Similarity of Matrices

Prerequisites:

- Matrix Representations of Linear Transformations

Definition 1: *Similarity, (Conjugacy)*

Suppose that $\boxed{\mathsf{A}}$ and $\widetilde{\boxed{\mathsf{A}}}$ are two $N \times N$ matrices. We say that $\boxed{\mathsf{A}}$ and $\widetilde{\boxed{\mathsf{A}}}$ are **similar** (or **conjugate**), and write

$$\boxed{\mathsf{A}} \sim \widetilde{\boxed{\mathsf{A}}}$$

if there is an invertible $N \times N$ matrix $\boxed{\mathsf{B}}$ so that

$$\widetilde{\boxed{\mathsf{A}}} = \boxed{\mathsf{B}}^{-1} \cdot \boxed{\mathsf{A}} \cdot \boxed{\mathsf{B}}$$

Example 2: $\boxed{\mathrm{A}} = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}$ is similar to $\widetilde{\boxed{\mathrm{A}}} = \begin{bmatrix} 5 & 1 \\ 2 & 0 \end{bmatrix}$. To see this, let $\boxed{\mathrm{B}} = \begin{bmatrix} \frac{1}{2} & \frac{-1}{2} \\ \frac{1}{2} & \frac{1}{2} \end{bmatrix}$. Then $\boxed{\mathrm{B}}^{-1} = \begin{bmatrix} 1 & 1 \\ -1 & 1 \end{bmatrix}$. Thus,

$$\begin{aligned}
\boxed{\mathrm{B}}^{-1} \cdot \boxed{\mathrm{A}} \cdot \boxed{\mathrm{B}} &= \begin{bmatrix} 1 & 1 \\ -1 & 1 \end{bmatrix} \cdot \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix} \cdot \begin{bmatrix} \frac{1}{2} & \frac{-1}{2} \\ \frac{1}{2} & \frac{1}{2} \end{bmatrix} \\
&= \begin{bmatrix} 1 & 1 \\ -1 & 1 \end{bmatrix} \cdot \begin{bmatrix} \frac{3}{2} & \frac{1}{2} \\ \frac{7}{2} & \frac{1}{2} \end{bmatrix} \\
&= \begin{bmatrix} 5 & 1 \\ 2 & 0 \end{bmatrix} \\
&= \widetilde{\boxed{\mathrm{A}}}
\end{aligned}$$

Remark 3:

1. In other words, $\boxed{\mathrm{A}} \sim \widetilde{\boxed{\mathrm{A}}}$ if there is an invertible $N \times N$ matrix $\boxed{\mathrm{B}}$ so that

$$\boxed{\mathrm{B}} \cdot \widetilde{\boxed{\mathrm{A}}} = \boxed{\mathrm{A}} \cdot \boxed{\mathrm{B}}$$

2. Notice that

- Any matrix $\boxed{\mathrm{A}}$ is similar to itself, because

$$\boxed{\mathrm{A}} = \boxed{\mathbf{Id}}^{-1} \cdot \boxed{\mathrm{A}} \cdot \boxed{\mathbf{Id}}$$

- If $\boxed{\mathrm{A}} \sim \widetilde{\boxed{\mathrm{A}}}$, then $\widetilde{\boxed{\mathrm{A}}} \sim \boxed{\mathrm{A}}$, because

$$\begin{aligned} \text{if } \widetilde{\boxed{\mathrm{A}}} &= \boxed{\mathrm{B}}^{-1} \cdot \boxed{\mathrm{A}} \cdot \boxed{\mathrm{B}} \\ \text{then } \boxed{\mathrm{B}} \cdot \widetilde{\boxed{\mathrm{A}}} \cdot \boxed{\mathrm{B}}^{-1} &= \boxed{\mathrm{A}} \end{aligned}$$

- If $\boxed{\mathrm{X}} \sim \boxed{\mathrm{Y}}$ and $\boxed{\mathrm{Y}} \sim \boxed{\mathrm{Z}}$, then $\boxed{\mathrm{X}} \sim \boxed{\mathrm{Z}}$:

$$\begin{aligned} \text{if } \boxed{\mathrm{Y}} &= \boxed{\mathrm{B}}^{-1} \cdot \boxed{\mathrm{X}} \cdot \boxed{\mathrm{B}} \\ \text{and } \boxed{\mathrm{Z}} &= \boxed{\mathrm{C}}^{-1} \cdot \boxed{\mathrm{Y}} \cdot \boxed{\mathrm{C}} \\ \text{then } \boxed{\mathrm{Z}} &= \boxed{\mathrm{C}}^{-1} \cdot \boxed{\mathrm{B}}^{-1} \boxed{\mathrm{X}} \cdot \boxed{\mathrm{B}} \cdot \boxed{\mathrm{C}} \\ &= \left(\boxed{\mathrm{B}} \cdot \boxed{\mathrm{C}}\right)^{-1} \boxed{\mathrm{X}} \cdot \left(\boxed{\mathrm{B}} \cdot \boxed{\mathrm{C}}\right). \end{aligned}$$

