Last time we saw direct sums and the primary decomposition of a vector space with respect to a linear transformation. If $f : V \to V$ is a linear transformation with minimal polynomial

$$\mu_f = p_1^{\epsilon_1} p_k^{\epsilon_k} \dots p_k^{\epsilon_k}$$

where each of the p_j is an irreducible polynomial, then we have

$$V = \mathrm{Ker}(p_1^{\epsilon_1}(f)) \oplus \dots \oplus \mathrm{Ker}(p_k^{\epsilon_k})f)).$$

In particular, if the characteristic polynomial χ_f has distinct linear factors, then f *diagonalizes*, i.e. there is a basis of eigenvectors $\vec{v}_1, \dots \vec{v}_n$ associated to the distinct eigenvalues $\lambda_1, \dots, \lambda_n$ and the matrix of f with respect to the standard basis A satisfies $S^{-1}AS = B$ where B is the diagonal matrix with eigenvalues λ_j on the diagonal and S is the change of basis matrix with columns $\vec{v}_1, \dots, \vec{v}_n$.

Example: Consider the linear transformation $f : \mathbb{R}^3 \to \mathbb{R}^3$ with matrix $A = \begin{bmatrix} -2 & 4 & -1 \\ -1 & 3 & -1 \\ 2 & -2 & 1 \end{bmatrix}$ in the standard basis. This matrix has characteristic polynomial

$$\begin{vmatrix} -2-\lambda & 4 & -1 \\ -1 & 3-\lambda & -1 \\ 2 & -2 & 1-\lambda \end{vmatrix} = -2 + \lambda + 2\lambda^2 - \lambda^3 = (-2+\lambda)(1-\lambda^2) = (-2+\lambda)(1-\lambda)(1+\lambda)$$

so the eigenvalues are $\lambda = 2, 1$ and -1. The eigenvectors are then

$$\begin{bmatrix} -2-2 & 4 & -1 \\ -1 & 3-2 & -1 \\ 2 & -2 & 1-2 \end{bmatrix} = \begin{bmatrix} -4 & 4 & -1 \\ -1 & 1 & -1 \\ 2 & -2 & -1 \end{bmatrix} \to \begin{bmatrix} 1 & -1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix} \Rightarrow \vec{v}_2 = (1,1,0),$$

$$\begin{bmatrix} -2-1 & 4 & -1 \\ -1 & 3-1 & -1 \\ 2 & -2 & 1-1 \end{bmatrix} = \begin{bmatrix} -3 & 4 & -1 \\ -1 & 2 & -1 \\ 2 & -2 & 0 \end{bmatrix} \to \begin{bmatrix} 1 & 0 & -1 \\ 0 & 1 & -1 \\ 0 & 0 & 0 \end{bmatrix} \Rightarrow \vec{v}_1 = (1,1,1)$$

and

$$\begin{bmatrix} -2+1 & 4 & -1 \\ -1 & 3+1 & -1 \\ 2 & -2 & 1+1 \end{bmatrix} = \begin{bmatrix} -1 & 4 & -1 \\ -1 & 4 & -1 \\ 2 & -2 & 2 \end{bmatrix} \to \begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix} \Rightarrow \vec{v}_{-1} = (-1,0,1)$$

Then we can verify that the matrix of f in the $\{\vec{v}_2, \vec{v}_1, \vec{v}_{-1}\}$ basis is the diagonal matrix

$$\begin{bmatrix} 2 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{bmatrix} = \begin{bmatrix} 1 & 1 & -1 \\ 1 & 1 & 0 \\ 0 & 1 & 1 \end{bmatrix}^{-1} \begin{bmatrix} -2 & 4 & -1 \\ -1 & 3 & -1 \\ 2 & -2 & 1 \end{bmatrix} \begin{bmatrix} 1 & 1 & -1 \\ 1 & 1 & 0 \\ 0 & 1 & 1 \end{bmatrix}$$

and we have

$$\mathbb{R}^3 = \mathrm{Ker}\left(\begin{bmatrix} -4 & 4 & -1 \\ -1 & 1 & -1 \\ 2 & -2 & -1 \end{bmatrix}\right) \oplus \mathrm{Ker}\left(\begin{bmatrix} -3 & 4 & -1 \\ -1 & 2 & -1 \\ 2 & -2 & 0 \end{bmatrix}\right) \oplus \mathrm{Ker}\left(\begin{bmatrix} -1 & 4 & -1 \\ -1 & 4 & -1 \\ 2 & -2 & 2 \end{bmatrix}\right).$$

