combination of the columns of A that gives $\mathbf{0}_V$
$\iff A = [\mathbf{0}_v]$ or $\exists j | V_j$ is linear combination of other columns of A
$\iff$ The first column of A is $\mathbf{0}_V$ or $\exists j | V_j$ is a linear combination some of $V_i, i < j$.

11. If atrix $A : \mathbb{R} \to V$ has a single column then it is injective if the column is not $\mathbf{0}_V$.

Properties 5.6. of elementary column operations for atrices. Elementary operations on the columns of F can be expressed as a composition of F and an appropriate elementary matrix E, $F \circ E$.

1. Swapping i^{th} and j^{th} columns: $F \circ E^{[i] \leftrightarrow [j]}$.

2. Scaling the j^{th} column by α: $F \circ E^{\alpha \cdot [j]}$.

3. Adjust the j^{th} column by adding to it $\alpha \times i^{th}$ column: $F \circ E^{[i] \overset{+}{\leftarrow} \alpha \cdot [j]}$.

4. Elementary column operations do not change the 'jectivity nor image of F.

5. If a column of F is a linear combination of some of the other columns then elementary column operations can turn it into $\mathbf{0}_V$.

6. Removing/Inserting null columns or columns that are linear combinations of other columns does not change the image of F.

7. Given atrix A, it is possible to eliminate (or not) columns of A to end up with an atrix B with $\mathrm{Im}(B) = \mathrm{Im}(A)$.

8. If B is obtained from insertion of columns into atrix A, then $\mathrm{Im}(B) = \mathrm{Im}(A) \iff$ the insert columns $\in \mathrm{Im}(A)$.

9. Inserting columns to a surjective A does not destroy surjectivity of A. (A is already having extra or just enough columns)

10. A surjective $\iff$ A is obtained by inserting columns (or not) into an invertible atrix $\iff$ deleting some columns of A (or not) gives an invertible matrix.

11. If A is injective, then column deletion does not destroy injectivity. (A is already "lacking" or having just enough columns)

12. The new atrix obtained from inserting columns from $\mathrm{Im}(A)$ into A is injective $\iff$ A is injective.

5.5 Linear Independence, Span, and Bases

5.5.1 Linear Independence

$$\begin{aligned}
& X_1 \ldots X_m \text{ are linearly independent} \\
\iff & F = [X_1 \ldots X_m] \text{ injective} \\
\iff & \textstyle\sum a_i X_i = 0 \iff a_i = 0 \forall i \\
\iff & \mathbf{0}_V \notin \{X_i\} \text{ and none are linear combinations of some of the others.} \\
\iff & X_1 \neq \mathbf{0}_V \text{ and } X_j \text{ is not a linear combination of any of } X_i\text{'s for } i < j.
\end{aligned}$$

Properties 5.7.

1. The singleton list is linearly independent if its entry is not the null element

2. Sublists of a linearly independent list are linearly independent

3. List operations cannot create/destroy linearly independence.

4. If a linear map $L : X \to W$ injective, then $X_1 \ldots X_m$ linearly independent $\iff$ $L(X_1) \ldots L(X_m)$ linearly independent.

5.5.2 Span

Properties 5.8.

1. Spans are subspaces. $\text{span}(X_1 \dots X_m), X_j \in V$ is a subspace of V.
2. $\text{span}(X_1 \dots X_j) \subseteq \text{span}(X_1 \dots X_k)$ if $j \leq k$.
3. Adding elements to a list that spans V produces a list that spans V.
4. The following list operations do not change the span of V: removing the null/linearly dependent element, inserting a linearly dependent element, scaling element(s), adding (multiples) of an element to another element.
5. It is possible to reduce a list that spans to a list that spans AND have linearly independent elements.
6. Consider $A : V \to W$. If X_i span W then $A(X_i)$ span $\text{Im}(A)$. $i = 1, 2, \dots, m$.
7. If $A : V \to W$ invertible, then X_i span $V \iff A(X_i)$ span W, $i = 1, 2, \dots, m$.

5.5.3 Bases

Properties 5.9.

1. A list is a basis of V if the elements are linearly independent and they span V.
2. A singleton linear space has no basis.
3. Re-ordering the elements of a basis gives another basis.
4. $\{X_1 \dots X_m\}$ is a basis of V if $(X_1 \dots X_m)$ is injective AND $\text{Im}(X_1 \dots X_m) = V$, i.e. $(X_1 \dots X_m)$ invertible.
5. If $\{X_i\}$ is a basis of V and $\{Y_i\}$ is a list of elements in W, then there exists a unique linear function $L : V \to W$ satisfying
$$L[X_i] = Y_i$$
6. If $A : V \to W$ bijective, then $\{V_i\}$ forms a basis of V and $\{A(X_i)\}$ forms a basis of W.
7. Elementary operations on bases give bases.

5.6 Linear Bijections and Isomorphisms

Definition 5.4. Let linear spaces V, W be given. V is **isomorphic** to W if $\exists F : V \to W$ bijective. We say $V \sim W$.

Properties 5.10.

