MATH-UA. 0148 Honors Linear Algebra

Lecture 28: More on linear transformations

I] Some more unusual vector spaces

* Let P_n be the set of all polynomials with real coefficients of degree n and smaller:

$$P \in P_n \iff \exists (a_0, a_1, \ldots, a_n) \in \mathbb{R}^{n+1} : P(x) = a_n x^n + \ldots + a_1 x + a_0$$

P_n is a vector space:

- $\forall P \in P_n, Q \in P_n, \quad P(x) = a_n x^n + \ldots + a_1 x + a_0$

$$Q(x) = b_n x^n + \ldots + b_1 x + b_0$$

$(P+Q)(x) = (a_n + b_n) x^n + \ldots + (a_1 + b_1) x + a_0 + b_0 \quad \in P_n$

- By commutativity of addition for real numbers $(P+Q)(x) = (Q+P)(x)$

- If $R(x) = c_n x^n + \ldots + c_1 x + c_0$,

$(P+Q)(x) + R(x) = P(x) + (Q+R)x$ by associativity of addition for real numbers

- $O(x) = 0 \in P_n$ is such that $(P+O)(x) = (O+P)(x) = P(x)$

- $\forall P \in P_n, \quad P(x) = a_n x^n + \ldots + a_1 x + a_0, \quad T(x) = (-a_n) x^n + \ldots + (-a_1) x + (-a_0)$ is such that $(P+T)(x) = (T+P)(x) = O(x) = 0$

T is the additive inverse of P.

- $\forall k \in \mathbb{R}, \forall P \in P_n, (kP)(x) = ka_n x^n + ka_{n-1}x^{n-1} + \dots + ka_1 x + ka_0 \in P_n$

- $\forall k \in \mathbb{R}, \forall h \in \mathbb{R}, \forall P \in P_n, (k+h)P(x) = kP(x) + kh(x)$
  by distributivity for real numbers

- Likewise, $\forall k \in \mathbb{R}, \forall h \in \mathbb{R}, \forall P \in P_n, (kh)P(x) = k.(hP(x))$
  by associativity of multiplication for real numbers.

- $1 \cdot P(x) = P(x) \quad \forall P \in P_n$

A basis for P_n is $\{1, x, x^2, \dots, x^n\}$

Indeed, it is clear that $\{1, x, x^2, \dots, x^n\}$ spans P_n. Furthermore, the vectors in the set are linearly independent:
If there exist $(c_0, c_1, \dots, c_n) \in \mathbb{R}^{n+1}$ such that $c_n x^n + \dots + c_1 x + c_0 \cdot 1 = 0$, then $c_n = 0, c_{n-1} = 0, \dots, c_1 = 0, c_0 = 0$.

We conclude that the dimension of P_n is $\mathbb{R}^{n+1}$.

Another way to see this is that to any element P of P_n, one can uniquely associate the vector $\begin{bmatrix} a_0 \\ a_1 \\ \vdots \\ a_n \end{bmatrix} \in \mathbb{R}^{n+1}$ of its coefficients.

* Let $V = \left\{ f : f : \mathbb{R} \longrightarrow \mathbb{R} \text{ such that } \frac{d}{dx} f(x) \text{ exists for all } x \in \mathbb{R} \right\}$

with vector addition defined as $(f+g)(x) = f(x) + g(x)$
and multiplication by a scalar defined by

$$(kf)(x) = k\, f(x)$$

V is also a vector space, by the properties of real addition, real multiplication, and differentiation.

It turns out that this vector space has infinite dimension.

II More unusual linear transformations

* Consider the function $L : P_n \longrightarrow P_n$

$$P \in P_n \longmapsto \frac{dP}{dx}$$

L is a linear transformation.
Indeed, for any $P \in P_n$, $Q \in P_n$,

$$L(P+Q) = \frac{d}{dx}\left[P(x) + Q(x)\right] = \frac{dP}{dx} + \frac{dQ}{dx} = LP + LQ$$

$\forall P \in P_n, \forall k \in \mathbb{R}, \; L(kP) = \frac{d}{dx}\left[k P(x)\right] = k \frac{dP}{dx} = k LP$

* Let us return to $V = \left\{ f : f : \mathbb{R} \longrightarrow \mathbb{R} \text{ such that } \frac{d}{dx} f(x) \text{ exists for all } x \in \mathbb{R} \right\}$

Let $a \in \mathbb{R}$, $b \in \mathbb{R}$ such that $a < b$.

