COORDINATES

MATH 196, SECTION 57 (VIPUL NAIK)

Corresponding material in the book: Section 3.4.

Note: Section 4 of the lecture notes, added Sunday December 8, is not included in the executive summary. It is related to the material in the advanced review sheet, Section 2, discussed Saturday December 7.

Executive summary

(1) Given a basis $\mathcal{B} = (\vec{v}_1, \vec{v}_2, \dots, \vec{v}_m)$ for a subspace $V \subseteq \mathbb{R}^n$ (note that this forces $m \le n$), every vector $\vec{x} \in V$ can be written in a unique manner as a linear combination of the basis vectors $\vec{v}_1, \vec{v}_2, \dots, \vec{v}_m$. The fact that there exists a way of writing it as a linear combination follows from the fact that $\mathcal{B}$ spans V. The uniqueness follows from the fact that $\mathcal{B}$ is linearly independent. The coefficients for the linear combination are called the *coordinates* of $\vec{x}$ in the basis $\mathcal{B}$.

(2) Continuing notation from point (1), finding the coordinates amounts to solving the linear system with coefficient matrix columns given by the basis vectors $\vec{v}_1, \vec{v}_2, \dots, \vec{v}_m$ and the augmenting column given by the vector $\vec{x}$. The linear transformation of the matrix is injective, because the vectors are linearly independent. The matrix, a $n \times m$ matrix, has full column rank m. The system is consistent if and only if $\vec{x}$ is actually in the span, and injectivity gives us uniqueness of the coordinates.

(3) A canonical example of a basis is the *standard* basis, which is the basis comprising the standard basis vectors, and where the coordinates are the usual coordinates.

(4) Continuing notation from point(1), in the special case that $m = n$, $V = \mathbb{R}^n$. So the basis is $\mathcal{B} = (\vec{v}_1, \vec{v}_2, \dots, \vec{v}_n)$ and it is an alternative basis for all of $\mathbb{R}^n$ (here, alternative is being used to contrast with the standard basis; we will also use “old basis” to refer to the standard basis and “new basis” to refer to the alternative basis). In this case, the matrix S whose columns are the basis vectors $\vec{v}_1, \vec{v}_2, \dots, \vec{v}_n$ is a $n \times n$ square matrix and is invertible. We will denote this matrix by S (following the book).

(5) Continuing notation from point (4), if we denote by $[\vec{x}]_{\mathcal{B}}$ the coordinates of $\vec{x}$ in the new basis, then $[\vec{x}]_{\mathcal{B}} = S^{-1}\vec{x}$ and $\vec{x} = S[\vec{x}]_{\mathcal{B}}$.

(6) For a linear transformation T with matrix A in the standard basis and matrix B in the new basis, then $B = S^{-1}AS$ or equivalently $A = SBS^{-1}$. The S on the right involves first converting from the new basis to the old basis, then we do the middle operation A on the old basis, and then we do S^{-1} to re-convert to the new basis.

(7) If A and B are $n \times n$ matrices such that there exists an invertible $n \times n$ matrix S satisfying $B = S^{-1}AS$, we say that A and B are *similar* matrices. Similar matrices have the same trace, determinant, and behavior with respect to invertibility and nilpotency. Similarity is an equivalence relation, i.e., it is reflexive, symmetric, and transitive.

(8) Suppose S is an invertible $n \times n$ matrix. The conjugation operation $X \mapsto SXS^{-1}$ from $\mathbb{R}^{n \times n}$ to $\mathbb{R}^{n \times n}$ preserves addition, scalar multiplication, multiplication, and inverses.

1. Coordinates

1.1. **Quick summary.** Suppose V is a subspace of $\mathbb{R}^n$ and $\mathcal{B}$ is a basis for V. Recall what this means: $\mathcal{B}$ is a linearly independent subset of V and the span of $\mathcal{B}$ is V. We will show that every element of V can be written in a *unique* manner as a linear combination of the vectors in $\mathcal{B}$. The coefficients used in that linear combination are the *coordinates* of the vector in the basis $\mathcal{B}$.

Although we will deal with finite-dimensional spaces here, the arguments used can be generalized to the infinite-dimensional setting.

1.2. **Justification for uniqueness.** Suppose the basis $\mathcal{B}$ has vectors $\vec{v}_1, \vec{v}_2, \ldots \vec{v}_m$. A vector $\vec{x} \in V$ is to be written as a linear combination of the basis vectors. Note that there is *at least* one way of writing $\vec{x}$ as a linear combination of the vectors because that is what it means for $\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_m$ to span the space. We now want to show that there is in fact *exactly one* such way.