1. A non-zero scalar multiple of an isomorphism is an isomorphism.
2. A composition of isomorphism is an isomorphism.
3. "Isomorphism" behaves like an equivalence relation:
   (a) Reflexivity: $V \sim V$.
   (b) Symmetry: if $V \sim W$ then $W \sim V$.
   (c) Transitivity: if $V \sim W$ and $W \sim Z$ then $V \sim Z$.
4. Consider $F : V \to W$ an isomorphism.
   (a) Isomorphisms preserve linear independence. V_i's are linearly independent in $V \iff F(V_i)$'s are linearly independent in W.
   (b) isomorphisms preserve spanning. V_i's span $V \iff F(V_i)$'s span W.
   (c) Isomorphisms preserve bases. $\{V_i\}$ is a basis of $V \iff F\{(V_i)\}$ is a basis of W.
5. If $V \sim W$ then $\dim(V) = \dim(W)$ (finite or infinite).
6. If a linear map $A : V \to W$ is given and $A(X_i)$'s are linearly independent, then A_i's are linearly independent.
7. If a linear map $A : V \to W$ is injective and $\{X_i\}$ is a basis of V then $\{A(X_i)\}$ is a basis of $\mathrm{Im}(A)$

5.7 Finite-Dimensional Linear Spaces

5.7.1 Dimension

Properties 5.11.

1. $\mathbb{R}^m \sim R^n \iff m = n$.
2. Isomorphisms $F : \mathbb{R}^n \to V$ are bijective atrices.
3. Isomorphisms $G : W \to \mathbb{R}^m$ are inverses of bijective atrices.
4. Consider a non-singleton linear space V
   (a) V has a basis with n elements.
   (b) $V \sim \mathbb{R}^n$.
   (c) $V \sim W$, where W is any linear space with a basis of n elements.

5. If V is a linear space with a basis with n elements, then any basis of V has n elements.

6. Linear space $V \sim W$ where W is n-dimensional if V is n-dimensional.

7. For a non-singleton linear space V, $V \sim \mathbb{R}^n \iff \dim(V) = n$.

8. $V \sim W \iff \dim(V) = \dim(W)$.

9. If W is a subspace of V, then $\dim(W) \leq \dim(V)$. Equality holds when $W = V$.

5.7.2 Rank-Nullity Theorem

Let finite-dimensional linear space V and linear map $F : V \to W$ be given. Then $\mathrm{Im}(F)$ is finite-dimensional and

$$\dim(\mathrm{Im}(F)) + \dim(\ker(F)) = \dim(V)$$

A stronger statement: If a linear map $F : V \to W$ has finite rank and finite nullity $\iff$ V is finite-dimensional, then

$$\mathrm{Im}(F) + \ker(F) = \dim(V).$$

5.8 Infinite-Dimensional Spaces

Consider a non-singleton linear space V. The following statements are equivalent:

1. V is infinite-dimensional.

2. Every linearly independent list in V can be enlarged to a strictly longer linearly independent set in V.

3. Every linearly independent list in V can be enlarged to an arbitrarily long (finite) linearly independent set in V.

4. There are arbitrarily long (finite) linearly independent lists in V.

5. There are linearly independent lists in V of any (finite) length.

6. No list of finitely many elements of V spans V.

6 Sums of Subspaces & Products of vector spaces

6.1 Direct Sums

Definition 6.1. Let U_j, $j = 1, 2, \ldots m$ are subspaces of V. $\sum_1^m U_j$ is a *direct sum* if each $u \in \sum U_j$ can be written in only one way as $u = \sum_1^m u_j$. The direct sum $\sum_i^m U_j$ is denoted as $U_1 \oplus \cdots \oplus U_m$.

Properties 6.1.

1. Condition for direct sum: If all U_j are subspaces of V, then $\sum_1^m U_j$ is a direct sum $\iff$ the only way to write 0 as $\sum_1^m u_j$, where $u_j \in U_j$ is to take $u_j = 0$ for all j.

2. If U, W are subspaces of V and $U \bigcap W = \{0\}$ then $U + W$ is a direct sum.

6.2 Products of Vector Spaces

Definition 6.2. Product of vectors spaces

$$V_1 \times \cdots \times V_m = \{(v_1, \ldots, v_m) : v_j \in V_j, j = 1, 2, \ldots, m\}.$$

Definition 6.3. Addition on $V_1 \times \cdots \times V_m$:

$$(u_1, \ldots, u_m) + (v_1, \ldots, v_m) = (u_1 + v_1, \ldots, v_m + u_m).$$

Definition 6.4. Scalar multiplication on $V_1 \times \cdots \times V_m$:

$$\lambda(v_1, \ldots, v_m) = (\lambda v_1, \ldots, \lambda v_m).$$

Properties 6.2.

1. Product of vectors spaces is a vector space.

   V_j are vectors spaces over $\mathbb{F} \implies V_1 \times \cdots \times V_m$ is a vector space over $\mathbb{F}$.

2. Dimension of a product is the sum of dimensions:

$$\dim(V_1 \times \cdots \times V_m) = \sum_1^m \dim(V_j)$$

3. Vector space products are NOT commutative:

$$W \times V \neq V \times W.$$

   However,

$$V \times W \sim W \times V.$$

4. Vector space products are NOT associative:

$$V \times (W \times Z) \neq (V \times W) \times Z$$

6.3 Products & Direct Sums

Properties 6.3.

1. Let $U_1, \ldots, U_m$ be subspaces of V. Define a linear map $\Gamma : U_1 \times \cdots \times U_m \to U_1 + \cdots + U_m$ by:
$$\Gamma(u_1, \ldots, u_m) = \sum_1^m u_j.$$
$U_1 + \cdots + U_m$ is a direct sum $\iff$ Γ is injective.

