Math 2270-004 Week 9 notes
We will not necessarily finish the material from a given day's notes on that day. We may also add or subtract some material as the week progresses, but these notes represent an in-depth outline of what we plan to cover. These notes cover material in 4.6-4.7, 5.1-5.2

Mon Mar 5

- 4.5 General theorems about finite dimensional vector spaces, bases, spanning sets, linearly independent sets and subspaces.

<u>Announcements:</u>

<u>Warm-up Exercise</u>:

Monday Review!

We've been studying *vector spaces*, which are a generalization of $\mathbb{R}^n$. They occur as *subspaces* of $\mathbb{R}^n$; also as vector spaces and subspaces of matrices, and of function spaces.

We've been studying *linear transformations* $T : V \rightarrow W$ between vector spaces, which are generalizations of matrix transformations $T : \mathbb{R}^n \rightarrow \mathbb{R}^m$ given as $T(\underline{\boldsymbol{x}}) = A\,\underline{\boldsymbol{x}}$.

For an $m \times n$ matrix A there are two interesting subspaces: $Nul\,A = \{\underline{\boldsymbol{x}} \in \mathbb{R}^n : A\,\underline{\boldsymbol{x}} = \underline{\boldsymbol{0}}\}$ and $Col\,A = \{\underline{\boldsymbol{b}} \in \mathbb{R}^m : \underline{\boldsymbol{b}} = A\,\underline{\boldsymbol{x}}, \underline{\boldsymbol{x}} \in \mathbb{R}^n\} = span\{\underline{\boldsymbol{a}}_1, \underline{\boldsymbol{a}}_2, \ldots \underline{\boldsymbol{a}}_n\}$. (Here we expressed $A = [\underline{\boldsymbol{a}}_1\ \underline{\boldsymbol{a}}_2\ \ldots\ \underline{\boldsymbol{a}}_n]$ in terms of its columns.) Through homework and food for thought questions we've understood the *rank+ nullity Theorem*, that $dim\ Nul\,A + dim\ Col\,A = n$. This theorem follows from considerations of the reduced row echelon form of A.

In the Friday food for thought questions this past Friday we realized there's a third interesting subspace associated to the matrix A, namely *Row A*, which is the subspace in $\mathbb{R}^n$ spanned by the rows of A. We'll see the fourth and final subspace associated with A tomorrow, and what how these four subspaces are connected to the domain and codomain geometry of the transformation $T(\underline{\boldsymbol{x}}) = A\,\underline{\boldsymbol{x}}$. (This is section 4.6, which we've been secretly thinking about for the past week. We'll utilize many of these ideas again in Chapter 6, e.g. section 6.5.)

We've defined *kernel T* and *range T* for linear transformations $T : V \rightarrow W$, generalizing *Nul A* and *Col A* for matrix transformations.

We've defined what it means for a linear transformation $T : V \rightarrow W$ to be an *isomorphism,* and checked that in this case the inverse function $T^{-1} : W \rightarrow V$ is also a linear transformation (isomorphism) - generalizing the notion of invertible matrix transformations $T : \mathbb{R}^n \rightarrow \mathbb{R}^n$ that are given as $T(\underline{\boldsymbol{x}}) = A\,\underline{\boldsymbol{x}}$, with $T^{-1}(\underline{\boldsymbol{y}}) = A^{-1}\,\underline{\boldsymbol{y}}$.

With a basis $\beta = \{\underline{\boldsymbol{v}}_1, \underline{\boldsymbol{v}}_2, \ldots \underline{\boldsymbol{v}}_n\}$ for a vector space V we can define the coordinate transformation isomorphism $T : V \rightarrow \mathbb{R}^n$

$$T\left(c_1\,\underline{\boldsymbol{v}}_1 + c_2\,\underline{\boldsymbol{v}}_2 + \ldots + c_n\,\underline{\boldsymbol{v}}_n\right) := \begin{bmatrix} c_1 \\ c_2 \\ \vdots \\ c_n \end{bmatrix} = [v]_\beta$$

and use these coordinate systems to answer questions about V.

There is a circle of ideas related to linear independence, span, and basis for vector spaces, which it is good to try and understand carefully. That's what we'll do today. These ideas generalize (and use) ideas we've already explored more concretely, and facts we already know to be true for the vector spaces $\mathbb{R}^n$. (A vector space that does not have a basis with a finite number of elements is said to be *infinite dimensional.* For example the space of all polynomials of arbitrarily high degree is an infinite dimensional vector space. We often study finite dimensional subspaces of infinite dimensional vector spaces.)

Theorem 1 (constructing a basis from a spanning set): Let V be a vector space of dimension at least one, and let $span\{\underline{\boldsymbol{v}}_1, \underline{\boldsymbol{v}}_2, \ldots \underline{\boldsymbol{v}}_p\} = V$.
Then a subset of the spanning set is a basis for V. (We followed a procedure like this to extract bases for *Col A*.)

Theorem 2 Let V be a vector space, with basis $\beta = \{\underline{\boldsymbol{b}}_1, \underline{\boldsymbol{b}}_2, \ldots \underline{\boldsymbol{b}}_n\}$. Then any set in V containing more than n elements must be linearly dependent. (We used reduced row echelon form to understand this in $\mathbb{R}^n$.)