Suppose there are two ways of writing $\vec{x}$ as a linear combination of these vectors:

$$\begin{aligned} \vec{x} &= a_1\vec{v}_1 + a_2\vec{v}_2 + \cdots + a_m\vec{v}_m \\ \vec{x} &= b_1\vec{v}_1 + b_2\vec{v}_2 + \cdots + b_m\vec{v}_m \end{aligned}$$

We thus get:

$$a_1\vec{v}_1 + a_2\vec{v}_2 + \cdots + a_m\vec{v}_m = b_1\vec{v}_1 + b_2\vec{v}_2 + \cdots + b_m\vec{v}_m$$

Rearranging gives us:

$$(a_1 - b_1)\vec{v}_1 + (a_2 - b_2)\vec{v}_2 + \cdots + (a_m - b_m)\vec{v}_m = 0$$

Since the vectors are linearly independent, this forces the linear relation above to be trivial, so that $a_1 - b_1 = a_2 - b_2 = \cdots = a_m - b_m = 0$. Thus, $a_1 = b_1$, $a_2 = b_2$, $\ldots$, $a_m = b_m$. In other words, the choice of coefficients for the linear combination is unique.

1.3. **Another way of thinking about uniqueness.** Finding the coefficients for the linear combination is tantamount to solving the linear system for $a_1, a_2, \ldots, a_m$:

$$\begin{bmatrix} \uparrow & \uparrow & \ldots & \uparrow \\ \vec{v}_1 & \vec{v}_2 & \ldots & \vec{v}_m \\ \downarrow & \downarrow & \ldots & \downarrow \end{bmatrix} \begin{bmatrix} a_1 \\ a_2 \\ \cdot \\ \cdot \\ \cdot \\ a_m \end{bmatrix} = \begin{bmatrix} \vec{x} \end{bmatrix}$$

The fact that the vectors are linearly independent tells us that the kernel of the linear transformation given by the matrix:

$$\begin{bmatrix} \uparrow & \uparrow & \ldots & \uparrow \\ \vec{v}_1 & \vec{v}_2 & \ldots & \vec{v}_m \\ \downarrow & \downarrow & \ldots & \downarrow \end{bmatrix}$$

is the zero subspace. That's because elements of the kernel of that linear transformation correspond to linear relations between the vectors.

This in turn is equivalent to saying that the matrix above has full column rank m, and that the linear transformation is *injective*. Thus, for *every* vector $\vec{x}$, there is at most one solution to the linear system. Note that there exists a solution if and only if $\vec{x}$ is in V, the span of $\mathcal{B}$. And if there exists a solution, it is unique. (Note: We saw a number of equivalent formulations of injectivity in the "linear dependence, bases and subspaces" lectures that correspond to Sections 3.2 and 3.3 of your book).

1.4. **Finding the coordinates.** Suppose V is a subspace of $\mathbb{R}^n$ and $\mathcal{B}$ is a basis for V. For convenience, we will take $\mathcal{B}$ as an *ordered basis*, and list the vectors of $\mathcal{B}$ as $\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_m$. For any vector $\vec{x}$ in V, we have noted above that there is a unique way of writing $\vec{x}$ in the form:

$$\vec{x} = a_1\vec{v}_1 + a_2\vec{v}_2 + \cdots + a_m\vec{v}_m$$

The coefficients $a_1, a_2, \ldots, a_m$ are called the *coordinates* of the vector $\vec{x}$ in terms of the basis $\mathcal{B} = \vec{v}_1, \vec{v}_2, \ldots, \vec{v}_m$. Verbally, they describe $\vec{x}$ by saying "how much" of each vector $\vec{v}_i$ there is in $\vec{x}$. Note that the coordinates in terms of the standard basis are just the usual coordinates.

Finding the coordinates is easy: we already spilled the beans on that a while ago. We simply have to solve the linear system:

$$\begin{bmatrix} \uparrow & \uparrow & \dots & \uparrow \\ \vec{v}_1 & \vec{v}_2 & \dots & \vec{v}_m \\ \downarrow & \downarrow & \dots & \downarrow \end{bmatrix} \begin{bmatrix} a_1 \\ a_2 \\ \cdot \\ \cdot \\ \cdot \\ a_m \end{bmatrix} = \begin{bmatrix} \uparrow \\ \vec{x} \\ \downarrow \end{bmatrix}$$

Note that if $\vec{v}_1, \vec{v}_2, \dots, \vec{v}_m$ are known in advance, we can perform the row reduction on that matrix in advance and store all the steps we did, then apply them to $\vec{x}$ when it is known to find the coordinates.

1.5. **The standard basis.** In the vector space $\mathbb{R}^n$, the standard basis vectors $\vec{e}_1, \vec{e}_2, \dots, \vec{e}_n$ form a basis for the *whole space* $\mathbb{R}^n$, and this basis is called the *standard basis*. Moreover, the coordinates of a vector in the standard basis are precisely its coordinates as it is usually written. For instance:

$$\begin{bmatrix} 2 \\ -1 \\ 4 \end{bmatrix} = 2\vec{e}_1 + (-1)\vec{e}_2 + 4\vec{e}_3$$

The coefficients used in the linear combination, which is what we would call the "coordinates" in the standard basis, are precisely the same as the usual coordinates: they are $2, -1, 4$ respectively.

That's why we call them the standard basis vectors! At the time we first started using the terminology "standard basis vectors" we did not have access to this justification, so it was just a phrase to remember. This also sheds some light on why we use the term *coordinates*: it generalizes to an arbitrary basis the role that the usual coordinates play with respect to the standard basis.

