Math S-21b – Lecture #12 Notes

Today's lecture focuses on what might be called the **structural analysis of linear transformations**. What are the intrinsic properties of a linear transformation? Are there any fixed directions? The discussion centers on the eigenvalues and eigenvectors associated with an $n \times n$ matrix – the definitions, calculations, and applications.

Invariant directions, eigenvectors, and eigenvalues

Let **A** be an $n \times n$ matrix representing a linear transformation $T_{\mathbf{A}} : \mathbf{R}^n \to \mathbf{R}^n$. Are there any ***invariant directions*** for this linear transformation? That is, can we find a vector **v** such that $T(\mathbf{v})$ is parallel to **v**? This is an example of an intrinsic property of the transformation – something that exists independent of what basis is used or the coordinates relative to that basis. For example, a rotation in $\mathbf{R}^3$ has an axis of rotation regardless what basis is used to describe the rotation. For an orthogonal projection onto some subspace $V \subseteq \mathbf{R}^n$, vectors in *V* remain unchanged, and vectors in its orthogonal complement are sent to the zero vector. Again, this has nothing to do with what basis is used to represent this linear transformation.

The question of whether we find a vector **v** such that $T(\mathbf{v})$ is parallel to **v** can be rephrased as whether there's a vector **v** such that $T(\mathbf{v}) = \mathbf{A}\mathbf{v} = \lambda\mathbf{v}$ for some scalar λ. This leads to the following definition:

Definition: If **A** is an $n \times n$ matrix, we call a vector **v** an **eigenvector** of **A** if $T(\mathbf{v}) = \mathbf{A}\mathbf{v} = \lambda\mathbf{v}$ for some scalar λ. This scalar λ is called the eigenvalue associated with the eigenvector.

Finding the eigenvalues and eigenvectors

We can rewrite $\mathbf{A}\mathbf{v} = \lambda\mathbf{v}$ as $\mathbf{A}\mathbf{v} = \lambda\mathbf{I}\mathbf{v}$ which is more conducive to using algebra. We can then write this as $\lambda\mathbf{I}\mathbf{v} - \mathbf{A}\mathbf{v} = \mathbf{0}$ or $\boxed{(\lambda\mathbf{I} - \mathbf{A})\mathbf{v} = \mathbf{0}}$. In order for a vector **v** to be an eigenvector, it must be in the kernel of $\lambda\mathbf{I} - \mathbf{A}$ for some appropriate choice of λ. This can only happen if this kernel is nontrivial which means that the matrix $\lambda\mathbf{I} - \mathbf{A}$ would have to <u>not</u> be invertible, and we know from our discussion of determinants that a necessary and sufficient condition for a matrix to not be invertible is that its determinant must be equal to 0. That is:

$$\mathbf{v} \text{ is an eigenvector of } \mathbf{A} \iff (\lambda\mathbf{I} - \mathbf{A})\mathbf{v} = \mathbf{0} \iff \det(\lambda\mathbf{I} - \mathbf{A}) = 0$$

As we'll see, if **A** is an $n \times n$ matrix $p_A(\lambda) = \det(\lambda\mathbf{I} - \mathbf{A})$ will be an *n*th degree polynomial in λ called the **characteristic polynomial** of **A**. So $\lambda\mathbf{I} - \mathbf{A}$ will have a nontrivial kernel if and only if λ is a root of this characteristic polynomial. The eigenvalues are therefore the roots of the characteristic polynomial.

Definition: The set of all eigenvalues of a matrix **A** is called the **spectrum of A**. Since the eigenvalues are the roots of an nth degree polynomial, the spectrum will consist of at most *n* values. These may be real numbers or complex numbers, possibly with repetition, and any complex eigenvalues must occur in complex conjugate pairs. [This follows from the Fundamental Theorem of Algebra – any polynomial with real coefficients can, in theory, always be factored into a product of linear factors an irreducible quadratic factors, and these irreducible quadratic factors will yield complex conjugate pairs (by the quadratic formula).]

If an eigenvalue λ occurs as a repeated root of the characteristic polynomial, we refer to the multiplicity of the root as the **algebraic multiplicity** of the eigenvalue.

Definition: If λ is an eigenvalue of **A**, then $\ker(\lambda\mathbf{I} - \mathbf{A})$ will be a subspace called the **eigenspace** of λ, or E_λ. As with any subspace it must be closed under scaling and vector addition. This yields the following two corollaries:

Corollary 1: If **v** is an eigenvector associated with an eigenvalue λ, then $t\mathbf{v}$ will also be an eigenvector for any scalar *t*.