Theorem 3 Let V be a vector space, with basis $\beta = \{\underline{\boldsymbol{b}}_1, \underline{\boldsymbol{b}}_2, \ \dots \underline{\boldsymbol{b}}_n\}$. Then no set $\alpha = \{\underline{\boldsymbol{a}}_1, \underline{\boldsymbol{a}}_2, \ \dots \underline{\boldsymbol{a}}_p\}$ with $p < n$ vectors can span V. (We know this for $\mathbb{R}^n$.)

Theorem 4 Let V be a vector space, with basis $\beta = \{\underline{\boldsymbol{b}}_1, \underline{\boldsymbol{b}}_2, \ \dots \underline{\boldsymbol{b}}_n\}$. Let $\alpha = \{\underline{\boldsymbol{a}}_1, \underline{\boldsymbol{a}}_2, \ \dots \underline{\boldsymbol{a}}_p\}$ be a set of independent vectors that don't span V. Then $p < n$, and additional vectors can be added to the set α to create a basis $\{\underline{\boldsymbol{a}}_1, \underline{\boldsymbol{a}}_2, \ \dots \underline{\boldsymbol{a}}_p, \dots \underline{\boldsymbol{a}}_n\}$ (We followed a procedure like this when we figured out all the subspaces of $\mathbb{R}^3$.)

Theorem 5 Let Let V be a vector space, with basis $\beta = \{\underline{\boldsymbol{b}}_1, \underline{\boldsymbol{b}}_2, \ \dots \ \underline{\boldsymbol{b}}_n\}$. Then every basis for V has exactly n vectors. (We know this for $\mathbb{R}^n$.)

Theorem 6 Let Let V be a vector space, with basis $\beta = \{\underline{\boldsymbol{b}}_1, \underline{\boldsymbol{b}}_2, \ \dots \ \underline{\boldsymbol{b}}_n\}$. If $\alpha = \{\underline{\boldsymbol{a}}_1, \underline{\boldsymbol{a}}_2, \ \dots \ \underline{\boldsymbol{a}}_n\}$ is another collection of exactly n vectors in V, and if $span\{\underline{\boldsymbol{a}}_1, \underline{\boldsymbol{a}}_2, \ \dots \ \underline{\boldsymbol{a}}_n\} = V$, then the set α is automatically linearly independent and a basis. Conversely, if the set $\{\underline{\boldsymbol{a}}_1, \underline{\boldsymbol{a}}_2, \ \dots \ \underline{\boldsymbol{a}}_n\}$ is linearly independent, then $span\{\underline{\boldsymbol{a}}_1, \underline{\boldsymbol{a}}_2, \ \dots \ \underline{\boldsymbol{a}}_n\} = V$ is guaranteed, and α is a basis. (We know all these facts for $\mathbb{R}^n$ from reduced row echelon form considerations.)

Corollary Let Let V be a vector space of dimension n. Then the subspaces of V have dimensions $0, 1, 2, \ldots n-1, n$. (We know this for $\mathbb{R}^n$.)

Remark We used the coordinate transformation isomorphism between a vector space V with basis $\beta = \{\underline{\boldsymbol{b}}_1, \underline{\boldsymbol{b}}_2, \ldots \underline{\boldsymbol{b}}_n\}$ for Theorem 2, but argued more abstractly for the other theorems. An alternate (quicker) approach is to just note that because the coordinate transformation is an isomorphism it preserves sets of independent vectors, and maps spans of vectors to spans of the image vectors, so maps subspaces to subspaces. Then every one of the theorems above follows from their special cases in $\mathbb{R}^n$, which we've already proven. But this shortcut shortchanges the conceptual ideas to some extent, which is why we've discussed the proofs more abstractly.

Tues Mar 6

- 4.6 The four subspaces associated with a matrix. the rank of a matrix.

Announcements:

Warm-up Exercise:

Let A be an $m \times n$ matrix. There are four subspaces associated with A. To keep them straight, keep in mind the associated linear transformation

$$T : \mathbb{R}^n \to \mathbb{R}^m \text{ given by } T(\underline{\mathbf{x}}) = A\,\underline{\mathbf{x}}.$$

And, as usual, we can express A in terms of its columns, $A = \begin{bmatrix} \underline{\boldsymbol{a}}_1 & \underline{\boldsymbol{a}}_2 & \ldots & \underline{\boldsymbol{a}}_n \end{bmatrix}$. Then the two subspaces we know well are

$$Col\,A = span\{\underline{\boldsymbol{a}}_1, \underline{\boldsymbol{a}}_2, \ \ldots\ \underline{\boldsymbol{a}}_n\} \subseteq \mathbb{R}^m$$

$$Nul\,A = \{\underline{\mathbf{x}} \in \mathbb{R}^n : A\,\underline{\mathbf{x}} = \underline{\mathbf{0}}\} \subseteq \mathbb{R}^n.$$

And, in your homework you already figured out the "rank + nullity" theorem, that

$$dim(Col\,A) + dim(Nul\,A) = n.$$

The reason for this is that if p is the number of pivots in the reduced row echelon form of A, then

$$dim(Col\,A) = p$$
$$dim(Nul\,A) = n - p.$$

The number of pivots, i.e. $dim(Col\,A)$ is called the *rank* of the matrix A. What are the other two subspaces and why do we care? Well,