1.6. **The special case of a basis for the whole space.** Suppose $\mathcal{B}$ is a basis for a subspace of $\mathbb{R}^n$. Then, $\mathcal{B}$ has size n if and only if the subspace is all of $\mathbb{R}^n$. In this case, every vector in $\mathbb{R}^n$ has unique coordinates with respect to $\mathcal{B}$, and we can go back and forth between coordinates in the standard basis. Let S be the matrix:

$$S = \begin{bmatrix} \uparrow & \uparrow & \dots & \uparrow \\ \vec{v}_1 & \vec{v}_2 & \dots & \vec{v}_n \\ \downarrow & \downarrow & \dots & \downarrow \end{bmatrix}$$

For a vector $\vec{x}$, if $[\vec{x}]_{\mathcal{B}}$ denotes the vector describing the coordinates for $\vec{x}$ in the new basis, then we have:

$$S[\vec{x}]_{\mathcal{B}} = \vec{x}$$

So we obtain $[\vec{x}]_{\mathcal{B}}$ be solving the linear system with augmented matrix:

$$\begin{bmatrix} S & | & \vec{x} \end{bmatrix}$$

In this case, since the matrix S is a full rank square matrix, we can write:

$$[\vec{x}]_{\mathcal{B}} = S^{-1}\vec{x}$$

2. Coordinate transformations: back and forth

2.1. **The setup.** Suppose $T : \mathbb{R}^n \to \mathbb{R}^n$ is a linear transformation with matrix A. Suppose $\mathcal{B}$ is an ordered basis for $\mathbb{R}^n$. $\mathcal{B}$ must have size n. Denote by S the matrix whose columns are the vectors of $\mathcal{B}$.

We want to write the matrix for T with respect to basis $\mathcal{B}$. What does this mean? Let's first recall in what sense A is the matrix for T: A is the matrix for T with respect to the standard basis (i.e., the ordered basis comprising the standard basis vectors). Explicitly the matrix A takes in a vector $\vec{x}$ in $\mathbb{R}^n$ written with coordinates in the standard basis, and outputs $T(\vec{x})$, again written with coordinates in the standard basis.

The matrix for T in the basis $\mathcal{B}$ should be capable of taking as input a vector expressed using coordinates in $\mathcal{B}$ and give an output that also uses coordinates in $\mathcal{B}$. In other words, it should be of the form:

$$[\vec{x}]_{\mathcal{B}} \mapsto [T(\vec{x})]_{\mathcal{B}}$$

Here's how we go about doing this. First, we convert $[\vec{x}]_{\mathcal{B}}$ to $\vec{x}$, i.e., we rewrite the vector in the standard basis. We know from before how this works. We have:

$$\vec{x} = S[\vec{x}]_{\mathcal{B}}$$

Now that the vector is in the standard basis, we can apply the linear transformation T using the matrix A, which is designed to operate in the standard basis. So we get:

$$T(\vec{x}) = A(S[\vec{x}]_{\mathcal{B}})$$

We have now obtained the obtained, but we need to re-convert to the new basis. So we multiply by S^{-1}, and get:

$$[T(\vec{x})]_{\mathcal{B}} = S^{-1}(A(S[\vec{x}]_{\mathcal{B}}))$$

By associativity of matrix multiplication, we can reparenthesize the right side, and we obtain:

$$[T(\vec{x})]_{\mathcal{B}} = (S^{-1}AS)[\vec{x}]_{\mathcal{B}}$$

Thus, the matrix for T in the new basis is:

$$B = S^{-1}AS$$

Explicitly:

- The right-most S, which is the operation that we do first, involves converting from the new basis to the old basis (the standard basis).
- We then apply the matrix A, which describes T in the standard basis.
- The left-most S^{-1} involves re-converting to the new basis.

We can also visualize this using the following diagram:

$$\begin{array}{ccc} \vec{x} & \stackrel{A}{\rightarrow} & T(\vec{x}) \\ \uparrow S & & \uparrow S \\ \vec{x}_{\mathcal{B}} & \stackrel{B}{\rightarrow} & [T(\vec{x})]_{\mathcal{B}} \end{array}$$

Our goal is to traverse the bottom B. To do this, we go up, right, and down, i.e., we do:

$$[\vec{x}]_{\mathcal{B}} \stackrel{S}{\rightsquigarrow} \vec{x} \stackrel{A}{\rightsquigarrow} T(\vec{x}) \stackrel{S^{-1}}{\rightsquigarrow} [T(\vec{x})]_{\mathcal{B}}$$

We are doing S, then A, then S^{-1}. But remember that we compose from right to left, so we get:

$$B = S^{-1}AS$$

2.2. **A real-world analogy.** Suppose you have a document in Chinese and you want to prepare an executive summary of it, again in Chinese. Unfortunately, you do not have access to any worker who can prepare executive summaries of documents in Chinese. However, you *do* have access to a person who can translate from Chinese to English, a person who can translate from English to Chinese, and a person who can do document summaries of English documents (with the summary also in English). The obvious solution is three-step:

- First, translate the document from Chinese to English
- Then, prepare the summary of the document in English, giving an English summary.
- Now, translate the summary from English to Chinese

This is analogous to the change-of-basis transformation idea: here, Chinese plays the role of the "new basis", English plays the role of the "old basis", the English summary person plays the role of the matrix A (performing the transformation in the old basis), and the overall composite process corresponds to the matrix B (performing the transformation in the new basis).

