**Definition 1.** A *vector space* is a set $V$ with the operation of scalar multiplication and addition defined (so $\lambda v_1 + v_2$ makes sense when $v_1, v_2 \in V$, $\lambda \in \mathbb{R}$) and so they obey the usual rules ($\lambda(v_1 + v_2) = \lambda v_1 + \lambda v_2$). An *independent set* $\{v_1, \ldots, v_n\}$ is a set so that $\alpha_1 v_1 + \ldots + \alpha_n v_n = 0 \Rightarrow \alpha_1, \alpha_2, \ldots, \alpha_n = 0$. A *spanning set* is a set $\{v_1, v_2, \ldots, v_n\}$ so that $V = \{\alpha_1 v_1 + \ldots + \alpha_n v_n : \alpha_1, \alpha_2, \ldots, \alpha_n \in \mathbb{R}\}$. A *basis* is a set which is both spanning and independent. To be formal at this point, one has to check that bases exist and so on. This is done properly with the *replacement theorem*, but we will skip that. Instead we have two useful corollaries of the replacement theorem. In practice these are all you need.

**Proposition 2.** [Useful Corollaries of the Replacement Theorem] All bases of a vector space $V$ are the same size, which we call the dimension of $V$ written $\dim(V)$. Every independent set has at most $\dim(V)$ elements, and can be augmented to a basis. Every spanning set has at least $\dim(V)$ elements and can be refined to a basis.

**Definition 3.** A *matrix* is a “box of numbers” that obeys certain rules about multiplication and so on.

**Definition 4.** A *linear transformation* is a linear map $A : V \to W$ where $V, W$ are vector spaces. Linear means $A(\lambda v_1 + v_2) = \lambda A(v_1) + A(v_2)$. Linear is important because it means we can know exactly what $A$ does everywhere, just by knowing what $A$ does to a basis (Why?). A choice of a basis $\{v_1, v_2, \ldots, v_n\}$ for $V$ and $\{w_1, w_2, \ldots, w_n\}$ for $W$ gives rise to a *matrix* by $A_{ij} = \text{the component of } w_i \text{ in } Av_j$. One could also say its defined by: $Av_j = A_{1j}v_1 + A_{2j}v_2 + \ldots + A_{nj}v_n$. Multiplication of matrices is composition of the linear transformations.

**Theorem 5.** [Rank-Nullity Theorem] Given a linear transformation $A : V \to W$, let $\text{Null}(A) = \{v : A(v) = 0\}$ be the subspace of $V$ that $A$ sends to zero. Let $\text{Range}(A) = A(V) = \{w : \exists v \text{ s.t. } A(v) = w\}$ be the subspace of $W$ that is is the image of $A$. Then:

$$
\dim(\text{Null}(A)) + \dim(\text{Range}(A)) = \dim(V)
$$

*Proof.* (Pf using Bases) Say $\dim(V) = n$. Take $\{v_1, \ldots, v_l\}$ to be a basis for $\text{Null}(A)$, so that $l = \dim(\text{Null}(A))$. Extend this to a basis for all of $V$ (replacement theorem), so that $\{v_1, \ldots, v_l, \tilde{v}_1, \ldots, \tilde{v}_{n-l}\}$ is a basis for $V$. Check that $\{A(\tilde{v}_1), \ldots, A(\tilde{v}_{n-l})\}$ is a basis for $\text{Range}(A)$. (Why is independent? Why is it spanning?). Hence $\dim(\text{Range}(A)) = n - l$ and the result follows.

**Remark 6.** [How to deal with matrices] Say $v_1, v_2, \ldots, v_n$ are a basis for $V$. A linear transformation is defined by where it sends $v_1, \ldots, v_n$. (Why?) Say we know $Av_1, \ldots, Av_n$. Then the matrix for $A$ is matrix whose columns are $Av_1, \ldots, Av_n$ (of course these have to be written in the basis $w_1, \ldots, w_l$:

$$
A = \begin{bmatrix}
\vdots & \vdots & \vdots \\
Av_1 & Av_2 & \ldots & Av_n \\
\vdots & \vdots & \vdots 
\end{bmatrix}
$$

In particular, its useful to remeber that $Ae_1$ = first column of $A$.