- First, recall the geometry fact that the dot product of two vectors in $\mathbb{R}^n$ is zero if and only if the vectors are perpendicular, i.e.

$$\underline{\boldsymbol{u}} \cdot \underline{\boldsymbol{v}} = 0 \quad \text{if and only if } \underline{\boldsymbol{u}} \perp \underline{\boldsymbol{v}}.$$

(Well, we really only know this in $\mathbb{R}^2$ or $\mathbb{R}^3$ so far, from multivariable Calculus class. But it's true for all $\mathbb{R}^n$, as we'll see in Chapter 6.) So for a vector $\underline{\boldsymbol{x}} \in Nul\ A$ we can interpret the equation

$$A\,\underline{\mathbf{x}} = \underline{\mathbf{0}}$$

as saying that $\underline{\boldsymbol{x}}$ is perpendicular to every row of A. Because the dot product distributes over addition, we see that each $\underline{\boldsymbol{x}} \in Nul\ A$ is perpendicular to every linear combination of the rows of A. This motivates the next subspace associated with A, namely the rowspace. In other words, if we express A in terms of its rows,

$$A = \begin{bmatrix} \underline{\boldsymbol{R}}_1 \\ \underline{\boldsymbol{R}}_2 \\ : \\ \underline{\boldsymbol{R}}_m \end{bmatrix}$$

then

$$Row\ A := span\{\underline{\boldsymbol{R}}_1, \underline{\boldsymbol{R}}_2, \ldots \underline{\boldsymbol{R}}_m\} \subseteq \mathbb{R}^n.$$

And, $Row\ A \perp Nul\ A$.

As we do elementary operations on the rows of A we don't change their span, so we get a great basis for $Row\ A$ by using the non-zero rows of $rref(A)$...as in your food for thought this past Friday, and this week's homework. So, the dimension of $Row(A)$ is p, the number of pivots in the reduced matrix. So in the domain $\mathbb{R}^n$, we have this picture:

$$dim(Nul\ A) = n - p$$
$$dim(Row\ A) = p$$
$$Nul\ A \perp Row\ A.$$

The final subspace lives in the codomain $\mathbb{R}^m$, along with $Col\ A$. Well, $Col\ A = Row\ A^T$. And so $Nul\ A^T$ is the final subspace. Since A^T has m columns an p pivots, there are $m - p$ free parameters when we solve $A^T\underline{\boldsymbol{y}} = 0$, so $dim\left(Nul\ A^T\right) = m - p$ and in the codomain $\mathbb{R}^m$ we have this picture:

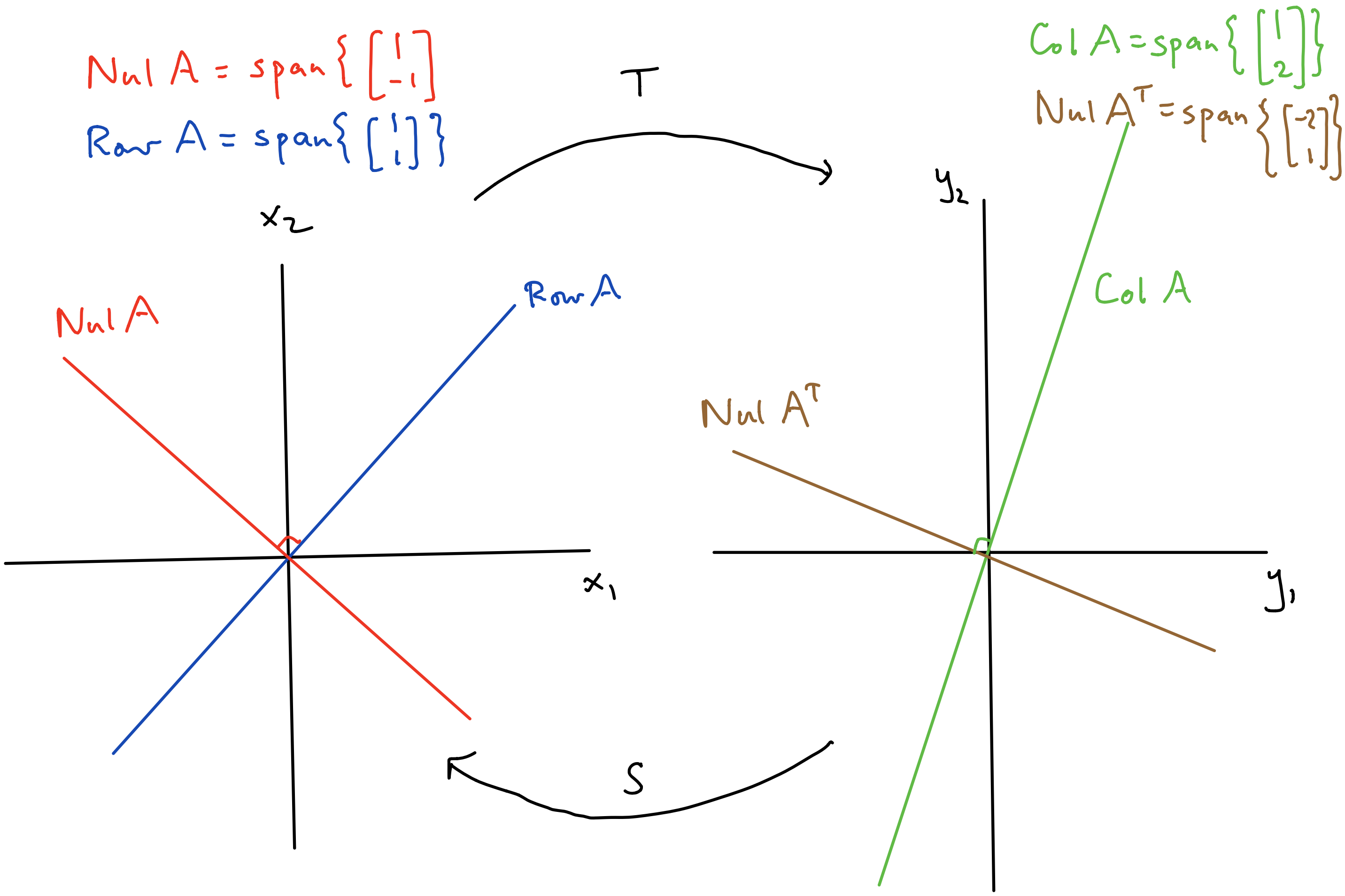
$$Col\ A = Row\left(A^T\right)$$
$$dim\left(Row\ A^T\right) = p$$
$$dim\left(Nul\ A^T\right) = m - p$$
$$Nul\ A^T \perp Row\ A^T.$$

