Our goal remains to find a good basis for $L : V \to V$, the first stage will be to find an easy to compute direct sum decomposition.

The idea we are motivated by is something which you have seen before:

Let $L : V \to V$ be a linear transformation.

Fix $\lambda \in \mathbb{R}$ (or $\lambda \in \mathbb{C}$) recall that if

$$\vec{v} \in \mathrm{Ker}(L - \lambda \mathrm{Id}_V)$$

then $L(\vec{v}) = \lambda \vec{v}$.

We call such a kernel (when it has positive dimension) an **Eigenspace**.
We call such a non-zero vector an **eigenvectors**.
We call the values λ for which these non-zero vectors exist an **eigenvalue**.

You have seen (hopefully!), though perhaps not in this language, that sometimes, these eigenspaces let us find basis where the matrix is really nice.

The result was that *when you had a basis of eigenvectors* you could use them to find an invertible P so that

$$P^{-1}AP$$

was diagonal. This was just a change of basis.
On the assignment you will prove a couple things about the case of eigenvectors independently of what we are about to do.

Generalized Eigenspaces

Let $L : V \to V$ be a linear transformation.

As before, fix $\lambda \in \mathbb{R}$ (or $\lambda \in \mathbb{C}$).

For any $k > 0$ if $\mathrm{Ker}\left((L - \lambda \mathrm{Id}_V)^k\right) \neq \{\vec{0}\}$ then we call it a **generalized eigenspace**.

Note that if $\mathrm{Ker}(L - \lambda \mathrm{Id}_V) = \{\vec{0}\}$ then for all k we have $\mathrm{Ker}\left((L - \lambda \mathrm{Id}_V)^k\right) = \{\vec{0}\}$ (Why?), and hence the values λ are exactly the roots of both the characteristic and minimal polynomial.

Note Because every polynomial factors completely over $\mathbb{C}$, every linear transformation of finite positive dimensional vector spaces has eigenvalues in $\mathbb{C}$.
In principal, whenever we are talking about L, we might need to 'extend scalars' to pretend we are working over $\mathbb{C}^n$ and with complex matricies rather than over $\mathbb{R}^n$ and real matricies.
In practice, you won't actually notice this is happening..

In the following we will talk alot about

$$\mathrm{Ker}(P(L))$$

for arbitrary polynomials L, these are just slight generalizations of generalized eigenspaces.

Generalized Eigenspaces are Invariant

Lemma
Suppose $L : V \to V$ is any linear transformation. For any polynomials $P(x)$ and $Q(x)$ all the subspaces

$$\mathrm{Ker}\,(P(L)) \qquad \text{and} \qquad \mathrm{Im}\,(P(L))$$

are $Q(L)$-invariant.
Note This includes the case $Q(x) = x$, so they are L-invariant.
Proof Idea Recalling that $P(L)$ and $Q(L)$ commute, this result follows from the assignment.

The important cases for us are when $P(x) = (x - \lambda)^k$

Lemma
Suppose $L : V \to V$ is any linear transformation. For any two polynomials $P_1(x)$ and $P_2(x)$ with no common factors we have:

$$\operatorname{Ker}(P_1(L)) \cap \operatorname{Ker}(P_2(L)) = \{\vec{0}\} \quad \text{and} \quad \operatorname{Im}(P_2(L)) \cap \operatorname{Ker}(P_1(L)) = \operatorname{Ker}(P_1(L))$$

Proof: For the first claim recall that we have polynomials $S_1(x)$ and $S_2(x)$ so that

$$S_1(x)P_1(x) + S_2(x)P_2(x) = 1 \qquad \Rightarrow \qquad S_1(L) \circ P_1(L) + S_2(L) \circ P_2(L) = \mathrm{Id}_V$$

From this it follows that if $\vec{v} \in \operatorname{Ker}(P_1(L)) \cap \operatorname{Ker}(P_2(L))$ then

$$\vec{v} = \mathrm{Id}(\vec{v}) = (S_1(L) \circ P_1(L) + S_2(L) \circ P_2(L))(\vec{v}) = S_1(L)(\vec{0}) + S_2(L)(\vec{0}) = \vec{0}$$

For the second claim, Now, for every $\vec{v} \in \operatorname{Ker}(P_1(L))$ we also get

$$\begin{aligned} \vec{v} &= \mathrm{Id}(\vec{v}) \\ &= (S_1(L) \circ P_1(L) + S_2(L) \circ P_2(L))(\vec{v}) \\ &= P_2(L) \circ S_2(L)(\vec{v}) \in \operatorname{Im}(P_2(L)) \end{aligned}$$

Which shows $\vec{v} \in \operatorname{Im}(P_2(L))$.

Generalized Eigenspaces are Complementary.