2. Let U_j be finite-dimensional and are subspaces of V.
$$U_1 \oplus \cdots \oplus U_m \iff \dim(U_1 + \cdots + U_m) = \sum_1^m \dim(U_j)$$

6.4 Rank-Nullity Theorem

Suppose Z_1 and Z_2 are subspaces of a finite-dimensional vector space W. Consider $z_1 \in Z_1$, $z_2 \in Z_2$, and a function $\phi : Z_1 \times Z_2 \to Z_1 + Z_2 \prec W$ defined by
$$\phi \begin{pmatrix} z_1 \\ z_2 \end{pmatrix} = z_1 + z_2.$$
First, ϕ is a linear function, as it satisfies the linearity condition:
$$\phi \left(\alpha \begin{pmatrix} z_1 \\ z_2 \end{pmatrix} + \beta \begin{pmatrix} z_1' \\ z_2' \end{pmatrix} \right) = \alpha\phi \begin{pmatrix} z_1 \\ z_2 \end{pmatrix} + \beta\phi\beta \begin{pmatrix} z_1' \\ z_2' \end{pmatrix}.$$
By rank-nullity theorem,
$$\dim(Z_1 \times Z_2) = \dim(Z_1 + Z_2) + \dim(\ker(\phi)).$$
But this is equivalent to
$$\dim(Z_1) + \dim(Z_2) = \dim(Z_1 + Z_2) + \dim(\ker(\phi))$$
The kernel of ϕ is:
$$\ker(\phi) = \left\{ \begin{pmatrix} v \\ -v \end{pmatrix} \middle| v \in z \in Z_1, z \in Z_2 \right\} = \left\{ \begin{pmatrix} v \\ -v \end{pmatrix} \middle| v \in z \in Z_1 \cap Z_2 \right\}$$
We can readily verify that $Z_1 \cap Z_2$ is a subspace of W. With this, $\dim(\ker(\phi)) = \dim(Z_1 \cap Z_2)$. So we end up with
$$\dim(Z_1 + Z_2) = \dim(Z_1) + \dim(Z_2) - \dim(Z_1 \cap Z_2).$$

Properties 6.4.

1. When $Z_1 \cap Z_2$ is trivial, then $Z_1 + Z_2$ is direct.

2. When $\dim(\ker(\phi)) = 0$, ϕ is injective. But ϕ is also surjective by definition, this implies ϕ is a bijection, in which case
$$Z_1 \oplus Z_2 \sim Z_1 + Z_2.$$

6.5 Nullspaces & Ranges of Operator Powers

1. Sequence of increasing null spaces: Suppose $T \in \mathcal{L}(V)$, i.e., T is some linear function mapping $V \to V$, then

$$\{0\} = \ker(T^0) \subset \ker(T^1) \subset \ker(T^2) \subset \cdots \subset \ker(T^k) \subset \ker(T^{k+1}) \subset \ldots.$$

Proof Outline. Let k be a nonnegative integer and $v \in \ker(T^k)$. Then $T^k v = 0$, so $T^{k+1}v = T(T^k v) = T(0) = 0$, so $v \in \ker T^{k+1}$. So $\ker(T^k) \subset \ker(T^{k+1})$. □

2. Equality in the sequence of null spaces: Suppose m is a nonnegative integer such that $\ker(T^m) = \ker(T^{m+1})$, then

$$\ker(T^m) = \ker(T^{m+1}) = \ker(T^{m+2}) = \ldots.$$

Proof Outline. We want to show

$$\ker(T^{m+k}) = (T^{m+k+1}).$$

We know that $\ker T^{m+k} \subset \ker T^{m+k+1}$. Suppose $v \in \ker T^{m+k+1}$, then

$$T^{m+1}(T^k v) = T^{m+k+1}v = 0.$$

So

$$T^k v \in \ker T^{m+1} = \ker T^m.$$

So

$$0 = T^m(T^k v) = T^{m+k}v,$$

i.e., $v \in \ker T^{m+k}$. So $\ker T^{m+k+1} \subset \ker T^{m+k}$. This completes the proof. □

3. Null spaces stop growing: If $n = \dim(V)$, then

$$\ker(T^n) = \ker(T^{n+1}) = \ker(T^{n+2}) = \ldots.$$

Proof Outline. To show:

$$\ker T^n = \ker T^{n+1}.$$

Suppose this is not true. Then the dimension of the kernel has to increase by at least 1 every step until $n+1$. Thus $\dim \ker T^{n+1} \geq n+1 > n = \dim(V)$. This is a contradiction. □

4. V is the direct sum of $\ker(T^{\dim(V)})$ and $\mathrm{Im}(T^{\dim(V)})$: If $n = \dim(V)$, then

$$V = \ker(T^n) \oplus \mathrm{Im}(T^n).$$

Proof Outline. To show:

$$\ker T^n \cap \operatorname{Im} T^n = \{0\}.$$

Suppose $v \in \ker T^n \cap \operatorname{Im} T^n$. Then $T^n v = 0$ and $\exists u \in V$ such that $v = T^n u$. So

$$T^n v = T^{2n} u = 0.$$

So

$$T^n u = 0.$$

But this means $v = 0$. $\square$

5. IT IS NOT TRUE THAT $V = \ker(T) \oplus \operatorname{Im}(T)$ in general.

6.6 Generalized Eigenvectors and Eigenspaces

Definition 6.5. Suppose $T \in \mathcal{L}(V)$ and Λ is an eigenvalue of T. A vector $v \in V$ is called a **generalized eigenvector** of T corresponding to λ if $v \neq 0$ and

$$(T - \lambda I)^j v = 0$$

for some positive integer j.

Definition 6.6. Generalized Eigenspace: Suppose $T \in \mathcal{L}(V)$ and $\lambda \in \mathbf{F}$. The **generalized eigenspace** of T corresponding to λ, denoted $G(\lambda, T)$, is defined to be the set of all generalized eigenvectors of T corresponding to λ, along with the 0 vector.

