Linear Transformations From Direct Sums

Recall that we say

$$V = V_1 \oplus V_2$$

if for every $\vec{v}$ in V we can write $\vec{v} = \vec{v_1} + \vec{v_2}$ with $\vec{v_1} \in V_1$ and $\vec{v_2} \in V_2$ in a unique way.

Lemma

Suppose $V = V_1 \oplus V_2$ and W is any other vector space. Suppose also $L_1 : V_1 \to W$ and $L_2 : V_2 \to W$ are linear transformations.

Then we can define a **function** $L : V \to W$ according to the rule

$$\forall \vec{v_1} \in V_1, \forall \vec{v_2} \in V_2, L(\vec{v_1} + \vec{v_2}) = L_1(\vec{v_1}) + L_2(\vec{v_2})$$

Explanation: To convince you this is a function I need to convince you of two things:

1. For every value in the domain, V, this rule associates at least one value in the codomain, W.
2. For every value in the domain, V, the value associated value in the codomain, W, is unique.

For the first point, we note that given an arbitrary $\vec{v}$ in V, that we can find $\vec{v_1} \in V_1$ and $\vec{v_2} \in V_2$ so that $\vec{v} = \vec{v_1} + \vec{v_2}$ and so the above rule tells us

$$L(\vec{v}) = L(\vec{v_1} + \vec{v_2}) = L_1(\vec{v_1}) + L_2(\vec{v_2})$$

and because the rules L_1 and L_2 (and $'+'$) are functions, this does give us a value in W. Importantly, this proceedure worked for an arbitrary element of V.

For the second point, we note that the above proceedure gives a unique value, because once we had fixed $\vec{v}$ the process involved **no** choices. That is $\vec{v_1}$ and $\vec{v_2}$ were uniquely determined, this is a condition of $V = V_1 \oplus V_2$, the values of $L_1(\vec{v_1})$ and $L_2(\vec{v_2})$ were uniquely determined because these are functions, as is the value of their sum.

Linear Transformations From Direct Sums

Lemma

Suppose $V = V_1 \oplus V_2$ and W is any other vector space. Suppose also $L_1 : V_1 \to W$ and $L_2 : V_2 \to W$ are linear transformations.

The function $L : V \to W$ given by the rule

$$\forall \vec{v}_1 \in V_1, \forall \vec{v}_2 \in V_2, L(\vec{v}_1 + \vec{v}_2) = L_1(\vec{v}_1) + L_2(\vec{v}_2)$$

is linear.

Proof Let $\vec{x}, \vec{y} \in V$ be arbitrary and $a, b \in \mathbb{R}$ be arbitrary.

We can write $\vec{x} = \vec{x}_1 + \vec{x}_2$ where $\vec{x}_i \in V_i$ and $\vec{y} = \vec{y}_1 + \vec{y}_2$ where $\vec{y}_i \in V_i$.

Denote $\vec{z}_1 = a\vec{x}_1 + b\vec{y}_1 \in V_1$ and $\vec{z}_2 = a\vec{x}_2 + b\vec{y}_2 \in V_2$ then we have

$$\begin{aligned}
L(a\vec{x} + b\vec{y}) &= L(a(\vec{x}_1 + \vec{x}_2) + b(\vec{y}_1 + \vec{y}_2)) && \\
&= L(\vec{z}_1 + \vec{z}_2) && \text{def of } \vec{z}_1, \vec{z}_2 \\
&= L_1(\vec{z}_1) + L_2(\vec{z}_2) && \text{rule for } L \\
&= L_1(a\vec{x}_1 + b\vec{y}_1) + L_2(a\vec{x}_2 + b\vec{y}_2)) && \text{def of } \vec{z}_1, \vec{z}_2 \\
&= aL_1(\vec{x}_1) + bL_1(\vec{y}_1) + aL_2(\vec{x}_2) + bL_2(\vec{y}_2)) && \text{Linearity of } L_1, L_2 \\
&= a(L_1(\vec{x}_1) + L_2(\vec{x}_2)) + b(L_1(\vec{y}_1) + L_2(\vec{y}_2))) && \text{rearrange terms} \\
&= aL(\vec{x}) + bL(\vec{y}) && \text{Definition of } L
\end{aligned}$$

which proves the function is linear.

Linear Transformations From Direct Sums

Recall If $U \subset V$ then $\mathrm{incl}_U : U \to V$ is the function $\mathrm{incl}_U(\vec{u}) = \vec{u}$.