Theorem

Suppose $L : V \to V$ is any linear transformation. If $P_1(x)$ and $P_2(x)$ are any polynomials with no common factors then:

$$\mathrm{Ker}(P_1(L) \circ P_2(L)) = \mathrm{Ker}(P_1(L)) \oplus \mathrm{Ker}(P_2(L))$$

Proof Idea:

We know that $\mathrm{Ker}(P_1(L))$ and $\mathrm{Ker}(P_2(L))$ are subspaces of $\mathrm{Ker}(P_1(L) \circ P_2(L))$.
We must show they satisfy conditions to give direct sum, as before write

$$P_1(x)S_1(x) + P_2(x)S_2(x) = 1$$

so that given any $\vec{v} \in \mathrm{Ker}(P_1(L) \circ P_2(L))$ we have

$$\vec{v} = P_2(L) \circ S_2(L)(\vec{v}) + P_1(L) \circ S_1(L)(\vec{v})$$

Then $P_2(L) \circ S_2(L)(\vec{v}) \in \mathrm{Ker}(P_1(L))$ because

$$P_1(L)(P_2(L) \circ S_2(L)(\vec{v})) = S_2(L)(P_1(L) \circ P_2(L)(\vec{v})) = S_2(L)(\vec{0}).$$

Likewise, $P_1(L) \circ S_1(L)(\vec{v}) \in \mathrm{Ker}(P_2(L))$.

The uniqueness of the expression for $\vec{v}$ follows from the fact that

$$\mathrm{Ker}(P_1(L)) \cap \mathrm{Ker}(P_2(L)) = \{\vec{0}\}.$$

using A2(proofs)Q4b.

Decomposition into Generalized Eigenspaces

Theorem
Suppose $L : V \to V$ is any linear transformation.
Suppose $\min_L(x) = P_1(x) \cdots P_r(x)$ is any factorization of the minimal polynomial such that the $P_i(x)$ have no common factors then

$$V = \mathrm{Ker}\,(P_1(L)) \oplus \cdots \oplus \mathrm{Ker}\,(P_r(L))$$

is an L-invariant direct sum decomposition.

Proof Idea
This follows from the previous theorem by induction, and the observation that

$$\begin{aligned} V = \mathrm{Ker}(\min_L(L)) &= \mathrm{Ker}(P_1(L) \circ \cdots \circ P_r(L)) \\ &= \mathrm{Ker}(P_1(L)) \oplus \mathrm{Ker}(P_2(L) \circ \cdots \circ P_r(L)) \\ &= \mathrm{Ker}(P_1(L)) \oplus \mathrm{Ker}(P_2(L)) \oplus \mathrm{Ker}(P_3(L) \circ \cdots \circ P_r(L)) \\ &\vdots \\ &= \mathrm{Ker}\,(P_1(L)) \oplus \cdots \oplus \mathrm{Ker}\,(P_r(L)) \end{aligned}$$

Which gives the result.

Generalized Eigenspaces Give Invariant Direct Sum Decomposition.

Theorem
Suppose $L : V \to V$ is any linear transformation of a finite dimensional vector space. Suppose $\lambda_1, \ldots, \lambda_r$ are the roots of the characteristic/minimial polynomial of L. Then

$$V = U_{\lambda_1} \oplus \cdots \oplus U_{\lambda_r}$$

is an invariant direct sum decomposition. where $U_{\lambda_i} = \mathrm{Ker}((L - \lambda_i)^{m_{\lambda_i}})$ and

$$\min_L(x) = \prod_{i=1}^{r} (x - \lambda_i)^{m_{\lambda_i}}$$

Proof This is a special case of the previous result because $(x - \lambda_i)^{m_{\lambda_i}}$ and $(x - \lambda_j)^{m_{\lambda_j}}$ have no common factors when $\lambda_i \neq \lambda_j$.

The above is the result we have been after, it is the first stage in getting a good basis for V!!!
Step 1 of finding the Jordan form will be working towards finding this kernels!!

Characteristic and Minimal Polynomials

Recalling that the characteristic polynomial for L is simply the product of the characteristic polynomial of each factor it would be useful to check what those are

Lemma

The characteristic polynomial for L acting on

$$\mathrm{Ker}\left((L-\lambda)^{m_\lambda}\right)$$

is

$$(x-\lambda)^{\mathrm{Dim}(\mathrm{Ker}((L-\lambda)^{m_\lambda}))}.$$

Proof: Indeed, we know that λ is the only root of the characteristic polynomial for L acting on this factor because for any other value $\mu \neq \lambda$ we have

$$\mathrm{Ker}(L-\mu\mathrm{Id}) \cap \mathrm{Ker}\left((L-\lambda)^{m_\lambda}\right) = \{\vec{0}\}$$

so $L-\mu\mathrm{Id}$ is injective.

Because the characteristic polynomial has degree equal to the dimension, and only a single root, this is the only form it could have.