Corollary 2: If $\mathbf{v}_1$ and $\mathbf{v}_2$ are eigenvectors associated with the same eigenvalue λ, then $c_1\mathbf{v}_1 + c_2\mathbf{v}_2$ will also be an eigenvector for any scalars c_1, c_2.

revised 7/21/2014

Definition: The **geometric multiplicity** of an eigenvalue λ of a matrix $\mathbf{A}$ is $\dim[\ker(\lambda\mathbf{I}-\mathbf{A})]$, i.e. the number of linearly independent eigenvectors associated with this eigenvalue.

Example: Find the eigenvalues and eigenvectors of the matrix $\mathbf{A}=\begin{bmatrix}3 & 1\\ 1 & 3\end{bmatrix}$.

Solution: We calculate $\lambda\mathbf{I}-\mathbf{A}=\lambda\begin{bmatrix}1 & 0\\ 0 & 1\end{bmatrix}-\begin{bmatrix}3 & 1\\ 1 & 3\end{bmatrix}=\begin{bmatrix}\lambda & 0\\ 0 & \lambda\end{bmatrix}-\begin{bmatrix}3 & 1\\ 1 & 3\end{bmatrix}=\begin{bmatrix}\lambda-3 & -1\\ -1 & \lambda-3\end{bmatrix}$, so the characteristic polynomial is $p_A(\lambda)=\det(\lambda\mathbf{I}-\mathbf{A})=\det\begin{bmatrix}\lambda-3 & -1\\ -1 & \lambda-3\end{bmatrix}=(\lambda-3)^2-1=\lambda^2-6\lambda+8$. This is easily factored to give $p_A(\lambda)=(\lambda-4)(\lambda-2)=0$, so the eigenvalues are $\lambda_1=4$ and $\lambda_2=2$. How you order these doesn't matter, but you should keep the indexing consistent. For each eigenvalue, we next its eigenvectors, i.e. $\ker(\lambda_i\mathbf{I}-\mathbf{A})$ for each eigenvalue λ_i:

$\lambda_1=4$ gives $\lambda\mathbf{I}-\mathbf{A}=\begin{bmatrix}4-3 & -1\\ -1 & 4-3\end{bmatrix}=\begin{bmatrix}1 & -1\\ -1 & 1\end{bmatrix}$, so $\ker(\lambda\mathbf{I}-\mathbf{A})$ is found by row reduction $\left[\begin{array}{cc|c}1 & -1 & 0\\ -1 & 1 & 0\end{array}\right]\to\left[\begin{array}{cc|c}\boxed{1} & -1 & 0\\ 0 & 0 & 0\end{array}\right]$. This gives $\left\{\begin{matrix}x_1=t\\ x_2=t\end{matrix}\right\}$ or $\mathbf{x}=t\begin{bmatrix}1\\ 1\end{bmatrix}$, so if we let $\mathbf{v}_1=\begin{bmatrix}1\\ 1\end{bmatrix}$, this spans the eigenspace E_4.

$\lambda_1=2$ gives $\lambda\mathbf{I}-\mathbf{A}=\begin{bmatrix}-1 & -1\\ -1 & -1\end{bmatrix}$, so $\ker(\lambda\mathbf{I}-\mathbf{A})$ is found by row reduction $\left[\begin{array}{cc|c}-1 & -1 & 0\\ -1 & -1 & 0\end{array}\right]\to\left[\begin{array}{cc|c}\boxed{1} & 1 & 0\\ 0 & 0 & 0\end{array}\right]$. This gives $\left\{\begin{matrix}x_1=-t\\ x_2=t\end{matrix}\right\}$ or $\mathbf{x}=t\begin{bmatrix}-1\\ 1\end{bmatrix}$, so if we let $\mathbf{v}_2=\begin{bmatrix}-1\\ 1\end{bmatrix}$, this spans the eigenspace E_2.