Lemma
Suppose $V = V_1 \oplus V_2$ and W is any other vector space and suppose also $L : V \to W$ is a linear transformation. Then with

$$L_1 = L \circ \mathrm{incl}_{V_1} : V_1 \to W \qquad L_2 = L \circ \mathrm{incl}_{V_2} : V_2 \to W$$

The map L satisfies

$$\forall \vec{v}_1 \in V_1, \forall \vec{v}_2 \in V_2, L(\vec{v}_1 + \vec{v}_2) = L_1(\vec{v}_1) + L_2(\vec{v}_2)$$

Proof:
Let $\vec{v}_1 \in V_1, \vec{v}_2 \in V_2$ be arbitrary then calculate the left hand side by linearity

$$L(\vec{v}_1 + \vec{v}_2) = L(\vec{v}_1) + L(\vec{v}_2)$$

Then we calculate the right hand side:

$$L_1(\vec{v}_1) + L_2(\vec{v}_2) = L \circ \mathrm{incl}_{V_1}(\vec{v}_1) + L \circ \mathrm{incl}_{V_2}(\vec{v}_2) = L(\vec{v}_1) + L(\vec{v}_2)$$

which proves the equality.

Linear Transformations From Direct Sums

Theorem

Every linear transformation $L : V_1 \oplus V_2 \to W$ is **uniquely** determined by two linear transformations

$$L_1 : V_1 \to W \qquad L_2 : V_2 \to W$$

according to the rule

$$\forall \vec{v}_1 \in V_1, \forall \vec{v}_2 \in V_2, L(\vec{v}_1 + \vec{v}_2) = L_1(\vec{v}_1) + L_2(\vec{v}_2)$$

Follows from previous three lemmas

Abstract Example

The projection maps:

$$\mathrm{Proj}_{V_1} : V_1 \oplus V_2 \to V_1 \qquad \mathrm{Proj}_{V_2} : V_1 \oplus V_2 \to V_2$$

where characterized by the rules

$$\mathrm{Proj}_{V_1}(\vec{v}_1 + \vec{v}_2) = \vec{v}_1 \qquad \mathrm{Proj}_{V_2}(\vec{v}_1 + \vec{v}_2) = \vec{v}_2$$

So $\mathrm{Proj}_{V_1} : V_1 \oplus V_2 \to V_1$ is associated to

- $\mathrm{Id}_{V_1} : V_1 \to V_1$, the identity map from V_1 to itself, and
- $O_{V_2,V_1} : V_2 \to V_1$, the zero map from V_2 to V_1.

Whereas $\mathrm{Proj}_{V_2} : V_1 \oplus V_2 \to V_2$ is associated to

- $O_{V_1,V_2} : V_1 \to V_2$, the zero map from V_1 to V_2.
- $\mathrm{Id}_{V_2} : V_2 \to V_2$, the identity map from V_2 to itself, and

Matrix Example

Recall the example $V = \mathbb{R}^5$ and

$$V_1 = \{(a, b, 0, 0, 0) \in V \mid a, b \in \mathbb{R}\} \qquad V_2 = \{(0, 0, c, d, e) \in V \mid c, d, e \in \mathbb{R}\}$$

then $V = V_1 \oplus V_2$, where $V_1 \simeq \mathbb{R}^2$ and $V_2 \simeq \mathbb{R}^3$
To give a linear transformation $L : V \to \mathbb{R}^2$ is to give a 2 by 5 matrix.
Consider for example:

$$\begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 5 & 4 & 3 & 2 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 2 & 0 & 0 & 0 \\ 5 & 4 & 0 & 0 & 0 \end{pmatrix} + \begin{pmatrix} 0 & 0 & 3 & 4 & 5 \\ 0 & 0 & 3 & 2 & 1 \end{pmatrix}$$

Under this decomposition when we evalute the linear transformation:

$$\left(\left(\begin{smallmatrix} 1 & 2 & 0 & 0 & 0 \\ 5 & 4 & 0 & 0 & 0 \end{smallmatrix}\right) + \left(\begin{smallmatrix} 0 & 0 & 3 & 4 & 5 \\ 0 & 0 & 3 & 2 & 1 \end{smallmatrix}\right)\right)\left(\begin{pmatrix} a \\ b \\ 0 \\ 0 \\ 0 \end{pmatrix} + \begin{pmatrix} 0 \\ 0 \\ c \\ d \\ e \end{pmatrix}\right) = \left(\begin{smallmatrix} 1 & 2 & 0 & 0 & 0 \\ 5 & 4 & 0 & 0 & 0 \end{smallmatrix}\right)\begin{pmatrix} a \\ b \\ 0 \\ 0 \\ 0 \end{pmatrix} + \left(\begin{smallmatrix} 1 & 2 & 0 & 0 & 0 \\ 5 & 4 & 0 & 0 & 0 \end{smallmatrix}\right)\begin{pmatrix} 0 \\ 0 \\ c \\ d \\ e \end{pmatrix} + \left(\begin{smallmatrix} 0 & 0 & 3 & 4 & 5 \\ 0 & 0 & 3 & 2 & 1 \end{smallmatrix}\right)\begin{pmatrix} a \\ b \\ 0 \\ 0 \\ 0 \end{pmatrix} + \left(\begin{smallmatrix} 0 & 0 & 3 & 4 & 5 \\ 0 & 0 & 3 & 2 & 1 \end{smallmatrix}\right)\begin{pmatrix} 0 \\ 0 \\ c \\ d \\ e \end{pmatrix}$$

$$= \left(\begin{smallmatrix} 1 & 2 & 0 & 0 & 0 \\ 5 & 4 & 0 & 0 & 0 \end{smallmatrix}\right)\begin{pmatrix} a \\ b \\ 0 \\ 0 \\ 0 \end{pmatrix} + \left(\begin{smallmatrix} 0 & 0 & 3 & 4 & 5 \\ 0 & 0 & 3 & 2 & 1 \end{smallmatrix}\right)\begin{pmatrix} 0 \\ 0 \\ c \\ d \\ e \end{pmatrix}$$

The linear transformation $L_1 : V_1 \simeq \mathbb{R}^2 \to \mathbb{R}^2$ is the map $\begin{pmatrix} 1 & 2 \\ 5 & 4 \end{pmatrix}$.