The above will be usefull for identifying characteristic polynomials by looking at matricies in Jordan form, but it is also useful because it tells us

what dimension we need each generalized eigenspace to be

while trying to find the Jordan form

Likewise, recalling that the minimal polynomial for L is simply the LCM of the minimal polynomial of each factor it would be useful to check what those are.

Recall that we are looking at vector spaces

$$\mathrm{Ker}\left((L - \lambda \mathrm{Id}_V)^{m_\lambda}\right)$$

where $(x - \lambda)^{m_\lambda}$ was a factor of the minimal polynomial.

Lemma

If V is finite dimensional then m_λ is the smallest integer such that

$$\forall n > m_\lambda, \mathrm{Ker}\left((L - \lambda \mathrm{Id}_V)^{n}\right) = \mathrm{Ker}\left((L - \lambda \mathrm{Id}_V)^{m_\lambda}\right).$$

and so the minimal polynomial for L acting on $\mathrm{Ker}\left((L - \lambda)^{m_\lambda}\right)$ is $(x - \lambda)^{m_\lambda}$.

Proof: If there was any $n_\lambda > m_\lambda$ for which $\mathrm{Ker}\left((L - \lambda \mathrm{Id}_V)^{n_\lambda}\right)$ has larger dimension than $\mathrm{Ker}\left((L - \lambda \mathrm{Id}_V)^{m_\lambda}\right)$ then the previous theorem would also imply

$$V = \bigoplus_\lambda \mathrm{Ker}\left((L - \lambda \mathrm{Id}_V)^{n_\lambda}\right)$$

But the dimension of V is the sum of the dimensions of the pieces... so these can't get any bigger.

The claim about the minimal polynomial follows from the previous lemma, as λ can be its only root, and the above definition which verifies the multiplicity.

Lemma

With notation as above, $m_{\lambda_i} \leq \dim(V_{\lambda_i})$.

Proving this directly is a bonus question on the assignment.

Characteristic Polynomials.

Lemma

Suppose $L : V \to V$ is any linear transformation of a finite dimensional vector space. Suppose $\lambda_1, \ldots, \lambda_r$ are the roots of the characteristic/minimial polynomial of L. Then:

$$\mathrm{char}_L(x) = \prod_{i=1}^{r}(x - \lambda_i)^{\dim(U_{\lambda_i})} \quad \text{and} \quad \min_L(x) = \prod_{i=1}^{r}(x - \lambda_i)^{m_{\lambda_i}}$$

where m_λ is the smallest value with $\forall n > m_\lambda, \mathrm{Ker}\left((L - \lambda \mathrm{Id}_V)^n\right) = \mathrm{Ker}\left((L - \lambda \mathrm{Id}_V)^{m_\lambda}\right)$.

Proof:

This follows from the formula for combining characteristic polynomials/minimial polynomials for invariant direct sum decompositions together with our the fact that for each factor the characteristic polynomial is $(x - \lambda_i)^{d_\lambda}$ and the minimal polynomial is $(x - \lambda_i)^{m_\lambda}$, hence have no common factors.

Theorem(Cayley-Hamilton)

Suppose $L : V \to V$ is any linear transformation of a finite dimensional vector space, then

$$\min_L(x) | \mathrm{char}_L(x)$$

Proof That $m_{\lambda_i} \leq \mathrm{Dim}(V_{\lambda_i})$ gives it to us immediately.

This is the second proof we have for this theorem.

What we have so far.

Because we have a **canonical** direct sum decomposition, in fact, a direct sum decomposition that we know how to compute!, we are now interested in studying canonical forms for each piece of that decomposition.

We must now look at the case:

$$\mathrm{char}_L(x) = (x-\lambda)^n \qquad \mathrm{char}_L(x) = (x-\lambda)^m$$

Lemma

If the linear transformation L has characteristic and minimal polynomials $P(x)$ and $Q(x)$ respectively then

$$L - \lambda \mathrm{Id}_V$$

has characteristic and minimal polynomials $P(x+\lambda)$ and $Q(x+\lambda)$.
This is on the assigment.

This reduces us to the case

$$\mathrm{char}_L(x) = x^n \qquad \mathrm{char}_L(x) = x^m$$

You will ultimiately prove the result for this case on the assignment, but we explain what it is.
If the matrix for $L - \lambda \mathrm{Id}_V$ in some basis is A, and A is really nice.
Then the matrix for L in the same basis is

$$A + \lambda \mathrm{Id}_n$$

Example

Consider the matrix

$$A = \begin{pmatrix} 0 & 6 & -5 & 3 \\ -2 & 7 & -2 & 4 \\ -1 & 2 & 2 & 2 \\ 1 & -2 & 1 & 1 \end{pmatrix}$$

Find a matrix P such that $P^{-1}AP$ is simple.