In the example, we had two distinct, real eigenvalues which produced two linearly independent eigenvectors which may be used as a basis for $\mathbf{R}^2$, an ***eigenbasis***. What is the matrix of this linear transformation relative to the special basis? The relations $\left\{\begin{matrix}\mathbf{A}\mathbf{v}_1=4\mathbf{v}_1\\ \mathbf{A}\mathbf{v}_2=2\mathbf{v}_2\end{matrix}\right\}\Rightarrow[\mathbf{A}]_{\mathscr{B}}=\begin{bmatrix}4 & 0\\ 0 & 2\end{bmatrix}=\begin{bmatrix}\lambda_1 & 0\\ 0 & \lambda_2\end{bmatrix}=\mathbf{D}$, a diagonal matrix. If we write $\mathbf{S}=\begin{bmatrix}\uparrow & \uparrow\\ \mathbf{v}_1 & \mathbf{v}_2\\ \downarrow & \downarrow\end{bmatrix}=\begin{bmatrix}1 & -1\\ 1 & 1\end{bmatrix}$, then $\mathbf{S}^{-1}=\frac{1}{2}\begin{bmatrix}1 & 1\\ -1 & 1\end{bmatrix}$, and $[\mathbf{A}]_{\mathscr{B}}=\mathbf{S}^{-1}\mathbf{A}\mathbf{S}=\mathbf{D}$. This will be the case for any matrix for which we can produce an entire basis consisting exclusively of eigenvectors. This motivates the following:

Definition: An $n\times n$ matrix $\mathbf{A}$ is called **diagonalizable** if it is possible to find a basis for $\mathbf{R}^n$ consisting of eigenvectors of $\mathbf{A}$.

If $\mathbf{A}$ is diagonalizable with eigenbasis $\mathscr{B}=\{\mathbf{v}_1,\cdots,\mathbf{v}_n\}$ and if we write $\mathbf{S}=\begin{bmatrix}\uparrow & & \uparrow\\ \mathbf{v}_1 & \cdots & \mathbf{v}_n\\ \downarrow & & \downarrow\end{bmatrix}$, then

$$\left\{\begin{matrix}\mathbf{A}\mathbf{v}_1=\lambda_1\mathbf{v}_1\\ \vdots\\ \mathbf{A}\mathbf{v}_n=\lambda_n\mathbf{v}_n\end{matrix}\right\}\Rightarrow[\mathbf{A}]_{\mathscr{B}}=\mathbf{S}^{-1}\mathbf{A}\mathbf{S}=\mathbf{D}=\begin{bmatrix}\lambda_1 & & 0\\ & \ddots & \\ 0 & & \lambda_n\end{bmatrix}.$$

Note: It is not always possible to diagonalize a matrix. We want to understand under what circumstances this will be possible.

Powers of a matrix: If a matrix $\mathbf{A}$ is diagonalizable, we can write $[\mathbf{A}]_{\mathcal{B}} = \mathbf{S}^{-1}\mathbf{A}\mathbf{S} = \mathbf{D}$ for some change of basis matrix $\mathbf{S}$. Therefore $\mathbf{A} = \mathbf{S}\mathbf{D}\mathbf{S}^{-1}$ and $\mathbf{A}^t = (\mathbf{S}\mathbf{D}\mathbf{S}^{-1})(\mathbf{S}\mathbf{D}\mathbf{S}^{-1})\cdots(\mathbf{S}\mathbf{D}\mathbf{S}^{-1}) = \mathbf{S}\mathbf{D}^t\mathbf{S}^{-1}$ where, if $\mathbf{D} = \begin{bmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{bmatrix}$, we'll have $\mathbf{D}^t = \begin{bmatrix} \lambda_1^{\,t} & & 0 \\ & \ddots & \\ 0 & & \lambda_n^{\,t} \end{bmatrix}$.

Example: For the matrix $\mathbf{A} = \begin{bmatrix} 3 & 1 \\ 1 & 3 \end{bmatrix}$, calculate $\mathbf{A}^t$ for any (positive integer) *t*.

Solution: For this matrix we found that $\mathbf{S}^{-1}\mathbf{A}\mathbf{S} = \mathbf{D}$ where $\mathbf{S} = \begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix}$, $\mathbf{S}^{-1} = \frac{1}{2}\begin{bmatrix} 1 & 1 \\ -1 & 1 \end{bmatrix}$, and $\mathbf{D} = \begin{bmatrix} 4 & 0 \\ 0 & 2 \end{bmatrix}$.

So $\mathbf{A} = \mathbf{S}\mathbf{D}\mathbf{S}^{-1}$ and $\mathbf{A}^t = \mathbf{S}\mathbf{D}^t\mathbf{S}^{-1} = \frac{1}{2}\begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix}\begin{bmatrix} 4^t & 0 \\ 0 & 2^t \end{bmatrix}\begin{bmatrix} 1 & 1 \\ -1 & 1 \end{bmatrix} = \frac{1}{2}\begin{bmatrix} 4^t & -2^t \\ 4^t & 2^t \end{bmatrix}\begin{bmatrix} 1 & 1 \\ -1 & 1 \end{bmatrix} = \frac{1}{2}\begin{bmatrix} 4^t + 2^t & 4^t - 2^t \\ 4^t - 2^t & 4^t + 2^t \end{bmatrix}$.

