MA110:
Lecture 11

Saurav Bhaumik
Department of Mathematics
IIT Bombay

Spring 2025

Similarity of Square Matrices

Definition

*Let $\mathbf{A}, \mathbf{B} \in \mathbb{K}^{n\times n}$. We say that $\mathbf{A}$ is **similar** to $\mathbf{B}$ (over $\mathbb{K}$) if there is an invertible $\mathbf{P} \in \mathbb{K}^{n\times n}$ such that $\mathbf{B} = \mathbf{P}^{-1}\mathbf{A}\mathbf{P}$, that is, $\mathbf{A}\mathbf{P} = \mathbf{P}\mathbf{B}$. In this case, we write $\mathbf{A} \sim \mathbf{B}$.*

One can easily check (i) $\mathbf{A} \sim \mathbf{A}$, (ii) if $\mathbf{A} \sim \mathbf{B}$ then $\mathbf{B} \sim \mathbf{A}$, and (iii) if $\mathbf{A} \sim \mathbf{B}$ and $\mathbf{B} \sim \mathbf{C}$, then $\mathbf{A} \sim \mathbf{C}$.

Examples:

(i) Let $\mathbf{A} := \begin{bmatrix} 4 & -3 \\ 2 & -1 \end{bmatrix}$. It is easily seen that $\mathbf{P} := \begin{bmatrix} 3 & 1 \\ 2 & 1 \end{bmatrix}$ is invertible and . $\mathbf{P}^{-1} = \begin{bmatrix} 1 & -1 \\ -2 & 3 \end{bmatrix}$. Hence $\mathbf{A}$ is similar to

$$\mathbf{B} := \begin{bmatrix} 1 & -1 \\ -2 & 3 \end{bmatrix} \begin{bmatrix} 4 & -3 \\ 2 & -1 \end{bmatrix} \begin{bmatrix} 3 & 1 \\ 2 & 1 \end{bmatrix} = \begin{bmatrix} 1 & -1 \\ -2 & 3 \end{bmatrix} \begin{bmatrix} 6 & 1 \\ 4 & 1 \end{bmatrix} = \begin{bmatrix} 2 & 0 \\ 0 & 1 \end{bmatrix},$$

which is a diagonal matrix.

More Examples and a Characterization of Similarity

(ii) Let $\mathbf{A} \in \mathbb{K}^{n\times n}$. Then $\mathbf{A} \sim \mathbf{I} \iff \mathbf{A} = \mathbf{I}$.

(iii) Let $\mathbf{A} \in \mathbb{K}^{n\times n}$, and let $\mathbf{E}$ be $n\times n$ an elementary matrix. Then $\mathbf{B} := \mathbf{EAE}^{-1}$ is similar to $\mathbf{A}$. Note: $\mathbf{EA}$ is obtained from $\mathbf{A}$ by an elementary row operation on $\mathbf{A}$, and $\mathbf{B}$ is obtained from $\mathbf{EA}$ by the 'reverse column operation' on $\mathbf{EA}$.

Diagonalizability

Definition

A matrix $\mathbf{A} \in \mathbb{K}^{n\times n}$ *is called* **diagonalizable** *(over* $\mathbb{K}$*) if* $\mathbf{A}$ *is similar to a diagonal matrix (over* $\mathbb{K}$*).*

Examples

(i) Let $\mathbf{A} := \begin{bmatrix} 4 & -3 \\ 2 & -1 \end{bmatrix}$. Let $\mathbf{P} := \begin{bmatrix} 3 & 1 \\ 2 & 1 \end{bmatrix}$, so that $\mathbf{P}^{-1} := \begin{bmatrix} 1 & -1 \\ -2 & 3 \end{bmatrix}$. Then $\mathbf{A}$ is similar to the matrix

$$\mathbf{B} := \begin{bmatrix} 1 & -1 \\ -2 & 3 \end{bmatrix} \begin{bmatrix} 4 & -3 \\ 2 & -1 \end{bmatrix} \begin{bmatrix} 3 & 1 \\ 2 & 1 \end{bmatrix} = \begin{bmatrix} 1 & -1 \\ -2 & 3 \end{bmatrix} \begin{bmatrix} 6 & 1 \\ 4 & 1 \end{bmatrix} = \begin{bmatrix} 2 & 0 \\ 0 & 1 \end{bmatrix},$$

which is a diagonal matrix.