Properties 6.5. 1. Suppose $T \in \mathcal{L}(V)$ and $\lambda \in \mathbf{F}$. Then

$$G(\lambda, T) = \ker(T - \lambda I)^{\dim(V)}.$$

Proof Outline. Suppose $v \in \ker(T_\lambda I)^{\dim(V)}$. Then $v \in G(\lambda, T)$. So, $\ker(T - \lambda I)^{\dim V} \subset G(\lambda, T)$. Next, suppose $v \in G(\lambda, T)$. Then these is a positive integer j such that

$$v \ in \ker(T - \lambda I)^j.$$

But if this is true, then

$$v \ in \ker(T - \lambda I)^{\dim V},$$

since $\ker(T - \lambda I)^{\dim V}$ is the largest possible kernel, in a sense. $\square$

2. Linearly independent generalized eigenvectors: Let $T \in \mathcal{L}(V)$. Suppose $\lambda_1, \ldots, \lambda_m$ are distinct eigenvalues of T and $v_1, \ldots, v_m$ are corresponding generalized eigenvectors. Then $v_1, \ldots, v_m$ is linearly independent.

6.7 Nilpotent Operators

Definition 6.7. An operator is called **nilpotent** if some power of it equals 0.

Properties 6.6.

Nilpotent operator raised to dimension of domain is 0: Suppose $N \in \mathcal{L}(V)$ is nilpotent. Then

$$N^{\dim(V)} = 0.$$

Matrix of a nilpotent operator: Suppose N is a nilpotent operator on V. Then there is a basis of V with espect to which the matrix of N has the form

$$\begin{pmatrix} 0 & & * \\ & \ddots & \\ 0 & & 0 \end{pmatrix};$$

here all entries on and below the diagonal are 0's.

6.8 Weyr Characteristic

7 Idempotents & Resolutions of Identity

8 Block-representations of operators

8.1 Direct sums of operators

8.2 Coordinatization & matricial representation of linear functions

Consider a finite-dimensional linear space V with basis $\{V_i\}$, $i = 1, 2, \ldots, m$. An element $\tilde{V}$ in V can be expressed in exactly one way:

$$\tilde{V} = \sum_{i=1}^{m} a_i V_i,$$

where $\{a_i\}$ is unique. We call $\{a_i\}$ the coordinate tuple of $\tilde{V}$ and a_i's the coordinates.

Properties 8.1.

1. Inverse of a bijective atrix outputs the coordinates. Suppose $A = [V_i]$. Then

$$\tilde{V} = \sum_{i=1}^{m} a_i V_i \iff A^{-1}(Z) = \begin{pmatrix} a_1 & \ldots & a_m \end{pmatrix}^\top$$

8.3 Equality of rank and trace for idempotents; Resolution of Identity Revisited

9 Invariant subspaces

9.1 Reducing subspaces

10 Polynomials applied to operators

10.1 Minimal polynomials of block-Δ^r operators

10.2 Minimal polynomials at a vector

11 Eigentheory

11.1 Spectral Mapping Theorem

12 Triangularization

12.1 Compression to invariant subspaces

12.2 Simultaneously Δ-ity of commuting families

13 Diagonalization

13.1 Spectral resolutions

13.2 Compressions to reducing subspaces

13.3 Simultaneous diagonalizability for commuting families

14 Primary decomposition over $\mathbb{C}$ and generalized eigenspaces

15 Cyclic decomposition and Jordan form

15.1 Square roots of operators

15.2 Similarity of a matrix and its transpose

15.3 Similarity of a matrix and its conjugate

15.4 Jordan forms of $\mathcal{AB}$ and $\mathcal{BA}$

15.5 Power-convergent operators

15.6 Power-bounded operators

15.7 Row-stochastic matrices

16 Determinant & Trace

16.1 Classical adjoints

16.2 Cayley-Hamilton theorem

17 Inner products and norms

17.1 Riesz representation theorem

17.2 Adjoints

17.3 Grammians

17.4 Orthogonal complements and orthogonal decompositions

17.5 Ortho-projections

17.6 Closest point solutions

17.7 Gram-Schmidt and orthonormal bases

18 Isometries and unitary operators

19 Ortho-triangularization

20 Spectral resolutions

21 Ortho-diagonalization; self-adjoint and normal operators; spectral theorems

22 Positive (semi-)definite operators

22.1 Classification of inner products

22.2 Positive square roots

23 Polar decomposition

24 Single value decomposition

24.1 Spectral/operator norm

24.2 Singular values and approximation

24.3 Singular values and eigenvalues

25 Problems and Solutions

25.1 Problem set 1

Problem. Prelim. 1.1. Suppose that $\mathbf{V}$ is a finite-dimensional vector space. Consider the following subspace $\mathbf{W}$ of the vector space $\mathbf{V} \times \mathbf{V}$:

$$\mathbf{W} = \left\{ \begin{pmatrix} v \\ -v \end{pmatrix} \middle| v \in \mathbf{V} \right\}.$$

Argue that

$$\dim(\mathbf{V}) = \dim(\mathbf{W}).$$

Solution.

We can equivalently argue that $\dim(V) = \dim(W)$ by showing $\mathbf{V}$ is isomorphic to $\mathbf{W}$, i.e., there exists a bijective linear function $\mathcal{L} : \mathbf{V} \to \mathbf{W} \prec \mathbf{V} \times \mathbf{V}$.