small example.

$$T:\mathbb{R}^2 \to \mathbb{R}^2$$

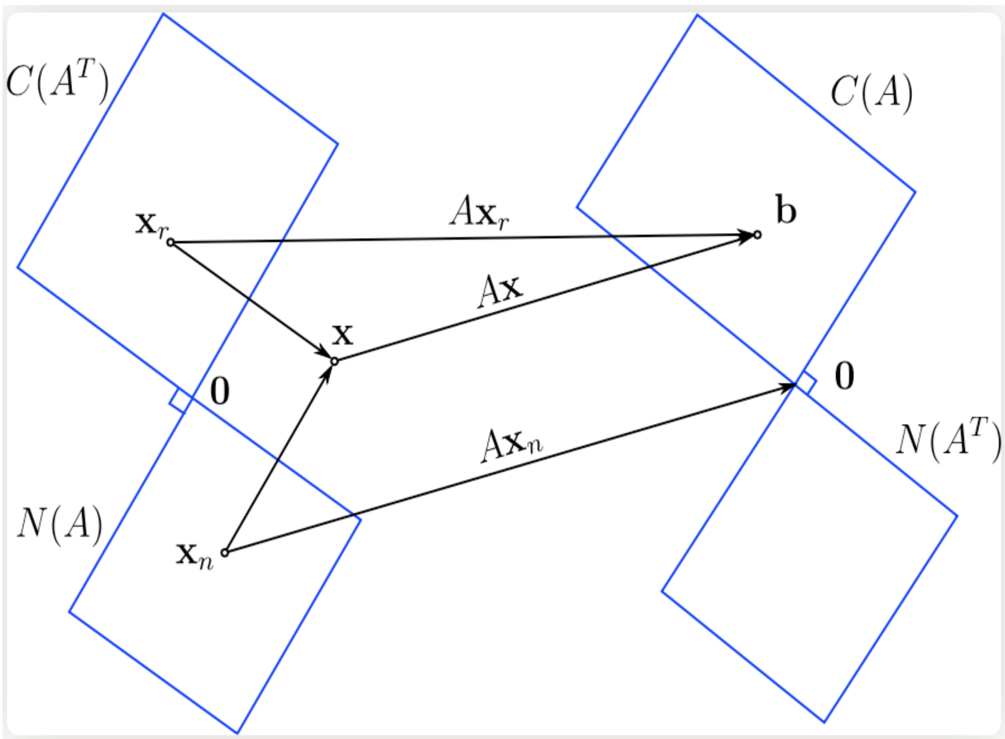
$$T\left(\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}\right) = \begin{bmatrix} 1 & 1 \\ 2 & 2 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$$

$$S\left(\begin{bmatrix} y_1 \\ y_2 \end{bmatrix}\right) = \begin{bmatrix} 1 & 2 \\ 1 & 2 \end{bmatrix}\begin{bmatrix} y_1 \\ y_2 \end{bmatrix}$$



Here's a schematic of what's going on, stolen from the internet. The web site I stole it from looks pretty good....

http://www.itshared.org/2015/06/the-four-fundamental-subspaces.html



More details on the decompositions .... In the domain $\mathbb{R}^n$, the two subspaces associated to A are *Row A* and *Nul A*. Notice that the only vector in their intersection is the zero vector, since

$$\underline{\mathbf{x}} \in Row\ A \cap Nul\ A \qquad \Rightarrow \underline{\mathbf{x}} \cdot \underline{\mathbf{x}} = 0 \qquad \Rightarrow \underline{\mathbf{x}} = \underline{\mathbf{0}}.$$