Application: Markov example

There are situations in which a fixed amount of some asset is distributed among a number of sites and where some iterated process simultaneously redistributes the amounts to other sites according to fixed percentages. For example, suppose you had a fixed number of beans distributed between two piles, A and B. A process simultaneously moves 50% of the beans in pile A to pile B (while retaining 50% in pile A) and moves 75% of the beans in pile B to pile A (while retaining 25% in pile B). We can describe the transition as follows:

If x_A is the number of beans in pile A and x_B is the number of beans in pile B, then the new values will be determined by $\left\{\begin{matrix} \text{new } x_A = .5x_A + .75x_B \\ \text{new } x_B = .5x_A + .25x_B \end{matrix}\right\}$. That is, the new values are determined by applying the matrix $\mathbf{A} = \begin{bmatrix} .5 & .75 \\ .5 & .25 \end{bmatrix}$. If we think of $\mathbf{x}_0 = \begin{bmatrix} x_A \\ x_B \end{bmatrix}$ as the initial distribution, then after one iteration we'll have $\mathbf{x}_1 = \mathbf{A}\mathbf{x}_0$, after two iterations $\mathbf{x}_2 = \mathbf{A}\mathbf{x}_1 = \mathbf{A}^2\mathbf{x}_0$, and so on. After *t* iterations the distribution will be given by $\mathbf{x}_t = \mathbf{A}^t\mathbf{x}_0$. The ability to calculate powers of a matrix using eigenvalues and eigenvectors greatly simplifies the analysis.

In this case, we have $\mathbf{A} = \begin{bmatrix} .5 & .75 \\ .5 & .25 \end{bmatrix}$, $\lambda\mathbf{I} - \mathbf{A} = \begin{bmatrix} \lambda - .5 & -.75 \\ -.5 & \lambda - .25 \end{bmatrix}$,

$p_{\mathbf{A}}(\lambda) = \det[\lambda\mathbf{I} - \mathbf{A}] = \lambda^2 - .75\lambda - .25 = (\lambda - 1)(\lambda + .25)$ and the eigenvalues are $\lambda_1 = 1$ and $\lambda_2 = -.25$.

These yield the eigenvectors $\mathbf{v}_1 = \begin{bmatrix} 3 \\ 2 \end{bmatrix}$ and $\mathbf{v}_2 = \begin{bmatrix} 1 \\ -1 \end{bmatrix}$.

If we began with any configuration $\mathbf{x}_0 = \begin{bmatrix} x_A \\ x_B \end{bmatrix}$ and expressed this in terms of the basis of eigenvectors $\mathcal{B} = \{\mathbf{v}_1, \mathbf{v}_2\}$ as $\mathbf{x}_0 = c_1\mathbf{v}_1 + c_2\mathbf{v}_2$, then we would have $\mathbf{x}_1 = c_1\lambda_1\mathbf{v}_1 + c_2\lambda_2\mathbf{v}_2$, $\mathbf{x}_2 = c_1\lambda_1^{\,2}\mathbf{v}_1 + c_2\lambda_2^{\,2}\mathbf{v}_2$, etc. After *t* iterations we would get $\mathbf{x}_t = c_1\lambda_1^{\,t}\mathbf{v}_1 + c_2\lambda_2^{\,t}\mathbf{v}_2$. But with $\lambda_1 = 1$ and $\lambda_2 = -.25$ we see that $\lambda_1^{\,t} = 1$ for all *t* and $\lambda_2^{\,t} \to 0$, so eventually $\mathbf{x}_t \to c_1\mathbf{v}_1$. In practical terms, this simply means that the number of beans in each pile will eventually be proportional to the components of the eigenvector $\mathbf{v}_1 = \begin{bmatrix} 3 \\ 2 \end{bmatrix}$. For example, if we began with 1000 beans initially configured in any way, eventually we'll find the number of beans to be approaching 600 in pile A and 400 in pile B.

Example: Find the eigenvalues and eigenvectors of the matrix $\mathbf{A}=\begin{bmatrix}3&0&-2\\-7&0&4\\4&0&-3\end{bmatrix}$ and diagonalize this matrix, if possible, by finding a basis consisting of eigenvectors.