The linear transformation $L_2 : V_2 \simeq \mathbb{R}^3 \to \mathbb{R}^2$ is the map $\begin{pmatrix} 3 & 4 & 5 \\ 3 & 2 & 1 \end{pmatrix}$.

Linear Transformations Into Direct Sums

Lemma

Suppose $W = W_1 \oplus W_2$ and V is any vectorspace.
If $L_1 : V \to W_1$ and $L_2 : V \to W_2$ are linear transformations then the function $L : V \to W$ given by:

$$L(\vec{v}) = L_1(\vec{v}) + L_2(\vec{v})$$

is a linear transformation.

Proof:

We let $\vec{v}_1, \vec{v}_2 \in V$ and $a, b \in \mathbb{R}$ be arbitrary.
Then

$$\begin{aligned}
L(a\vec{v}_1 + b\vec{v}_2) &= L_1(a\vec{v}_1 + b\vec{v}_2) + L_2(a\vec{v}_1 + b\vec{v}_2) && \text{definition of } L \\
&= aL_1(\vec{v}_1) + bL_1(\vec{v}_2) + aL_2(\vec{v}_1) + bL_2(\vec{v}_2) && \text{Linearity of } L_1, L_2 \\
&= a(L_1(\vec{v}_1) + L_2(\vec{v}_1)) + b(L_1(\vec{v}_2) + L_2(\vec{v}_2)) && \\
&= aL(\vec{v}_1) + bL(\vec{v}_2) && \text{definition of } L
\end{aligned}$$

which shows the map is linear.

Linear Transformations Into Direct Sums

Lemma

Suppose $W = W_1 \oplus W_2$ and V is any vectorspace and suppose $L : V \to W$ is any linear transformation. If we define $L_1 : V \to W_1$ and $L_2 : V \to W_2$ by:

$$L_1 = \mathrm{Proj}_{W_1} \circ L \qquad L_2 = \mathrm{Proj}_{W_2} \circ L$$

Then the function L satisfies

$$L(\vec{v}) = L_1(\vec{v}) + L_2(\vec{v})$$

Proof:

Let $\vec{v} \in V$ be arbitrary and for notational convenience write $\vec{w} = L(\vec{v})$.
Because we have a direct sum decomposition $\vec{w} = \vec{w}_1 + \vec{w}_2$ with $\vec{w}_1 \in W_1$ and $\vec{w}_2 \in W_2$.
By definition of projection $\mathrm{Proj}_{W_1}(\vec{w}) = \vec{w}_1$ and $\mathrm{Proj}_{W_2}(\vec{w}) = \vec{w}_2$. It follows that

$$\begin{aligned} L_1(\vec{v}) + L_2(\vec{v}) &= \mathrm{Proj}_{W_1} \circ L(\vec{v}) + \mathrm{Proj}_{W_2} \circ L(\vec{v}) \\ &= \mathrm{Proj}_{W_1}(\vec{w}) + \mathrm{Proj}_{W_2}(\vec{w}) \\ &= \vec{w}_1 + \vec{w}_2 = \vec{w} \\ &= L(\vec{v}) \end{aligned}$$

Linear Transformations Into Direct Sums

Theorem
Every linear transformation $L : V \to W_1 \oplus W_2$ is of the form

$$L(\vec{v}) = L_1(\vec{v}) + L_2(\vec{v})$$

for unique linear transformations $L_1 : V \to W_1$ and $L_2 : V \to W_2$.
Follows from previous two lemmas

Abstract Example

The inclusion maps:

$$\mathrm{incl}_{V_1} \to V_1 \oplus V_2 \qquad \mathrm{incl}_{V_2} \to V_1 \oplus V_2$$

are constructed as:

$$\mathrm{incl}_{V_1} = \mathrm{Id}_{V_1} + \mathrm{O}_{V_1,V_2}$$

and

$$\mathrm{incl}_{V_2} = \mathrm{O}_{V_2,V_1} + \mathrm{Id}_{V_2}$$

So these satisfy

$$\mathrm{incl}_{V_1}(\vec{v}_1) = \vec{v}_1 \qquad \mathrm{incl}_{V_2}(\vec{v}_2) = \vec{v}_2$$

Moreover we have:

$$\mathrm{proj}_{V_1} \circ \mathrm{incl}_{V_1} = \mathrm{Id}_{V_1} \qquad \mathrm{proj}_{V_2} \circ \mathrm{incl}_{V_2} = \mathrm{Id}_{V_2}$$

whereas

$$\mathrm{proj}_{V_2} \circ \mathrm{incl}_{V_1} = 0_{V_1 V_2} \qquad \mathrm{proj}_{V_1} \circ \mathrm{incl}_{V_2} = 0_{V_2 V_1}$$

are the zero maps.

