Linear transformations:

Recall: \( T: V \to W \) is linear if for all \( x, y \in V \) and \( \alpha \in \mathbb{F} \), we have \( T(x + y) = T(x) + T(y) \) and \( T(\alpha x) = \alpha T(x) \).

The nullspace of \( T \) is \( \text{N}(T) = \{ x \in V \mid T(x) = 0 \} \), a subspace of \( V \).

The range of \( T \) is \( \text{R}(T) = \{ T(x) \mid x \in V \} \), a subspace of \( W \).

- \( T \) is injective \( \iff \) \( \text{N}(T) = \{ 0 \} \iff \text{nullity}(T) = 0 \)
- \( T \) is surjective \( \iff \) \( \text{R}(T) = W \iff \text{rank}(T) = \text{dim}(W) \)

Assume that \( \mathcal{B} = \{ v_1, \ldots, v_n \} \) is a basis of \( V \).

Then:

1. \( T \) is injective if and only if \( \{ T(v_1), \ldots, T(v_n) \} \) is linearly independent.
2. \( T \) is surjective if and only if \( \{ T(v_1), \ldots, T(v_n) \} \) spans \( W \).

Proof:

1. Assume \( T \) is injective, and let \( \lambda_1, \ldots, \lambda_n \in \mathbb{F} \) be such that \( \lambda_1 T(v_1) + \cdots + \lambda_n T(v_n) = 0 \).

Since \( T \) is linear, \( T(\lambda_1 v_1 + \cdots + \lambda_n v_n) = 0 \). So \( \lambda_1 v_1 + \cdots + \lambda_n v_n \in \text{N}(T) \).

But \( \text{N}(T) = \{ 0 \} \) since \( T \) is injective. Hence \( \lambda_1 v_1 + \cdots + \lambda_n v_n = 0 \).

Since \( \{ v_1, \ldots, v_n \} \) is linearly independent, we deduce that \( \lambda_1 = 0, \ldots, \lambda_n = 0 \).

In conclusion, \( \{ T(v_1), \ldots, T(v_n) \} \) is linearly independent.

Conversely, assume \( \{ T(v_1), \ldots, T(v_n) \} \) is linearly independent. To prove that \( T \) is injective, we show that \( \text{N}(T) = \{ 0 \} \). It is clear that \( \{ 0 \} \subset \text{N}(T) \) since \( \text{N}(T) \) is a subspace.

Let \( x \in \text{N}(T) \). Since \( \{ v_1, \ldots, v_n \} \) generates \( W \), there exists \( a_1, \ldots, a_n \in \mathbb{F} \) such that \( x = a_1 v_1 + \cdots + a_n v_n \). Then:

\[
0 = T(x) = T(a_1 v_1 + \cdots + a_n v_n) = a_1 T(v_1) + \cdots + a_n T(v_n).
\]

Since \( \{ T(v_1), \ldots, T(v_n) \} \) are independent, we have \( a_1 = 0, \ldots, a_n = 0 \).

So \( x = 0 v_1 + \cdots + 0 v_n = 0 \). In conclusion, \( \text{N}(T) = \{ 0 \} \) and \( T \) is injective. Q.E.D.
2. Assume \( T \) is surjective. Let \( y \in W \). Since \( T \) is surjective, there exists \( x \in V \) such that \( y = T(x) \). Since \( \{v_1, \ldots, v_n\} \) spans \( V \), we have \( x = \alpha_1 v_1 + \cdots + \alpha_n v_n \) for some \( \alpha_1, \ldots, \alpha_n \in F \).

Hence \( y = T(x) = T(\alpha_1 v_1 + \cdots + \alpha_n v_n) = \alpha_1 T(v_1) + \cdots + \alpha_n T(v_n) \) since \( T \) is linear.

So \( \{T(v_1), \ldots, T(v_n)\} \) spans \( W \).

Conversely, assume \( \{T(v_1), \ldots, T(v_n)\} \) spans \( W \). Let \( y \in W \). Since \( \{T(v_1), \ldots, T(v_n)\} \) spans \( W \), there exist \( \beta_1, \ldots, \beta_n \in F \) such that \( y = \beta_1 T(v_1) + \cdots + \beta_n T(v_n) \).

Since \( T \) is linear \( y = T(\beta_1 v_1 + \cdots + \beta_n v_n) \).

Hence \( y \in \text{R}(T) \), and \( T \) is surjective. \( \square \)

Applications of the rank-nullity theorem.
Recall that if \( V \) is finite dimensional and \( T : V \to W \) is linear then \( \dim(V) = \text{rank}(T) + \text{nullity}(T) \).

\[
\text{dim}(\text{R}(T)) = \text{dim}(\text{N}(T))
\]

Consequence: if \( \dim(V) = \dim(W) \) then \( T \) is injective if and only if \( T \) is surjective.

Example of application: let \( V \) be finite dimensional, and \( T, U : V \to V \) linear such that \( TU \) is injective. Show that \( T \) is injective.

Proof: since \( TU : V \to V \) is injective, it is also surjective.