Let $\mathcal{L} : \mathbf{V} \to \mathbf{W}$ be defined as

$$\mathcal{L}[v] = \begin{pmatrix} v \\ -v \end{pmatrix},$$

where $v \in \mathbf{V}$. Let $v_1, v_2 \in \mathbf{V}$ and $a, b \in \mathbb{C}$ be given, $\mathcal{L}$ is a linear function† because it satisfies the linearity condition because

$$\mathcal{L}[av_1 + bv_2] = \begin{pmatrix} av_1 + bv_2 \\ -(av_1 + bv_2) \end{pmatrix} = a \begin{pmatrix} v_1 \\ -v_1 \end{pmatrix} + b \begin{pmatrix} v_2 \\ -v_2 \end{pmatrix} = a\mathcal{L}[v_1] + b\mathcal{L}[v_2],$$

where the second equality follows from the fact that $\mathbf{W}$ is a subspace.

Now, $\mathcal{L}$ is surjective$^\Delta$, by definition:

$$\operatorname{Im}(\mathcal{L}) = \left\{ \begin{pmatrix} v \\ -v \end{pmatrix} \middle| v \in \mathbf{V} \right\} = \mathbf{W}.$$

$\mathcal{L}$ is also injective$^\square$, since $\ker(\mathcal{L}) = \mathbf{0}_V$:

$$\mathcal{L}[v] = \mathbf{0_W} \iff \begin{pmatrix} v \\ -v \end{pmatrix} = \mathbf{0_W} = \begin{pmatrix} \mathbf{0_V} \\ \mathbf{0_V} \end{pmatrix} \iff v = \mathbf{0_V}.$$

By $\dagger, \Delta$, and $\square$, $\mathcal{L}$ is an isomorphism. It follows that $\mathbf{V} \sim \mathbf{W} \iff \dim(V) = \dim(W)$.

Problem. Prelim. 1.2. Argue that $\mathbf{V} + \{\mathbf{0_W}\} = \mathbf{V}$, for any subspace $\mathbf{V}$ of a vector space $\mathbf{W}$.

Solution.

We shall first argue that $\mathbf{0_W} = \mathbf{0_V}$. Let $v_+ \in \mathbf{V}$ be given, and let the unique additive antipodal element of v_+ in $\mathbf{V}$ be v_-. Because $\mathbf{V}$ is a subspace of $\mathbf{W}$, $v_+ \in \mathbf{W}$ and $v_- \in \mathbf{W}$,

$$\mathbf{0_V} = v_+ \overset{\mathbf{V}}{+} v_- = v_+ \overset{\mathbf{W}}{+} v_- = \mathbf{0_W}.$$

Therefore, $\{\mathbf{0_W}\} = \{\mathbf{0_V}\}$. It follows that

$$\mathbf{V} + \{\mathbf{0_W}\} = \mathbf{V} + \{\mathbf{0_V}\}. \tag{1}$$

Now, it suffices to show: (i) $\mathbf{V} \subseteq \mathbf{V} + \{\mathbf{0_V}\}$, and (ii) $\mathbf{V} + \{\mathbf{0_V}\} \subseteq \mathbf{V}$.

Let $v \in \mathbf{V}$ be given, then $v \in \mathbf{V} + \{\mathbf{0_V}\}$, so $\mathbf{V} \subseteq \mathbf{V} + \{\mathbf{0_V}\}$. But because both $\mathbf{V}$ and $\mathbf{V} + \{\mathbf{0_V}\}$ are subspaces, $\mathbf{V}$ is a subspace of $\mathbf{V} + \{\mathbf{0_V}\}$. Hence, (i) is true.

Consider $v_0 \in \mathbf{V} + \{\mathbf{0_V}\}$, then there are some $v_{\mathbf{V}} \in \mathbf{V}$ and $v_{\mathbf{0}} \in \{\mathbf{0_V}\}$ such that

$$v_0 = v_{\mathbf{V}} + v_{\mathbf{0}}.$$

But $v_{\mathbf{0}}$ is identically $\mathbf{0_V}$ since $v_{\mathbf{0}} \in \{\mathbf{0_V}\}$, so

$$v_0 = v_{\mathbf{V}} + \mathbf{0_V} = v_{\mathbf{V}} \in \mathbf{V},$$

which implies $\mathbf{V} + \{\mathbf{0_V}\} \subseteq \mathbf{V}$. Therefore $\mathbf{V} + \{\mathbf{0_V}\}$ is a subspace of $\mathbf{V}$ because they are subspaces. Hence, (ii) is true.

Therefore, by Eq. (1), facts (i), and (ii),

$$\mathbf{V} + \{\mathbf{0_V}\} = \mathbf{V}.$$

Problem. 1. Equivalent formulations of directness of a subspace sum: Suppose that $\mathbf{V}_1, \mathbf{V}_2 \ldots, \mathbf{V}_{315}$ are subspaces of a vector space $\mathbf{W}$. Argue that the following claims are equivalent.

1. The subspace sum $\mathbf{V}_1 + \mathbf{V}_2 + \cdots + \mathbf{V}_{315}$ is direct.
2. If $x_i \in \mathbf{V}_i$ and $x_1 + x_2 + \cdots + x_{315} = \mathbf{0_W}$, then $x_i = \mathbf{0_W}$, for every i.
3. If $x_i, y_i \in \mathbf{V}_i$ and $x_1 + x_2 + x_3 + \cdots + x_{315} = y_1 + y_2 + y_3 + \cdots + y_{315}$, then $x_i = y_i$, for every i.
4. For any i, no non-null element of $\mathbf{V}_i$ can be express as a sum of the elements of the other V_j's.
5. For any i, no non-null element of $\mathbf{V}_i$ can be expressed as a sum of the elements of the preceding V_j's.