3. Similar matrices and the conjugation operation

This section will be glossed over quickly in class, but may be relevant to a better understanding of quizzes and class-related material.

3.1. **Similarity as an equivalence relation.** Suppose A and B are $n \times n$ matrices. We say that A and B are *similar matrices* if there exists an invertible $n \times n$ matrix S such that the following equivalent conditions are satisfied:

- $AS = SB$
- $B = S^{-1}AS$
- $A = SBS^{-1}$

Based on the preceding discussion, this means that there is a linear transformation $T : \mathbb{R}^n \to \mathbb{R}^n$ whose matrix in the standard basis is A and whose matrix in the basis given by the columns of S is B.

Similarity defines an *equivalence relation*. In particular, the following are true:

- *Reflexivity*: Every matrix is similar to itself. In other words, if A is a $n \times n$ matrix, then A is similar to A. The matrix S that we can use for similarity is the identity matrix.

  The corresponding "real-world" statement would be that a document summary person who works in English is similar to himself.
- *Symmetry*: If A is similar to B, then B is similar to A. Indeed, the matrices we use to go back and forth are inverses of each other. Explicitly, if $B = S^{-1}AS$, then $A = SBS^{-1} = (S^{-1})^{-1}BS^{-1}$.

  The corresponding "real-world" statement would involve noting that since translating between English and Chinese allows us to do Chinese document summaries using the English document summary person, we can also go the other way around: if there is a person who does document summaries in Chinese, we can use that person and back-and-forth translators to carry out document summaries in English.
- *Transitivity*: If A is similar to B, and B is similar to C, then A is similar to C. Explicitly, if $S^{-1}AS = B$ and $T^{-1}BT = C$, then $(ST)^{-1}A(ST) = C$.

  The corresponding "real-world" statement would note that since Chinese document summary persons are similar to English document summary persons, and English document summary persons are similar to Spanish document summary persons, the Chinese and Spanish document summary persons are similar to each other.

3.2. **Conjugation operation preserves addition and multiplication.** For an invertible $n \times n$ matrix S, define the following mapping:

$$\mathbb{R}^{n \times n} \to \mathbb{R}^{n \times n}$$

by:

$$X \mapsto SXS^{-1}$$

This is termed the *conjugation mapping* by S. Intuitively, it involves moving from the new basis to the old basis (it's the reverse of moving from the old basis to the new basis). It satisfies the following:

- It preserves addition: $S(X+Y)S^{-1} = SXS^{-1} + SYS^{-1}$ for any two $n \times n$ matrices X and Y. We can check this algebraically, but conceptually it follows from the fact that the sum of linear transformations should remain the sum regardless of what basis we use to express the linear transformations.
- It preserves multiplication: $(SXS^{-1})(SYS^{-1}) = S(XY)S^{-1}$. We can check this algebraically, but conceptually it follows from the fact that the composite of linear transformations remains the composite regardless of what basis we write them in.
- It preserves scalar multiplication: $S(\lambda X)S^{-1} = \lambda(SXS^{-1})$ for any real number λ and any $n \times n$ matrix X.
- It preserves inverses: If X is an invertible $n \times n$ matrix, then $SX^{-1}S^{-1} = (SXS^{-1})^{-1}$.
- It preserves powers: For any $n \times n$ matrix X and any positive integer r, $(SXS^{-1})^r = SX^rS^{-1}$. If X is invertible, the result holds for negative powers as well.

Intuitively, the reason why it preserves everything is the same as the reason that translating between languages is supposed to preserve all the essential features and operations. We're just using a different language and different labels, but the underlying structures remain the same. More technically, we can think of conjugation mappings as isomorphisms of the additive and scalar-multiplicative structure.

A quick note: Real-world language translation fails to live up to these ideas, because languages of the world are not structurally isomorphic. There are some ideas expressible in a certain way in English that have no equivalent expression in Chinese, and conversely, there are some ideas expressible in a certain way in Chinese that have no equivalent expression in English. Further, the translation operations between the languages, as implemented in human or machine translators, are *not* exact inverses: if you translate a document from English to Chinese and then translate that back to English, you are unlikely to get precisely the same document. So while language translation is a good example to help build real-world intuition for why we need both the S and the S^{-1} sandwiching the original transformation, we should not take the language analogy too literally. The real-world complications of language translation far exceed the complications arising in linear algebra.[1] Thus, language translation is unlikely to be a useful guide in thinking about similarity of linear transformations. The only role is the initial pedagogy, and we have (hopefully!) accomplished that.

3.3. **What are the invariants under similarity?** Conjugation can change the appearance of a matrix a lot. But there are certain attributes of matrices that remain invariant under conjugation. In fact, we can provide a complete list of invariants under conjugation, but this will take us too far afield. So, instead, we note a few attributes that remain invariant under conjugation.