So, let

$$\{\underline{\boldsymbol{u}}_1, \underline{\boldsymbol{u}}_2, \ \dots \underline{\boldsymbol{u}}_p\} \qquad \text{be a basis for} \quad Row\ A$$

$$\{\underline{\boldsymbol{v}}_1, \underline{\boldsymbol{v}}_2, \ \dots \underline{\boldsymbol{v}}_{n-p}\} \qquad \text{be a basis for } Nul\ A.$$

Then we can check that set of n vectors obtained by taking the union of the two sets,

$$\{\underline{\boldsymbol{u}}_1, \underline{\boldsymbol{u}}_2, \ \dots \underline{\boldsymbol{u}}_p, \underline{\boldsymbol{v}}_1, \underline{\boldsymbol{v}}_2, \ \dots \underline{\boldsymbol{v}}_{n-p}\}$$

is actually a basis for $\mathbb{R}^n$. This is because we can show that the n vectors in the set are linearly independent, so they automatically span $\mathbb{R}^n$ and are a basis: To check independence,, let

$$c_1\underline{\boldsymbol{u}}_1 + c_2\underline{\boldsymbol{u}}_2 + \ \dots + c_p\underline{\boldsymbol{u}}_p + d_1\underline{\boldsymbol{v}}_1 + d_2\underline{\boldsymbol{v}}_2 + \ \dots + d_{n-p}\underline{\boldsymbol{v}}_{n-p} = \underline{\mathbf{0}}.$$

then

$$c_1\underline{\boldsymbol{u}}_1 + c_2\underline{\boldsymbol{u}}_2 + \ \dots + c_p\underline{\boldsymbol{u}}_p = -d_1\underline{\boldsymbol{v}}_1 - d_2\underline{\boldsymbol{v}}_2 - \ \dots - d_{n-p}\underline{\boldsymbol{v}}_{n-p}.$$

Since the vector on the left is in *Row A* and the one that it equals on the right is in *Nul A*, this vector is the zero vector:

$$c_1\underline{\boldsymbol{u}}_1 + c_2\underline{\boldsymbol{u}}_2 + \ \dots + c_p\underline{\boldsymbol{u}}_p = \underline{0} = -d_1\underline{\boldsymbol{v}}_1 - d_2\underline{\boldsymbol{v}}_2 - \ \dots - d_{n-p}\underline{\boldsymbol{v}}_{n-p}.$$

Since $\{\underline{\boldsymbol{u}}_1, \underline{\boldsymbol{u}}_2, \ \dots \underline{\boldsymbol{u}}_p\}$ and $\{\underline{\boldsymbol{v}}_1, \underline{\boldsymbol{v}}_2, \ \dots \underline{\boldsymbol{v}}_{n-p}\}$ are linearly independent sets, we deduce from these two equations that

$$c_1 = c_2 = \dots = c_p = 0, \qquad d_1 = d_2 = \dots = d_{n-p} = 0\ .$$

Q.E.D.

So the picture on the previous page is completely general, also for the decomposition of the codomain. One can check that the transformation $T(\underline{x}) = A\,\underline{\mathbf{x}}$ restricts to an isomorphism from *Row A* to *Col A*, because it is $1 - 1$ on these subspaces of equal dimension, so must also be onto. So, T squashes *Nul A*, and maps every translation of *Nul A* to a point in *Col A*. More precisely, Each

$$\underline{\boldsymbol{x}} \in \mathbb{R}^n$$

can be written uniquely as

$$\underline{\boldsymbol{x}} = \underline{\boldsymbol{u}} + \underline{\boldsymbol{v}} \qquad \text{with } \underline{\boldsymbol{u}} \in Row\ A,\ \underline{\boldsymbol{v}} \in Nul\ A.$$

and

$$T(\underline{\boldsymbol{u}} + \underline{\boldsymbol{v}}) = T(\underline{\boldsymbol{u}}) + T(\underline{\boldsymbol{v}}) = T(\underline{u}) \in Col\ (A).$$

As sets,

$$T(\{\underline{\boldsymbol{u}} + Nul\ A\}) = T(\underline{\boldsymbol{u}}).$$

Wed Mar 7

- 4.7 Change of basis

Announcements:

Warm-up Exercise:

The setup: Let V be a finite dimensional vector space, with two bases,

$$B = \{\underline{\boldsymbol{b}}_1, \underline{\boldsymbol{b}}_2, \; .... \; \underline{\boldsymbol{b}}_n\}$$
$$C = \{\underline{\boldsymbol{c}}_1, \underline{\boldsymbol{c}}_2, \; .... \; \underline{\boldsymbol{c}}_n\}$$

How do we change from the coordinate system of the B basis to that of the C basis? If we can express the B vectors in terms of the C vectors it's straightforward:

Example Let $B = \{\underline{\boldsymbol{b}}_1, \underline{\boldsymbol{b}}_2\}$, $C = \{\underline{\boldsymbol{c}}_1, \underline{\boldsymbol{c}}_2\}$ be bases for the two-dimensional vector space V. Suppose

$$\underline{\boldsymbol{b}}_1 = 4\,\underline{\boldsymbol{c}}_1 + \underline{\boldsymbol{c}}_2$$
$$\underline{\boldsymbol{b}}_2 = -6\,\underline{\boldsymbol{c}}_1 + \underline{\boldsymbol{c}}_2.$$

Let $[\underline{\boldsymbol{v}}]_B = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$. Find $[\underline{\boldsymbol{v}}]_C$.