Solution: Before getting started, note that the column of 0's means that $\mathbf{Ae}_2=\mathbf{0}$, so $\mathbf{e}_2$ is actually an eigenvectors with eigenvalue $\lambda=0$. Indeed, the kernel of any $n\times n$ matrix is just the eigenspace E_0.

$$\lambda\mathbf{I}-\mathbf{A}=\begin{bmatrix}\lambda-3&0&2\\7&\lambda&-4\\-4&0&\lambda+3\end{bmatrix}, \text{ so } p_{\mathbf{A}}(\lambda)=\det[\lambda\mathbf{I}-\mathbf{A}]=(\lambda-3)(\lambda^2+3\lambda)+2(4\lambda)=\lambda^3-\lambda=\lambda(\lambda-1)(\lambda+1)=0.$$

This yields three distinct, real eigenvalues $\lambda_1=1$, $\lambda_2=0$ and $\lambda_3=-1$. (Order doesn't matter, but be consistent.)

$$\lambda_1=1\Rightarrow\left[\begin{array}{ccc|c}-2&0&2&0\\7&1&-4&0\\-4&0&4&0\end{array}\right]\Rightarrow\left[\begin{array}{ccc|c}\boxed{1}&0&-1&0\\0&\boxed{1}&3&0\\0&0&0&0\end{array}\right]\Rightarrow\left\{\begin{array}{c}x_1=t\\x_2=-3t\\x_3=t\end{array}\right\}\Rightarrow t\begin{bmatrix}1\\-3\\1\end{bmatrix}\Rightarrow\mathbf{v}_1=\begin{bmatrix}1\\-3\\1\end{bmatrix}$$

$$\lambda_2=0\Rightarrow\left[\begin{array}{ccc|c}-3&0&2&0\\7&0&-4&0\\-4&0&3&0\end{array}\right]\Rightarrow\left[\begin{array}{ccc|c}\boxed{1}&0&0&0\\0&0&\boxed{1}&0\\0&0&0&0\end{array}\right]\Rightarrow\left\{\begin{array}{c}x_1=0\\x_2=t\\x_3=0\end{array}\right\}\Rightarrow t\begin{bmatrix}0\\1\\0\end{bmatrix}\Rightarrow\mathbf{v}_2=\begin{bmatrix}0\\1\\0\end{bmatrix}$$

$$\lambda_3=-1\Rightarrow\left[\begin{array}{ccc|c}-4&0&2&0\\7&-1&-4&0\\-4&0&2&0\end{array}\right]\Rightarrow\left[\begin{array}{ccc|c}\boxed{1}&0&-\frac{1}{2}&0\\0&\boxed{1}&\frac{1}{2}&0\\0&0&0&0\end{array}\right]\Rightarrow\left\{\begin{array}{c}x_1=t\\x_2=-t\\x_3=2t\end{array}\right\}\Rightarrow t\begin{bmatrix}1\\-1\\2\end{bmatrix}\Rightarrow\mathbf{v}_3=\begin{bmatrix}1\\-1\\2\end{bmatrix}$$

Once again, we were fortunate to be able to produce a basis of eigenvectors $\mathcal{B}=\{\mathbf{v}_1,\mathbf{v}_2,\mathbf{v}_3\}$.

Theorem: Eigenvectors corresponding to distinct eigenvalues are linearly independent.

Proof: We prove this fact using an inductive argument in which each successive step uses the result of the previous step. For a finite set of eigenvalues, there will be a finite number of steps.

(1) If there is just one eigenvalue λ_1, then there must be a corresponding nonzero eigenvector $\mathbf{v}_1$. This is a linearly independent set.

(2) Suppose there are two distinct eigenvalues $\lambda_1\neq\lambda_2$ with corresponding eigenvectors $\{\mathbf{v}_1,\mathbf{v}_2\}$. We want to show that these must necessarily be linearly independent. To this end, let $c_1\mathbf{v}_1+c_2\mathbf{v}_2=\mathbf{0}$. If we multiply by the matrix **A**, we get $\mathbf{A}(c_1\mathbf{v}_1+c_2\mathbf{v}_2)=c_1\mathbf{Av}_1+c\mathbf{A}_2\mathbf{v}_2=c_1\lambda_1\mathbf{v}_1+c_2\lambda_2\mathbf{v}_2=\mathbf{A}(\mathbf{0})=\mathbf{0}$. The original relation $c_1\mathbf{v}_1+c_2\mathbf{v}_2=\mathbf{0}$ gives that $c_2\mathbf{v}_2=-c_1\mathbf{v}_1$, so $c_1\lambda_1\mathbf{v}_1+\lambda_2(-c_1\mathbf{v}_1)=c_1(\lambda_1-\lambda_2)\mathbf{v}_1=\mathbf{0}$. Because $\lambda_1\neq\lambda_2$ and $\mathbf{v}_1\neq\mathbf{0}$, we must have $c_1=0$. But therefore $c_2\mathbf{v}_2=\mathbf{0}$, so necessarily $c_2=0$. Therefore $\{\mathbf{v}_1,\mathbf{v}_2\}$ are linearly independent.

