Similarity and Diagonalisation

Similarity:

Given two matrices A and B, if there is an invertible matrix S such that $\boxed{A = SBS^{-1}}$, A and B are similar.
It follows that $AS = SB$, which means that S transforming B gives the same result as A transforming S.

In real numbers, $as = sb$ means $a = b$. We cannot same the same for matrices as their multiplication is not commutative, but $AS = SB$ hints that A and B while not equal would share some common properties.
e.g. Their determinants are equal, so they are both invertable or both not; They are both diagonalisable or both not; Their eigenvalues are the same (with different eigenspaces connected through S).

If $A \sim B$, $A = SBS^{-1}$, $B = S^{-1}AS = (S^{-1})A(S^{-1})^{-1}$, then $B \sim A$ (commutative law).
If $A \sim B$ and $B \sim C$, $AS_1 = S_1B$, $BS_2 = S_2C$, $A(S_1S_2) = S_1BS_2 = S_1S_2C = (S_1S_2)C$, then $A \sim C$ (transitive law).
Because $A = IAI^{-1}$, $A \sim A$ (reflexive law).

If λ is an eigenvalue of A, then $\lambda I - A = \lambda I - SBS^{-1} = \lambda SIS^{-1} - SBS^{-1} = S(\lambda I - B)S^{-1}$.
$\therefore \det(\lambda I - B) = \det(S(\lambda I - B)S^{-1}) = \det(\lambda I - A) = 0$. So λ is an eigenvalue of B, and vice versa.

If $\mathbf{v}$ is an eigenvector of B, $B\mathbf{v} = \lambda\mathbf{v}$, $AS\mathbf{v} = SB\mathbf{v} = \lambda S\mathbf{v}$, $A(S\mathbf{v}) = \lambda(S\mathbf{v})$, i.e. $S\mathbf{v}$ is an eigenvector of A, and vice versa.

Diagonisation: If a matrix A is similar to a diagonal matrix D, then A is said to be diagnoisable. Since matrix similarity is transitive, all matrices similar to a diagonal matrix D form a "family" of mutually similar matrices.

Let $D = \begin{bmatrix} d_1 & 0 & \dots & 0 \\ 0 & d_2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & d_n \end{bmatrix}$, $\lambda I - D = \begin{bmatrix} \lambda - d_1 & 0 & \dots & 0 \\ 0 & \lambda - d_2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \lambda - d_n \end{bmatrix}$, $\det(\lambda I - D) = \prod_{k=1}^{n}(\lambda - d_k)$.

It is obvious that any d_r $(r = 1, 2, \dots, n)$ are eigenvalues of D, as $\det(d_rI - D) = \prod_{k=1}^{n}(d_r - d_k) = 0$.

If A is an invertible matrix ($\det(A) \neq 0$) and is similar to D, the eigenvalues of A are d_r and they are all distinct.
Let $P = \begin{bmatrix}\mathbf{v_1 v_2} \dots \mathbf{v_n}\end{bmatrix}$ where $\mathbf{v_r}$ are A's eigenvectors, which must be mutually independent. i.e. $\det(P) \neq 0$ and P^{-1} exists, and $A = PDP^{-1}$. $A \sim D$ means $A = SDS^{-1}$, which means $S = P$ and its columns are eigenvectors.

You may take similarity as "having the same eigenvalues". If A and B have the same eigenvalue matrix D, $A = P_ADP_A^{-1}$ and $B = P_BDP_B^{-1}$, and there is a matrix S such that $P_A = SP_B$, then $P_A^{-1} = P_B^{-1}S^{-1}$ and $A = P_ADP_A^{-1} = (SP_B)D(P_B^{-1}S^{-1}) = S(P_BDP_B^{-1})S^{-1} = SBS^{-1}$. After all, A and B have the same characteristic polynomial because $\det(xI - A) = \det(xI - B)$. The roots of the polynomial are eigenvalues of both A and B. So they must have the same eigenvalues.

Not all matrices are diagonalisable. But if a matrix has n eigenvectors (including identical ones), it is diagonalisable. In fact, as long as P is invertible (i.e. there are n mutually independent eigenvectors, A does not need to be invertible (i.e. eigenvalues distinct). If there are identical eigenvalues, D is not invertible, neither is A, but it may still have n mutually independent eigenvectors, so P^{-1} exists and therefore is diagonalisable (i.e. similar to D).