We first need to find kernels of $(A - \lambda Id)^m$
Step 1 find characteristic polynomial and factor:

$$x^4 - 10x^3 + 37x^2 - 60x + 36 = (x-2)^2(x-3)^2$$

So the λ we need are 2 and 3.

Example - Continued

Step 2 Find matricies to find Kernels of:

$$(A-2I)=\begin{pmatrix}-2 & 6 & -5 & 3\\ -2 & 5 & -2 & 4\\ -1 & 2 & 0 & 2\\ 1 & -2 & 1 & -1\end{pmatrix} \qquad (A-3I)=\begin{pmatrix}-3 & 6 & -5 & 3\\ -2 & 4 & -2 & 4\\ -1 & 2 & -1 & 2\\ 1 & -2 & 1 & -2\end{pmatrix}$$

Making a mistake at this step sucks!!

Example - Continued

Step 3a Row reduce and find a basis for kernel of $(A - 2I)$:

$$\begin{pmatrix} -2 & 6 & -5 & 3 \\ -2 & 5 & -2 & 4 \\ -1 & 2 & 0 & 2 \\ 1 & -2 & 1 & -1 \end{pmatrix} \quad \begin{array}{c} R4 \leftrightarrow R1 \\ R2 + 2R4 \rightarrow R2 \\ R3 + R4 \rightarrow R3 \\ R1 + 2R4 \rightarrow R4 \end{array}$$

$$\Rightarrow \begin{pmatrix} 1 & -2 & 1 & -1 \\ 0 & 1 & 0 & 2 \\ 0 & 0 & 1 & 1 \\ 0 & 2 & -3 & 1 \end{pmatrix} \quad \begin{array}{c} R1 + 2R2 - R3 \rightarrow R1 \\ R2 \rightarrow -R2 \\ \\ R4 - 2R2 \rightarrow R4 \end{array}$$

$$\Rightarrow \begin{pmatrix} 1 & 0 & 0 & 2 \\ 0 & 1 & 0 & 2 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & -3 & -3 \end{pmatrix} \quad \begin{array}{c} \\ \\ \\ R4 + 3R3 \rightarrow R4 \end{array}$$

$$\Rightarrow \begin{pmatrix} 1 & 0 & 0 & 2 \\ 0 & 1 & 0 & 2 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

a basis for this kernel is $(-2, -2, -1, 1)$.

If you find that you have no non-trivial solutions you made a mistake.

Example - Continued

Step 3 b Become a little bit sad that you didn't find enough vectors in the basis, we wanted 2, but only found 1.

Step 3 c Compute $(A - 2I)^2$ and row reduce

$$(A-2I)^2 = \begin{pmatrix} 0 & 2 & 1 & 5 \\ 0 & 1 & 4 & 6 \\ 0 & 0 & 3 & 3 \\ 0 & 0 & -2 & -2 \end{pmatrix} \qquad \begin{matrix} R1 \leftrightarrow R2 \\ R1 - 2R2 \to R2 \\ R3 + R4 \to R3 \end{matrix}$$

$$\Rightarrow \begin{pmatrix} 0 & 1 & 4 & 6 \\ 0 & 0 & -7 & -7 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & -2 & -2 \end{pmatrix} \qquad \begin{matrix} R1 - 4R3 \to R1 \\ R2 \leftrightarrow R \\ R2 + 7R3 \to R3 \\ R4 + 2R3 \to R4 \end{matrix}$$

$$\Rightarrow \begin{pmatrix} 0 & 1 & 0 & 2 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

a basis for this kernel is $(1,0,0,0), (0,-2,-1,1)$. **The original vector $(-2,-2,-1,1)$ will be in this kernel, but it may not be one of your basis vectors!**

Example - Continued

Step 3 d Rejoice at having found enough vectors!

Step 3 e Pick a good basis for the kernel We will actually do this later.

Step 3 f Repeat Step 3 a for each eigenvector.

Example - Continued

Step 3a Row reduce and find a basis for kernel of $(A - 3I)$:

$$\begin{pmatrix} -3 & 6 & -5 & 3 \\ -2 & 4 & -2 & 4 \\ -1 & 2 & -1 & 2 \\ 1 & -2 & 1 & -2 \end{pmatrix} \qquad \begin{matrix} R4 \leftrightarrow R1 \\ R2 + 2R4 \rightarrow R2 \\ R3 + R4 \rightarrow R3 \\ R1 + 3R4 \rightarrow R4 \end{matrix}$$

$$\begin{pmatrix} 1 & -2 & 1 & -2 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & -2 & -3 \end{pmatrix}$$

a basis for this kernel is: $(2, 1, 0, 0), (7, 0, -3, 2)$.

Step 3b Rejoice at having 2 vectors already!

