SIMILAR MATRICES

1. Similar Matrices

Fix a linear transformation

$$T : \mathbb{R}^n \to \mathbb{R}^n$$

and an ordered basis, $\mathcal{B} = (\vec{v}_1, \ldots, \vec{v}_n)$, of $\mathbb{R}^n$. The standard matrix $[T]$ and the matrix of T with respect to $\mathcal{B}$, $[T]_{\mathcal{B}}$, are related by

$$[T]S = S[T]_{\mathcal{B}} \text{ and } [T] = S[T]_{\mathcal{B}}S^{-1} \text{ and } [T]_{\mathcal{B}} = S^{-1}[T]S.$$

where here

$$S = \begin{bmatrix} \vec{v}_1 & | & \cdots & | & \vec{v}_n \end{bmatrix}$$

is the change of basis matrix of the basis. In order to understand this relationship better, it is convenient to take it as a definition and then study it abstractly.

Given two $n \times n$ matrices, A and B, we say that A is *similar* to B if there exists an invertible $n \times n$ matrix, S, so that

$$AS = SB.$$

Observe, this is equivalently, to

$$A = SBS^{-1} \text{ or } B = S^{-1}AS.$$

EXAMPLE: If A is similar to I_n, then $A = I_n$.

EXAMPLE: $A = \begin{bmatrix} 2 & -3 \\ 1 & -2 \end{bmatrix}$ is similar $B = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}$. To see this, let

$$S = \begin{bmatrix} 3 & 1 \\ 1 & 1 \end{bmatrix},$$

and compute

$$AS = \begin{bmatrix} 3 & -1 \\ 1 & -1 \end{bmatrix} = SB.$$

EXAMPLE: If A is similar to B, then A^2 is similar to B^2. To see this observe, that, by definition, $A = SBS^{-1}$ for some invertible S. Hence,

$$A^2 = (SBS^{-1})(SBS^{-1}) = SB(S^{-1}S)BS^{-1} = SBI_nBS^{-1} = SB^2S^{-1}$$

and so A^2 is similar to B^2.

EXAMPLE: If A is similar to B and one is invertible, then both are and A^{-1} is similar to B^{-1}. Indeed, suppose B is invertible, then $A = SBS^{-1}$ for invertible S and so A is also invertible as it is the product of three invertible matrices. A similar argument shows that B is invertible if A is. Finally, one has

$$A^{-1} = (SBS^{-1})^{-1} = (S^{-1})^{-1}B^{-1}S^{-1} = SB^{-1}S^{-1},$$

which completes the proof. In general, A^t is similar to B^t for any integer t (need A, B invertible when $t < 0$).

EXAMPLE: $A = \begin{bmatrix} 2 & -3 \\ 1 & -2 \end{bmatrix}$ has $A^{100} = I_2$. This is because A is similar to $B = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}$. As B is diagonal,

$$B^{100} = \begin{bmatrix} 1^{100} & 0 \\ 0 & (-1)^{100} \end{bmatrix} = I_2,$$

and so A is similar to I_2 and hence is equal to I_2. Similarly,

$$A^{101} = SB^{101}S^{-1} = SBS^{-1} = A.$$

2. Properties of Similar matrices

Being similar is a *equivalence relation*. That is:

(1) (Reflexive): A is similar to A.
(2) (Symmetric): A is similar to $B \iff B$ is similar to A
(3) (Transitive): if A is similar to B and B is similar to C, then A is similar to C.

This is relatively straightforward to check:

(1) Follows by taking $S = I_n$.
(2) Follows by observing $AS = SB$ means $BS^{-1} = S^{-1}A$, i.e. $BS' = S'A$ for $S' = S^{-1}$.
(3) Follows by observing $AS = SB$ and $BT = TC$ means that $AST = SBT = STC$ and so A is similar to C by using $S' = ST$. Here we used that S' is the product of two invertible matrices and so is invertible.

It is not easy, in general, to tell whether two matrices are similar and this is a question we will return to later in the class. It can be easy to tell when they are *not* similar.

Theorem 2.1. *If A and B are similar, then* $\mathrm{null}(A) = \mathrm{null}(B)$ *(and so* $\mathrm{rank}(A) = \mathrm{rank}(B)$*).*

Proof. As A is similar to B, we have $AS = SB$ for some invertible S. Now suppose that $\vec{z} \in \ker(B)$. It is clear that $S\vec{z} \in \ker(A)$. Indeed,

$$A(S\vec{z}) = (AS)\vec{z} = (SB)\vec{z} = S(B\vec{z}) = S\vec{0} = \vec{0}.$$

Moreover, if $\vec{z}_1, \ldots, \vec{z}_p$ are a basis of $\ker(B)$, then $S\vec{z}_1, \ldots, S\vec{z}_p$ are linearly independent. Indeed,

$$c_1S\vec{z}_1 + \cdots + c_pS\vec{z}_p = \vec{0} \Rightarrow S(c_1\vec{z}_1 + \cdots c_p\vec{z}_p)$$

Hence,

$$c_1\vec{z}_1 + \cdots c_p\vec{z}_p \in \ker(S) \Rightarrow c_1\vec{z}_1 + \cdots c_p\vec{z}_p = \vec{0}.$$

This must be a trivial linear relation (as $z_1, \ldots, z_p$ are a basis) which shows the only linear relations among $S\vec{z}_1, \ldots, S\vec{z}_p$ is the trivial one which proves the claim.