Matrix similarity is important because of its relationship to the *representation* of **linear transformations....**

Proposition 4: *Similarity of Matrix Representations for Linear Transformations*

Let $\mathbb{V}$ be a finite-dimensional vector spaces, and $f : \mathbb{V} \longrightarrow \mathbb{V}$ a linear transformation.

1. Suppose that $\mathcal{A}$ and $\widetilde{\mathcal{A}}$ are two bases for $\mathbb{V}$. Let $\boxed{\mathsf{F}}$ be the matrix representation of f with respect to $\mathcal{A}$, and $\widetilde{\boxed{\mathsf{F}}}$ be the matrix representation of f with respect to $\widetilde{\mathcal{A}}$. Then $\boxed{\mathsf{F}}$ and $\widetilde{\boxed{\mathsf{F}}}$ are similar matrices.
2. Suppose $\mathcal{A}$ is any basis of $\mathbb{V}$, and $\boxed{\mathsf{F}}$ is the matrix of f with respect to $\mathcal{A}$. If $\widetilde{\boxed{\mathsf{F}}}$ is *any* matrix similar to $\boxed{\mathsf{F}}$, then there is a basis $\widetilde{\mathcal{A}}$ for $\mathbb{V}$ so that $\widetilde{\boxed{\mathsf{F}}}$ is the matrix of f with respect to $\widetilde{\mathcal{A}}$.

Proof:

Proof of Part 1: Recall: if $\boxed{\mathrm{B}}$ is the **change-of-basis** matrix from $\mathcal{A}$ to $\widetilde{\mathcal{A}}$, then

$$\boxed{\widetilde{\mathrm{F}}} = \boxed{\mathrm{B}} \cdot \boxed{\mathrm{F}} \cdot \boxed{\mathrm{B}}^{-1}$$

is the matrix representation of f with respect to $\mathcal{B}$. Thus, $\boxed{\widetilde{\mathrm{F}}}$ and $\boxed{\mathrm{F}}$ are similar

Proof of Part 2: Suppose that $\boxed{\widetilde{\mathrm{F}}} = \boxed{\mathrm{B}} \cdot \boxed{\mathrm{F}} \cdot \boxed{\mathrm{B}}^{-1}$ for some matrix $\boxed{\mathrm{B}}$. Thus, we want to find a basis $\widetilde{\mathcal{A}}$ so that that $\boxed{\mathrm{B}}$ is the change-of-basis matrix from $\mathcal{A}$ to $\widetilde{\mathcal{A}}$. Thus, $\boxed{\mathrm{B}}^{-1}$ should be the change-of-basis matrix from $\widetilde{\mathcal{A}}$ to $\mathcal{A}$. Thus means that the column-vectors of $\boxed{\mathrm{B}}^{-1}$ should be the *coordinates* of the elements of $\widetilde{\mathcal{A}}$, relative to $\mathcal{A}$.

So, suppose $\mathcal{A} = \{\mathbf{a}_1, \ldots, \mathbf{a}_N\}$, and suppose $\boxed{\mathrm{B}}^{-1} = \begin{bmatrix} c_{11} & c_{12} & \cdots & c_{1N} \\ c_{21} & c_{22} & \cdots & c_{2N} \\ \vdots & \vdots & \ddots & \vdots \\ c_{N1} & c_{N2} & \cdots & c_{NN} \end{bmatrix}$,

then, for all $k \in [1...N]$, simply define

$$\widetilde{\mathbf{a}}_k = \sum_{n=1}^{N} c_{nk} \mathbf{a}_n$$

and let $\widetilde{\mathcal{A}} = \{\widetilde{\mathbf{a}}_1, \ldots, \widetilde{\mathbf{a}}_N\}$. Then $\widetilde{\mathcal{A}}$ is the basis we seek.