Intermission - What we know so far

$$\mathrm{Ker}(A - 2Id)^2 = \mathrm{Span}((1,0,0,0),(0,-2,-1,1))$$
$$\mathrm{Ker}(A - 3Id) = \mathrm{Span}((2,1,0,0),(7,0,-3,2))$$
$$\mathbb{R}^4 = \mathrm{Span}((1,0,0,0),(0,-2,-1,1)) \oplus \mathrm{Span}((2,1,0,0),(7,0,-3,2))$$

and if we use this basis to describe A, the change of basis matrix is

$$P = \begin{pmatrix} 1 & 0 & 2 & 7 \\ 0 & -2 & 1 & 0 \\ 0 & -1 & 0 & -3 \\ 0 & 1 & 0 & 2 \end{pmatrix}$$

and so we know

$$P^{-1}AP = \begin{pmatrix} ? & ? & 0 & 0 \\ ? & ? & 0 & 0 \\ 0 & 0 & * & * \\ 0 & 0 & * & * \end{pmatrix}$$

We find

$$\begin{pmatrix} ? & ? \\ ? & ? \end{pmatrix} \qquad \text{and respectively} \qquad \begin{pmatrix} * & * \\ * & * \end{pmatrix}$$

By looking at what A does to $(1,0,0,0),(0,-2,-1,1)$ (respectively $(2,1,0,0),(7,0,-3,2)$).

Intermission - what are the blocks

We can check that

$$A(1,0,0,0) = (0,-2,-1,1) = 0(1,0,0,0) + 1(0,-2,-1,1)$$

and that

$$A(0,-2,-1,1) = (4,-8,-4,4) = -4(1,0,0,0) + 4(0,-2,-1,1)$$

so

$$\begin{pmatrix} ? & ? \\ ? & ? \end{pmatrix} = \begin{pmatrix} 0 & -4 \\ 1 & 4 \end{pmatrix}$$

This isn't quite what we wanted, because we didn't carefully pick the basis!

We can also check that

$$A(2,1,0,0) = 3(2,1,0,0)$$

and that

$$A(7,0,-3,2) = 3(7,0,-3,2)$$

and so

$$\begin{pmatrix} * & * \\ * & * \end{pmatrix} = \begin{pmatrix} 3 & 0 \\ 0 & 3 \end{pmatrix}$$

This is what we wanted.

Return to Step 3 e

Step 3 e Pick a good basis for the kernel.

How do we do that?

The observation is that the matrix for A acting on $\mathrm{Span}((1,0,0,0),(0,-2,-1,1))$ was

$$\begin{pmatrix} ? & ? \\ ? & ? \end{pmatrix} = \begin{pmatrix} 0 & -4 \\ 1 & 4 \end{pmatrix}$$

We notice that the characteristic polynomial of this matrix is

$$(x-2)^2$$

However, if we look at $(A-2I)$ it will be the matrix

$$\begin{pmatrix} 0 & -4 \\ 1 & 4 \end{pmatrix} - 2\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} -2 & -4 \\ 1 & 2 \end{pmatrix}$$

and the characteristic polynomial of this matrix is x^2!!

We will return to this in a moment, but that case of a simple characterist polynomial is important.

We say a linear transformation is **nilpotent** if there exists m so that $L^m = 0$.
From the assignment we know L is nilpotent if and only if $\mathrm{char}_L(x) = x^n$.

Theorem
If $L : V \to V$ is nilpotent, so if $\mathrm{char}_L(x) = x^n$, then there exists a basis for V such that the matrix for L has the form:

$$\begin{pmatrix} A_1 & 0 & \cdots & & 0 \\ 0 & A_2 & \ddots & & \\ \vdots & \ddots & \ddots & & \vdots \\ & & 0 & A_{r-1} & 0 \\ 0 & & \cdots & 0 & A_r \end{pmatrix}$$

Where each of that A_i has the form:

$$\begin{pmatrix} 0 & 1 & 0 & & \cdots & 0 \\ 0 & 0 & 1 & 0 & & \\ \vdots & & \ddots & \ddots & \ddots & \vdots \\ & & & 0 & 1 & 0 \\ & & & & 0 & 1 \\ 0 & & \cdots & & & 0 \end{pmatrix}$$

This is on the assigment

Corollary
If the characteristic polynomial of L is instead $(x - \lambda)^n$ then the blocks A_i above instead have λ on the diagonal.
This is also on the assigment