Hence, we have p linearly independent vectors in $\ker(A)$ and so

$$\mathrm{null}(A) = \dim\ker(A) \geq p = \dim(\ker(B)) = \mathrm{null}(B).$$

As this argument is symmetric in A and B we conclude also that $\mathrm{null}(B) \geq \mathrm{null}(A)$ which proves the result. Finally, by the rank-nullity theorem

$$\mathrm{rank}(A) = n - \mathrm{null}(A) = n - \mathrm{null}(B) = \mathrm{rank}(B)$$

where n is the number of columns in A and B. □

EXAMPLE: $\begin{bmatrix} 1 & -1 \\ 2 & -2 \end{bmatrix}$ is not similar to $\begin{bmatrix} 1 & 2 \\ 0 & -1 \end{bmatrix}$. By inspection, the first matrix has rank $= 1$ and second has rank $= 2$.

3. Diagonal Matrices

A matrix is *diagonal* if its only non-zero entries are on the diagonal. For instance,

$$B = \begin{bmatrix} k_1 & 0 & 0 \\ 0 & k_2 & 0 \\ 0 & 0 & k_3 \end{bmatrix},$$

is a 3×3 diagonal matrix. Geometrically, a diagonal matrix acts by "stretching" each of the standard vectors. Algebraically, this means B is diagonal if and only if

$$B\vec{e}_i = k_i\vec{e}_i$$

for each standard vector $\vec{e}_i$. As we have seen diagonal matrices and matrices that are similar to diagonal matrices are extremely useful for computing large powers of the matrix. As such, it is natural to ask when a given matrix is similar to a diagonal matrix.

We have the following complete answer:

Theorem 3.1. *A matrix A is similar to a diagonal matrix if and only if there is an ordered basis $\mathcal{B} = (\vec{v}_1, \ldots, \vec{v}_n)$ so that*

$$A\vec{v}_i = k_i\vec{v}_i$$

for some $k_i \in \mathbb{R}$.

That is A stretches the $\vec{v}_i$ by a factor k_i. It is worth mentioning that the $\vec{v}_i$ are examples of *eigenvectors of* A (a topic we will study later).

Proof. Suppose first that A is similar to a diagonal matrix B. That is, $AS = SB$ for a diagonal matrix B and invertible matrix S. As B is diagonal $B\vec{e}_i = k_i\vec{e}_i$ where k_i is the ith entry on the diagonal.

If $\vec{v}_i = S\vec{e}_i$, then $\mathcal{B} = (\vec{v}_1, \ldots, \vec{v}_n)$ is an ordered basis. Moreover,

$$A\vec{v}_i = AS\vec{e}_i = SB\vec{e}_i = S(B\vec{e}_i) = S(k_i\vec{e}_i) = k_iS\vec{e}_i = k_i\vec{v}_i.$$

This verifies one direction of the theorem.

Conversely, if one has the ordered basis $\mathcal{B} = (\vec{v}_1, \ldots, \vec{v}_n)$ so that $A\vec{v}_i = k_i\vec{v}_i$, and S is the change of basis matrix of $\mathcal{B}$, then

$$(S^{-1}AS)\vec{e}_i = S^{-1}(A(S\vec{e}_i)) = S^{-1}(A(\vec{v}_i)) = S^{-1}(k_i\vec{v}_i) = k_iS^{-1}\vec{v}_i = k_i\vec{e}_i.$$

Hence, $B = S^{-1}AS$ is diagonal and is similar to A by definition. $\square$

This may be rephrased in terms of linear transformations as follows:

Theorem 3.2. *A linear transformation $T : \mathbb{R}^n \to \mathbb{R}^n$ given by $T(\vec{x}) = A\vec{x}$ has a basis $\mathcal{B} = (\vec{v}_1, \ldots, \vec{v}_n)$ so that $[T]_{\mathcal{B}}$ is diagonal if and only if $T(\vec{v}_i) = \lambda_i\vec{v}_i$.*

Proof. Take $A = [T]$ and $B = [T]_{\mathcal{B}}$ and apply the preceding theorem. $\square$

EXAMPLE: Let $R_{\frac{\pi}{2}} : \mathbb{R}^2 \to \mathbb{R}^2$ be rotation counter clockwise by $90°$. For any $\vec{v} \neq \vec{0}$, $R_{\frac{\pi}{2}}(\vec{v})$ is perpendicular to $\vec{v}$, so there are no non-zero vectors that get stretched. Hence, there is no basis $\mathcal{B}$ for which $[R_{\frac{\pi}{2}}]_{\mathcal{B}}$ is diagonal.

EXAMPLE: Let $P : \mathbb{R}^2 \to \mathbb{R}^2$ be projection onto the line $y = x$. This means

$$P\left(\begin{bmatrix} 1 \\ 1 \end{bmatrix}\right) = \begin{bmatrix} 1 \\ 1 \end{bmatrix} \text{ and } P\left(\begin{bmatrix} 1 \\ -1 \end{bmatrix}\right) = \begin{bmatrix} 0 \\ 0 \end{bmatrix}.$$

Hence, if

$$\mathcal{B} = \left(\begin{bmatrix} 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ -1 \end{bmatrix}\right),$$

then

$$[P]_{\mathcal{B}} = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}.$$

As a consequence, if

$$S = \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}$$

is the change of basis matrix of $\mathcal{B}$, then

$$[P] = S[P]_{\mathcal{B}}S^{-1} = S \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} S^{-1}.$$