Solution. It suffices to show [1] $\implies$ [2] $\implies$ [3] $\implies$ [1] and [2] $\iff$ [4] $\iff$ [5].

- [1] implies [2] : If $\mathbf{V}_1 + \mathbf{V}_2 + \cdots + \mathbf{V}_{315}$ is direct, then by definition there is a unique way to write

$$\mathbf{0_W} = x_1 + x_2 + \cdots + x_{315},$$

  where $x_i \in \mathbf{V}_i$. Since taking $x_i = \mathbf{0_W}$ for every i is one such way, it is also *the only way* to express $\mathbf{0_W}$ as $x_1 + x_2 + \cdots + x_{315}$, $x_i \in \mathbf{V}_i$. Therefore [1] implies [2].

- [2] implies [3] : Let $x_i, y_i \in \mathbf{V}_i$ be given such that

$$x_1 + x_2 + x_3 + \cdots + x_{315} = y_1 + y_2 + y_3 + \cdots + y_{315}.$$

  Then rearranging gives

$$(x_1 - y_1) + (x_2 - y_2) + (x_3 - y_3) + \cdots + (x_{315} - y_{315}) = \mathbf{0_W}.$$

  By [2], $x_i - y_i = \mathbf{0_W}$ for every i, i.e., $x_i = y_i$ for every i. Therefore, [2] implies [3].

- [3] implies [1] : [3] implies that any expression of $h \in \mathbf{V}$ as $x_1+x_2+\cdots+x_{315}$ is unique, which implies $\mathbf{V}_1 + \mathbf{V}_2 + \cdots + \mathbf{V}_{315}$ is direct by definition.

- [4] implies [2]: Suppose [2] is false. Without loss of generality, suppose that for $x_i \in \mathbf{V}_i$ and $x_1 \neq \mathbf{0_W}$,

$$x_1 + x_2 + \cdots + x_{315} = \mathbf{0_W}$$

  still holds. It follows that

$$x_1 = -(x_2 + x_3 + \cdots + x_{315}) \neq \mathbf{0_W},$$

  which implies [4] is false. By contraposition, [4] implies [2].

- [2] implies [4] : Suppose [4] is false. Without loss of generality, suppose that $x_1 = -\sum_{j=2}^{315} x_j \neq \mathbf{0_W}$, where no $x_j \in \mathbf{V}_j$ is necessarily non-null. Then

$$x_1 + x_2 + \dots x_{315} = -\sum_{j=2}^{315} x_j + \sum_{j=2}^{315} x_j = \mathbf{0_W}.$$

  But $x_1 \neq \mathbf{0_W}$ implies [2] is false. By contraposition, [2] implies [4].

- [4] and [5] are equivalent: [4] implies [5] evidently since any "sum of the elements of the preceding $\mathbf{V}_j$'s" is a "sum of the elements of the other $\mathbf{V}_j$'s." Conversely, the non-existence of the former (one that expresses an $x_i \in \mathbf{V}_i$) implies the non-existence of the latter, so [5] implies [4].

Problem. 2. Sub-sums of direct sums are direct:
Suppose that $\mathbf{V}_1, \mathbf{V}_2, \ldots, \mathbf{V}_{315}$ are subspaces of a vector space $\mathbf{W}$, and the subspace sum $\mathbf{V}_1 + \mathbf{V}_2 + \cdots + \mathbf{V}_{315}$ is direct.

1. Suppose that for each i, $\mathbf{Z}_i$ is a subspace of $\mathbf{V}_i$. Argue that the subspace sum $\mathbf{Z}_1 + \mathbf{Z}_2 + \ldots \mathbf{Z}_{315}$ is direct.

2. Argue that the sum $\mathbf{V}_2 + \mathbf{V}_5 + \mathbf{V}_7 + \mathbf{V}_{12}$ is also direct.

Solution.

1. By Problem 1., because $\mathbf{V}_1 + \mathbf{V}_2 + \cdots + \mathbf{V}_{315}$ is direct, for any i, no non-null element of $\mathbf{V}_i$ can be expressed as a sum of the elements of the other $\mathbf{V}_j$'s. Since $\mathbf{Z}_i$ is a subspace of $\mathbf{V}_i$, no $z_i \in \mathbf{Z}_i \prec \mathbf{V}_i$, $z_i \neq \mathbf{0_W}$, can be expressed as a sum of the elements of the other $\mathbf{Z}_j$'s $\prec \mathbf{V}_j$'s. Hence, the subspace sum $\mathbf{Z}_1 + \mathbf{Z}_2 + \ldots \mathbf{Z}_{315}$ is also direct.

2. This is true by [1]. Since each $\mathbf{V}_j$ is a subspace of itself, the subspace sum of distinct $\mathbf{V}_j$'s is direct. Therefore, the subspace sum $\mathbf{V}_2 + \mathbf{V}_5 + \mathbf{V}_7 + \mathbf{V}_{12}$ is direct.

Problem. 3. Associativity of directness of subspace sums: Suppose that

$$\mathbf{Y}, \mathbf{V}_1, \mathbf{V}_2, \ldots, \mathbf{V}_5, \mathbf{U}_1, \mathbf{U}_2, \ldots, \mathbf{U}_{12}, \mathbf{Z}_1, \mathbf{Z}_2, \ldots, \mathbf{Z}_4, \mathbf{X}_1, \mathbf{X}_2, \ldots, \mathbf{X}_{53}$$

are subspaces of a vector space $\mathbf{W}$. Argue that the following claims are equivalent.