So \( \text{R}(TU) = V \). But \( \text{R}(T) \supseteq \text{R}(TU) = V \), so \( \text{R}(T) = V \). Hence \( T \) is surjective. So it is also injective. \( \square \)
Useful properties of composition:
1. $R(\text{TU}) \subseteq R(\tau)$
2. $N(\tau) \subseteq N(\text{TU})$

Important example of linear transformation: projections

**Def:** a projection of $V$ is a linear transformation $P: V \to V$ such that $P^2 = P$

**Example:** orthogonal projection on a plane in $\mathbb{R}^3$

```
\begin{figure}[h]
\centering
\begin{tikzpicture}
\draw[->, thick] (0,0) -- (4,0) node[anchor=north] {$R(P)$};
\draw[->, thick] (0,0) -- (0,4) node[anchor=east] {$n(P)$};
\draw[->, thick] (0,0) -- (3,3) node[anchor=south] {$P(x)$};
\end{tikzpicture}
\end{figure}
```

**Theorem:** Let $P$ be a projection of $V$. Then:
1. $V = N(P) \oplus R(P)$
2. for all $x \in V$, $x = y + z$ with $y \in N(P)$, $z \in R(P)$ we have $P(x) = z$.

**Pf:**
1. Let us prove that $N(P) \cap R(P) = \{0\}$.
   - Let $x \in N(P) \cap R(P)$. Then $x = P(y)$ for some $y \in V$.
   - Also $P(x) = 0$ so $P^2(y) = 0$. But $P = P^2$ so $P(y) = P^2(y) = 0$.
   - Hence $x = P(y) = 0$. So $N(P) \cap R(P) = \{0\}$.
   - Now $\dim (N(P) + R(P)) = \dim(N(P)) + \dim(R(P)) + \dim(N(P) \cap R(P)) = 0$.
   - So $V = N(P) + R(P)$.

In conclusion, $V = N(P) \oplus R(P)$

2. If $x = y + z$ with $y \in N(P)$, $z \in R(P)$ then $P(x) = P(z)$.
But \( y = P(x) \) for some \( x \in V \Rightarrow P(y) = P^2(x) = P(x) = y \).

Hence \( P(x) = y \) \( \Box \).

Remark: other proof of \( V = N(P) + R(P) \):

Let \( x \in V \), then \( x = x - P(x) + P(x) \).

\( P(x) \in R(P) \)

\( P(x - P(x)) = P(x) - P^2(x) = 0 \), so \( x - P(x) \in N(P) \).

\( \Rightarrow x \in N(P) + R(P) \).
Matrix of linear transformation:

Recall, assume \( T : V \rightarrow W \) is linear, \( \beta = (v_1, \ldots, v_m) \) is an ordered basis of \( V \) and \( \gamma = (w_1, \ldots, w_n) \) is an ordered basis of \( W \).

For each \( j \in \{1, \ldots, m\} \), there exist unique scalars \( a_{ij} \), \( i \in \{1, \ldots, n\} \) such that \( T(v_j) = a_{i1}w_1 + \cdots + a_{in}w_n \).

Then the matrix of \( T \) in the bases \( \beta, \gamma \) is \( [T]_{\beta}^{\gamma} = [a_{ij}]_{i=1}^{n}j=1}^{m} \).

If \( V = W \) and \( \beta = \gamma \), we denote it by \( [T]_{\beta} \).

Example: \( T : P_2(\mathbb{R}) \rightarrow P_3(\mathbb{R}) \)

\[ f(x) \mapsto \int_{0}^{x} f(t) \, dt + f(2) \cdot x \]

Check: \( T \) is linear.

With \( \beta = (1, x, x^2) \) and \( \gamma = (1, x, x^2, x^3) \), let us find \( [T]_{\beta}^{\gamma} \).

- \( T(v_1) = \int_{0}^{x} 1 \, dt + 1 \cdot x = 2x = 0w_1 + 2w_2 + 0w_3 + 0w_4 \)
  \( \Rightarrow \) first column of \( [T]_{\beta}^{\gamma} \) is \( \begin{bmatrix} 0 \\ 2 \\ 0 \\ 0 \end{bmatrix} \)

- \( T(v_2) = \int_{0}^{x} t \, dt + 2x = \frac{x^2}{2} + 2x = 0w_1 + 2w_2 + \frac{1}{2}w_3 + 0w_4 \)
  \( \Rightarrow \) second column of \( [T]_{\beta}^{\gamma} \) is \( \begin{bmatrix} 0 \\ 2 \\ \frac{1}{2} \\ 0 \end{bmatrix} \)

- \( T(v_3) = \int_{0}^{x} t^2 \, dt + 4x = \frac{x^3}{3} + 4x = 0w_1 + 4w_2 + 0w_3 + \frac{1}{3}w_4 \)
  \( \Rightarrow \) third column of \( [T]_{\beta}^{\gamma} \) is \( \begin{bmatrix} 0 \\ 0 \\ 0 \\ \frac{1}{3} \end{bmatrix} \)

Conclusion: \( [T]_{\beta}^{\gamma} = \begin{bmatrix} 0 & 2 & 0 & 0 \\ 0 & \frac{1}{2} & 4 & 0 \\ 0 & 0 & 0 & \frac{1}{3} \end{bmatrix} \)
2. Assume that \( \dim(V) = 2 \), \( \beta = (e_1, e_2) \) is an ordered basis of \( V \) and \( T: V \rightarrow V \) has matrix \( [T]_\beta = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} \).