Back to the misbehaving 2x2 block

The idea behind this was to use a decomposition with basis of the forms

$$L^{\ell}(\vec{v}), \dots, L^2(\vec{v}), L(\vec{v}), \vec{v}$$

so if we apply this to

$$A - 2I = \begin{pmatrix} -2 & -4 \\ 1 & 2 \end{pmatrix}$$

pick any vector $\vec{v}$, how about $(1,0)$, then $(A-2I)(\vec{v}) = (-2,1)$ and so the matrix in the basis

$$(-2,1), (1,0)$$

In that order is

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

So the matrix for A in the basis

$$(-2,1), (1,0)$$

is

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

Example - Continued

So what we should have done originally when picking the basis for

$$\mathrm{Ker}(A - 2Id)^2 = \mathrm{Span}((1,0,0,0),(0,-2,-1,1))$$

Was use a basis like

$$(A - 2Id)\vec{v}, \vec{v}$$

Where

$$\vec{v} \in \mathrm{Ker}(A - 2Id)^2 \setminus \mathrm{Ker}(A - 2Id)$$

so lets say:

$$(A - 2Id)(1,0,0,0),(1,0,0,0)$$

which is the basis

$$(-2,-2,-1,1),(1,0,0,0)$$

Taking this into account we find

$$\begin{pmatrix} -2 & 1 & 2 & 7 \\ -2 & 0 & 1 & 0 \\ -1 & 0 & 0 & -3 \\ 1 & 0 & 0 & 2 \end{pmatrix}^{-1} \begin{pmatrix} 0 & 6 & -5 & 3 \\ -2 & 7 & -2 & 4 \\ -1 & 2 & 2 & 2 \\ 1 & -2 & 1 & 1 \end{pmatrix} \begin{pmatrix} -2 & 1 & 2 & 7 \\ -2 & 0 & 1 & 0 \\ -1 & 0 & 0 & -3 \\ 1 & 0 & 0 & 2 \end{pmatrix} = \begin{pmatrix} 2 & 1 & 0 & 0 \\ 0 & 2 & 0 & 0 \\ 0 & 0 & 3 & 0 \\ 0 & 0 & 0 & 3 \end{pmatrix}$$

We will work through completely some examples shortly, but for now I want to walk us through the process we will use to go from arbitrary matrix

$$A$$

to finding a basis with respect to which the matrix is just a bunch of blocks:

$$\begin{pmatrix} A_1 & 0 & \cdots & & 0 \\ 0 & A_2 & \ddots & & \\ \vdots & \ddots & \ddots & & \vdots \\ & & 0 & A_{r-1} & 0 \\ 0 & & \cdots & 0 & A_r \end{pmatrix}$$

where each A_i is of the form:

$$\begin{pmatrix} \lambda_i & 1 & 0 & & \cdots & 0 \\ 0 & \lambda_i & 1 & 0 & & \\ \vdots & & \ddots & \ddots & \ddots & \vdots \\ & & & \lambda_i & 1 & 0 \\ & & & & \lambda_i & 1 \\ 0 & & \cdots & & & \lambda_i \end{pmatrix}$$

where some or all of the λ_i may be the same or different.

Finding the Decomposition

1. Compute the characteristic polynomial $\mathrm{char}_L(x)$.
2. Factor $\mathrm{char}_L(x) = \prod(x-\lambda)^{d_\lambda}$, the multiplicities are important.
3. Find the kernels $\mathrm{Ker}((L-\lambda\mathrm{Id})^j)$ until you get dimension d_λ.
   The power j when this first happens is the multiplicity m_λ for minimal polynomial.
4. If $m_\lambda = d_\lambda$ or if $m_\lambda = 1$ then you are done!! Expect this in questions you get.
   If not, you need to carefully look at how the ranks change as you increase j to figure out how many *Jordan Blocks* of each length there are.

Finding the Basis - Do this for each $\mathrm{Ker}((L-\lambda\mathrm{Id})^{m_\lambda})$ - There are a few cases:

(a) If $m_\lambda = 1$, pick any basis $\vec{v}_1, \ldots, \vec{v}_k$ for $\mathrm{Ker}((L-\lambda\mathrm{Id}))$, done!.

(b) If $m_\lambda = d_\lambda$ then pick any element $\vec{v} \in \mathrm{Ker}((L-\lambda\mathrm{Id})^{m_\lambda}) \setminus \mathrm{Ker}((L-\lambda\mathrm{Id})^{m_\lambda-1})$.

The basis is $(L-\lambda\mathrm{Id})^{m_\lambda-1}(\vec{v}), (L-\lambda\mathrm{Id})^{m_\lambda-2}(\vec{v}), \ldots, \vec{v}$ done!.

(c) Otherwise...

Keep repeating (b) on largest ℓ where

$$\mathrm{Ker}((L-\lambda\mathrm{Id})^{\ell}) \setminus \mathrm{Span}(\mathrm{Ker}((L-\lambda\mathrm{Id})^{\ell-1}), \text{ vectors already picked})$$

is not empty.

keeping doing this until you have a basis you will get k lists like in (b):

$$(L-\lambda\mathrm{Id})^{m_\lambda-1}(\vec{v}_1), \ldots, \vec{v}_1, (L-\lambda\mathrm{Id})^{\ell_2-1}(\vec{v}_2), \ldots, \vec{v}_2, \ldots\ldots, (L-\lambda\mathrm{Id})^{\ell_k}(\vec{v}_k), \ldots, \vec{v}_k$$

This list will have d_λ vectors total!! and each $\ell_i \leq m_\lambda$.