The function
$$L : V \longrightarrow \mathbb{R}$$
$$f \in V \longmapsto L(f) = \int_a^b f(x)dx$$

is a linear transformation, by the properties of integration: for any $f \in V$, $g \in V$, $k \in \mathbb{R}$,

$$L(k(f+g)) = \int_a^b k(f(x) + g(x))\, dx = k\int_a^b f(x)dx + k\int_a^b g(x)dx$$

III) Kernel, range, nullity, rank

1) Range

Let $f : S \longrightarrow T$ be a function (not necessarily a linear transformation at this point).
S is called the domain of f, and T the codomain of f.

The range, or image, of f is the set:

$$\operatorname{ran}(f)\ (= \operatorname{im}(f)) = f(S) = \{ f(s) : s \in S \} \subset T$$

The preimage of any subset U of the codomain T is:
$$f^{-1}(U) = \{ s \in S : f(s) \in U \} \subset S$$

2) Injectivity, surjectivity, bijectivity

- A function $f: S \longrightarrow T$ is $\begin{cases}\text{one-to-one} \\ \text{injective}\end{cases}$ if $\forall (x,y) \in S^2$,

$$x \neq y \implies f(x) \neq f(y)$$

Example: $f: \mathbb{R} \longrightarrow \mathbb{R}$

$\qquad x \longmapsto e^x$

is injective

$f: \mathbb{R} \longrightarrow \mathbb{R}$

$\qquad x \longmapsto x^2$

is not injective (or one-to-one).

- A function $f: S \longrightarrow T$ is $\begin{cases}\text{onto} \\ \text{surjective}\end{cases}$ if $\forall t \in T$, there exists $s \in S$ such that $f(s) = t$

Example: $f: \mathbb{R} \longrightarrow \mathbb{R}^+$

$\qquad x \longmapsto e^x$

is surjective (or onto).

$f: \mathbb{R} \longrightarrow \mathbb{R}$

$\qquad x \longmapsto e^x$

is not surjective.

- A function which is both one-to-one and onto, i.e both injective and surjective, is said to be bijective.

Example: $f: \mathbb{R} \longrightarrow \mathbb{R}^+$

$$x \longmapsto e^x$$

is bijective.

Theorem: A function $f: S \longrightarrow T$ has an inverse function $g: T \longrightarrow S$ (such that $\begin{cases} f\circ g(t) = t & \forall t \in T \\ g\circ f(s) = s & \forall s \in S \end{cases}$) if and only if it is bijective.

Proof: • Suppose f is bijective. Then, any $t \in T$ has a unique preimage $s \in S$. Let g be the function which for any $t \in T$ assigns the unique preimage $s = f^{-1}(t)$ of t.

For any $t \in T$, $f(g(t)) = f(s) = f(f^{-1}(t)) = t$

• Conversely, suppose f has an inverse function g.
Let $x \in S$, $y \in S$ such that $f(x) = f(y)$
Then $g(f(x)) = g(f(y)) \Leftrightarrow x = y$, f is one-to-one.

Let $t \in T$. $f(g(t)) = t$ so $g(t)$ is a preimage of t in S. f is surjective.

It is now time to focus on the particular case for which f is a linear transformation between two vector spaces. The linearity of the function and the vector spaces will allow to derive many more results than for general functions.

3) Kernel of a linear transformation

Definition: Let $L: V \longrightarrow W$ be a linear transformation. The set of all vectors $\vec{v} \in V$ such that $L\vec{v} = \vec{0}$ is called the kernel of L, written $\ker L$.

Note 1: When we write $L\vec{v} = \vec{0}$, the $\vec{0}$ here is the zero element for vector addition in W. One sometimes writes $\vec{0}_W$ for clarity.

Note 2: We have seen in lecture 27 that if L is a linear transformation, then $\vec{0}_V \in \ker L$.