From this you can see how matrix multiplication on the *left* works on columns. If $B$ is a matrix whose columns are $b_1, b_2, \ldots, b_n$, i.e. $B = \begin{bmatrix} b_1 & b_2 & \ldots & b_n \end{bmatrix}$ then:

$$
A \begin{bmatrix} b_1 & b_2 & \ldots & b_n \end{bmatrix} = \begin{bmatrix} Ab_1 & Ab_2 & \ldots & Ab_n \end{bmatrix}
$$

(How do we see this from the definition/ideas above?) We can also see how matrix multiplication on the *right* works on rows:

$$
\begin{bmatrix} \ldots & a_1 & \ldots \\
\ldots & a_2 & \ldots \\
\vdots \\
\ldots & a_m & \ldots 
\end{bmatrix} B = \begin{bmatrix} \ldots & a_1B & \ldots \\
\ldots & a_2B & \ldots \\
\vdots \\
\ldots & a_mB & \ldots 
\end{bmatrix}
$$
This gives rise to two little formulas that are sometimes useful:

\[
\begin{bmatrix}
  \vdots & \vdots & \vdots \\
  b_1 & b_2 & \cdots & b_n \\
  \vdots & \vdots & \ddots & \vdots \\
\end{bmatrix}
\begin{bmatrix}
  \lambda_1 \\
  \lambda_2 \\
  \vdots \\
\end{bmatrix}
= 
\begin{bmatrix}
  \vdots & \vdots & \vdots \\
  \lambda_1 b_1 & \lambda_2 b_2 & \cdots & \lambda_n b_n \\
  \vdots & \vdots & \ddots & \vdots \\
\end{bmatrix}
\begin{bmatrix}
  \lambda_1 \\
  \lambda_2 \\
  \vdots \\
\end{bmatrix}
= 
\begin{bmatrix}
  \lambda_1 a_1 \\
  \lambda_2 a_2 \\
  \vdots \\
\end{bmatrix}
\]

Theorem 7. [Change of Basis Formula] Let \( T \) be a linear transformation from \( \mathbb{R}^n \) to \( \mathbb{R}^n \). By using the ordinary basis for \( \mathbb{R}^n \), we can think of \( T \) as a matrix too. Let \( B = \{ v_1, \ldots, v_n \} \) be a basis of \( \mathbb{R}^n \). Let \( T_B \) be the matrix that represents the same transformation as \( T \), but in the basis \( B \). Then: (Why is the inverse ok in the below formula?)

\[
T_B = \left[ \begin{array}{ccc}
  \vdots & \vdots & \vdots \\
  v_1 & v_2 & \cdots & v_n \\
  \vdots & \vdots & \ddots & \vdots \\
\end{array} \right]^{-1} \left[ \begin{array}{ccc}
  \vdots & \vdots & \vdots \\
  v_1 & v_2 & \cdots & v_n \\
  \vdots & \vdots & \ddots & \vdots \\
\end{array} \right]
\]

Example 8. The change of basis formula is useful, because if there is a basis of \( V \) where the action of \( A \) is known, then we can recover the matrix for \( A \). As an example, suppose we know that \( A \) does this:

\[
\left[ \begin{array}{ccc}
  1 & 0 & 0 \\
  0 & 1 & 0 \\
  0 & 0 & 1 \\
\end{array} \right] \rightarrow 
\left[ \begin{array}{ccc}
  1 & 1 & 0 \\
  0 & 1 & 1 \\
  0 & 0 & 1 \\
\end{array} \right]
\]

Then in the basis \( \left[ \begin{array}{cc}
  1 & 1 \\
  1 & -1 \\
\end{array} \right] \) we know that \( A \) looks like \([Av_1 Av_2] = \left[ \begin{array}{cc}
  0 & 1 \\
  1 & 0 \\
\end{array} \right] \). So then, in the standard basis of \( \mathbb{R}^n \) we have:

\[
A = \left[ \begin{array}{ccc}
  1 & 1 & 0 \\
  0 & 1 & 1 \\
  0 & 0 & 1 \\
\end{array} \right]^{-1} = \left[ \begin{array}{cc}
  1 & 0 \\
  0 & -1 \\
\end{array} \right]
\]

(How could we have seen that differently?) You can use these ideas to easily construct the matrices for rotations, reflections etc.

Definition 9. An eigenvector of eigenvalue \( \lambda \) is a vector \( v \) so that \( Av = \lambda v \). All the possible eigenvalues are given by the roots of the \( n \)-th degree polynomial \( Det(A - \lambda I) \) (This is called the characteristic polynomial). The roots of this give values of \( \lambda \) where \( Null(A - \lambda I) \) is non-trivial. The spaces \( Null(A - \lambda k I) \) are called eigenspaces. If we have a basis of eigenvectors, then we can write:

\[
A = \left[ \begin{array}{ccc}
  \vdots & \vdots & \vdots \\
  v_1 & v_2 & \cdots & v_n \\
  \vdots & \vdots & \ddots & \vdots \\
\end{array} \right] 
\left[ \begin{array}{ccc}
  \lambda_1 \\
  \lambda_2 \\
  \vdots \\
\end{array} \right]^{-1} = 
\left[ \begin{array}{ccc}
  \vdots & \vdots & \vdots \\
  v_1 & v_2 & \cdots & v_n \\
  \vdots & \vdots & \ddots & \vdots \\
\end{array} \right] 
\left[ \begin{array}{ccc}
  \lambda_1 \\
  \lambda_2 \\
  \vdots \\
\end{array} \right]
\]

Remark 10. Under what circumstances can we find a basis of eigenvectors? One condition is that all the eigenvalues are distinct (no repeated roots of the characteristic polynomial) The lemma below shows us why:

Lemma 11. For \( \lambda_1 \neq \lambda_2 \) we have \( Null(A - \lambda_1 I) \cap Null(A - \lambda_2 I) = \{ 0 \} \). As a corollary, we see that if \( v_1 \in Null(A - \lambda_1 I) \), \( v_2 \in Null(A - \lambda_2 I) \), \ldots \( v_k \in Null(A - \lambda_n I) \) then \( \{ v_1 \ldots v_k \} \) are independent.