Not every matrix is diagonalizable. If $\mathbb{F}$ is an algebraicially closed field like $\mathbb{C}$, then the minimal polynomial always factors into linear polynomials p_j, but these polynomials may appear in μ_f with multiplicity higher than 1. This means that the kernels of the maps $A - \lambda I$ do not span all of V, and V does not have a basis of eigenvectors. To get the primary decomposition, we must then find a basis for each $\mathrm{Ker}(A - \lambda_k)^{\epsilon_k}$. The resulting primary decomposition is not diagonal, but merely block diagonal. We can do better than just block-diagonal though; we can prove

Theorem: (Triangular Form Theorem) Let $f : V \to V$ be a linear transformation with minimal polynomial μ_f which factors into linear polynomials,

$$\mu_f = p_1^{\epsilon_1} \dots p_k^{\epsilon_k} = (x - \lambda_1)^{\epsilon_1} \dots (x - \lambda_k)^{\epsilon_k}.$$

Then there exists a basis of V such that the matrix of f with respect to this basis has the form

$$\begin{bmatrix} A_1 & 0 & \dots & 0 \\ 0 & A_2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & A_n \end{bmatrix} \quad \text{where} \quad A_i = \begin{bmatrix} * & * & \dots & * \\ 0 & * & \dots & * \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & * \end{bmatrix}.$$

That is, we can make the blocks upper triangular.

To see this, we'll need some new defintions.

Defintion: A linear transformation $f : V \to V$ is *idempotent* if $f^2 = f$, i.e. if $f(f(\vec{v})) = f(\vec{v})$ for all $\vec{v} \in V$. A linear transformation is *nilpotent* if $f^n = 0$ for some $n > 1$. In terms of matrices, a matrix A is the matrix of an idempotent lienar transformation if $A^2 = A$; a matrix is the matrix of a nilpotent linear transformation if $A^n = 0$ for some $n > 1$.

Example: The matrix $M = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}$ is idempotent and the matrix $N = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}$ is nilpotent:

$$M^2 = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} = \begin{bmatrix} 1+0 & 0+0 \\ 0+0 & 0+0 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} = M$$

and

$$N^2 = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} = \begin{bmatrix} 0+0 & 0+0 \\ 0+0 & 0+0 \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix} = 0$$

The idea behind the proof of the triangular form theorem is to note that every linear transformation $f : V \to V$ whose minimal polynomial factors into linear factors can be expressed as a sum $f = f_d + f_n$ of a diagonalizable linear transformation f_d and a nilpotent linear transformation f_n, known as the *Jordan decompositon*. More precisely, if we write

$$V = \mathrm{Ker}(p_1^{\epsilon_1}(f)) \oplus \dots \oplus \mathrm{Ker}(p_k^{\epsilon_k}(f))$$

then the projection maps $\pi_j : V \to \mathrm{Ker}(p_j^{\epsilon_j}(f))$ are idempotent maps and satisfy

$$\pi_1 + \dots + \pi_k = \mathrm{Id}.$$

Then the map $f_d : V \to V$ defined by

$$f_d = \lambda_1 \pi_1 + \dots + \lambda_k \pi_k$$

can be shown to be diagonalizable and the map $f_n = f - f_d$ can be shown to be nilpotent. Moreover, we can show that there exist polynomials p_d and p_n such that $f_d = p_d(f)$ and $f_n = p_n(f)$.