Theorem: A linear transformation L is injective if and only if $\ker L = \{\vec{0}_V\}$

Proof: Let L be a linear transformation such that $\ker L = \{\vec{0}_V\}$, and let $\vec{x}$ and $\vec{y}$ such that $L\vec{x} = L\vec{y}$

$$\text{Then } L(\vec{x} - \vec{y}) = \vec{0}_W \overset{\ker L = \{\vec{0}_V\}}{\Longrightarrow} \vec{x} - \vec{y} = \vec{0}_V \Leftrightarrow \vec{x} = \vec{y}$$

L is injective.

Conversely, if L is one-to-one, then we know that the only vector such that $L\vec{x} = \vec{0}_W$ is $\vec{x} = \vec{0}_V$:

$$\ker L = \{\vec{0}_V\}$$

This completes our proof.

Observe that if L has a matrix representation M in some basis, then finding the kernel of L is equivalent to finding the nullspace of M.

In this context, the following theorem makes sense.

Theorem: Let $L: V \longrightarrow W$ be a linear transformation. $\operatorname{Ker} L$ is a subspace of V.

Proof: • As mentioned, $\vec{0}_V \in \operatorname{Ker} L$.

• Let $\vec{u} \in \operatorname{Ker} L$, $\vec{v} \in \operatorname{Ker} L$, and $k \in \mathbb{R}$, $h \in \mathbb{R}$.
$L(k\vec{u} + h\vec{v}) = k L(\vec{u}) + h L(\vec{v}) = \vec{0}_W$ ■

4) Rank and nullity of a linear transformation

Theorem: Let $L: V \longrightarrow W$ be a linear transformation. The image $L(V)$ is a subspace of W.

Proof: • $\vec{0}_V \in V$ and $L(\vec{0}_V) = \vec{0}$ so $\vec{0} \in W$
• Let $\vec{u} \in L(V)$, $\vec{v} \in L(V)$. There exists $\vec{x} \in V$ and $\vec{y} \in V$ such that $L(\vec{x}) = \vec{u}$ and $L(\vec{y}) = \vec{v}$
For any $k \in \mathbb{R}$ and $h \in \mathbb{R}$, $k\vec{u} + h\vec{v} = kL(\vec{x}) + hL(\vec{y})$
$= L(k\vec{x} + h\vec{y})$

$k\vec{x} + h\vec{y} \in V$ since V is a vector space, so $k\vec{u} + h\vec{v} \in W$ ■

<u>Definition:</u> The rank of a linear transformation L is the dimension of its image $L(V)$:

$$\text{rank } L = \dim L(V)$$

The nullity of a linear transformation is the dimension of the kernel of L:

$$\text{null } L = \dim \text{Ker}(L)$$

<u>Theorem</u> (Dimension Formula):
Let $L: V \longrightarrow W$ be a linear transformation, where V is a finite dimensional space. Then

$$\dim V = \dim(\text{Ker } L) + \dim L(V)$$
$$= \text{null } L + \text{rank } L$$

<u>Proof:</u> Let $n = \dim V$, and $n-r = \dim(\text{Ker } L)$, with $r \geq 0$.
Let $(\vec{u}_1, \vec{u}_2, \dots, \vec{u}_{n-r})$ be a basis of $\text{Ker } L$, which we complete as $(\vec{u}_1, \vec{u}_2, \dots, \vec{u}_{n-r}, \vec{v}_1, \dots, \vec{v}_r)$ for a basis of V.

We want to show that rank $L = r$. To do so, we will show that $\{L(\vec{v}_1), \dots, L(\vec{v}_r)\}$ is a basis of $L(V)$

Let us first prove that $L(V) = \text{span}\{L(\vec{v}_1), \dots, L(\vec{v}_r)\}$

Consider $\vec{w} \in L(V)$. We can write

$$\vec{w} = L(c_1\vec{u}_1 + \dots + c_{n-r}\vec{u}_{n-r} + d_1\vec{v}_1 + \dots + d_r\vec{v}_r)$$
$$= d_1 L(\vec{v}_1) + \dots + d_r L(\vec{v}_r)$$

which proves our point.

Now we need to prove that $\{L(\vec{v}_1), \ldots, L(\vec{v}_r)\}$ is linearly independent.