For the discussion below, we assume A and B are similar $n \times n$ matrices, and that S is a $n \times n$ matrix such that $SBS^{-1} = A$.

- *Trace*: Recall that for any two matrices X and Y, XY and YX have the same trace. Thus, in particular, $(SB)S^{-1} = A$ and $S^{-1}(SB) = B$ have the same trace. So, A and B have the same trace.
- *Invertibility*: A is invertible if and only if B is invertible. In fact, their inverses are similar via the same transformation, as mentioned earlier.
- *Nilpotency*: A is nilpotent if and only B is nilpotent. Further, if r is the smallest positive integer such that $A^r = 0$, then r is also the smallest positive integer such that $B^r = 0$.
- *Other stuff related to powers*: A is idempotent if and only if B is idempotent. A has a power equal to the identity matrix if and only if B has a power equal to the identity matrix.
- *Determinant*: The determinant is a complicated invariant that controls the invertibility and some other aspects of the matrix. We have already seen the explicit description of the determinant for 2×2 matrices. We briefly discuss the significance of the determinant in the next section.

3.4. **The determinant: what it means.** *Not covered in class, but relevant for some quizzes.*

The determinant is a number that can be computed for any $n \times n$ (square) matrix, with the following significance:

- *Whether or not the determinant is zero* determines whether the linear transformation is invertible. In particular:
  - If the determinant is zero, the linear transformation is non-invertible.
  - If the determinant is nonzero, the linear transformation is invertible.
- The *sign* of the determinant determines whether the linear transformation preserves orientation. In particular:
  - If the determinant is positive, the linear transformation is orientation-preserving.
  - If the determinant is negative, the linear transformation is orientation-reversing.
- The *magnitude* of the determinant is the factor by which volumes get scaled.

Let's consider a method to compute the determinant. Recall the various row operations that we perform in order to row reduce a matrix. These operations include:

(1) Multiply or divide a row by a nonzero scalar.
(2) Add or subtract a multiple of a row to another row.

[1] The claim may seem strange if you're finding linear algebra a lot harder than English or Chinese, but you've probably spent a lot more time in total mastering your first language than you have mastering linear algebra. Linear algebra is also easier to systematize and abstract than natural language.

(3) Swap two rows.

We now keep track of some rules for the determinant:

(1) Each time we multiply a row of a matrix by a scalar λ, the determinant gets multiplied by λ.
(2) Adding or subtracting a multiple of a row to another row preserves the determinant.
(3) Swapping two rows multiplies the determinant by -1.

Finally, the following two all-important facts:

- The determinant of a non-invertible matrix is 0.
- The determinant of the identity matrix is 1.

The procedure is now straightforward:

- Convert the matrix to rref. Keep track of the operations and note what the determinant overall gets multiplied by.
- If the rref is non-invertible, the original matrix is also non-invertible and its determinant is 0.
- If the rref is invertible, it must be the identity matrix. We thus get 1 equals (some known number) times (the unknown determinant). The unknown determinant is thus the reciprocal of that known number. For instance, if we had to multiply by $1/2$, then by -1, then by 3, to get to rref, then we have overall multiplied by $-3/2$ and the determinant is thus $-2/3$.

Here is how the determinant interacts with addition, mutliplication, and inversion of mtarices:

(1) There is no formula to find the determinant of the sum in terms of the individual determinants. Rather, it is necessary to know the actual matrices. In fact, as you know, a sum of non-invertible matrices may be invertible or non-invertible, so that rules out the possibility of a formula.
(2) The determinant of a product of two $n \times n$ matrices is the product of the determinants. In symbols, if we use det to denote the determinant, then:

$$\det(AB) = \det(A)\det(B)$$

(3) The determinant of the inverse of an invertible $n \times n$ matrix is the inverse (i.e., the reciprocal) of the determinant. In symbols:

$$\det(A^{-1}) = \frac{1}{\det A}$$

We can reconcile the observations about the product and inverse with each other, and also reconcile them with earlier observations about the significance of the magnitude and sign of the determinant.

3.5. Similarity and commuting with the change-of-basis matrix.

This subsection was added December 8.

In the special case that the change-of-basis matrix S commutes with the matrix A, we have that $S^{-1}AS = S^{-1}SA = A$.

Two other further special cases are worth noting:

- The case that S is a scalar matrix: In this case, S commutes with all possible choices of A, so this change of basis does not affect any of the matrices. Essentially, the matrix S is causing the same scaling on both the inputs and the outputs, so it does not affect the description of the linear transformation.
- The case that A is a scalar matrix: In this case, S commutes with A for all possible choices of S, so the change of basis does not affect A. Intuitively, this is because a scalar multiplication looks the same regardless of the basis.

  Another way of formulating this is that a scalar matrix is the *only* matrix in its similarity class, i.e., a scalar matrix can be similar to no *other* matrix (scalar or non-scalar).

4. Constructing similar matrices

This section was added later, and is related to material covered in the Saturday (December 7) review session.

4.1. **Similarity via the identity matrix.** This might seem too obvious to mention, but it is still worth mentioning when trying to construct examples and counterexamples: for any $n \times n$ matrix A, A is similar to itself, and we can take the change-of-basis matrix S to be the identity matrix I_n.