(ii) Let $\mathbf{A} \in \mathbb{K}^{n \times n}$. Then $\mathbf{A} \sim \mathbf{I} \iff \mathbf{A} = \mathbf{I}$.

(iii) Let $\mathbf{A} \in \mathbb{K}^{n \times n}$, and let $\mathbf{E}$ be $n \times n$ an elementary matrix. Then $\mathbf{B} := \mathbf{E}\mathbf{A}\mathbf{E}^{-1}$ is similar to $\mathbf{A}$. Note: $\mathbf{EA}$ is obtained from $\mathbf{A}$ by an elementary row operation on $\mathbf{A}$, and $\mathbf{B}$ is obtained from $\mathbf{EA}$ by the 'reverse column operation' on $\mathbf{EA}$.

Similarity of matrices has the following characterisation.

Proposition

Let $\mathbf{A}, \mathbf{B} \in \mathbb{K}^{n\times n}$. Then $\mathbf{A} \sim \mathbf{B}$ if and only if there is an ordered basis E for $\mathbb{K}^{n\times 1}$ such that $\mathbf{B}$ is the matrix of the linear transformation $T_{\mathbf{A}} : \mathbb{K}^{n\times 1} \to \mathbb{K}^{n\times 1}$ with respect to E.

In fact, $\mathbf{B} = \mathbf{P}^{-1}\mathbf{AP}$ if and only if the columns of $\mathbf{P}$ form an ordered basis, say E, for $\mathbb{K}^{n\times 1}$ and $\mathbf{B} = \mathbf{M}_E^E(T_{\mathbf{A}})$.

Proof. Let $\mathbf{B} := [b_{jk}]$. Now $\mathbf{A} \sim \mathbf{B} \iff$ there is an invertible matrix $\mathbf{P}$ such that $\mathbf{AP} = \mathbf{PB}$. Let $\mathbf{x}_k$ be the k-th column of P. Then $E := (\mathbf{x}_1, \ldots, \mathbf{x}_n)$ is an ordered basis for $\mathbb{K}^{n\times 1}$. Conversely, if $E = (\mathbf{x}_1, \ldots, \mathbf{x}_n)$ is an ordered basis then the matrix $P = [\mathbf{x_1}, \ldots, \mathbf{x_n}]$ is invertible. Now $\mathbf{AP} = \mathbf{PB}$ implies:

$$\mathbf{A}\begin{bmatrix}\mathbf{x}_1 & \cdots & \mathbf{x}_n\end{bmatrix} = \begin{bmatrix}\mathbf{x}_1 & \cdots & \mathbf{x}_n\end{bmatrix}\begin{bmatrix} b_{11} & \cdots & b_{1n} \\ \vdots & \vdots & \vdots \\ b_{n1} & \cdots & b_{nn}\end{bmatrix}.$$

The kth column of LHS is $\mathbf{A}\mathbf{x}_k$ and the kth column of RHS is the linear combination of $\mathbf{x}_1, \ldots, \mathbf{x}_n$ with coefficients from the k column of $\mathbf{B}$. Thus $\mathbf{A}\mathbf{x}_k = b_{1k}\mathbf{x}_1 + \cdots + b_{nk}\mathbf{x}_n$ for $k = 1, \ldots, n$. This means the kth column of $\mathbf{M}_E^E(T_\mathbf{A})$ is the kth column $\begin{bmatrix} b_{1k} & \cdots & b_{nk} \end{bmatrix}^\mathsf{T}$ of $\mathbf{B}$, $k = 1, \ldots, n$, that is, $\mathbf{B} = \mathbf{M}_E^E(T_\mathbf{A})$.