Solution:

$$\underline{\boldsymbol{v}} = x_1\,\underline{\boldsymbol{b}}_1 + x_2\,\underline{\boldsymbol{b}}_2$$

$$\Rightarrow [\underline{\boldsymbol{v}}]_C = [x_1\,\underline{\boldsymbol{b}}_1 + x_2\,\underline{\boldsymbol{b}}_2]_C$$

$$= x_1\,[\underline{\boldsymbol{b}}_1]_C + x_2\,[\underline{\boldsymbol{b}}_2]_C$$

$$[\underline{\boldsymbol{v}}]_C = \left[[\underline{\boldsymbol{b}}_1]_C \; [\underline{\boldsymbol{b}}_2]_C\right] \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$$

$$[\underline{\boldsymbol{v}}]_C = \begin{bmatrix} 4 & -6 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}.$$

Note that the coordinate transition matrix $P_{C \leftarrow B}$ would always given by

$$\left[[\underline{\boldsymbol{b}}_1]_C \; [\underline{\boldsymbol{b}}_2]_C\right].$$

no matter what the particular coordinate vectors $[\underline{\boldsymbol{b}}_1]_C$, $[\underline{\boldsymbol{b}}_2]_C$ are.

Exercise 1 Consider $V = \{a + b\,t\}$, the space of polynomials in t of degree ≤ 1. Let $C = \{1, t\}$ be the "standard basis". and let $B = \{1 + t, 1 - t\}$. Be an alternate basis.

1a) Find the transition matrix $P_{C \leftarrow B}$.

1b) Suppose $q(t) \in V$ with $[q]_B = \begin{bmatrix} 2 \\ 3 \end{bmatrix}$. Find $[q]_C$ by using the transition matrix $P_{C \leftarrow B}$. Compare to the direct method (which should be just as easy in this simple case).

1c) The transition matrix in the reverse direction must be the inverse of the original transition matrix. Find $P_{B \leftarrow C} = (P_{C \leftarrow B})^{-1}$.

1d) Suppose $r(t) = 1 + 7\,t$. Find $[r]_B$ and check your work.

Change of coordinate transition matrices work the same in every dimension.

Let V be a finite dimensional vector space, with two bases,

$$B = \{\underline{\boldsymbol{b}}_1, \underline{\boldsymbol{b}}_2, \; .... \underline{\boldsymbol{b}}_n\}$$
$$C = \{\underline{\boldsymbol{c}}_1, \underline{\boldsymbol{c}}_2, \; .... \underline{\boldsymbol{c}}_n\}.$$

Then for

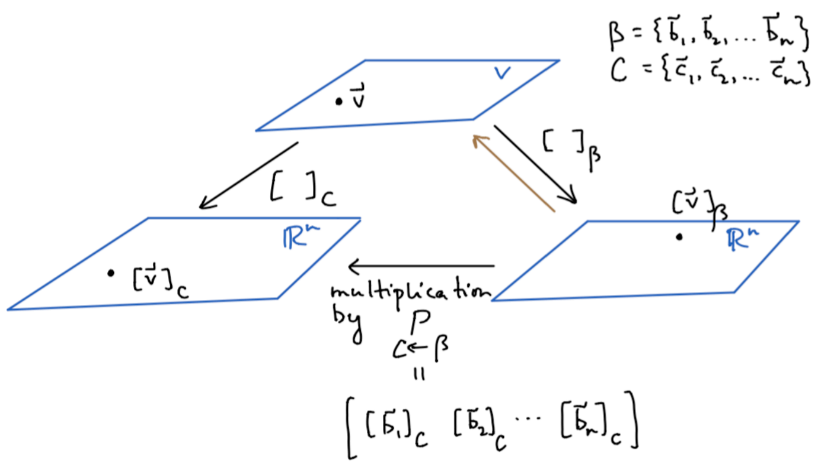
$$\underline{\boldsymbol{v}} = x_1\, \underline{\boldsymbol{b}}_1 + x_2\, \underline{\boldsymbol{b}}_2 + \ldots + x_n\, \underline{\boldsymbol{b}}_n$$

$$[\underline{\boldsymbol{v}}]_C = [\, x_1\, \underline{\boldsymbol{b}}_1 + x_2\, \underline{\boldsymbol{b}}_2 + \ldots + x_n\, \underline{\boldsymbol{b}}_n]_C$$

$$[\underline{\boldsymbol{v}}]_C = x_1\, [\underline{\boldsymbol{b}}_1]_C + x_2\, [\underline{\boldsymbol{b}}_2]_C + \;\ldots\; + x_n\, [\underline{\boldsymbol{b}}_n]_C$$

$$[\underline{\boldsymbol{v}}]_C = \left[\, [\underline{\boldsymbol{b}}_1]_C \; [\underline{\boldsymbol{b}}_2]_C \cdots [\underline{\boldsymbol{b}}_n]_C \,\right] \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}.$$

$$P_{C \leftarrow B} = \left[\, [\underline{\boldsymbol{b}}_1]_C \; [\underline{\boldsymbol{b}}_2]_C \cdots [\underline{\boldsymbol{b}}_n]_C \,\right].$$



A special case of change of coordinates is when the vector space V is $\mathbb{R}^n$ itself. In that case there are two ways to find the coordinate transition matrices. Let

$$B = \{\underline{b}_1, \underline{b}_2, \; .... \underline{b}_n\}$$
$$C = \{\underline{c}_1, \underline{c}_2, \; .... \underline{c}_n\}.$$

be two bases for $\mathbb{R}^n$.

