Lecture 3s

Real Canonical Form

(pages 420-423)

Okay, so our eigenvalue $\lambda = a + bi$ for A leads to an eigenvector $\vec{z} = \vec{x} + i\vec{y}$, and we now know that $\text{Span}\{\vec{x}, \vec{y}\}$ is an invariant subspace under A. What next? We will end up using the real vectors $\vec{x}$ and $\vec{y}$ to form a matrix (instead of the actual eigenvectors, as before), but to see why, we want to start by looking at the case when A is a 2×2 matrix. Why? Well, not only is $\text{Span}\{\vec{x}, \vec{y}\}$ a subspace of $\mathbb{R}^n$, but it is specifically a two-dimensional subspace of $\mathbb{R}^n$, with $\mathcal{B} = \{\vec{x}, \vec{y}\}$ as a basis. In the case when $n = 2$, the only two-dimensional subspace of $\mathbb{R}^2$ is $\mathbb{R}^2$ itself, so we get that the set $\mathcal{B}$ is a basis for $\mathbb{R}^2$. This means that the matrix $P = \begin{bmatrix} \vec{x} & \vec{y} \end{bmatrix}$ can be thought of as a change of coordinates matrix, from standard coordinates to $\mathcal{B}$ coordinates. And, thus, P is invertible, but moreover we know that $P^{-1}AP$ will be the matrix for the linear mapping $A\vec{r}$, but with respect to $\mathcal{B}$ coordinates.

To look at this further, let's step back a bit, and define $L : \mathbb{R}^2 \to \mathbb{R}^2$ by $L(\vec{r}) = A\vec{r}$. So $[L] = A$, and $[L]_{\mathcal{B}}$ is $\begin{bmatrix} [L(\vec{x})]_{\mathcal{B}} & [L(\vec{y})]_{\mathcal{B}} \end{bmatrix} = \begin{bmatrix} [A\vec{x}]_{\mathcal{B}} & [A\vec{y}]_{\mathcal{B}} \end{bmatrix}$. So, we need to find the $\mathcal{B}$ coordinates of $A\vec{x}$ and $A\vec{y}$. But, we can recall from our work in showing that $\text{Span}\mathcal{B}$ is an invariant subspace, that

$$A\vec{x} = a\vec{x} - b\vec{y} \qquad\qquad A\vec{y} = b\vec{x} + a\vec{y}$$

Then we see that $[A\vec{x}]_{\mathcal{B}} = \begin{bmatrix} a \\ -b \end{bmatrix}$ and $[A\vec{y}]_{\mathcal{B}} = \begin{bmatrix} b \\ a \end{bmatrix}$. And so we see that $[L]_{\mathcal{B}} = \begin{bmatrix} a & b \\ -b & a \end{bmatrix}$. But remember that $[L]_{\mathcal{B}}$ is the matrix for the linear mapping $A\vec{x}$ with respect to $\mathcal{B}$ coordinates, so we have that $[L]_{\mathcal{B}} = P^{-1}AP$. And so we have ended up with a situation similar to diagonalization: we use the eigenvectors to find an invertible matrix P, and $P^{-1}AP$ is a matrix built using the eigenvalues. Our matrix $\begin{bmatrix} a & b \\ -b & a \end{bmatrix}$ is known as a real canonical form for A.

<u>Definition</u>: Let A be a 2×2 real matrix with eigenvalue $\lambda = a + ib$, $b \neq 0$. The matrix $\begin{bmatrix} a & b \\ -b & a \end{bmatrix}$ is called a **real canonical form** for A.

Example: In lecture 3q, we found that $\lambda = 1 + 2i$ is an eigenvalue for $A = \begin{bmatrix} 3 & 4 \\ -2 & -1 \end{bmatrix}$, with corresponding eigenvector $\begin{bmatrix} -1-i \\ 1 \end{bmatrix} = \begin{bmatrix} -1 \\ 1 \end{bmatrix} + i \begin{bmatrix} -1 \\ 0 \end{bmatrix}$. So, in this case we have $a = 1$, $b = 2$, $\vec{x} = \begin{bmatrix} -1 \\ 1 \end{bmatrix}$ and $\vec{y} = \begin{bmatrix} -1 \\ 0 \end{bmatrix}$, so

we must have that $C = \begin{bmatrix} 1 & 2 \\ -2 & 1 \end{bmatrix}$ is a real canonical form for A, and that $P = \begin{bmatrix} -1 & -1 \\ 1 & 0 \end{bmatrix}$ is such that $P^{-1}AP = C$.

Note: You can easily calculate that $P^{-1} = \begin{bmatrix} 0 & 1 \\ -1 & -1 \end{bmatrix}$, and then compute the product $P^{-1}AP$ to verify that it is in fact $\begin{bmatrix} 1 & 2 \\ -2 & 1 \end{bmatrix}$. But I doubt that this verification would be very interesting in print, so I'll leave it to you to do that on your own.