Suppose it is not: there exists $(d_1, \ldots, d_r) \in \mathbb{R}^r$ such that $d_1 L(\vec{v}_1) + \ldots + d_r L(\vec{v}_r) = \vec{0}_W$

not all zero

$\iff L(d_1 \vec{v}_1 + \ldots + d_r \vec{v}_r) = \vec{0}_W$

$\iff d_1 \vec{v}_1 + \ldots + d_r \vec{v}_r \in \operatorname{Ker} L$

Now $\{\vec{v}_1, \ldots, \vec{v}_r\}$ is linearly independent, so $d_1 \vec{v}_1 + \ldots + d_r \vec{v}_r \neq \vec{0}$

Hence $d_1 \vec{v}_1 + \ldots + d_r \vec{v}_r$ is a linear combination of $\vec{u}_1, \vec{u}_2, \ldots, \vec{u}_{n-r}$. This contradicts the fact that $\{\vec{u}_1, \vec{u}_2, \ldots, \vec{u}_{n-r}, \vec{v}_1, \ldots, \vec{v}_r\}$ is a basis of V

We conclude that $\{L(\vec{v}_1), \ldots, L(\vec{v}_r)\}$ is linearly independent.

Hence, $\{L(\vec{v}_1), \ldots, L(\vec{v}_r)\}$ is a basis of $L(V)$, so rank $L = r$.

The theorem is not a surprise for us. If L has a matrix representation M in some basis, we know that:

$$\dim(N(M)) = n-r \quad , \quad \dim(C(M)) = r$$

Hence $\dim(N(M)) + \dim(C(M)) = n - r + r = n$

The theorem has an important corollary:
Let $L: V \longrightarrow W$ be a linear transformation, where V is a finite dimensional space, and $\dim(V) = \dim(W)$.
Then L is injective if and only if it is surjective.
In other words, L is injective if and only if L is bijective and has an inverse L^{-1}.

Proof: L is injective if and only if $\dim(\operatorname{Ker} L) = 0$, i.e
if and only if $\dim(L(V)) = \dim V = \dim W$, i.e
if and only if L is surjective.

From this corollary, we conclude that there is no one-to-one linear transformation from V to W if $\dim W < \dim V$

There is no onto linear transformation from V to W if $\dim V < \dim W$

IV) Sum of subspaces

1) Definition

Suppose S_1 and S_2 are two subspaces of a vector space V.
We define:

$$S_1 + S_2 = \left\{ \vec{s_1} + \vec{s_2} : \vec{s_1} \in S_1 \text{ and } \vec{s_2} \in S_2 \right\}$$

It is a good exercise to prove that $S_1 + S_2$ is a subspace of V (proof left for the reader), which is called the sum of S_1 and S_2.

Example: Let $V = \mathbb{R}^2$, $S_1 = \{(x,x) : x \in \mathbb{R}\} \subset \mathbb{R}^2$,

$S_2 = \{(x,-x) : x \in \mathbb{R}\} \subset \mathbb{R}^2$.

Then $S_1 + S_2 = \mathbb{R}^2$, since $\forall (x,y) \in \mathbb{R}^2$,

$$(x,y) = \left(\frac{x+y}{2}, \frac{x+y}{2}\right) + \left(\frac{x-y}{2}, -\frac{x-y}{2}\right)$$

<u>Theorem:</u> If S_1 and S_2 are subspaces of a vector space V, all with finite dimension, then

$$\dim(S_1 + S_2) = \dim(S_1) + \dim(S_2) - \dim(S_1 \cap S_2)$$

<u>Proof:</u> Let $B = (\vec{b}_1, \vec{b}_2, \ldots, \vec{b}_r)$ be a basis of $S_1 \cap S_2$.

B can be complemented to a basis $B_1 = (\vec{b}_1, \ldots, \vec{b}_r, \vec{c}_1, \ldots, \vec{c}_s)$ of S_1 and to a basis $B_2 = (\vec{b}_1, \ldots, \vec{b}_r, \vec{d}_1, \ldots, \vec{d}_t)$ of S_2.

$\dim(S_1 \cap S_2) = r$, $\dim(S_1) = r+s$, $\dim(S_2) = r+t$.

Let $\mathcal{B} = (\vec{b}_1, \vec{b}_2, \ldots, \vec{b}_r, \vec{c}_1, \ldots, \vec{c}_s, \vec{d}_1, \ldots, \vec{d}_t)$. If $\mathcal{B}$ is a basis of $S_1 + S_2$, then

$$\begin{aligned}\dim(S_1+S_2) &= r+s+t \\ &= (r+s)+(r+t)-r \\ &= \dim(S_1) + \dim(S_2) - \dim(S_1 \cap S_2)\end{aligned}$$

Let us prove $\mathcal{B}$ is a basis of $S_1 + S_2$.