□ [Proposition 4]

Similarity Invariants

In general, it is quite difficult to tell if two matrices $\boxed{\mathrm{A}}$ and $\boxed{\widetilde{\mathrm{A}}}$ are similar. We need a matrix $\boxed{\mathrm{B}}$ so that $\boxed{\mathrm{B}} \cdot \boxed{\widetilde{\mathrm{A}}} = \boxed{\mathrm{A}} \cdot \boxed{\mathrm{B}}$. But how can we find such a $\boxed{\mathrm{B}}$, assuming it exists? We might spend a long time looking, only to realize that the two matrices are *not* similar; thus, there *is* no such $\boxed{\mathrm{B}}$, and we are wasting our time.

It would be nice if there was a quick way to tell when two matrices are *not* similar. This is the purpose of a *similarity invariant*.

Definition 5: *Similarity Invariant*

A **similarity invariant** is a function $f : \mathcal{M}_{N \times N} \longrightarrow \mathbb{S}$ (where $\mathbb{S}$ is some set) so that, for any matrices $\boxed{\mathrm{A}}$ and $\boxed{\widetilde{\mathrm{A}}}$,

$$\left(\boxed{\mathrm{A}} \text{ and } \boxed{\widetilde{\mathrm{A}}} \text{ are similar} \right) \Longrightarrow \left(f\left(\boxed{\mathrm{A}}\right) = f\left(\boxed{\widetilde{\mathrm{A}}}\right) \right)$$

Hence, if $f\left(\boxed{\mathrm{A}}\right) \neq f\left(\widetilde{\boxed{\mathrm{A}}}\right)$ then we know right away that $\boxed{\mathrm{A}}$ and $\widetilde{\boxed{\mathrm{A}}}$ are *not* similar; there is no point looking for $\boxed{\mathrm{B}}$.

Note: If $f\left(\boxed{\mathrm{A}}\right) = f\left(\widetilde{\boxed{\mathrm{A}}}\right)$, this does *not* automatically mean that $\boxed{\mathrm{A}}$ and $\widetilde{\boxed{\mathrm{A}}}$ are similar. It only means that they *might* be similar.

Theorem 6: *Some Similarity Invariants*
Suppose $\boxed{\mathrm{A}}$ and $\widetilde{\boxed{\mathrm{A}}}$ are $N \times N$ matrices. If $\boxed{\mathrm{A}}$ is **similar** to $\widetilde{\boxed{\mathrm{A}}}$, then:

1. $\mathrm{rank}\left[\boxed{\mathrm{A}}\right] = \mathrm{rank}\left[\widetilde{\boxed{\mathrm{A}}}\right]$.
2. $\mathrm{nullity}\left[\boxed{\mathrm{A}}\right] = \mathrm{nullity}\left[\widetilde{\boxed{\mathrm{A}}}\right]$.
3. $\det\left(\boxed{\mathrm{A}}\right) = \det\left(\widetilde{\boxed{\mathrm{A}}}\right)$.
4. $\boxed{\mathrm{A}}$ and $\widetilde{\boxed{\mathrm{A}}}$ have the same **characteristic polynomial**.

In other words, the **rank**, **nullity**, **determinant**, and **characteristic polynomial** of a matrix are all *similarity invariants*.

Proof: **Part 1** and **Part 2** are left as exercises.

Proof of Part 3: Suppose $\boxed{\mathrm{A}}$ and $\widetilde{\boxed{\mathrm{A}}}$ are similar; thus, there is an invertible matrix $\boxed{\mathrm{B}}$ so that $\widetilde{\boxed{\mathrm{A}}} = \boxed{\mathrm{B}}^{-1} \cdot \boxed{\mathrm{A}} \cdot \boxed{\mathrm{B}}$. Thus,

$$
\begin{aligned}
\det\left(\widetilde{\boxed{\mathrm{A}}}\right) &= \det\left(\boxed{\mathrm{B}}^{-1} \cdot \boxed{\mathrm{A}} \cdot \boxed{\mathrm{B}}\right) \\
&= \det\left(\boxed{\mathrm{B}}^{-1}\right) \cdot \det\left(\boxed{\mathrm{A}}\right) \cdot \det\left(\boxed{\mathrm{B}}\right) \\
&= \det\left(\boxed{\mathrm{B}}\right)^{-1} \cdot \det\left(\boxed{\mathrm{B}}\right) \cdot \det\left(\boxed{\mathrm{A}}\right) \\
&= \det\left(\boxed{\mathrm{A}}\right)
\end{aligned}
$$