Conversely, if $\mathbf{B} = \mathbf{M}_E^E(T_\mathbf{A})$ then again $\mathbf{A}\mathbf{x}_k = T_A(\mathbf{x_k}) = b_{1k}\mathbf{x}_1 + \cdots + b_{nk}\mathbf{x}_n$ for $k = 1, \ldots, n$, so $\mathbf{AP} = \mathbf{PB}$ where $P = [\mathbf{x_1}, \ldots, \mathbf{x_n}]$. $\square$

Proposition

A matrix $\mathbf{A} \in \mathbb{K}^{n\times n}$ is diagonalizable if and only if there is a basis for $\mathbb{K}^{n\times 1}$ consisting of eigenvectors of $\mathbf{A}$.

In fact, $\mathbf{X}^{-1}\mathbf{A}\mathbf{X} = \mathbf{D}$, where $\mathbf{X} := \begin{bmatrix}\mathbf{x}_1 & \cdots & \mathbf{x}_n\end{bmatrix}$ and $\mathbf{D} := \mathrm{diag}(\lambda_1, \ldots, \lambda_n) \iff \{\mathbf{x}_1, \ldots, \mathbf{x}_n\}$ is a basis for $\mathbb{K}^{n\times 1}$ and $\mathbf{A}\mathbf{x}_k = \lambda_k \mathbf{x}_k$ for $k = 1, \ldots, n$.

Proof. $\mathbf{A}$ is diagonalizable $\iff$ there is an invertible matrix $\mathbf{X}$ and a diagonal matrix $\mathbf{D}$ such that $\mathbf{A}\mathbf{X} = \mathbf{X}\mathbf{D}$. This is the case if and only if there is a basis $\{\mathbf{x}_1, \ldots, \mathbf{x}_n\}$ for $\mathbb{K}^{n\times 1}$ and there are $\lambda_1, \ldots, \lambda_n \in \mathbb{K}$ such that

$$\mathbf{A}\begin{bmatrix}\mathbf{x}_1 & \cdots & \mathbf{x}_n\end{bmatrix} = \begin{bmatrix}\mathbf{x}_1 & \cdots & \mathbf{x}_n\end{bmatrix}\mathrm{diag}(\lambda_1, \ldots, \lambda_n),$$

The LHS is just $[\mathbf{A}\mathbf{x_1}, \cdots, \mathbf{A}\mathbf{x_n}]$ and the RHS is just $[\lambda\mathbf{x}_1, \cdots, \lambda_n\mathbf{x}_n]$. Thus the proposition follows. $\square$

If $\mathbf{A} = \mathbf{X}\mathbf{D}\mathbf{X}^{-1}$, then for any $m \in \mathbb{N}$,

$$\mathbf{A}^m = (\mathbf{X}\mathbf{D}\mathbf{X}^{-1}) \cdots (\mathbf{X}\mathbf{D}\mathbf{X}^{-1}) = \mathbf{X}\mathbf{D}^m\mathbf{X}^{-1} = \mathbf{X}\,\mathrm{diag}(\lambda_1^m, \ldots, \lambda_n^m)\mathbf{X}^{-1}.$$

Example

We have seen that

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

and $\begin{bmatrix} 3 & 1 \\ 2 & 1 \end{bmatrix}^{-1} = \begin{bmatrix} 1 & -1 \\ -2 & 3 \end{bmatrix}$. Hence

$$\begin{aligned} \mathbf{A}^m &= \begin{bmatrix} 3 & 1 \\ 2 & 1 \end{bmatrix} \begin{bmatrix} 2^m & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & -1 \\ -2 & 3 \end{bmatrix} = \begin{bmatrix} 2^m 3 & 1 \\ 2^m 2 & 1 \end{bmatrix} \begin{bmatrix} 1 & -1 \\ -2 & 3 \end{bmatrix} \\ &= \begin{bmatrix} 2^m 3 - 2 & -2^m 3 + 3 \\ 2^m 2 - 2 & -2^m 2 + 3 \end{bmatrix} \quad \text{for } m \in \mathbb{N}. \end{aligned}$$