1. The subspace sum
$$\begin{aligned} &\mathbf{Y} + \mathbf{V}_1 + \mathbf{V}_2 + \cdots + \mathbf{V}_5 + \mathbf{U}_1 + \mathbf{U}_2 + \cdots + \mathbf{U}_{12} \\ &\quad + \mathbf{Z}_1 + \mathbf{Z}_2 + \cdots + \mathbf{Z}_4 + \mathbf{X}_1 + \mathbf{X}_2 + \cdots + \mathbf{X}_{53} \end{aligned}$$
is direct.

2. The subspace sums
$$\begin{aligned} &[\mathbf{V} :=]\, \mathbf{V}_1 + \mathbf{V}_2 + \cdots + \mathbf{V}_5 \\ &[\mathbf{U} :=]\, \mathbf{U}_1 + \mathbf{U}_2 + \cdots + \mathbf{U}_{12} \\ &[\mathbf{Z} :=]\, \mathbf{Z}_1 + \mathbf{Z}_2 + \cdots + \mathbf{Z}_4 \\ &[\mathbf{X} :=]\, \mathbf{X}_1 + \mathbf{X}_2 + \cdots + \mathbf{X}_{53} \\ &\mathbf{Y} + \mathbf{V} + \mathbf{U} + \mathbf{X} + \mathbf{Z} \end{aligned}$$
are all direct.

Solution.

1. [1] implies [2] : If the subspace sum
$$\begin{aligned} &\mathbf{Y} + \mathbf{V}_1 + \mathbf{V}_2 + \cdots + \mathbf{V}_5 + \mathbf{U}_1 + \mathbf{U}_2 + \cdots + \mathbf{U}_{12} \\ &\quad + \mathbf{Z}_1 + \mathbf{Z}_2 + \cdots + \mathbf{Z}_4 + \mathbf{X}_1 + \mathbf{X}_2 + \cdots + \mathbf{X}_{53} \end{aligned}$$
is direct, then by part 2 of Problem 2., the subspace sums
$$\begin{aligned} &[\mathbf{V} :=]\, \mathbf{V}_1 + \mathbf{V}_2 + \cdots + \mathbf{V}_5 \\ &[\mathbf{U} :=]\, \mathbf{U}_1 + \mathbf{U}_2 + \cdots + \mathbf{U}_{12} \\ &[\mathbf{Z} :=]\, \mathbf{Z}_1 + \mathbf{Z}_2 + \cdots + \mathbf{Z}_4 \\ &[\mathbf{X} :=]\, \mathbf{X}_1 + \mathbf{X}_2 + \cdots + \mathbf{X}_{53} \\ &\mathbf{Y} + \mathbf{V} + \mathbf{U} + \mathbf{X} + \mathbf{Z} \end{aligned}$$
are all direct.

2. [2] implies [1] : The subspace sum
$$\mathbf{Y} \oplus \mathbf{V} \oplus \mathbf{U} \oplus \mathbf{X} \oplus \mathbf{Z}$$
can be written as
$$\begin{aligned} &\mathbf{Y} \oplus (\mathbf{V}_1 \oplus \mathbf{V}_2 \oplus \cdots \oplus \mathbf{V}_5) \oplus (\mathbf{U}_1 \oplus \mathbf{U}_2 \oplus \cdots \oplus \mathbf{U}_{12}) \\ &\quad \oplus (\mathbf{Z}_1 \oplus \mathbf{Z}_2 \oplus \cdots \oplus \mathbf{Z}_4) \oplus (\mathbf{X}_1 \oplus \mathbf{X}_2 \oplus \cdots \oplus \mathbf{X}_{53}), \end{aligned}$$

which is equivalent to

$$\begin{aligned}\mathbf{Y} \oplus \mathbf{V}_1 \oplus \mathbf{V}_2 \oplus \cdots \oplus \mathbf{V}_5 \oplus \mathbf{U}_1 \oplus \mathbf{U}_2 \oplus \cdots \oplus \mathbf{U}_{12} \\ \oplus \mathbf{Z}_1 \oplus \mathbf{Z}_2 \oplus \cdots \oplus \mathbf{Z}_4 \oplus \mathbf{X}_1 \oplus \mathbf{X}_2 \oplus \cdots \oplus \mathbf{X}_{53}\end{aligned}$$

by associativity of vector space addition, i.e., the subspace sum in the first statement is direct.

Problem. 4. Direct sums preserve linear independence: Suppose that $\mathbf{U}, \mathbf{V}, \mathbf{W}$ are subspaces of a vector space $\mathbf{Z}$, and the sum $\mathbf{U}, \mathbf{V}, \mathbf{W}$ is direct.

1. Suppose that $U_1, U_2, \ldots, U_{13}$ is a linearly independent list in $\mathbf{U}$, $V_1, V_2, \ldots, V_6$ is a linearly independent list in $\mathbf{V}$, and $W_1, W_2, \ldots, W_{134}$ is a linearly independent list in $\mathbf{W}$. Argue that the concatenated list

$$U_1, U_2, \ldots, U_{13}, V_1, V_2, \ldots, V_6, W_1, W_2, \ldots, W_{134}$$

is linearly independent.

2. Suppose that $U_1, U_2, \ldots, U_{13}$ is a basis of $\mathbf{U}$, $V_1, V_2, \ldots, V_6$ is a basis of $\mathbf{V}$, and $W_1, W_2, \ldots, W_{134}$ is a basis of $\mathbf{W}$. Argue that the concatenated list

$$U_1, U_2, \ldots, U_{13}, V_1, V_2, \ldots, V_6, W_1, W_2, \ldots, W_{134}$$

is a basis of $\mathbf{U} \oplus \mathbf{V} \oplus \mathbf{W}$.

Solution.