So, things seem to work quite nicely in the 2×2 case, but we definitely don't have that $\{\vec{x}, \vec{y}\}$ is a basis for $\mathbb{R}^3$. However, we do know that $\{\vec{v}, \vec{x}, \vec{y}\}$ is a basis for $\mathbb{R}^3$, where $\vec{v}$ is the eigenvector for the real eigenvalue. For, if A is a 3×3 matrix, then its characteristic polynomial is a degree three polynomial, which must have at least one real root. In fact, since we know that complex roots of this polynomial will come in conjugate pairs, either A will have three real eigenvalues (counting multiplicity), or one real eigenvalue and two complex eigenvalues (that are conjugates). Having already looked at the case where A has only real eigenvalues, let's now see we what happens when A has one real eigenvalue μ with eigenvector $\vec{v}$, and complex eigenvalues $a \pm ib$ with eigenvectors $\vec{x} \pm i\vec{y}$. Theorem 9.4.2 still applies, so we know that $\text{Span}\{\vec{x}, \vec{y}\}$ is a two-dimensional subspace of $\mathbb{R}^3$. But this is where we make use of the fact that $\text{Span}\{\vec{x}, \vec{y}\}$ does not contain any real eigenvectors, and thus specifically does not contain $\vec{v}$. So, if we recall the technique for expanding a linearly independent set to a basis, we can start with the linearly independent set $\{\vec{x}, \vec{y}\}$, and add the vector $\vec{v} \notin \text{Span}\{\vec{x}, \vec{y}\}$, and we know that the resulting set $\{\vec{v}, \vec{x}, \vec{y}\}$ is linearly independent. And since we have a linearly independent set with three vectors, by the two-out-of-three rule, we know that this set is a basis for $\mathbb{R}^3$.

Since $\mathcal{B} = \{\vec{v}, \vec{x}, \vec{y}\}$ is a basis for $\mathbb{R}^3$, the matrix $P = \begin{bmatrix} \vec{v} & \vec{x} & \vec{y} \end{bmatrix}$ is the change of coordinates matrix from standard coordinates to $\mathcal{B}$ coordinates. And this means that we still have that $P^{-1}AP$ is the matrix for the linear mapping $A\vec{r}$ with respect to $\mathcal{B}$ coordinates. As we did in the 2×2 case, let's use our knowledge of A to figure out what $P^{-1}AP$ is.

So, let's let $L : \mathbb{R}^3 \to \mathbb{R}^3$ be defined by $L(\vec{r}) = A\vec{r}$, so that $[L] = A$. Then

$$\begin{aligned} [L]_{\mathcal{B}} &= \begin{bmatrix} [L(\vec{v})]_{\mathcal{B}} & [L(\vec{x})]_{\mathcal{B}} & [L(\vec{y})]_{\mathcal{B}} \end{bmatrix} \\ &= \begin{bmatrix} [A\vec{v}]_{\mathcal{B}} & [A\vec{x}]_{\mathcal{B}} & [A\vec{y}]_{\mathcal{B}} \end{bmatrix} \end{aligned}$$

So now we need to find $[A\vec{v}]_{\mathcal{B}}$, $[A\vec{x}]_{\mathcal{B}}$, and $[A\vec{y}]_{\mathcal{B}}$. We still know that $A\vec{x} = a\vec{x} - b\vec{y}$,

so $[A\vec{x}]_{\mathcal{B}} = \begin{bmatrix} 0 \\ a \\ -b \end{bmatrix}$, and $A\vec{y} = b\vec{x} + a\vec{y}$, so $[A\vec{y}]_{\mathcal{B}} = \begin{bmatrix} 0 \\ b \\ a \end{bmatrix}$. And since $A\vec{v} = \mu\vec{v}$, we see that $[A\vec{v}]_{\mathcal{B}} = \begin{bmatrix} \mu \\ 0 \\ 0 \end{bmatrix}$. So we have that

$$P^{-1}AP = \begin{bmatrix} \mu & 0 & 0 \\ 0 & a & b \\ 0 & -b & a \end{bmatrix}$$

And this is a real canonical form for a 3×3 matrix A with one real eigenvalue and two complex eigenvalues.

Example: In Lecture 3q, we found that the matrix $A = \begin{bmatrix} 3 & 0 & 1 \\ 2 & 1 & 2 \\ -1 & 0 & 3 \end{bmatrix}$ had eigenvalue 1 with corresponding eigenvector $\begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}$, and complex eigenvalues $3 \pm i$ with corresponding eigenvectors $\begin{bmatrix} \mp i \\ 2 \mp 6i \\ 5 \end{bmatrix} = \begin{bmatrix} 0 \\ 2 \\ 5 \end{bmatrix} \pm i \begin{bmatrix} -5 \\ -6 \\ 0 \end{bmatrix}$. Then we know that

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 3 & 1 \\ 0 & -1 & 3 \end{bmatrix}$$

is a real canonical form for A, and that the matrix

$$P = \begin{bmatrix} 0 & 0 & -5 \\ 1 & 2 & -6 \\ 0 & 5 & 0 \end{bmatrix}$$

is a change of coordinates matrix that can bring A into this form.

Once we get into larger matrices, we end up with more potential combinations of real and complex eigenvalues, including the possibility of repeated complex roots. The mathematics involved with these new complications are beyond the scope of this course, so we will only look at 2×2 and 3×3 matrices.