Note: You can always use option (c), and if you are going to you can skip steps 3/4 in the finding the decomposition part above.

Result

Having done what is on the previous slide, for each root λ you will have d_λ many vectors (multiplicity from characteristic polynomial).

Combining these for all λ that are roots you get a basis in which your matrix will be in Jordan form.

For questions I ask you it will basically be the case that either $m_\lambda = 1$ and the block looks like:

$$\lambda \mathrm{Id}_{d_\lambda}$$

or $d_\lambda = m_\lambda$ and the block looks like

$$\begin{pmatrix} \lambda & 1 & 0 & & \cdots & 0 \\ 0 & \lambda & 1 & 0 & & \\ \vdots & & \ddots & \ddots & \ddots & \vdots \\ & & & \lambda & 1 & 0 \\ & & & & \lambda & 1 \\ 0 & & \cdots & & & \lambda \end{pmatrix}$$

Though in principal, otherwise for each of the chains $(L - \lambda \mathrm{Id})^{\ell_i - 1}(\vec{v}_i), \ldots, \vec{v}_i$ you get such a block So you could get something like

$$\begin{pmatrix} \lambda & 1 & 0 & 0 & 0 \\ 0 & \lambda & 0 & 0 & 0 \\ 0 & 0 & \lambda & 1 & 0 \\ 0 & 0 & 0 & \lambda & 0 \\ 0 & 0 & 0 & 0 & \lambda \end{pmatrix}$$

which would correspond to the ordered basis

$$(L - \lambda \mathrm{Id})(\vec{v}_1), \vec{v}_1, (L - \lambda \mathrm{Id})(\vec{v}_2), \vec{v}_2, \vec{v}_3$$

Find the Jordan form for A, and a basis of $\mathbb{R}^3$ for which the matrix is in this Jordan form when

$$A = \begin{pmatrix} 3 & -2 & -2 \\ -1 & 4 & -1 \\ -3 & 6 & 2 \end{pmatrix}$$

Note: the characteristic polynomial of A is

$$\mathrm{char}_A(x) = (x-2)^2(x-5)$$

Change of Basis Matrix

If A is the original matrix and B is the matrix in Jordan form.
Then

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

where P is the matrix whose columns are the vectors $\vec{e}_1, \ldots, \vec{e}_n$ which put the matrix in Jordan form.

The order of the vectors $\vec{e}_1, \ldots, \vec{e}_n$ is essential to actually getting the matrix in the same shape as what you said B would be.

Things to Look out for in the Process

Finding the Jordan form is a long process, it is easy to make mistakes, most of these can be detected by something going wrong.

- If λ is a root of $\mathrm{char}_L(x)$ it is **not** possible that

  $$\mathrm{Ker}(L - \lambda \mathrm{Id}) = \{\vec{0}\}$$

  there must be a non-zero vector.
- If $(x-\lambda)^2$ divides $\mathrm{char}_L(x)$ it is **not** possible that

  $$\mathrm{Dim}(\mathrm{Ker}((L - \lambda \mathrm{Id})^2)) = 1$$

  because more generally if $(x-\lambda)^k$ divides $\mathrm{char}_L(x)$ then

  $$\mathrm{Dim}(\mathrm{Ker}((L - \lambda \mathrm{Id})^k)) \geq k$$
- If you pick a vector $\vec{v} \in \mathrm{Ker}((L - \lambda \mathrm{Id})^2)$, then it will always be the case that

  $$(L - \lambda \mathrm{Id})(\vec{v}) \in \mathrm{Ker}(L - \lambda \mathrm{Id})$$

Reading off Minimal and Characterisitic Polynomials

Though our proofs already tell us how to read off the characteristic and minimal polynomials from the Jordan form of a matrix, it is useful to quickly go over a good way to remind yourself.

- Group blocks by diagonal entry.
- The total number of each diagonal entry gives characteristic polynomial.
- The largest block for each diagonal entry gives the minimal polynomial.

Can you read off the answers for:

$$\begin{pmatrix} 2 & 1 & 0 & 0 & 0 \\ 0 & 2 & 0 & 0 & 0 \\ 0 & 0 & 3 & 1 & 0 \\ 0 & 0 & 0 & 3 & 0 \\ 0 & 0 & 0 & 0 & 4 \end{pmatrix} \quad \begin{pmatrix} 3 & 1 & 0 & 0 & 0 \\ 0 & 3 & 0 & 0 & 0 \\ 0 & 0 & 2 & 1 & 0 \\ 0 & 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 0 & 2 \end{pmatrix} \quad \begin{pmatrix} 2 & 1 & 0 & 0 & 0 \\ 0 & 2 & 0 & 0 & 0 \\ 0 & 0 & 2 & 1 & 0 \\ 0 & 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 0 & 4 \end{pmatrix} \quad \begin{pmatrix} 2 & 1 & 0 & 0 & 0 \\ 0 & 2 & 1 & 0 & 0 \\ 0 & 0 & 2 & 1 & 0 \\ 0 & 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 0 & 2 \end{pmatrix}$$