1. By Problem 1., because $\mathbf{U} \oplus \mathbf{V} \oplus \mathbf{W} = \mathbf{Z}$, for any $U \in \mathbf{U}, V \in \mathbf{V}$, and $W \in \mathbf{W}$,

$$U + V + W = \mathbf{0_Z} \iff U = V = W = \mathbf{0_Z}.$$

Therefore U, V, W are linearly independent. It follows, that if $U_1, U_2, \ldots, U_{13}$ is a linearly independent list in $\mathbf{U}$, $V_1, V_2, \ldots, V_6$ is a linearly independent list in $\mathbf{V}$, and $W_1, W_2, \ldots, W_{134}$ is a linearly independent list in $\mathbf{W}$, then the list

$$U_1, U_2, \ldots, U_{13}, V_1, V_2, \ldots, V_6, W_1, W_2, \ldots, W_{134}$$

is linearly independent.

2. Since $U_1, U_2, \ldots, U_{13}$ is a basis of $\mathbf{U}$, $V_1, V_2, \ldots, V_6$ is a basis of $\mathbf{V}$, $W_1, W_2, \ldots, W_{134}$ is a basis of $\mathbf{W}$, and that any list $U \in \mathbf{U}$, $V \in \mathbf{V}$, and $W \in \mathbf{W}$ is linearly independent (as shown in 1.), the list

$$U_1, U_2, \ldots, U_{13}, V_1, V_2, \ldots, V_6, W_1, W_2, \ldots, W_{134}$$

is linearly independent. It, then, suffices to show that this list also spans $\mathbf{U} \oplus \mathbf{V} \oplus \mathbf{W}$.

Consider $H \in \mathbf{U} \oplus \mathbf{V} \oplus \mathbf{W}$, which can be expressed uniquely as

$$H = U_h + V_h + W_h,$$

where $U_h \in \mathbf{U}, V_h \in \mathbf{V}, W_h \in \mathbf{W}$. U_h can be expressed as a linear combination of $U_1, U_2, \ldots, U_{13}$ because $U_1, U_2, \ldots, U_{13}$ form a basis of $\mathbf{U}$. Similarly, V_h, W_h can be expressed uniquely as linear combinations of basis elements of their respective subspaces. It follows that H can be expressed as a linear combination of the given elements of the concatenated list, i.e, the concatenated list spans $\mathbf{U} \oplus \mathbf{V} \oplus \mathbf{W}$.

Therefore,

$$U_1, U_2, \ldots, U_{13}, V_1, V_2, \ldots, V_6, W_1, W_2, \ldots, W_{134}$$

is a basis of $\mathbf{U} \oplus \mathbf{V} \oplus \mathbf{W}$.

Problem. 5. Suppose that $x_1, x_2, \ldots, x_{14}$ are non-null elements of a linear space $\mathbf{W}$.

1. Argue that the following claims are equivalent.

   (a) $x_1, x_2, \ldots, x_{14}$ are linearly independent.

   (b) The subspace sum

   $$\text{span}(x_1) + \text{span}(x_2) + \cdots + \text{span}(x_{14})$$

   is direct.

2. Argue that the following claims are equivalent.

   (a) $x_1, x_2, \ldots, x_{14}$ is a basis of $\mathbf{W}$.

   (b) $\mathbf{W} = \text{span}(x_1) \oplus \text{span}(x_2) \oplus \cdots \oplus \text{span}(x_{14})$

Solution.

1. (a) $[a]$ implies $[b]$: If $x_1, x_2, \ldots, x_{14}$ are linearly independent, then any x_i is neither null nor can be expressed as a linear combination of the other x_j's. Therefore, no non-null $x_i \in \text{span}(x_i)$ can be written as a sum of the other x_j's $\in \text{span}(x_j)$, i.e.,

   $$\text{span}(x_1) + \text{span}(x_1) + \cdots + \text{span}(x_{14})$$

   is direct, by Problem 1.

   (b) $[b]$ implies $[a]$: By problem 1., if

   $$\text{span}(x_1) + \text{span}(x_2) + \cdots + \text{span}(x_{14})$$

   is direct, then no non-null $x_i \in \text{span}(x_i)$ can be written as a sum of the other x_j's $\in \text{span}(x_j)$'s. Therefore, the list $x_1, x_2, \ldots, x_{14}$ is linearly independent.

2. Because x_i is a basis of $\text{span}(x_i)$ and the subspace sum $\text{span}(x_1)+\text{span}(x_2)+\cdots+\text{span}(x_{14})$ is direct, the concatenated list $x_1, x_2, \ldots, x_{14}$ is a basis of $\text{span}(x_1)\oplus\text{span}(x_2)\oplus\cdots\oplus\text{span}(x_{14})$, by Problem 4 and part 1 of Problem 5. (†)

   (a) $[a]$ implies $[b]$: By (†) and assumption, $\mathbf{W}$ and $\text{span}(x_1)\oplus\text{span}(x_2)\oplus\cdots\oplus\text{span}(x_{14})$ have a common basis, namely the list $x_1, x_2, \ldots, x_{14}$. Therefore, $\mathbf{W} = \text{span}(x_1) \oplus \text{span}(x_2) \oplus \cdots \oplus \text{span}(x_{14})$.

   (b) $[b]$ implies $[a]$: It follows from fact (†) that if $\mathbf{W} = \text{span}(x_1) \oplus \text{span}(x_2) \oplus \cdots \oplus \text{span}(x_{14})$, then $x_1, x_2, \ldots, x_{14}$ is a basis of $\mathbf{W}$.

25.2 Problem set 2

25.3 Problem set 3