Similarity and Eigenvalues

Recall that $\lambda \in \mathbb{K}$ is an eigenvalue of $\mathbf{A} \in \mathbb{K}^{n\times n}$ if $\mathbf{A}\mathbf{x} = \lambda\mathbf{x}$ for some $\mathbf{x} \in \mathbb{K}^{n\times 1}$ with $\mathbf{x} \neq \mathbf{0}$. It turns out that similar matrices have the same eigenvalues. In fact, more is true.

Proposition

Let $\mathbf{A}$, $\mathbf{A}' \in \mathbb{K}^{n\times n}$ be similar. Then $p_{\mathbf{A}}(t) = p_{\mathbf{A}'}(t)$. In particular, $\lambda \in \mathbb{K}$ is an eigenvalue of $\mathbf{A}$ if and only if λ is an eigenvalue of $\mathbf{A}'$. Consequently, the algebraic multiplicity of λ as an eigenvalue of $\mathbf{A}$ is equal to the algebraic multiplicity of λ as an eigenvalue of $\mathbf{A}'$. Furthermore, the geometric multiplicity of λ as an eigenvalue of $\mathbf{A}$ is equal to the geometric multiplicity of λ as an eigenvalue of $\mathbf{A}'$.

Proof: SInce $\mathbf{A} \sim \mathbf{A}'$, there is an invertible $\mathbf{P} \in \mathbb{K}^{n\times n}$ such that $\mathbf{A}' = \mathbf{P}^{-1}\mathbf{A}\mathbf{P}$. Writing $\mathbf{I} = \mathbf{P}^{-1}\mathbf{P}$, we see that

$$p_{\mathbf{A}'}(t) = \det(\mathbf{A}' - t\mathbf{I}) = \det(\mathbf{P}^{-1}\mathbf{A}\mathbf{P} - t\mathbf{I}) = \det(\mathbf{A} - t\mathbf{I}) = p_{\mathbf{A}}(t).$$

Proof Contd. It remains to prove the assertion about geometric multiplicity. Let λ be an eigenvalue of $\mathbf{A}$ and $\mathbf{x}$ be an eigenvector of $\mathbf{A}$ corresponding to λ. Then $\mathbf{x} \neq \mathbf{0}$ and $\mathbf{A}\mathbf{x} = \lambda\mathbf{x}$. Since $\mathbf{A}' = \mathbf{P}^{-1}\mathbf{A}\mathbf{P}$, we see that $\mathbf{x}' := \mathbf{P}^{-1}\mathbf{x}$ satisfies $\mathbf{A}'\mathbf{x}' = \lambda\mathbf{x}'$. Also $\mathbf{x}' \neq \mathbf{0}$ since $\mathbf{P}$ is invertible. Thus $\mathbf{x}'$ is an eigevector of $\mathbf{A}'$ corresponding to λ. Also, it is easy to check that $\{\mathbf{x}_1, \ldots, \mathbf{x}_g\}$ is a basis for $\mathcal{N}(\mathbf{A} - \lambda\mathbf{I})$ if and only if $\{\mathbf{P}^{-1}\mathbf{x}_1, \ldots, \mathbf{P}^{-1}\mathbf{x}_g\}$ is a basis for $\mathcal{N}(\mathbf{A}' - \lambda\mathbf{I})$. Hence the geometric multiplicity of λ as an eigenvalue of $\mathbf{A}$ is equal to the geometric multiplicity of λ as an eigenvalue of $\mathbf{A}'$. $\square$