Proof of Part 4: Suppose $\boxed{\mathrm{A}}$ and $\widetilde{\boxed{\mathrm{A}}}$ are similar; thus, there is an invertible matrix $\boxed{\mathrm{B}}$ so that $\widetilde{\boxed{\mathrm{A}}} = \boxed{\mathrm{B}}^{-1} \cdot \boxed{\mathrm{A}} \cdot \boxed{\mathrm{B}}$. But also notice, for any fixed $\lambda \in \mathbb{R}$,

$$\lambda.\boxed{\mathbf{Id}} = \boxed{\mathrm{B}}^{-1} \cdot \lambda.\boxed{\mathbf{Id}} \cdot \boxed{\mathrm{B}}.$$

Thus, for any fixed λ,

$$\begin{aligned} \widetilde{\boxed{\mathrm{A}}} - \lambda.\boxed{\mathbf{Id}} &= \left(\boxed{\mathrm{B}}^{-1} \cdot \boxed{\mathrm{A}} \cdot \boxed{\mathrm{B}}\right) - \left(\boxed{\mathrm{B}}^{-1} \cdot \lambda.\boxed{\mathbf{Id}} \cdot \boxed{\mathrm{B}}\right) \\ &= \boxed{\mathrm{B}}^{-1} \cdot \left(\boxed{\mathrm{A}} - \lambda.\boxed{\mathbf{Id}}\right) \cdot \boxed{\mathrm{B}} \end{aligned}$$

Thus, by **Part 3**,

$$\widetilde{c}(\lambda) = \det\left(\widetilde{\boxed{\mathrm{A}}} - \lambda.\boxed{\mathbf{Id}}\right) = \det\left(\boxed{\mathrm{A}} - \lambda.\boxed{\mathbf{Id}}\right) = c(\lambda)$$

where c and $\widetilde{c}$ are the **characteristic polynomials** of $\boxed{\mathrm{A}}$ and $\widetilde{\boxed{\mathrm{A}}}$, respectively.

Since this is true for *all* λ, the functions $c(x)$ and $\widetilde{c}(x)$ are equal *everywhere* —they must be the same polynomial:

$$c(x) = \widetilde{c}(x).$$

□ [Theorem 6]

Example 7: Recall from example 2 on page 1 that $\boxed{\mathrm{A}} = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix} \sim \begin{bmatrix} 5 & 1 \\ 2 & 0 \end{bmatrix} = \widetilde{\boxed{\mathrm{A}}}$. Thus

$$\begin{aligned} \mathsf{nullity}\left[\begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}\right] &= 0 = \mathsf{nullity}\left[\begin{bmatrix} 5 & 1 \\ 2 & 0 \end{bmatrix}\right] \\ \mathsf{rank}\left[\begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}\right] &= 2 = \mathsf{rank}\left[\begin{bmatrix} 5 & 1 \\ 2 & 0 \end{bmatrix}\right] \\ \det\begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix} &= -2 = \det\begin{bmatrix} 5 & 1 \\ 2 & 0 \end{bmatrix} \end{aligned}$$

and both matrices have characteristic polynomial $x^2 - 5x + 2$.

However, if $\boxed{\mathrm{C}} = \begin{bmatrix} 1 & 2 \\ 3 & 6 \end{bmatrix}$, then

$$\text{nullity}\left[\begin{bmatrix} 1 & 2 \\ 3 & 6 \end{bmatrix}\right] = 1$$
$$\text{rank}\left[\begin{bmatrix} 1 & 2 \\ 3 & 6 \end{bmatrix}\right] = 1$$
$$\det\begin{bmatrix} 1 & 2 \\ 3 & 6 \end{bmatrix} = 0$$

and the characteristic polynomial of $\boxed{\text{C}}$ is $x^2 - 7x$. Any single one of these four facts would be enough to prove that $\boxed{\text{C}}$ could not be similar to $\boxed{\text{A}}$.

Corollary 8: The **spectrum** of a matrix is also a similarity invariant.

Proof: If $\boxed{\text{A}}$ and $\widetilde{\boxed{\text{A}}}$ are similar, then they have the same characteristic polynomials. The **spectrum** of $\boxed{\text{A}}$ is simply the list of all roots of its characteristic polynomial; thus, $\boxed{\text{A}}$ and $\widetilde{\boxed{\text{A}}}$ must have the same spectrum.

□ [Corollary 8]

Corollary 9: *The characteristic polynomial of a transformation*

Let $\mathbb{V}$ be a finite-dimensional vector space, and $f : \mathbb{V} \longrightarrow \mathbb{V}$ a linear transformation.