Matrix Example

Consider $W = \mathbb{R}^4$ and

$$W_1 = \{(a, b, 0, 0) \in W \mid a, b \in \mathbb{R}\} \qquad W_2 = \{(0, 0, c, d) \in W \mid c, d \in \mathbb{R}\}$$

So $W = W_1 \oplus W_2$ and $W_1 \simeq \mathbb{R}^2$ and $W_2 \simeq \mathbb{R}^2$.
To give a linear transformation $L : \mathbb{R}^3 \to W$ is to give a 4 by 3 matrix.
Consider the example

$$\begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \\ 1 & 1 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} + \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 7 & 8 & 9 \\ 1 & 1 & 1 \end{pmatrix}$$

Under this decomposition when we evaluate the linear transformation:

$$\left(\begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} + \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 7 & 8 & 9 \\ 1 & 1 & 1 \end{pmatrix} \right) \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} + \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 7 & 8 & 9 \\ 1 & 1 & 1 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix}$$

The linear transformation $L_1 : \mathbb{R}^3 \to W_1$ is the map $\begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{pmatrix}$.

The linear transformation $L_2 : \mathbb{R}^3 \to W_2$ is the map $\begin{pmatrix} 7 & 8 & 9 \\ 1 & 1 & 1 \end{pmatrix}$.

Theorem

To give a linear transformation $L : V_1 \oplus V_2 \to W_1 \oplus W_2$ it is equivalent to give four maps

$$L_{11} : V_1 \to W_1, \quad L_{21} : V_1 \to W_2, \quad L_{12} : V_2 \to W_1, \quad L_{22} : V_2 \to W_2$$

according to the rule

$$\forall \vec{v}_1 \in V_1, \forall \vec{v}_2 \in V_2, L(\vec{v}_1 + \vec{v}_2) = L_{11}(\vec{v}_1) + L_{21}(\vec{v}_1) + L_{12}(\vec{v}_2) + L_{22}(\vec{v}_2)$$

Proof Idea:

This follows from the previous two theorems.

We can first obtain two maps

$$V_1 \to W_1 \oplus W_2 \quad \text{and} \quad V_2 \to W_1 \oplus W_2$$

each of which splits up to gives us two maps, for a total of four maps.

Alternatively we can first get two maps

$$V_1 \oplus V_2 \to W_1 \quad \text{and} \quad V_1 \oplus V_2 \to W_2$$

each of which splits up to gives us two maps, for a total of four maps.

In either case, the rule which will be satisfied is

$$\forall \vec{v}_1 \in V_1, \forall \vec{v}_2 \in V_2, L(\vec{v}_1 + \vec{v}_2) = L_{11}(\vec{v}_1) + L_{21}(\vec{v}_1) + L_{12}(\vec{v}_2) + L_{22}(\vec{v}_2)$$

one can verify the uniqueness conditions from this rule aswell so that both proceedures give the same result.

Recall $V = \mathbb{R}^5$ and

$$V_1 = \{(a, b, 0, 0, 0) \in V \mid a, b \in \mathbb{R}\} \qquad V_2 = \{(0, 0, c, d, e) \in V \mid c, d, e \in \mathbb{R}\}$$

then $V = V_1 \oplus V_2$, where $V_1 \simeq \mathbb{R}^2$ and $V_2 \simeq \mathbb{R}^3$ and $W = \mathbb{R}^4$ and

$$W_1 = \{(a, b, 0, 0) \in W \mid a, b \in \mathbb{R}\} \qquad W_2 = \{(0, 0, c, d) \in W \mid c, d \in \mathbb{R}\}$$

so that $W = W_1 \oplus W_2$ and $W_1 \simeq \mathbb{R}^2$ and $W_2 \simeq \mathbb{R}^2$.

A map from $V_1 \oplus V_2 \to W_1 \oplus W_2$ is just a 4 by 5 matrix, for example

$$\begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 5 & 4 & 3 & 2 & 1 \\ 1 & 1 & 1 & 1 & 1 \\ 2 & 2 & 3 & 3 & 3 \end{pmatrix}$$

The map $L_{11} : V_1 \to W_1$ is $\begin{pmatrix} 1 & 2 \\ 5 & 4 \end{pmatrix}$.

The map $L_{12} : V_2 \to W_1$ is $\begin{pmatrix} 3 & 4 & 5 \\ 3 & 2 & 1 \end{pmatrix}$.

The map $L_{21} : V_1 \to W_2$ is $\begin{pmatrix} 1 & 1 \\ 2 & 2 \end{pmatrix}$.