Example

$$\mathbf{A} := \begin{bmatrix} \lambda & 1 \\ 0 & \lambda \end{bmatrix} \Rightarrow p_{\mathbf{A}}(t) = \det \begin{bmatrix} \lambda - t & 1 \\ 0 & \lambda - t \end{bmatrix} = (\lambda - t)^2.$$

Hence λ is the only eigenvalue of **A**, and its algebraic multiplicity is 2. But its geometric multiplicity is 1 since

$$\mathbf{A} - \lambda \mathbf{I} = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} \implies \operatorname{rank}(\mathbf{A} - \lambda \mathbf{I}) = 1 \implies \operatorname{nullity}(\mathbf{A} - \lambda \mathbf{I}) = 1.$$

Note that in the above example, if **A** were similar to a diagonal matrix **D**, then we must have $\mathbf{D} = \operatorname{diag}(\lambda, \lambda)$, since eigenvalues and their algebraic multiplicities of **A** and **D** have to be the same. But the geometric multiplicity of λ as an eigenvalue of **D** is 2. This shows that **A** is not diagonalizable.

Example

$$\mathbf{A} := \begin{bmatrix} 3 & 0 & 0 \\ -2 & 4 & 2 \\ -2 & 1 & 5 \end{bmatrix} \Rightarrow p_{\mathbf{A}}(t) = \det \begin{bmatrix} 3-t & 0 & 0 \\ -2 & 4-t & 2 \\ -2 & 1 & 5-t \end{bmatrix}.$$

Computing the determinant, we find $p_{\mathbf{A}}(t) = (3-t)^2(6-t)$. Hence 3 is an eigenvalue of $\mathbf{A}$ of algebraic multiplicity 2, and 6 is an eigenvalue of $\mathbf{A}$ of algebraic multiplicity 1. Also,

$$\mathbf{A} - 3\mathbf{I} = \begin{bmatrix} 0 & 0 & 0 \\ -2 & 1 & 2 \\ -2 & 1 & 2 \end{bmatrix} \Rightarrow \operatorname{rank}(\mathbf{A} - 3\mathbf{I}) = 1.$$

So nullity$(\mathbf{A} - 3\mathbf{I}) = 2$. In fact, $\left\{ \begin{bmatrix} 1 & 0 & 1 \end{bmatrix}^{\mathsf{T}}, \begin{bmatrix} 1 & 2 & 0 \end{bmatrix}^{\mathsf{T}} \right\}$ is a basis of the eigenspace of $\mathbf{A}$ corresponding to eigenvalue 3, and so its geometric multiplicity is equal to 2.

Relating geometric and algebraic multiplicities

Proposition

Let $\mathbf{A} \in \mathbb{K}^{n \times n}$, and let λ be an eigenvalue of $\mathbf{A}$. Then the geometric multiplicity of λ is less than or equal to its algebraic multiplicity.

Proof. Let g be the geometric multiplicity of λ. Let $(\mathbf{v}_1, \ldots, \mathbf{v}_g)$ be an ordered basis of the eigenspace of λ; extend it to an ordered basis $(\mathbf{v}_1, \ldots, \mathbf{v}_g, \mathbf{v}_{g+1}, \ldots, \mathbf{v}_n)$ of $\mathbb{K}^{n \times 1}$. Define $\mathbf{P} := \begin{bmatrix} \mathbf{v}_1 & \cdots & \mathbf{v}_n \end{bmatrix}$. Then $\mathbf{P}$ is invertible since its n columns $\mathbf{v}_1, \ldots, \mathbf{v}_n$ are linearly independent. Consider $\mathbf{A}' := \mathbf{P}^{-1}\mathbf{A}\mathbf{P}$. Since $\mathbf{A}\mathbf{v}_j = \lambda \mathbf{v}_j$ and $\mathbf{P}\mathbf{e}_j = \mathbf{v}_j$ for $j = 1, \ldots, g$, we see that the j th column of $\mathbf{A}'$ is given by

$$\mathbf{A}'\mathbf{e}_j = \mathbf{P}^{-1}\mathbf{A}\mathbf{P}\mathbf{e}_j = \mathbf{P}^{-1}\mathbf{A}\mathbf{v}_j = \lambda \mathbf{P}^{-1}\mathbf{v}_j = \lambda \mathbf{e}_j.$$