Note that if $(A - \lambda I)\vec{v} = \vec{0}$, then $(A - \lambda I)^2 \vec{v} = \vec{0}$ and $(A - \lambda I)^3 \vec{v} = \vec{0}$, etc. A vector $\vec{v}$ satisfying

$$(A - \lambda I)^j \vec{v} = \vec{0} \quad \text{and} \quad (A - \lambda I)^{j-1} \vec{v} \neq \vec{0}$$

is a *generalized eigenvector of rank j*. For each factor $p_j^{\epsilon_j}$ with $\epsilon_1 > 1$ we can find generalized eigenvectors $\vec{v}_1, \dots, \vec{v}_{\epsilon_j}$ forming a basis for $\mathrm{Ker}(A - \lambda I)^{\epsilon_1}$ by finding a basis B for the λ-eigenspace and completing B to a basis $\mathrm{Ker}(A - \lambda I)^{\epsilon_1}$ with vectors $\vec{u}$ satisfying $(A - \lambda I)\vec{u} = \vec{v}$ for $\vec{v} \in B$. Then the matrices of f_d and f_n with respect to this basis are diagonal and strictly upper triangular respectively.

Example: Find the triangular form for $A = \begin{bmatrix} 1 & -1 & -1 \\ -1 & 0 & -2 \\ 1 & 2 & 4 \end{bmatrix}$.

First, find the characteristic polynomial:

$$\chi_A = \begin{vmatrix} 1-\lambda & -1 & -1 \\ -1 & -\lambda & -2 \\ 1 & 2 & 4-\lambda \end{vmatrix} = (1-\lambda)(2-\lambda)^2$$

Then there are only two eigenvalues, $\lambda = 1$ and $\lambda = 2$.

$$\begin{bmatrix} 1-1 & -1 & -1 \\ -1 & -1 & -2 \\ 1 & 2 & 4-1 \end{bmatrix} = \begin{bmatrix} 0 & -1 & -1 \\ -1 & -1 & -2 \\ 1 & 2 & 3 \end{bmatrix} \to \begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & 0 \end{bmatrix} \Rightarrow \vec{v}_1 = (-1,-1,1)$$

and

$$\begin{bmatrix} 1-2 & -1 & -1 \\ -1 & -2 & -2 \\ 1 & 2 & 4-2 \end{bmatrix} = \begin{bmatrix} -1 & -1 & -1 \\ -1 & -2 & -2 \\ 1 & 2 & 2 \end{bmatrix} \Rightarrow \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 1 \\ 0 & 0 & 0 \end{bmatrix} \vec{v}_2 = (0,-1,1)$$

So, $\mathrm{Ker}(A-2I)$ is only 1-dimensional, and $\mathbb{R}^3$ does not have a basis of eigenvectors for A. We can simply find a basis for $\mathrm{Ker}(A-2I)^2$:

$$\begin{bmatrix} -1 & -1 & -1 \\ -1 & -2 & -2 \\ 1 & 2 & 2 \end{bmatrix}^2 = \begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ -1 & -1 & -1 \end{bmatrix} \Rightarrow \begin{bmatrix} 1 & 1 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \Rightarrow \mathrm{Ker}(A-2I)^2 = \mathrm{Span}((-1,1,0),(-1,0,1)).$$

Then the change-of-basis matrix S yields block-diagonal decomposition

$$S = \begin{bmatrix} -1 & -1 & -1 \\ -1 & 1 & 0 \\ 1 & 0 & 1 \end{bmatrix} \quad \Rightarrow \quad S^{-1}AS = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & -1 \\ 0 & 1 & 3 \end{bmatrix}.$$

To get a basis which puts A in triangular form, we need to find a generalized eigenvector $\vec{v}_{2'}$ satisfying $A\vec{v}_{2'} = \vec{v}$ to form a completed basis for $\mathrm{Ker}(A-2I)^2$.

$$\left[\begin{array}{ccc|c} -1 & -1 & -1 & 0 \\ -1 & -2 & -2 & -1 \\ 1 & 2 & 2 & 1 \end{array}\right] \to \left[\begin{array}{ccc|c} 1 & 0 & 0 & -1 \\ 0 & 1 & 1 & 1 \\ 0 & 0 & 0 & 0 \end{array}\right] \to \vec{v}_{2'} = (-1,1,0)$$

So with the new basis, we have

$$S = \begin{bmatrix} -1 & 0 & -1 \\ -1 & -1 & 1 \\ 1 & 1 & 0 \end{bmatrix} \Rightarrow S^{-1}AS = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 2 & 1 \\ 0 & 0 & 2 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 2 \end{bmatrix} + \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix}.$$