It is clear that $S_1 + S_2 \subset \text{span}(\mathcal{B})$.

Indeed, let $\vec{w} \in S_1 + S_2$. There exist $\vec{w}_1 = \ell_1 \vec{b}_1 + \dots + \ell_r \vec{b}_r + \ell_{r+1} \vec{c}_1 + \dots + \ell_{r+s} \vec{c}_s$

and $\vec{w}_2 = h_1 \vec{b}_1 + \dots + h_r \vec{b}_r + h_{r+1} \vec{d}_1 + \dots + h_{r+t} \vec{d}_t$ such that

$$\vec{w} = \vec{w}_1 + \vec{w}_2 = (\ell_1 + h_1)\vec{b}_1 + \dots + (\ell_r + h_r)\vec{b}_r + \ell_{r+1}\vec{c}_1 + \dots + \ell_{r+s}\vec{c}_s + h_{r+1}\vec{d}_1 + \dots + h_{r+t}\vec{d}_t \in \text{span}(\mathcal{B})$$

Let us now show $\mathcal{B}$ is linearly independent. Consider coefficients $\ell_1, \dots, \ell_{r+s}, h_1, \dots, h_t$ such that

$$\ell_1 \vec{b}_1 + \ell_2 \vec{b}_2 + \dots + \ell_r \vec{b}_r + \ell_{r+1} \vec{c}_1 + \dots + \ell_{r+s} \vec{c}_s + h_1 \vec{d}_1 + \dots + h_t \vec{d}_t = \vec{0}$$

$$\iff \underbrace{\underbrace{\ell_1 \vec{b}_1 + \ell_2 \vec{b}_2 + \dots + \ell_r \vec{b}_r + \ell_{r+1} \vec{c}_1 + \dots + \ell_{r+s} \vec{c}_s}_{\in S_1}}_{\Rightarrow\ \in S_1 \cap S_2} = \underbrace{-(h_1 \vec{d}_1 + \dots + h_t \vec{d}_t)}_{\in S_2}$$

There exists $p_1, \dots, p_r$ such that $h_1 \vec{d}_1 + \dots + h_t \vec{d}_t = p_1 \vec{b}_1 + \dots + p_r \vec{b}_r$

$$\iff p_1 \vec{b}_1 + \dots + p_r \vec{b}_r + h_1 \vec{d}_1 + \dots + h_t \vec{d}_t = \vec{0}$$

$$\underset{\substack{B_2 \text{ is a} \\ \text{basis of} \\ S_2}}{\iff} \begin{cases} p_i = 0 & \text{for } i = 1, 2, \dots, r \\ h_i = 0 & \text{for } i = 1, 2, \dots, t \end{cases}$$

Hence $\ell_1 \vec{b}_1 + \ell_2 \vec{b}_2 + \dots + \ell_r \vec{b}_r + \ell_{r+1} \vec{c}_1 + \dots + \ell_{r+s} \vec{c}_s = \vec{0}$

$$\overset{\substack{B_1 \text{ is a} \\ \text{basis of } S_1}}{\iff} \ell_i = 0 \text{ for } i = 1, \dots, r+s$$

so $\mathcal{B}$ is linearly independent, and indeed a basis of $S_1 + S_2$ ■

2) Direct sum of subspaces

Definition: The sum $W_1 + W_2$ is called direct if $W_1 \cap W_2 = \{\vec{0}_V\}$.

A vector space V is said to be the direct sum of two subspaces W_1 and W_2 if $V = W_1 + W_2$ and $W_1 \cap W_2 = \{\vec{0}_V\}$.

When V is a direct sum of W_1 and W_2, we write $V = W_1 \oplus W_2$.