We can express \( T(e_1) \) and \( T(e_2) \) as linear combinations of \( e_1 \) and \( e_2 \) by looking at the columns of \( [T]_\beta \).

1st column \( \begin{bmatrix} 1 \\ 0 \end{bmatrix} = [T(e_1)]_\beta \) so \( T(e_1) = 1 e_1 + 0 e_2 = e_1 \).

2nd column \( \begin{bmatrix} 0 \\ 1 \end{bmatrix} = [T(e_2)]_\beta \) so \( T(e_2) = 0 e_1 + 1 e_2 = e_2 \).

Question: given \( T: V \rightarrow W \), can we find ordered bases \( \beta \) of \( V \) and \( \gamma \) of \( W \) such that \( [T]_\beta^{\gamma} \) is "simple"?

Theorem: assume that \( V, W \) are finite dimensional and \( T: V \rightarrow W \) is a linear transformation. Then there exists bases \( \beta, \gamma \) such that:

\[
[T]_\beta^{\gamma} = \begin{bmatrix} 1 & \cdots & 0 \\ & \ddots & \vdots \\ 0 & \cdots & 1 \end{bmatrix}
\]

the number of non-zero entries is \( \text{rank}(T) \).

Proof: pick a basis \( v_1, \ldots, v_n \) of \( \text{N}(T) \). Complete into a basis \( v_1, \ldots, v_n, u_1, \ldots, u_r \) of \( V \).

Claim: \( T(u_1), \ldots, T(u_r) \) is a basis of \( \text{R}(T) \).

Proof of claim: let \( y \in \text{R}(T) \). Then \( y = T(x) \) for some \( x \in V \).

There are scalars \( \alpha_1, \ldots, \alpha_n, \beta_1, \ldots, \beta_r \) such that \( x = \alpha_1 v_1 + \cdots + \alpha_n v_n + \beta_1 u_1 + \cdots + \beta_r u_r \).

Thus \( y = \alpha_1 T(v_1) + \cdots + \alpha_n T(v_n) + \beta_1 T(u_1) + \cdots + \beta_r T(u_r) \).

Hence \( T(u_1), \ldots, T(u_r) \) span \( \text{R}(T) \).
Assume that $\lambda_1 T(u_1) + \cdots + \lambda_r T(u_r) = 0$ for scalars $\lambda_1, \ldots, \lambda_r$.

Then $T(\lambda_1 u_1 + \cdots + \lambda_r u_r) = 0$, so $\lambda_1 u_1 + \cdots + \lambda_r u_r \in \ker(T)$.

Since $v_1, \ldots, v_n$ span $\ker(T)$, there are scalars $\mu_1, \ldots, \mu_n$ of $\lambda_1 u_1 + \cdots + \lambda_r u_r = \mu_1 v_1 + \cdots + \mu_n v_n$.

So $\lambda_1 u_1 + \cdots + \lambda_r u_r - \mu_1 v_1 - \cdots - \mu_n v_n = 0$.

But $u_1, \ldots, u_r, v_1, \ldots, v_n$ are linearly independent so $\lambda_1 = 0, \ldots, \lambda_r = 0$.

Hence $T(u_1), \ldots, T(u_r)$ are linearly independent.

This ends the proof of the claim.

Back to the proof of the theorem:

$u_1, \ldots, u_r, v_1, \ldots, v_n$ basis of $V$  

$\lambda_1 u_1 + \cdots + \lambda_r u_r$ basis of $\ker(T)$

$T(u_1), \ldots, T(u_r)$ basis of $\operatorname{R}(T) = \ker(T)$ complete into a basis $T(u_1), \ldots, T(u_r), w_1, \ldots, w_k$ of $W$.

Then let $\beta = (u_1, \ldots, u_r, v_1, \ldots, v_n)$  

$\gamma = (T(u_1), \ldots, T(u_r), w_1, \ldots, w_k)$.

We have $T(v_i) = 0$ so $[T(v_i)]_\gamma = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix}$.

$T(w_i) = \lambda_1 T(u_1) + \cdots + \lambda_r T(u_r) + 0 w_1 + \cdots + 0 w_k$.

So $[T(w_i)]_\gamma = \begin{bmatrix} \lambda_1 \\ \lambda_2 \\ \vdots \\ \lambda_r \end{bmatrix}$ in $\mathbb{R}$ now.

Hence $[T]_\beta^\gamma$ has the desired form $\square$.

Much harder (and interesting) question: given $T : V \to V$, can we find a basis $\beta$ such that $[T]_\beta$ is "as simple as possible"? ("simple" = most entries are 0).

$\implies$ diagonalization, Jordan canonical form

\(\implies\) later

\(\implies\) Math 115B