(3) Suppose $\lambda_1,\lambda_2,\lambda_3$ are distinct eigenvalues (hence $\lambda_1\neq\lambda_2,\lambda_1\neq\lambda_3,\lambda_2\neq\lambda_3$), with corresponding eigenvectors $\{\mathbf{v}_1,\mathbf{v}_2,\mathbf{v}_3\}$. Once again, we write $c_1\mathbf{v}_1+c_2\mathbf{v}_2+c_3\mathbf{v}_3=\mathbf{0}$. Multiplication by **A** gives $c_1\mathbf{Av}_1+c_2\mathbf{Av}_2+c_3\mathbf{Av}_3=c_1\lambda_1\mathbf{v}_1+c_2\lambda_2\mathbf{v}_2+c_3\lambda_3\mathbf{v}_3=\mathbf{0}$, and the original relation allows us to solve for $c_3\mathbf{v}_3=-c_1\mathbf{v}_1-c_2\mathbf{v}_2$. Substitution gives $c_1\lambda_1\mathbf{v}_1+c_2\lambda_2\mathbf{v}_2+\lambda_3(-c_1\mathbf{v}_1-c_2\mathbf{v}_2)=c_1(\lambda_1-\lambda_3)\mathbf{v}_1+c_2(\lambda_2-\lambda_3)\mathbf{v}_2=\mathbf{0}$. The previous step established the linear independence of $\{\mathbf{v}_1,\mathbf{v}_2\}$, so necessarily $c_1(\lambda_1-\lambda_3)=0$ and

$c_2(\lambda_2 - \lambda_3) = 0$. Because the eigenvalues are all distinct, this implies that $c_1 = 0$ and $c_2 = 0$. Therefore $c_3\mathbf{v}_3 = \mathbf{0}$, so $c_3 = 0$ as well. So $\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$ are linearly independent.

The argument continues in the same fashion so that if $\lambda_1, \cdots, \lambda_k$ are distinct eigenvalues with corresponding eigenvectors $\{\mathbf{v}_1, \cdots, \mathbf{v}_k\}$, these must be linearly independent.

Corollary: If **A** is an $n \times n$ matrix with distinct, real eigenvalues, then **A** is diagonalizable.
Proof: If the roots of the nth degree characteristic polynomial are $\lambda_1, \cdots, \lambda_n$, each will yield a corresponding eigenvector so we'll have a collection $\{\mathbf{v}_1, \cdots, \mathbf{v}_n\}$ of linearly independent eigenvectors. This will constitute a basis for $\mathbf{R}^n$, so the matrix **A** will be diagonalizable.

Note: This means that for a matrix **A** to fail to be diagonalizable, its spectrum must contain either repeated eigenvalues, complex eigenvalues, or possibly both. However, it is quite possible for a matrix with repeated eigenvalues to still be diagonalizable. The best example is the $n \times n$ identity matrix which has only the eigenvalue 1 but this eigenvalue has algebraic multiplicity *n*. The identity matrix is clearly diagonalizable because it's already diagonal! All vectors are eigenvectors of the identity matrix.

Example: If we compare the three matrices $\mathbf{A} = \begin{bmatrix} 2 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 2 \end{bmatrix}$, $\mathbf{B} = \begin{bmatrix} 2 & 1 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 2 \end{bmatrix}$, and $\mathbf{C} = \begin{bmatrix} 2 & 1 & 0 \\ 0 & 2 & 1 \\ 0 & 0 & 2 \end{bmatrix}$, we'll see that they each have the same characteristic polynomial $p(\lambda) = (\lambda - 2)^3$, so they each have just the one eigenvalue $\lambda = 2$ with algebraic multiplicity 3. However, a quick calculation with each of these matrices reveals that the geometric multiplicity of **A** is 3 (every vector is an eigenvector), the geometric multiplicity of **B** is 2, and the geometric multiplicity of **C** is 1. Neither matrix **B** nor matrix **C** is diagonalizable.

Notes by Robert Winters