Theorem: Suppose W_1 and W_2 are subspaces of a vector space V such that $V = W_1 + W_2$. Then $V = W_1 \oplus W_2$ if and only if every vector $\vec{v} \in V$ can be written in a unique way as $\vec{v} = \vec{w}_1 + \vec{w}_2$, with $\vec{w}_1 \in W_1$ and $\vec{w}_2 \in W_2$

Proof: Suppose $V = W_1 \oplus W_2$ and assume there exist $\vec{w}_1$ and $\vec{w}_1'$ in W_1 and $\vec{w}_2$ and $\vec{w}_2'$ in W_2 such that $\vec{v} \in V$ can be written as $\vec{v} = \vec{w}_1 + \vec{w}_2 = \vec{w}_1' + \vec{w}_2'$.
Then,
$$\underbrace{\vec{w}_1 - \vec{w}_1'}_{\in W_1} = \underbrace{\vec{w}_2' - \vec{w}_2}_{\in W_2}$$
Since $V = W_1 \oplus W_2$, this implies $\vec{w}_1 - \vec{w}_1' = \vec{0}_V = \vec{w}_2' - \vec{w}_2$, so that $\vec{w}_1 = \vec{w}_1'$, $\vec{w}_2 = \vec{w}_2'$.

Conversely, suppose there is a unique way to write $\vec{v} = \vec{w}_1 + \vec{w}_2$, with $\vec{v} \in V$, $\vec{w}_1 \in W_1$, and $\vec{w}_2 \in W_2$.
Let us assume there exists $\vec{u} \neq \vec{0}_V$ such that $\vec{u} \in W_1 \cap W_2$.
Then $\vec{u} \in V$, and $\vec{u} = \underset{\in W_1}{\vec{u}} + \underset{\in W_2}{\vec{0}_V} = \underset{\in W_1}{\vec{0}_V} + \underset{\in W_2}{\vec{u}}$, which contradicts the uniqueness of the decomposition.
Thus, the only vector in $W_1 \cap W_2$ is $\vec{0}_V$, and $V = W_1 \oplus W_2$ ∎

Examples:

* Let M be an $m \times n$ matrix.
We saw in Lecture 14 that $C(M^T) \oplus N(M) = \mathbb{R}^n$ and that $C(M) \oplus N(M^T) = \mathbb{R}^m$

* Let V be the vector space of all $n \times n$ matrices with real coefficients. Let W_1 be the subspace of all $n \times n$ symmetric matrices, and W_2 the subspace of antisymmetric matrices.
You showed in Problem Set 3 that $V = W_1 \oplus W_2$

* Let V be the vector space of all $n \times n$ matrices with real coefficients. Let W_1 be the subspace of upper triangular matrices, and W_2 the subspace of lower triangular matrices.
Can we write $V = W_1 \oplus W_2$?

* Let M be an $n \times n$ diagonalizable matrix with real coefficients, and let $\lambda_1, \lambda_2, \dots, \lambda_p$ be the eigenvalues

of M, with $r \leq n$.
We define the subspaces spanned by the eigenvectors for each distinct eigenvalue λ_i:

$$E_{\lambda_i} = \{ \vec{x} \in \mathbb{R}^n : M\vec{x} = \lambda_i \vec{x} \}$$

Since M is diagonalizable, we can write

$$\mathbb{R}^n = E_{\lambda_1} \oplus E_{\lambda_2} \oplus \dots \oplus E_{\lambda_r}$$

Direct sums and dimension

At the beginning of this section, we showed that if $V = W_1 + W_2$, then
$\dim V = \dim(W_1) + \dim(W_2) - \dim(W_1 \cap W_2) \leq \dim(W_1) + \dim(W_2)$

We easily conclude that when $V = W_1 \oplus W_2$,

$\dim V = \dim W_1 + \dim W_2$

V) Direct sums and projections

Proposition 1: If V is a vector space and $P : V \longrightarrow V$ is a projection, then

$$V = \operatorname{Im} P \oplus \operatorname{Ker} P$$

Proof: • Let us first show that $V = \operatorname{Im} P + \operatorname{Ker} P$
Clearly, $\operatorname{Ker} P \subseteq V$ and $\operatorname{Im} P \subset V$, so $\operatorname{Im} P + \operatorname{Ker} P \subseteq V$ since V is a vector space.

