Lecture 3r

Invariant Subspace

(pages 419)

While its nice that we can diagonalize more matrices over $\mathbb{C}$ than we could over $\mathbb{R}$, when we do so we end up with a diagonal matrix with complex entries instead of real entries. There is clearly no magic solution to this problem, as if we could have diagonalized the matrix over $\mathbb{R}$ we would have. So, instead of expanding the numbering system we use in our matrices, what if we loosen the restriction that we need to end up with a diagonal matrix. Perhaps it is time we search for the next best thing. Well, if that's our conclusion, why did we bother with the complex numbers in the first place? Because we will still be making use of the complex eigenvalues and eigenvectors of our real matrix–simply in a different way.

The first thing we need to do is start thinking of our complex vectors as having a real and imaginary part, much as we think of a complex number as having a real and imaginary part. So, for any vector $\vec{z} \in \mathbb{C}^n$, there are vectors $\vec{x}$ and $\vec{y}$ in $\mathbb{R}^n$ such that

$$\vec{z} = \begin{bmatrix} z_1 \\ \vdots \\ z_n \end{bmatrix} = \begin{bmatrix} x_1 + y_1 i \\ \vdots \\ x_n + y_n i \end{bmatrix} = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} + i \begin{bmatrix} y_1 \\ \vdots \\ y_n \end{bmatrix} = \vec{x} + i\vec{y}$$

Next, we want to notice that if $\vec{z}$ is an eigenvector for a matrix A with real entries, with corresponding eigenvalue $\lambda = a + bi$, then we have $A\vec{z} = \lambda\vec{z}$, which means that

$$A\vec{z} = A(\vec{x} + i\vec{y}) = A\vec{x} + iA\vec{y}$$

AND

$$A\vec{z} = \lambda\vec{z} = (a + bi)(\vec{x} + i\vec{y}) = a\vec{x} + ai\vec{y} + bi\vec{x} + bi^2\vec{y} = (a\vec{x} - b\vec{y}) + i(b\vec{x} + a\vec{y})$$

So we see that $A\vec{x} + iA\vec{y} = (a\vec{x} - b\vec{y}) + i(b\vec{x} + a\vec{y})$. Now, since $a, b \in \mathbb{R}$ and $\vec{x}, \vec{y} \in \mathbb{R}^n$, the vectors $a\vec{x} - b\vec{y}$ and $b\vec{x} + a\vec{y}$ are in $\mathbb{R}^n$, which means that we have really have written $A\vec{z}$ as its real and imaginary parts in two ways. Since these ways must be the same, we get that

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

I know you were hoping we would get a result like $A\vec{x} = a\vec{x}$, but I did warn you that we were going to have to pull back from an ideal situation. If we look past this initial disappointment, we can still find something useful from this result. And that is the fact that both $A\vec{x}$ and $A\vec{y}$ are linear combinations of $\vec{x}$ and $\vec{y}$. That is to say, we have that $A\vec{x}, A\vec{y} \in \text{Span}\{\vec{x}, \vec{y}\}$. Moreover, since multiplication by A is linear, given any $\vec{w} \in \text{Span}\{\vec{x}, \vec{y}\}$, we get that $A\vec{w} \in \text{Span}\{\vec{x}, \vec{y}\}$. So, while the ideal situation would have been to find a real vector $\vec{x}$ such that $A\vec{x} \in \text{Span}\{\vec{x}\}$, instead we will be able to use our complex eigenvalues and eigenvectors to find a pair of real vectors such that for any $\vec{w} \in \text{Span}\{\vec{x}, \vec{y}\}$, we have $A\vec{w} \in \text{Span}\{\vec{x}, \vec{y}\}$. We will state this result as a theorem, but first we want to introduce a new term.

<u>Definition</u>: If $T : \mathbb{V} \to \mathbb{V}$ is a linear operator and $\mathbb{U}$ is a subspace of $\mathbb{V}$ such that $T(\mathbf{u}) \in \mathbb{U}$ for all $\mathbf{u} \in \mathbb{U}$, then $\mathbb{U}$ is called an **invariant subspace** of T.

<u>Theorem 9.4.2</u> Suppose that $\lambda = a + bi$, $b \neq 0$, is an eigenvalue of an $n \times n$ real matrix A with corresponding eigenvector $\vec{z} = \vec{x} + i\vec{y}$. Then $\text{Span}\{\vec{x}, \vec{y}\}$ is a two-dimensional subspace of $\mathbb{R}^n$ that is invariant under A and contains no real eigenvectors of A.