The map $L_{22} : V_2 \to W_2$ is $\begin{pmatrix} 1 & 1 & 1 \\ 3 & 3 & 3 \end{pmatrix}$.

Recall Matricies For Abstract Transformations

Suppose U has basis $\vec{e}_1, \ldots, \vec{e}_n$ and V has basis $\vec{f}_1, \ldots, \vec{f}_m$.
The matrix for a linear transformation $L : U \to V$ has the form:

$$\begin{array}{cc}
\begin{array}{ccccccc}
L(\vec{e}_1) & L(\vec{e}_2) & L(\vec{e}_3) & \cdots & L(\vec{e}_j) & \cdots & L(\vec{e}_n) \\
\downarrow & \downarrow & \downarrow & & \downarrow & & \downarrow
\end{array} & \\
\begin{pmatrix}
a_{11} & a_{12} & a_{13} & \cdots & a_{1j} & \cdots & a_{1n} \\
a_{21} & a_{22} & a_{23} & \cdots & a_{2j} & \cdots & a_{2n} \\
a_{31} & a_{32} & a_{33} & \cdots & a_{3j} & \cdots & a_{3n} \\
\vdots & \vdots & \vdots & & \vdots & & \vdots \\
a_{i1} & a_{i2} & a_{i3} & \cdots & \color{red}{a_{ij}} & \cdots & a_{in} \\
\vdots & \vdots & \vdots & & \vdots & & \vdots \\
a_{m1} & a_{m2} & a_{m3} & \cdots & a_{mj} & \cdots & a_{mn}
\end{pmatrix} &
\begin{array}{c}
\leftarrow \vec{f}_1 \\ \leftarrow \vec{f}_2 \\ \leftarrow \vec{f}_3 \\ \vdots \\ \leftarrow \vec{f}_i \\ \vdots \\ \leftarrow \vec{f}_m
\end{array}
\end{array}$$

According to the rule

$$L(\vec{e}_j) = \overbrace{a_{1j}\vec{f}_1 + a_{2j}\vec{f}_2 + a_{3j}\vec{f}_3 + \cdots + \color{red}{a_{ij}}\vec{f}_i + \cdots + a_{mj}\vec{f}_m}^{\text{gives entries } j\text{th column}}$$

Because $\vec{f}_i$ are a basis, this expression is **unique**!
Because $\vec{e}_j$ are a basis, this expression **determines** L (A3Q5a)

Suppose U has basis $\vec{e}_1, \ldots, \vec{e}_r, \vec{f}_1, \ldots, \vec{f}_n$ and V has basis $\vec{g}_1, \ldots, \vec{g}_\ell, \vec{h}_1, \ldots, \vec{h}_k$. The matrix for a linear transformation $L : U \to V$ has the form:

$$
\begin{array}{cc}
\begin{matrix}
L(\vec{e}_1) & \cdots & L(\vec{e}_j) & \cdots & L(\vec{e}_r) & L(\vec{f}_1) & \cdots & L(\vec{f}_j) & \cdots & L(\vec{f}_n) \\
\downarrow & & \downarrow & & \downarrow & \downarrow & & \downarrow & & \downarrow
\end{matrix} & \\
\begin{pmatrix}
a_{11} & \cdots & a_{1j} & \cdots & a_{1r} & b_{11} & \cdots & b_{1j} & \cdots & b_{1n} \\
a_{21} & \cdots & a_{2j} & \cdots & a_{2r} & b_{21} & \cdots & b_{2j} & \cdots & b_{2k} \\
\vdots & & \vdots & & \vdots & \vdots & & \vdots & & \vdots \\
a_{i1} & \cdots & a_{ij} & \cdots & a_{ir} & b_{i1} & \cdots & b_{ij} & \cdots & b_{in} \\
\vdots & & \vdots & & \vdots & \vdots & & \vdots & & \vdots \\
a_{\ell 1} & \cdots & a_{\ell j} & \cdots & a_{\ell r} & b_{\ell 1} & \cdots & b_{\ell j} & \cdots & b_{\ell n} \\
c_{11} & \cdots & c_{1j} & \cdots & c_{1r} & d_{11} & \cdots & d_{1j} & \cdots & d_{1k} \\
\vdots & & \vdots & & \vdots & \vdots & & \vdots & & \vdots \\
c_{i1} & \cdots & c_{ij} & \cdots & c_{ir} & d_{i1} & \cdots & d_{ij} & \cdots & d_{1n} \\
\vdots & & \vdots & & \vdots & \vdots & & \vdots & & \vdots \\
c_{k1} & \cdots & c_{kj} & \cdots & c_{kr} & d_{k1} & \cdots & d_{kj} & \cdots & d_{kn}
\end{pmatrix} &
\begin{matrix}
\leftarrow \vec{g}_1 \\
\leftarrow \vec{g}_2 \\
\vdots \\
\leftarrow \vec{g}_i \\
\vdots \\
\leftarrow \vec{g}_\ell \\
\leftarrow \vec{h}_1 \\
\vdots \\
\leftarrow \vec{h}_i \\
\vdots \\
\leftarrow \vec{h}_k
\end{matrix}
\end{array}
$$