Hence

$$\mathbf{A}' = \left[\begin{array}{ccc|c} \lambda & \cdots & 0 & \\ \vdots & \ddots & \vdots & \mathbf{C} \\ 0 & \cdots & \lambda & \\ \hline & \mathbf{O} & & \mathbf{D} \end{array}\right]$$

where $\mathbf{C} \in \mathbb{K}^{g\times(n-g)}$, $\mathbf{O} \in \mathbb{K}^{(n-g)\times g}$ and $\mathbf{D} \in \mathbb{K}^{(n-g)\times(n-g)}$. Expanding by the first column, we see that

$$\det(\mathbf{A}' - t\mathbf{I}) = (\lambda - t)^g q(t),$$

where $q(t)$ is a polynomial of degree $n - g$. Thus

$$p_{\mathbf{A}}(t) = p_{\mathbf{A}'}(t) = \det(\mathbf{A}' - t\mathbf{I}) = (\lambda - t)^g q(t).$$

Thus $(\lambda - t)^g$ divides the characteristic polynomial $p_{\mathbf{A}}(t)$ of $\mathbf{A}$. Since the algebraic multiplicity of λ is the largest natural number m such that $(\lambda - t)^m$ divides $p_{\mathbf{A}}(t)$, we obtain $g \leq m$. $\square$

Eigenvectors corresponding to distinct eigenvalues

Our next result is about the linear independence of eigenvectors corresponding to distinct eigenvalues of a matrix.

Lemma

Let $\mathbf{A} \in \mathbb{K}^{n \times n}$, and let $\lambda_1, \ldots, \lambda_k$ be distinct eigenvalues of $\mathbf{A}$. Let $\mathbf{x}_1, \ldots, \mathbf{x}_k$ belong to the eigenspaces of $\mathbf{A}$ corresponding to $\lambda_1, \ldots, \lambda_k$ respectively. Then

$$\mathbf{x}_1 + \cdots + \mathbf{x}_k = \mathbf{0} \iff \mathbf{x}_1 = \cdots = \mathbf{x}_k = \mathbf{0}.$$

In particular, if $\mathbf{x}_1, \ldots, \mathbf{x}_k$ are eigenvectors of $\mathbf{A}$ corresponding to $\lambda_1, \ldots, \lambda_k$ respectively, then the set $\{\mathbf{x}_1, \ldots, \mathbf{x}_k\}$ is linearly independent.

Proof. We use induction on the number k of distinct eigenvalues of $\mathbf{A}$. Clearly, the result holds for $k = 1$.

Let $k \geq 2$ and assume that the result holds for $k - 1$.

Suppose $\mathbf{x} := \mathbf{x}_1 + \cdots + \mathbf{x}_{k-1} + \mathbf{x}_k = \mathbf{0}$. Then $\mathbf{A}\mathbf{x} = \mathbf{0}$, that is, $\lambda_1 \mathbf{x}_1 + \cdots + \lambda_{k-1}\mathbf{x}_{k-1} + \lambda_k \mathbf{x}_k = \mathbf{0}$. Also, multiplying the first equation by λ_k, we obtain $\lambda_k \mathbf{x}_1 + \cdots + \lambda_k \mathbf{x}_{k-1} + \lambda_k \mathbf{x}_k = \mathbf{0}$. Subtraction gives $(\lambda_1 - \lambda_k)\mathbf{x}_1 + \cdots + (\lambda_{k-1} - \lambda_k)\mathbf{x}_{k-1} = \mathbf{0}$.