4.2. **Similarity via coordinate interchange.** We begin by considering the matrix:

$$S = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$$

The matrix S describes the linear transformation that interchanges the standard basis vectors $\vec{e}_1$ and $\vec{e}_2$. Explicitly, this linear transformation interchanges the coordinates, i.e., it is described as:

$$\begin{bmatrix} x \\ y \end{bmatrix} \mapsto \begin{bmatrix} y \\ x \end{bmatrix}$$

Note that $S = S^{-1}$, or equivalently, S^2 is the identity matrix.

Now, consider two matrices A and B that are similar via S, so that:

$$A = SBS^{-1}, B = S^{-1}AS$$

We have $S = S^{-1}$, and we obtain:

$$A = SBS, B = SAS$$

Let's understand carefully what left and right multiplication by S are doing. Consider:

$$A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$$

We have:

$$AS = \begin{bmatrix} a & b \\ c & d \end{bmatrix} \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} = \begin{bmatrix} b & a \\ d & c \end{bmatrix}$$

We thereby obtain:

$$SAS = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} b & a \\ d & c \end{bmatrix} = \begin{bmatrix} d & c \\ b & a \end{bmatrix}$$

We could also do the multiplication in the other order.

$$SA = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} a & b \\ c & d \end{bmatrix} = \begin{bmatrix} c & d \\ a & b \end{bmatrix}$$

We thereby obtain:

$$SAS = \begin{bmatrix} c & d \\ a & b \end{bmatrix} \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} = \begin{bmatrix} d & c \\ b & a \end{bmatrix}$$

The upshot is that:

$$SAS = \begin{bmatrix} d & c \\ b & a \end{bmatrix}$$

Conceptually:

- Right multiplication by S interchanges the columns of the matrix. Conceptually, AS takes as input $[\vec{x}]_{\mathcal{B}}$ ($\vec{x}$ written in the *new* basis) and gives as output the vector $T(\vec{x})$ (written in the old, i.e. standard, basis). AS has the same columns as A but in a different order, signifying that which standard basis vector goes to which column vector gets switched around.
- Left multiplication by S (conceptually, S^{-1}, though $S = S^{-1}$ in this case) interchanges the rows of the matrix. Conceptually, $S^{-1}A$ takes as input $\vec{x}$ and outputs $[T(\vec{x})]_{\mathcal{B}}$. The left multiplication by $S = S^{-1}$ signifies that after we obtain the output, we swap its two coordinates.

- Left and right multiplication by S signifies that we interchange the rows and interchange the columns. In other words, we change the row and column of each entry. This can also be thought of as "reflecting" the entries of the square matrix about the center: the entries on the main diagonal a and d get interchanged, and the entries on the main *anti*-diagonal b and c get interchanged.

To summarize, the matrices:

$$A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}, SAS = \begin{bmatrix} d & c \\ b & a \end{bmatrix}$$

are similar via S.

We can now do some sanity checks for similarity.

- *Same trace?*: The trace of A is $a + d$ and the trace of SAS is $d + a$. Since addition is commutative, these are equal.
- *Same determinant?*: The matrix A has determinant $ad - bc$ whereas the matrix SAS has determinant $da - cb$. Since multiplication is commutative, these are the same.

Note that these are *sanity checks*: they don't prove anything new. Rather, their purpose is to make sure that things are working as they "should" if our conceptual framework and computations are correct.

4.3. **Linear transformations, finite state automata, and similarity.** Recall the setup of linear transformations and finite state automata described in your quizzes as well as in the lecture notes for matrix multiplication and inversion.

We will now discuss how we can judge similarity of the linear transformations given by these matrices.

4.3.1. *The case $n = 2$: the nonzero nilpotent matrices.* Consider the functions f and g given as follows:

$$f(0) = 0, f(1) = 2, f(2) = 0, \qquad g(0) = 0, g(1) = 0, g(2) = 1$$

The finite state automaton diagrams for f and g are as follows:

$$f : 1 \to 2 \to 0, \qquad g : 2 \to 1 \to 0$$

(there are loops at 0 that I did not write above).

The matrices are as follows:

$$M_f = \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}, M_g = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}$$

Both of these square to zero, i.e., we have $M_f^2 = 0$ and $M_g^2 = 0$.

The following three equivalent observations can be made:

(1) If we interchange 1 and 2 on the input side and *also* interchange 1 and 2 on the output side for f, then we obtain the function g. Equivalently, if we do the same starting with g, we obtain f.
(2) If we take the finite state automaton diagram for f, and interchange the labels 1 and 2, we obtain the finite state automaton diagram for g. Equivalently, if we interchange the labels starting with the diagram for g, we obtain the diagram for f.
(3) If we consider the matrix S of Section 4.2:

$$S = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$$

then (recall that $S = S^{-1}$):

$$M_g = SM_fS = S^{-1}M_fS = SM_fS^{-1}$$

Note that the matrices M_f and M_g being similar fits in well with other observations about the matrices, namely:

(a) *Same trace*: Both matrices have trace 0.
(b) *Same determinant*: Both matrices have determinant 0 (note that this follows from their not being full rank).