According to the rules

$$L(\vec{e}_j) = a_{1j}\vec{g}_1 + \cdots + a_{ij}\vec{g}_i + \cdots + a_{\ell j}\vec{g}_\ell + c_{1j}\vec{h}_1 + \cdots + c_{ij}\vec{h}_i + \cdots + c_{kj}\vec{h}_k$$

and

$$L(\vec{f}_j) = b_{1j}\vec{g}_1 + \cdots + b_{ij}\vec{g}_i + \cdots + b_{\ell j}\vec{g}_\ell + d_{1j}\vec{h}_1 + \cdots + d_{ij}\vec{h}_i + \cdots + d_{kj}\vec{h}_k$$

By splitting the basis up into multiple pieces we can cut the matrix up into pieces. The same things works if we use even more pieces.

Recall that cutting up a basis is equivalent to giving a direct sum decomposition and so if we write

$$V_1 = \mathrm{Span}(\vec{e}_1, \ldots, \vec{e}_r) \qquad V_2 = \mathrm{Span}(\vec{f}_1, \ldots, \vec{f}_n)$$

$$W_1 = \mathrm{Span}(\vec{g}_1, \ldots, \vec{g}_\ell) \qquad W_2 = \mathrm{Span}(\vec{h}_1, \ldots, \vec{h}_k)$$

So that

$$V = V_1 \oplus V_2 \qquad \text{and} \qquad W = W_1 \oplus W_2.$$

Then giving a linear transformation $L : V \to W$ with the formula

$$L(\vec{e}_j) = \left(\sum_{i=1}^{s} a_{ij}\vec{g}_i\right) + \left(\sum_{i=1}^{t} c_{ij}\vec{h}_i\right) \qquad L(\vec{f}_j) = \left(\sum_{i=1}^{s} b_{ij}\vec{g}_i\right) + \left(\sum_{i=1}^{t} d_{ij}\vec{h}_i\right)$$

is equivalent to giving four linear transformations

$$L_{11} : V_1 \to W_1, \;\; L_{21} : V_1 \to W_2, \;\; L_{12} : V_2 \to W_1, \;\; L_{22} : V_2 \to W_2$$

With the formulas:

$$L_{11}(\vec{e}_j) = \left(\sum_{i=1}^{\ell} a_{ij}\vec{g}_i\right) \;\; L_{21}(\vec{e}_j) = \left(\sum_{i=1}^{k} c_{ij}\vec{h}_i\right) \;\; L_{12}(\vec{f}_j) = \left(\sum_{i=1}^{\ell} b_{ij}\vec{g}_i\right) \;\; L_{22}(\vec{f}_j) = \left(\sum_{i=1}^{k} d_{ij}\vec{h}_i\right)$$

Then the matrix for L is of the form:

$$\begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix}$$

Where A_{ij} is the matrix for L_{ij} with respect to the relevant basis.

Zero-Blocks

If we happen to know that $L(V_1) \subset W_1$, then this says that in the formula

$$L(\vec{e}_j) = \left(\sum_{i=1}^{s} a_{ij}\vec{g}_i \right) + \left(\sum_{i=1}^{t} b_{ij}\vec{h}_i \right)$$

all of the $b_{ij} = 0$ (Why?), and consequently $L_{21} = 0$ and so $A_{21} = 0$, and so the matrix is actually of the form:

$$\begin{pmatrix} A_{11} & A_{12} \\ 0 & A_{22} \end{pmatrix}.$$

Which is a nicer matrix to work with than the general case.

If instead we know $L(V_2) \subset W_2$ we obtain instead $A_{12} = 0$.

Consequently, if we know both $L(V_1) \subset W_1$, and $L(V_2) \subset W_2$ then we know

$$A_{12} = 0 \qquad A_{21} = 0$$

and so the matrix for L is of the shape:

$$\begin{pmatrix} A_{11} & 0 \\ 0 & A_{22} \end{pmatrix}.$$

Which is significantly nicer than the more general case.
There are also conditions that would make the other blocks zero.

Notes about Block Decompositions

One key feature of block decompositions is the following formulas hold so long as the matricies A_{ij} and B_{ij} have matching sizes.

$$\begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix} \begin{pmatrix} B_{11} & B_{12} \\ B_{21} & B_{22} \end{pmatrix} = \begin{pmatrix} A_{11}B_{11} + A_{12}B_{21} & A_{11}B_{12} + A_{12}B_{22} \\ A_{21}B_{11} + A_{22}B_{21} & A_{21}B_{12} + A_{22}B_{22} \end{pmatrix}$$

$$\begin{pmatrix} A_{11} & A_{12} \\ 0 & A_{22} \end{pmatrix} \begin{pmatrix} B_{11} & B_{12} \\ 0 & B_{22} \end{pmatrix} = \begin{pmatrix} A_{11}B_{11} & A_{11}B_{12} + A_{12}B_{22} \\ 0 & A_{22}B_{22} \end{pmatrix}$$