Proof. If \( v \in Null(A - \lambda_1 I) \cap Null(A - \lambda_2 I) \) then \( \lambda_1 v = Av = \lambda_2 v \) so \( v = 0 \). The corollary follows by induction on \( k \). It is clear when \( k = 1 \). Suppose it holds for \( k \). If \( \alpha_{k+1}v_{k+1} + \sum_{i=1}^k \alpha_i v_i = 0 \), then apply \( A - \lambda_{k+1}I \) to both sides to get \( \alpha_{k+1} + \sum_{i=1}^k \alpha_i (\lambda_i - \lambda_{k+1}) v_i = 0 \) By the induction hypothesis now, \( \alpha_i (\lambda_i - \lambda_{k+1}) = 0 \) for \( 1 \leq i \leq k \) and the result follows.

Proposition 12. If \( A \) has \( n \) distinct eigenvalues, then \( A \) is diagonalizable.

Proof. Let \( \lambda_1, \ldots, \lambda_n \) be the eigenvalues. By the definition of these, we can find non-trivial \( v_1 \in Null(A - \lambda_1 I) \), \( v_2 \in Null(A - \lambda_2 I) \), \ldots \( v_k \in Null(A - \lambda_n I) \). By the lemma, this is an independent set. Since \( \dim = n \) here, this is a basis! Hence we have a basis of eigenvalues and we are done.

Remark 13. If there are repeated roots of the minimal polynomial, then \( A \) might not be diagonalizable. In this case define the generalized eigenspaces as \( G_\lambda = \{ v : \exists k \text{ s.t. } (A - \lambda I)^k v = 0 \} = Null(A - \lambda I) + Null ((A - \lambda I)^2) + \ldots \).
In this case, a lemma almost identical to the above shows that the generalized eigenspaces don't overlap. We then take a basis for each generalized eigenspace individually, so that in this basis we have $A$ decomposed into blocks.

$$A = \begin{bmatrix} M_{\lambda_1} & & \\ & M_{\lambda_2} & \\ & & \ddots \\ & & & M_{\lambda_k} \end{bmatrix}$$

Now, restricting our attention to the generalized eigenspaces, we have (Why does the chain end eventually?):

$$\text{Null}((A - \lambda I)) \subset \text{Null}((A - \lambda I)^2) \subset \ldots \subset \text{Null}((A - \lambda I)^r)$$

Take any $v \in \text{Null}((A - \lambda I)^r)$ s.t. $v \notin \text{Null}((A - \lambda I)^{r-1})$. Then $(A - \lambda I)^r v \in \text{Null}((A - \lambda I)^{r-s})$ and one can check that \{v_1, v_2, \ldots, v_r\} = \{(A - \lambda I)^{r-1}v, (A - \lambda I)^{r-2}v, \ldots, (A - \lambda I)v, v\} is an independent set from the definitions (hint: induction). This set up is called a Jordan chain. Notice that:

$$Av_s = v_{s-1} + \lambda v_s$$
$$Av_1 = \lambda v_1$$

So on this set of vectors, $A$ looks like:

$$\begin{bmatrix} \lambda & 1 \\ \lambda & 1 \\ \lambda & \lambda \end{bmatrix}$$

This is called a Jordan Block. If we start off with another independent vector $v \in \text{Null}((A - \lambda I)^r)$, we can make another Jordan chain and get another Jordan block. In general every matrix can be written in Jordan Canonical Form with only Jordan blocks.

**Definition 14.** The minimal polynomial is the “smallest” polynomial $p(x) = a_n x^n + \ldots + a_0$ so that:

$$p_A(A) = a_n A^n + \ldots + a_1 A + a_0 = 0$$

**Remark 15.** Factor $p_A = \prod (x - \lambda_s)^k$. By our work with the Jordan form, we know that the the only possible roots of $p_A$ are eigenvalues. (Otherwise each Jordan block cannot vanish). Moreover, the exponent $k$ corresponds to the size of the Largest Jordan Block.

**Problem 16.** Categories to look a from Written’s Wiki: “Rank Nullity Theorem”, “Jordan Canonical Form”, “Change of Basis”, “Projections, Rotations, Reflections”