Method 1: Let

$$E = \{\underline{e}_1, \underline{e}_2, \; .... \underline{e}_n\}$$

be the standard basis. As we discussed previously and as a special case of our current discussion, since for $\underline{v} \in \mathbb{R}^n$, $[\underline{v}]_E = \underline{v}$,

$$P_{E \leftarrow B} = [\underline{b}_1, \underline{b}_2, \; .... \underline{b}_n]$$

$$P_{E \leftarrow C} = [\underline{c}_1, \underline{c}_2, \; .... \underline{c}_n].$$

Since composition of matrix transformations corresponds to matrix multiplication, the transition matrix from B to C coordinates can be computed via the standard coordinate transiton matrices:

$$P_{C \leftarrow B} = P_{C \leftarrow E} P_{E \leftarrow B} = (P_{E \leftarrow C})^{-1} P_{E \leftarrow B}$$

Method 2: Direct method. We know

$$P_{C \leftarrow B} = [[\underline{b}_1]_C \; [\underline{b}_2]_C \cdots [\underline{b}_n]_C].$$

Consider the columns of the transition matrix as unknowns - as when we were finding the columns of inverses matrices by a multi-augmented matrix procedure to solve $A X = I$. In this case, and illustrating with $n = 2$ for simplicity,

$$P_{C \leftarrow B} = [[\underline{b}_1]_C \; [\underline{b}_2]_C].$$

The first column

$$[\underline{b}_1]_C = \begin{bmatrix} y_1 \\ y_2 \end{bmatrix}$$

satisfies

$$y_1 \underline{c}_1 + y_2 \underline{c}_2 = \underline{b}_1$$

$$[\underline{\boldsymbol{c}}_1\ \underline{\boldsymbol{c}}_2]\begin{bmatrix} y_1 \\ y_2 \end{bmatrix} = \underline{\boldsymbol{b}}_1$$

and the second column

$$[\underline{\boldsymbol{b}}_2]_C = \begin{bmatrix} z_1 \\ z_2 \end{bmatrix}$$

satisfies

$$[\underline{\boldsymbol{c}}_1\ \underline{\boldsymbol{c}}_2]\begin{bmatrix} z_1 \\ z_2 \end{bmatrix} = \underline{\boldsymbol{b}}_2$$

We solve for the two columns with a double augmented matrix reduction:

$$\left[\begin{array}{cc|cc} \underline{\boldsymbol{c}}_1 & \underline{\boldsymbol{c}}_2 & \underline{\boldsymbol{b}}_1 & \underline{\boldsymbol{b}}_2 \end{array}\right] \to \begin{bmatrix} 1 & 0 & y_1 & z_1 \\ 0 & 1 & y_2 & z_2 \end{bmatrix}$$

(And this generalizes to $\mathbb{R}^n$.)

Exercise 2 Test the two methods for finding

$$P_{C \leftarrow B}$$

where

$$B = \left\{ \begin{bmatrix} -3 \\ 1 \end{bmatrix}, \begin{bmatrix} 2 \\ -1 \end{bmatrix} \right\} \qquad C = \left\{ \begin{bmatrix} 1 \\ -1 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \end{bmatrix} \right\}$$

Fri Mar 2

- 5.1-5.2 Eigenvectors and eigenvalues for square matrices

Announcements:

Warm-up Exercise:

Eigenvalues and eigenvectors for square matrices.

To introduce the idea of eigenvalues and eigenvectors we'll first think geometrically.

Example   Consider the matrix transformation $T : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ with formula

$$T\left(\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}\right) = \begin{bmatrix} 3 & 0 \\ 0 & 1 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = x_1\begin{bmatrix} 3 \\ 0 \end{bmatrix} + x_2\begin{bmatrix} 0 \\ 1 \end{bmatrix}.$$

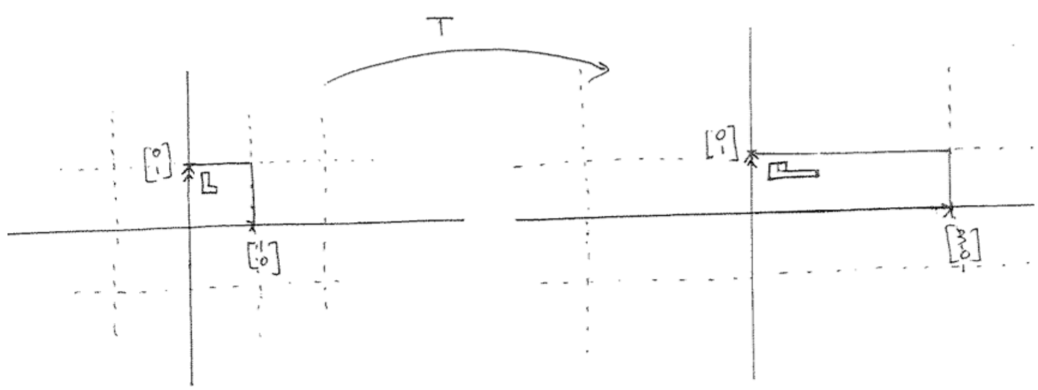
Notice that for the standard basis vectors $\underline{e}_1 = [1, 0]^T, \underline{e}_2 = [0, 1]^T$

$$T(\underline{e}_1) = 3\underline{e}_1$$
$$T(\underline{e}_2) = \underline{e}_2 \ .$$