$$\begin{pmatrix} A_{11} & 0 \\ 0 & A_{22} \end{pmatrix} \begin{pmatrix} B_{11} & 0 \\ 0 & B_{22} \end{pmatrix} = \begin{pmatrix} A_{11}B_{11} & 0 \\ 0 & A_{22}B_{22} \end{pmatrix}$$

$$\begin{pmatrix} A_{11} & 0 \\ 0 & A_{22} \end{pmatrix}^{-1} = \begin{pmatrix} A_{11}^{-1} & 0 \\ 0 & A_{22}^{-1} \end{pmatrix}$$

and if both A_{11} and A_{22} are square matricies then:

$$\mathrm{Det}\left(\begin{pmatrix} A_{11} & A_{12} \\ 0 & A_{22} \end{pmatrix}\right) = \mathrm{Det}\,(A_{11})\,\mathrm{Det}\,(A_{22})$$

Checking these formulas is an excercise.

One can also prove abstract versions of these relating to composing $L : V_1 \oplus V_2 \to W_1 \oplus W_2$ and $M : W_1 \oplus W_2 \to U_1 \oplus U_2$.

Putting this into practice - Canonical Forms

Let V and W be any vector spaces, and $L : V \to W$ be any linear transformation.
Our Goal find a basis for V and one for W so the matrix for L is simple!
Pick any basis $\vec{f}_1, \ldots, \vec{f}_n$ for $\mathrm{Ker}(L)$, so in particular $\mathrm{Ker}(L) = \mathrm{Span}(\vec{f}_1, \ldots, \vec{f}_n)$.
Because $\vec{f}_1, \ldots, \vec{f}_n$ are linearly independent, there is a basis B for V of the form:

$$\vec{e}_1, \ldots, \vec{e}_r, \vec{f}_1, \ldots, \vec{f}_n$$

Now recall that we proved that we can write:

$$V = \mathrm{Span}(\vec{e}_1, \ldots, \vec{e}_r) \oplus \mathrm{Span}(\vec{f}_1, \ldots, \vec{f}_n) = \mathrm{Span}(\vec{e}_1, \ldots, \vec{e}_r) \oplus \mathrm{Ker}(L)$$

Lemma
The collection of vectors $\vec{g}_1 = L(\vec{e}_1), \ldots, \vec{g}_r = L(\vec{e}_r)$ are linearly independent.
Proof:
First, we know that because we have a direct sum decomposition we have

$$\mathrm{Span}(\vec{e}_1, \ldots, \vec{e}_r) \cap \mathrm{Span}(\vec{f}_1, \ldots, \vec{f}_n) = \{\vec{0}\}$$

but this means

$$\mathrm{Span}(\vec{e}_1, \ldots, \vec{e}_r) \cap \mathrm{Ker}(L) = \{\vec{0}\}$$

so by the assigment, we know

$$L(\vec{e}_1), \ldots, L(\vec{e}_r)$$

are linearly independent.

Lemma
The collection of vectors

$$\vec{g}_1 = L(\vec{e}_1), \ldots, \vec{g}_r = L(\vec{e}_r)$$

are a generating set for $\mathrm{Im}(L)$, hence a basis for $\mathrm{Im}(L)$.

Proof: From the assignment we know because $\vec{e}_1, \ldots, \vec{e}_r, \vec{f}_1, \ldots, \vec{f}_n$ is a basis for V that

$$L(\vec{e}_1), \ldots, L(\vec{e}_r), L(\vec{f}_1), \ldots, L(\vec{f}_n)$$

is a generating set for $\mathrm{Im}(L)$, but because $\vec{f}_1, \ldots, \vec{f}_n \in \mathrm{Ker}(L)$ we see that

$$\begin{aligned}\mathrm{Im}(L) &= \mathrm{Span}(L(\vec{e}_1), \ldots, L(\vec{e}_r), L(\vec{f}_1), \ldots, L(\vec{f}_n)) \\ &= \mathrm{Span}(L(\vec{e}_1), \ldots, L(\vec{e}_r), \vec{0} \ldots, \vec{0}) \\ &= \mathrm{Span}(L(\vec{e}_1), \ldots, L(\vec{e}_r))\end{aligned}$$

which shows that $\vec{g}_1, \ldots, \vec{g}_r$ are a generating set.

Corollary[Rank-Nullity Theorem]
For any linear transformation $L : V \to W$, the rank r plus the nullity n, is equal to the dimension of the domain, $\dim(V)$

$$\mathrm{Dim}(V) = \mathrm{Rank}(L) + \mathrm{Null}(L)$$

Proof In the above we can see that, r is the rank, n is the nullity, and $r + n$ is the number of elements in a basis for V.

This result is useful, we have seen it before for matricies.