<u>Proof of Theorem 9.4.2</u>: We've already done the work to show that $\text{Span}\{\vec{x}, \vec{y}\}$ is invariant under A, but our theorem states two additional facts about $\text{Span}\{\vec{x}, \vec{y}\}$:

(1) $\text{Span}\{\vec{x}, \vec{y}\}$ is two-dimensional

(2) $\text{Span}\{\vec{x}, \vec{y}\}$ contains no real eigenvectors of A

To see (1), we will show that $\vec{x}$ and $\vec{y}$ are linearly independent, so that the set $\{\vec{x}, \vec{y}\}$ is actually a basis for $\text{Span}\{\vec{x}, \vec{y}\}$. We will prove this by contradiction, so assume (by way of contradiction) that the set $\{\vec{x}, \vec{y}\}$ is linearly dependent. Then we must have that $\vec{x} = s\vec{y}$ for some $s \in \mathbb{R}$.

From our earlier work, this would mean that $A\vec{x} = a\vec{x} - b\vec{y} = as\vec{y} - b\vec{y} = (as - b)\vec{y}$. But we also get that $A\vec{x} = A(s\vec{y}) = sA\vec{y} = s(b\vec{x} + a\vec{y}) = s(bs\vec{y} + a\vec{y}) = (bs^2 + as)\vec{y}$. Combining these facts, we get that $(as - b)\vec{y} = (bs^2 + as)\vec{y}$. These vectors are equal if either $as - b = bs^2 + as$, or if $\vec{y} = \vec{0}$. Let's examine the first possibility. If $as - b = bs^2 + as$, then we must have that $-b = bs^2$. Since $b \neq 0$ is one of the assumptions of our theorem, we can divide both sides by b, and get that $s^2 = -1$. But $s \in \mathbb{R}$, so this is not possible.

So we must have that $\vec{y} = \vec{0}$. But since $\vec{x} = s\vec{y}$, this would also mean that $\vec{x} = 0$, which means that our eigenvector was $\vec{0} + i\vec{0} = \vec{0}$. But $\vec{0}$ is not allowed to be an eigenvector, so we again have a contradiction.

And this means that our original assumption (that $\{\vec{x}, \vec{y}\}$ is linearly dependent) must be wrong. So we have that $\{\vec{x}, \vec{y}\}$ is linearly independent, as desired.

To see (2), we will again employ a proof by contradiction. So, assume (by way of contradiction) that there is a vector $\vec{w} \in \text{Span}\{\vec{x}, \vec{y}\}$ that is an eigenvector for A, with corresponding eigenvalue λ. The first thing we want to note is that $\lambda \in \mathbb{R}$, since $\vec{w} \in \mathbb{R}^n$ and $\lambda\vec{w} = A\vec{w} \in \mathbb{R}^n$. Next, since we've already shown that

$\{\vec{x}, \vec{y}\}$ is a basis for Span$\{\vec{x}, \vec{y}\}$, we know that there are scalars $s, t \in \mathbb{R}$ such that $\vec{w} = s\vec{x} + t\vec{y}$. So, we have that

$$A\vec{w} = \lambda\vec{w} = \lambda(s\vec{x} + t\vec{y}) = \lambda s\vec{x} + \lambda t\vec{y}$$

AND

$$A\vec{w} = A(s\vec{x}+t\vec{y}) = sA\vec{x}+tA\vec{y} = s(a\vec{x}-b\vec{y})+t(b\vec{x}+a\vec{y}) = (sa+tb)\vec{x}+(ta-sb)\vec{y}$$

So we see that $\lambda s\vec{x} + \lambda t\vec{y} = (sa + tb)\vec{x} + (ta - sb)\vec{y}$. Recalling that $\{\vec{x}, \vec{y}\}$ is a basis for Span$\{\vec{x}, \vec{y}\}$, we know that the elements of Span$\{\vec{x}, \vec{y}\}$ can be written UNIQUELY as a linear combination of $\vec{x}$ and $\vec{y}$. This means that we have

$$\lambda s = sa + tb \quad \text{and} \quad \lambda t = ta - sb$$

Solving both of these equations for λ, we get

$$\lambda = \frac{sa + tb}{s} = \frac{ta - sb}{t}$$

Cross multiplying gives us

$$tsa + t^2b = tsa - s^2b$$

This means that $t^2b = -s^2b$, and since $b \neq 0$, we can divide by b to get $t^2 = -s^2$. This statement can only be true of real numbers s and t if $s = t = 0$. But, as before, this means that our eigenvector is $\vec{0}$, which can never be an eigenvector. And so we have reached our contradiction. This means that our original assumption was incorrect, and so it must be that Span$\{\vec{x}, \vec{y}\}$ does not contain any real eigenvectors of A.