By the induction hypothesis, $(\lambda_1 - \lambda_k)\mathbf{x}_1 = \cdots = (\lambda_{k-1} - \lambda_k)\mathbf{x}_{k-1} = \mathbf{0}$. Since $\lambda_1 \neq \lambda_k, \ldots, \lambda_{k-1} \neq \lambda_k$, we obtain $\mathbf{x}_1 = \cdots = \mathbf{x}_{k-1} = \mathbf{0}$, and so $\mathbf{x}_k = \mathbf{0}$ as well.

Now let $\mathbf{x}_1, \ldots, \mathbf{x}_k$ be eigenvectors. If $\alpha_1 \mathbf{x}_1 + \cdots + \alpha_k \mathbf{x}_k = \mathbf{0}$, then $\alpha_1 \mathbf{x}_1 = \cdots = \alpha_k \mathbf{x}_k = \mathbf{0}$. But $\mathbf{x}_1 \neq \mathbf{0}, \ldots, \mathbf{x}_k \neq \mathbf{0}$, so that $\alpha_1 = \cdots = \alpha_k = 0$. Thus $\{\mathbf{x}_1, \ldots, \mathbf{x}_k\}$ is linearly independent.

Theorem

Let $\mathbf{A} \in \mathbb{K}^{n \times n}$, and let $\lambda_1, \ldots, \lambda_k$ be the distinct eigenvalues of $\mathbf{A}$. Let g_j be the geometric multiplicity of λ_j for $j = 1, \ldots, k$. Then $g_1 + \cdots + g_k \leq n$. Further, $\mathbf{A}$ is diagonalizable if and only if $g_1 + \cdots + g_k = n$.

Proof. Let $p_A(t) = \prod_{j=1}^{k}(\lambda_j - t)^{n_j}$. Then n_j is the algebraic multiplicity of λ_j. We have already seen that $g_j \leq n_j$. Thus,

$$\sum_{j=1}^{k} g_j \leq \sum_{j=1}^{k} n_j = n.$$

If A is diagonalizable, that is, similar to a diagonal matrix D, then the distinct eigenvalues of D are exactly $\lambda_1, \ldots, \lambda_k$ and for each $j = 1, \ldots, k$, the geometric and algebraic multiplicities of λ_j are same for A and D. However as D is a diagonal matrix, the algebraic and geometric multiplicities for each eigenvalue are same, hence $g_j = n_j$ for each j, thus $\sum_{j=1}^{n} g_j = n$.

Conversely, suppose $\sum_{j=1}^{n} g_j = n$. Let V_j denote the eigenspace $\mathcal{N}(\mathbf{A} - \lambda_j \mathbf{I})$ of $\mathbb{K}^{n \times 1}$, and let E_j be a basis for V_j consisting of g_j eigenvectors of $\mathbf{A}$ corresponding to λ_j for $j = 1, \ldots, k$.

Proof contd.

We claim that the set $E := E_1 \cup \cdots \cup E_k$ is linearly independent. Let $\mathbf{x}$ be a linear combination of elements of E. Collate the elements of E_j for each $j = 1, \ldots, k$ and write $\mathbf{x} = \mathbf{x}_1 + \cdots + \mathbf{x}_k$, where $\mathbf{x}_j \in V_j$ for $j = 1, \ldots, k$. Suppose $\mathbf{x} = \mathbf{0}$. Then $\mathbf{x}_j = \mathbf{0}$ for $j = 1, \ldots, k$ by the previous lemma. For $j \in \{1, \ldots, k\}$, $\mathbf{x}_j$ is a linear combination of elements of the set E_j, and since the set E_j is linearly independent, every coefficient in this linear combination must be 0. Since this holds for each $j = 1, \ldots, k$, we see that every coefficient in the linear combination $\mathbf{x}$ of elements of E must be 0.

Now E_j has g_j many elements, so E has $\sum_j g_j = n$ many elements. As the dimension of $\mathbb{K}^n$ is n, E must be a basis. As every element of E is an eigenvector of A, by results already proved, A is diagonalizable. $\square$