Now, using that $\vec{g}_1, \ldots, \vec{g}_r$ are linearly independent in W, we can as before, extend to a basis

$$\vec{g}_1, \ldots, \vec{g}_r, \vec{h}_1, \ldots, \vec{h}_\ell$$

for W, so that again we have

$$W = \mathrm{Span}(\vec{g}_1, \ldots, \vec{g}_r) \oplus \mathrm{Span}(\vec{h}_1, \ldots, \vec{h}_\ell) = \mathrm{Im}(L) \oplus \mathrm{Span}(\vec{h}_1, \ldots, \vec{h}_\ell)$$

Lemma

The matrix A for L with respect to the basis

$$\vec{e}_1, \ldots, \vec{e}_r, \vec{f}_1, \ldots, \vec{f}_n$$

for V and

$$\vec{g}_1, \ldots, \vec{g}_r, \vec{h}_1, \ldots, \vec{h}_\ell$$

for W is of the form:

$$\begin{pmatrix} \mathrm{Id}_r & 0 \\ 0 & 0 \end{pmatrix}$$

Proof Sketch:

The first r columns come from the equations

$$L(\vec{e}_i) = \vec{g}_i = 0\vec{g}_1 + 0\vec{g}_2 + \cdots + 1\vec{g}_i + \cdots + 0\vec{g}_r + 0\vec{h}_1 + \cdots + 0\vec{h}_\ell$$

the last n columns come from the equations.

$$L(\vec{f}_j) = \vec{0} = 0\vec{g}_1 + \cdots + 0\vec{g}_r + 0\vec{h}_1 + \cdots + 0\vec{h}_\ell$$

putting all this together gives the result.

Theorem[Smith Normal Form]
If $L : V \to W$ is any linear transformation between finite dimensional vector spaces, there exists a choice of basis for V and a choice of basis for W such that with respect to this basis the matrix is of the form:

$$\begin{pmatrix} \mathrm{Id}_r & 0 \\ 0 & 0 \end{pmatrix}$$

where r is the rank of L.
Proof Immediate from previous few lemmas.

Recall: On assignment one, you took a matrix A, and found P and Q so that

$$PAQ$$

was upper and lower triangular.
P was a change of basis on the codomain.
Q was a change of basis on the domain.
The result was almost (and with decent probability was) the Smith Normal Form.

The above also suggests that for Linear transformations $L : V \to W$, the really only defining features are the dimension of V, the dimension of W and the rank.

What we just saw is the best that we can do for the situation $L : V \to W$ in terms of finding a simple matrix representation for a single linear transformation.
There are other questions one could ask in this situation:

- If I have two linear transformations, $L_1 : V \to W$ and $L_2 : V \to W$ can I make both nice?
- If I have $L_1 : V \to W$ and $L_2 : U \to V$ can I make both nice?
- Other more elaborate setups.

Some of the above are not so hard, others lead to very hard problems. We will not focus on this though.

What we will focus on next is the situation $L : V \to V$, because we must use the same basis for the domain and the codomain, the above strategy largely fails...
Except the part about using direct sum decompositions!!!

Canonical Forms $L : V \to V$

The above strategy fails for maps $L : V \to V$, because in the above you had to pick a basis for each of V and W, whereas for a map from $V \to V$, you only get to pick one basis.

Given two square matrices A and B we say that A and B are **similar** if there exists an invertible matrix P so that:

$$P^{-1}AP = B$$

We have seen that being similar means that the matricies can be seen to describe the same linear transformation, just with respect to different choices of basis.

The goal of finding a canonical (that is nice) form for the matrix of $L : V \to V$, is the same as the goal of finding: *the simplest matrix B which is similar to A*

This is useful because

- The simpler form will be easier to study.
- In applications the interesting information the linear transformation encodes is typically easier to see in the simple form and the choice of basis that gives rise to it.

Preview of what we will see

Fix a linear transformation $L : V \to V$.

We can always find a basis so that the matrix A associated to L has the shape

$$A = \begin{pmatrix} A_1 & & & \\ & A_2 & & \\ & & \ddots & \\ & & & A_r \end{pmatrix}$$

where each of the A_i are square matricies (maybe not the same size, maybe all 1 by 1, or maybe only one of them), and all the other entries are 0.

Each of the A_i have some simple structure, our two main simple shapes are:

$$A_i = \begin{pmatrix} \lambda & 1 & 0 & & \cdots & & 0 \\ 0 & \lambda & 1 & 0 & & & \\ 0 & 0 & \lambda & 1 & & & \\ \vdots & & & \ddots & \ddots & & \vdots \\ & & & & \lambda & 1 & 0 \\ & & & & 0 & \lambda & 1 \\ 0 & & \cdots & & & 0 & \lambda \end{pmatrix} \quad \text{or} \quad A_i = \begin{pmatrix} 0 & & \cdots & & 0 & -a_0 \\ 1 & 0 & & & 0 & -a_1 \\ 0 & 1 & 0 & & 0 & -a_2 \\ \vdots & & \ddots & & \vdots & \vdots \\ 0 & & & 1 & 0 & -a_{\ell-2} \\ 0 & & \cdots & & 1 & -a_{\ell-1} \end{pmatrix}$$

The form on the left is part of **Jordan Canonical Form**, the form on the right is part of **Rational Canonical Form**, our focus will mostly be on the form on the left.