(c) *Same rank*: Both matrices have rank 1.
(d) *Both are nilpotent with the same nilpotency*: Both matrices square to zero.

4.3.2. *The case $n = 2$: idempotent matrices.* Consider the functions f and g given as follows:

$$f(0) = 0, f(1) = 1, f(2) = 0, \qquad g(0) = 0, g(1) = 0, g(2) = 2$$

The finite state automaton diagram for f loops at 1, loops at 0, and sends 2 to 0. The finite state automaton for g loops at 2, loops at 0, and sends 1 to 0. The corresponding matrices are:

$$M_f = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}, M_g = \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}$$

For this f and g, observations of similarity similar to the preceding case (the numbered observations (1)-(3)) apply.

Consider now two other functions h and j:

$$h(0) = 0, h(1) = 1, h(2) = 1, \qquad j(0) = 0, j(1) = 2, j(2) = 2$$

The corresponding matrices are:

$$M_h = \begin{bmatrix} 1 & 1 \\ 0 & 0 \end{bmatrix}, M_j = \begin{bmatrix} 0 & 0 \\ 1 & 1 \end{bmatrix}$$

It is clear that observations analogous to the numbered observations (1)-(3) apply to the functions h and j. Thus, M_h and M_j are similar via the change of basis matrix S.

However, the somewhat surprising fact that if the matrices M_f, M_g, M_h, and M_j are *all* similar. We have already established the similarity of M_f and M_g (via S) and of M_h and M_j (via S). To establish the similarity of all four, we need to show that M_f and M_h are similar. This is a little tricky. The idea is to use:

$$S = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}$$

We obtain:

$$S^{-1} = \begin{bmatrix} 1 & -1 \\ 0 & 1 \end{bmatrix}$$

We then get:

$$S^{-1} M_f S = M_h$$

Thus, M_f and M_h are similar matrices.

Intuition behind the discovery of S: S should send $\vec{e}_1$ to $\vec{e}_1$ because that is a basis vector for the unique fixed line of M_f and M_h. So, the first column of S is $\vec{e}_1$. S should send $\vec{e}_2$ to a vector for which the second column of M_h applies, i.e., a vector that gets sent to $\vec{e}_1$ under M_f. One such vector is $\vec{e}_1 + \vec{e}_2$. Thus, the second column of S is $\vec{e}_1 + \vec{e}_2$. The matrix for S looks like:

$$\begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}$$

We can make the following observations similar to the earlier observations in the previous subsubsection, but now for all *four* of the matirces M_f, M_g, M_h, and M_j.

(a) *Same trace*: All four matrices have trace 1.
(b) *Same determinant*: All four matrices have determinant 0. This follows from their not being of full rank.
(c) *Same rank*: All four matrices have rank 1.
(d) *All are idempotent*: $M_f^2 = M_f$, $M_g^2 = M_g$, $M_h^2 = M_h$, $M_j^2 = M_j$.

4.3.3. *The case $n = 3$: nilpotent matrices of nilpotency three.* For $n = 3$, we can construct many different choices of f such that $M_f^2 \neq 0$ but $M_f^3 = 0$. A typical example is a function f with the following finite state automaton diagram:

$$3 \to 2 \to 1 \to 0$$

The corresponding matrix M_f is:

$$\begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix}$$

We can obtain other such matrices that are similar to this matrix via coordinate permutations, i.e., via interchanging the labels 1, 2, and 3 (there is a total of 6 different matrices we could write down of this form in this similarity class).

4.3.4. *The case $n = 3$: 3-cycles.* The 3-cycles are permutations that cycle the three coordinates around. There are two such 3-cycles possible, namely the functions f and g described below.

$$f(0) = 0, f(1) = 2, f(2) = 3, f(3) = 1, \qquad g(0) = 0, g(1) = 3, g(2) = 1, g(3) = 2$$

The corresponding matrices are:

$$M_f = \begin{bmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix}, M_g = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{bmatrix}$$

These matrices are similar to each other via any interchange of two coordinates. Explicitly, for instance, M_f and M_g are similar via the following matrix S (that equals its own inverse):

$$S = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix}$$

Note also that $M_f^3 = M_g^3 = I_3$, and $M_g = M_f^2 = M_f^{-1}$ while at the same time $M_f = M_g^2 = M_g^{-1}$.

4.4. **Counterexample construction for similarity using finite state automata and other methods.** Below are some situations where you need to construct examples of matrices, and how finite state automata can be used to motivate the construction of relevant examples. Finite state automata are definitely not *necessary* for the construction of the examples, and in fact, in many cases, randomly written examples are highly likely to work. But using examples arising from finite state automata allows us to see exactly *where, why, and how* things fail.