$$\begin{pmatrix} \color{red}{2} & \color{red}{1} & 0 & 0 & 0 \\ \color{red}{0} & \color{red}{2} & 0 & 0 & 0 \\ 0 & 0 & \color{blue}{3} & \color{blue}{1} & 0 \\ 0 & 0 & \color{blue}{0} & \color{blue}{3} & 0 \\ 0 & 0 & 0 & 0 & \color{purple}{4} \end{pmatrix} \quad \begin{pmatrix} \color{red}{3} & \color{red}{1} & 0 & 0 & 0 \\ \color{red}{0} & \color{red}{3} & 0 & 0 & 0 \\ 0 & 0 & \color{blue}{2} & \color{blue}{1} & 0 \\ 0 & 0 & \color{blue}{0} & \color{blue}{2} & 0 \\ 0 & 0 & 0 & 0 & \color{purple}{2} \end{pmatrix} \quad \begin{pmatrix} \color{red}{2} & \color{red}{1} & 0 & 0 & 0 \\ \color{red}{0} & \color{red}{2} & 0 & 0 & 0 \\ 0 & 0 & \color{blue}{2} & \color{blue}{1} & 0 \\ 0 & 0 & \color{blue}{0} & \color{blue}{2} & 0 \\ 0 & 0 & 0 & 0 & \color{purple}{4} \end{pmatrix} \quad \begin{pmatrix} \color{red}{2} & \color{red}{1} & \color{red}{0} & \color{red}{0} & 0 \\ \color{red}{0} & \color{red}{2} & \color{red}{1} & \color{red}{0} & 0 \\ \color{red}{0} & \color{red}{0} & \color{red}{2} & \color{red}{1} & 0 \\ \color{red}{0} & \color{red}{0} & \color{red}{0} & \color{red}{2} & 0 \\ 0 & 0 & 0 & 0 & \color{purple}{2} \end{pmatrix}$$

Criterion for Diagonalizability

Theorem
Suppose $L : V \to V$ is a linear transformation of a finite dimensional vector space then L is diagonalizable if and only if V has a basis which consists of eigenvectors which occurs if and only if all roots of $\min_L(x)$ have multiplicity 1.
This follows from our work on the Jordan form

The easiest case to idenfity when this will happen is given by the following. (It is easy to identify because the characteristic polynomial is easier to compute than the minimial polynomial).

Theorem
If the roots of $\mathrm{char}_L(x)$ all have multiplicity 1, then so to do the roots of $\min_L(x)$ and hence L will be diagonalizable.
This follows from our work on the Jordan form

What invariants distinguish matricies?

You will never need to actually make use of this, though in principal it helps compute Jordan decompositions in the most general case.
The main theoretic result here is it tells you what extra invariants fully determine the Jordan form.

Theorem

If L is nilpotent, then the collection of numbers:

$$r_k = \mathrm{rank}(L^k)$$

determine the dimensions of the blocks A_i appearing.

Proof Idea

For $k > 0$ the difference between r_k and r_{k+1} tells you the exact number of the blocks (the matricies A_i) which have dimension at least k.
This is because every time you take a power of the standard unipotent block the rank drops by 1 until it reaches 0.

Theorem

The numbers r_k satisfy:

$$n = r_0 \geq r_1 \geq \cdots \geq r_n = 0$$

moreover

$$n \geq \mathrm{null}(L) = r_1 - r_0 \geq r_2 - r_1 \geq \cdots \geq r_n - r_{n+1} = 0$$

and any set of r_i satisfying the above can occur.
The first claim follows from the above, the converse is essentially by construction.

Natural Questions About Jordan Canonical Form

- How do I identify the Jordan canonical form of a matrix/linear transformation?
  You should know the process,
  for large matricies it is not typical to do it by hand
  You should expect either $d_\lambda = m_\lambda$ or $m_\lambda = 1$ for examples I have you do.
- How do I find a basis for the Jordan canonical form?
  As above, You should expect either $d_\lambda = m_\lambda$ or $m_\lambda = 1$ for examples I have you do.
- How can I identify the Characteristic/Minimal polynomial of a matrix in Jordan form?
  The connection to the sizes of the blocks is key
- How can I write down a matrix in Jordan form that has specific Characteristic/Minimal polynomial?
  The connection to the sizes of the blocks is key
- What information is hidden in the Jordan canonical form?
  This is open ended, a number of applications have been on the assignments.
- What information is hidden in the basis that realizes it?
  This is open ended, a number of applications have been on the assignments.