Now, let $\vec{x} \in V$. $\vec{y} = P(\vec{x}) \in \text{Im}\, P$. $\vec{z} = \vec{x} - \vec{y}$ is such that $P(\vec{z}) = P(\vec{x}) - P(\vec{y}) = \vec{y} - \vec{y} = \vec{0}$, so $\vec{z} \in \text{Ker}\, P$, and we constructed $\vec{x} = \underbrace{\vec{y}}_{\in \text{Im}\, P} + \underbrace{\vec{z}}_{\in \text{Ker}\, P}$

$V \subseteq \text{Ker}\, P + \text{Im}\, P$, so $V = \text{Ker}\, P + \text{Im}\, P$

Finally, let $\vec{x} \in \text{Im}\, P$. There exists $\vec{u} \in V$ such that $\vec{x} = P(\vec{u})$. Hence, $P(\vec{x}) = P^2(\vec{u}) = P(\vec{u}) = \vec{x}$

$\vec{x} \in \text{Im}\, P \cap \text{Ker}\, P \Leftrightarrow \vec{x} = P(\vec{x}) \underset{\vec{x} \in \text{Ker}\, P}{=} \vec{0}$

so $\vec{x} = \vec{0}_V$

We conclude that $\text{Ker}\, P \oplus \text{Im}\, P = V$, as desired.

Proposition 2: Let V be a vector space, and W_1 and W_2 two subspaces such that $V = W_1 \oplus W_2$. Then there exists a unique projection $P: V \to V$ such that $\text{Im}\, P = W_1$ and $\text{Ker}\, P = W_2$

Proof: $\forall \vec{v} \in V$, there is a unique decomposition $\vec{v} = \underbrace{\vec{w_1}}_{\in W_1} + \underbrace{\vec{w_2}}_{\in W_2}$

Let us then consider the transformation

$$P: V \longrightarrow W_1$$
$$\vec{v} = \vec{w_1} + \vec{w_2} \longmapsto P(\vec{v}) = \vec{w_1}$$

Let $(\vec{v_1}, \vec{v_2}) \in V^2$, $(k, h) \in \mathbb{R}^2$. $\vec{v_1} = \vec{w_1} + \vec{w_2}$, $\vec{v_2} = \vec{w_1}' + \vec{w_2}'$

$k\vec{v_1} + h\vec{v_2} = k\vec{w_1} + h\vec{w_1}' + k\vec{w_2} + h\vec{w_2}'$

$P(k\vec{v_1} + h\vec{v_2}) = P(k\vec{w_1} + h\vec{w_1}' + k\vec{w_2} + h\vec{w_2}') = k\vec{w_1} + h\vec{w_1}' = kP(\vec{w_1}) + hP(\vec{w_1}') = kP(\vec{v_1}) + hP(\vec{v_2})$

and in particular $P(\vec{0}_V) = \vec{0}_V$
We conclude that P is a linear transformation

Let $\vec{v} \in V$, $\vec{v} = \vec{w}_1 + \vec{w}_2$
$P(\vec{v}) = \vec{w}_1$
$P^2(\vec{v}) = P(\vec{w}_1) = P(\vec{w}_1 + \vec{0}) = \vec{w}_1 = P(\vec{v})$
$P^2 = P$, so P is a projection.

By the definition of P, $\operatorname{Im} P \subseteq W_1$
Now, let $\vec{w}_1 \in W_1$. Then $\vec{w}_1 + \vec{0} \in V$, and
$P(\vec{w}_1 + \vec{0}) = \vec{w}_1 \Rightarrow \vec{w}_1 \in \operatorname{Im} P$, and $W_1 \subseteq \operatorname{Im} P$
Hence $W_1 = \operatorname{Im} P$

Likewise, by definition of P, $W_2 \subseteq \operatorname{Ker} P$
Now, let $\vec{v} \in \operatorname{Ker} P$. $\vec{v} = \underbrace{\vec{w}_1}_{\in W_1} + \underbrace{\vec{w}_2}_{\in W_2}$

$P(\vec{v}) = \vec{0}_V = P(\vec{w}_1 + \vec{w}_2) = \vec{w}_1$
Hence $\vec{w}_1 = \vec{0}_V$, $\vec{v} = \vec{w}_2 \in W_2 \Rightarrow \operatorname{Ker} P \subseteq W_2$
$W_2 = \operatorname{Ker} P$

Example: For $V = W_1 \oplus W_2$, where V is the vector space of all $n \times n$ matrices with real coefficients, W_1 the subspace of all symmetric matrices and W_2 the subspace of all antisymmetric matrices, you showed in Problem Set 2 that

$$P(M) = \frac{M + M^T}{2} \text{ is the projection.}$$