If $\boxed{\text{F}}$ and $\widetilde{\boxed{\text{F}}}$ are two different **matrix representations** of f (with respect to two different bases $\mathcal{B}$ and $\widetilde{\mathcal{B}}$ of $\mathbb{V}$), then $\boxed{\text{F}}$ and $\widetilde{\boxed{\text{F}}}$ are *similar*, and therefor have the *same determinant* and the *same characteristic polynomial*.

In other words, the *determinant* and *characteristic polynomial* of a linear transformation on an abstract vector space are well-defined, *independent* of the choice of basis.

Proof: Exercise

□ [Corollary 9]

Another similarity invariant is a function called the **trace**

Definition 10: *Trace*

Let $\boxed{\text{A}}$ be an $N \times N$ matrix. The **trace** of $\boxed{\text{A}}$ is the *sum of the diagonal entries in* $\boxed{\text{A}}$. In other words, if $\boxed{\text{A}} = \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1N} \\ a_{21} & a_{22} & \dots & a_{2N} \\ \vdots & \vdots & \ddots & \vdots \\ a_{N1} & a_{N2} & \dots & a_{NN} \end{bmatrix}$, then

$$\text{trace}\left[\boxed{\mathsf{A}}\right] = a_{11} + a_{22} + \ldots + a_{nn}$$

Example 11:

- trace $\begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix} = 1 + 4 = 5.$
- The trace of the $N \times N$ identity matrix $\boxed{\mathbf{Id}_N}$ is N.

Proposition 12: *Properties of the Trace*

1. **trace** : $\mathcal{M}_{N\times N} \longrightarrow \mathbb{R}$ is a linear function. In other words, trace $\left[\boxed{\mathsf{A}} + \boxed{\mathsf{B}}\right] =$ trace $\left[\boxed{\mathsf{A}}\right]$ + trace $\left[\boxed{\mathsf{B}}\right]$, and trace $\left[r.\boxed{\mathsf{A}}\right] = r$.trace $\left[\boxed{\mathsf{A}}\right]$.
2. If $\boxed{\mathsf{A}}$ and $\boxed{\mathsf{B}}$ are two $N \times N$ matrices, then

$$\text{trace}\left[\boxed{\mathsf{A}}\cdot\boxed{\mathsf{B}}\right] = \text{trace}\left[\boxed{\mathsf{B}}\cdot\boxed{\mathsf{A}}\right]$$

Proof: **Part 1** is an exercise. To see **Part 2**, suppose that

$$\boxed{\mathrm{A}}\cdot\boxed{\mathrm{B}} = \begin{bmatrix} c_{11} & c_{12} & \ldots & c_{1N} \\ c_{21} & c_{22} & \ldots & c_{2N} \\ \vdots & \vdots & \ddots & \vdots \\ c_{N1} & c_{N2} & \ldots & c_{NN} \end{bmatrix},$$

$$\text{and } \boxed{\mathrm{B}}\cdot\boxed{\mathrm{A}} = \begin{bmatrix} \widetilde{c}_{11} & \widetilde{c}_{12} & \ldots & \widetilde{c}_{1N} \\ \widetilde{c}_{21} & \widetilde{c}_{22} & \ldots & \widetilde{c}_{2N} \\ \vdots & \vdots & \ddots & \vdots \\ \widetilde{c}_{N1} & \widetilde{c}_{N2} & \ldots & \widetilde{c}_{NN} \end{bmatrix}.$$

Notice that,

$$\text{for any } n \in [1...N], \quad c_{nn} = \sum_{k=1}^{N} a_{nk} b_{kn},$$

$$\text{and, for any } k \in [1...N], \quad \widetilde{c}_{kk} = \sum_{n=1}^{N} b_{kn} a_{nk},$$

$$\text{Thus, } \operatorname{trace}\left[\boxed{A}\cdot\boxed{B}\right] = \sum_{n=1}^{N} c_{nn} = \sum_{n=1}^{N}\sum_{k=1}^{N} a_{nk}b_{kn} = \sum_{k=1}^{N}\sum_{n=1}^{N} b_{kn}a_{nk} = \sum_{k=1}^{N} \widetilde{c}_{kk} = \operatorname{trace}\left[\boxed{B}\cdot\boxed{A}\right]$$

□ [Proposition 12]

Corollary 13: The trace is a **similarity invariant**.

Proof: Exercise. Use the previous theorem

□ [Corollary 13]

.