4.4.1. *Similarity does not behave well with respect to addition, subtraction, and multiplication.* Our goal here is to construct matrices A_1, A_2, B_1, and B_2 such that A_1 is similar to B_1 and A_2 is similar to B_2, but the following failures of similarity hold:

(a) $A_1 + A_2$ is not similar to $B_1 + B_2$
(b) $A_1 - A_2$ is not similar to $B_1 - B_2$
(c) $A_1 A_2$ is not similar to $B_1 B_2$

Note that we need to crucially make sure that we *must* use *different* change-of-basis matrices for the change of basis from A_1 to B_1 and the change of basis from A_2 to B_2. Denote by S_1 the change-of-basis matrix that we use between A_1 and B_1 and denote by S_2 the change-of-basis matrix that we use between A_2 and B_2. The simplest strategy will be to choose matrices such that:

$$n = 2, S_1 = I_2 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, S_2 = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$$

In other words, we choose $A_1 = B_1$, and B_2 is obtained by swapping rows *and* swapping columns in A_2. For more on how the change of basis given by S_2 works, see Section 4.2.

The following examples *each* work for *all* the points (a)-(c) above:

(1) $A_1 = A_2 = B_1 = \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}$, $B_2 = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}$. These example matrices are discussed in Section 4.3.1. You can verify the conditions (a)-(c) manually (trace, determinant, and rank can be used to rule out similarity).

(2) $A_1 = A_2 = B_1 = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}$, $B_2 = \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}$. These example matrices are discussed in Section 4.3.2. You can verify the conditions (a)-(c) manually (trace, determinant, and rank can be used to rule out similarity).

4.4.2. *Similarity behaves well with respect to matrix powers but not with respect to matrix roots.* Suppose A and B are similar $n \times n$ matrices, and p is a polynomial. Then, $p(A)$ and $p(B)$ are similar matrices, and in fact, the same matrix S for which $A = SBS^{-1}$ also satisfies $f(A) = Sf(B)S^{-1}$. In particular, for any positive integer r, $A^r = SB^rS^{-1}$. (Proof-wise, we *start* with proving the case of positive integer powers, then combine with the case for scalar multiples and addition, and obtain the statement for arbitrary polynomials).

In the case that $r = 1$, the implication works both ways, but for $r > 1$, A^r and B^r being similar does not imply that A and B are similar.

We discuss some examples:

- For r even, we can choose $n = 1$ and take $A = [1]$ and $B = [-1]$.
- *Non-invertible example based on finite state automata*: For all $r > 1$, we can choose $A = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$ and $B = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}$. In this case, $A^r = B^r = 0$, but A is not similar to B because the zero matrix is not similar to any nonzero matrix. Here, B is a matrix of the type discussed in Section 4.3.1.
- *Invertible example*: For all $r > 1$, we can choose $n = 2$ and take A to be the identity matrix and B to be the matrix that is counter-clockwise rotation by an angle of $2\pi/r$. In this case, $A^r = B^r = I_2$, but A is not similar to B because the identity matrix is not similar to any non-identity matrix.
- *Invertible example based on finite state automata*: For $r = 3$, we can construct examples using finite state automata. We take A as the identity matrix and to take B as one of the automata based on the 3-cycle (as described in Section 4.3.4). $A^3 = B^3 = I_3$ but A is not similar to B because the identity matrix is not similar to any non-identity matrix.

4.5. **Similarity via negation of the second coordinate.** Another example of a change-of-basis matrix S that is worth keeping handy is:

$$S = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}$$

This matrix fixes $\vec{e}_1$ (in more advanced jargon, we would say that $\vec{e}_1$ is an eigenvector with eigenvalue 1) and sends $\vec{e}_2$ to its negative (in more advanced jargon, we would say that $\vec{e}_2$ is an eigenvector with eigenvalue -1).

Note that $S = S^{-1}$ in this case.

Multiplying on the left by S means multiplying the second row by -1. Multiplying on the right by S means multiplying the second column by -1. Multiplying on both the left and the right by S multiplies both the off-diagonal entries by -1 (note that the bottom right entry gets multiplied by -1 twice, so the net effect is that it returns to its original value). Explicitly, the map sending a matrix A to the matrix SAS is:

$$\begin{bmatrix} a & b \\ c & d \end{bmatrix} \mapsto \begin{bmatrix} a & -b \\ -c & d \end{bmatrix}$$

To summarize, the matrices:

$$A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}, SAS = \begin{bmatrix} a & -b \\ -c & d \end{bmatrix}$$

are similar.

We can now do some sanity checks for similarity.

- *Same trace*: The matrices A and SAS have the same diagonal (diagonal entries a and d). Both matrices therefore have the same trace, namely $a + d$.
- *Same determinant*: The determinant of A is $ad - bc$. The determinant of SAS is $ad - (-b)(-c)$. The product $(-b)(-c)$ is bc, so we obtain that the determinant is $ad - bc$.

Note that these are *sanity checks*: they don't prove anything new. Rather, their purpose is to make sure that things are working as they "should" if our conceptual framework and computations are correct.

Similarity via reflection plays a crucial role in explaining why the counter-clockwise rotation matrix and clockwise rotation matrix for the same angle are similar. Explicitly, if we denote the rotation matrix for θ as $R(\theta)$, then the matrices:

$$R(\theta) = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}, R(-\theta) = \begin{bmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{bmatrix}$$

are similar via the matrix S discussed above. Note that both these matrices have trace $2\cos\theta$ and determinant 1.