The facts that T is linear and that it transforms $\underline{e}_1, \underline{e}_2$ by scalar multiplying them, lets us understand the geometry of this transformation completely:

$$T\left(\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}\right) = T(x_1\underline{e}_1 + x_2\underline{e}_2) = x_1 T(\underline{e}_1) + x_2 T(\underline{e}_2)$$
$$= x_1(3\underline{e}_1) + x_2(1\underline{e}_2) \ .$$

In other words, T stretches by a factor of 3 in the $\underline{e}_1$ direction, and by a factor of 1 in the $\underline{e}_2$ direction, transforming a square grid in the domain into a parallel rectangular grid in the image:



Definition: If $A_{n \times n}$ and if $A\,\underline{\boldsymbol{v}} = \lambda\,\underline{\boldsymbol{v}}$ for a scalar λ and a vector $\underline{\boldsymbol{v}} \neq \underline{\mathbf{0}}$ then $\underline{\boldsymbol{v}}$ is called an eigenvector of A, and λ is called the eigenvalue of $\underline{\boldsymbol{v}}$. (In some texts the words characteristic vector and characteristic value are used as synonyms for these words.)

• In the example above, the standard basis vectors (or multiples of them) are eigenvectors, and the corresponding eigenvalues are the diagonal matrix entries. A non-diagonal matrix may still have eigenvectors and eigenvalues, and this geometric information can still be important to find. But how do you find eigenvectors and eigenvalues for non-diagonal matrices? ...

Exercise 2) Try to find eigenvectors and eigenvalues for the non-diagonal matrix, by just trying random input vectors $\underline{\boldsymbol{x}}$ and computing $A\,\underline{\boldsymbol{x}}$.

$$A = \begin{bmatrix} 3 & 2 \\ 1 & 2 \end{bmatrix}.$$

How to find eigenvalues and eigenvectors (including eigenspaces) systematically:

If

$$A\,\underline{\boldsymbol{v}} = \lambda\,\underline{\boldsymbol{v}}$$

$$\Leftrightarrow A\,\underline{\boldsymbol{v}} - \boldsymbol{\lambda}\,\underline{\boldsymbol{v}} = \underline{\boldsymbol{0}}$$

$$\Leftrightarrow A\,\underline{\boldsymbol{v}} - \lambda\,I\,\underline{\boldsymbol{v}} = \underline{\boldsymbol{0}}$$

where I is the identity matrix.

$$\Leftrightarrow (A - \lambda\,I)\underline{\boldsymbol{v}} = \underline{\boldsymbol{0}}\,.$$

As we know, this last equation can have non-zero solutions $\underline{\boldsymbol{v}}$ if and only if the matrix $(A - \lambda\,I)$ is not invertible, i.e.

$$\Leftrightarrow det(A - \lambda\,I) = 0\,.$$

So, to find the eigenvalues and eigenvectors of matrix you can proceed as follows:

- Compute the polynomial in λ

$$p(\lambda) = det(A - \lambda\,I)\,.$$

If $A_{n\,\times\,n}$ then $p(\lambda)$ will be degree n. This polynomial is called the <u>characteristic polynomial</u> of the matrix A.

- λ_j can be an eigenvalue for some non-zero eigenvector $\underline{\boldsymbol{v}}$ if and only if it's a root of the characteristic polynomial, i.e. $p(\lambda_j) = 0$. For each such root, the homogeneous solution space of vectors $\underline{\boldsymbol{v}}$ solving

$$(A - \lambda_j\,I)\underline{\boldsymbol{v}} = \underline{\boldsymbol{0}}$$

will be eigenvectors with eigenvalue λ_j. This subspace of eigenvectors will be at least one dimensional, since $(A - \lambda_j\,I)$ does not reduce to the identity and so the explicit homogeneous solutions will have free parameters. Find a basis of eigenvectors for this subspace. Follow this procedure for each eigenvalue, i.e. for each root of the characteristic polynomial.

<u>Notation</u>: The subspace of eigenvectors for eigenvalue λ_j is called the <u>λ_j eigenspace</u>, and we'll denote it by $E_{\lambda = \lambda_j}$. The basis of eigenvectors is called an <u>eigenbasis</u> for E_{λ_j}.

Exercise 3) a) Use the systematic algorithm to find the eigenvalues and eigenbases for the non-diagonal matrix of Exercise 2.

$$A = \begin{bmatrix} 3 & 2 \\ 1 & 2 \end{bmatrix}.$$

b) Use your work to describe the geometry of the linear transformation in terms of directions that get stretched:

$$T\left(\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}\right) = \begin{bmatrix} 3 & 2 \\ 1 & 2 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}.$$

Exercise 4) Find the eigenvalues and eigenspace bases for

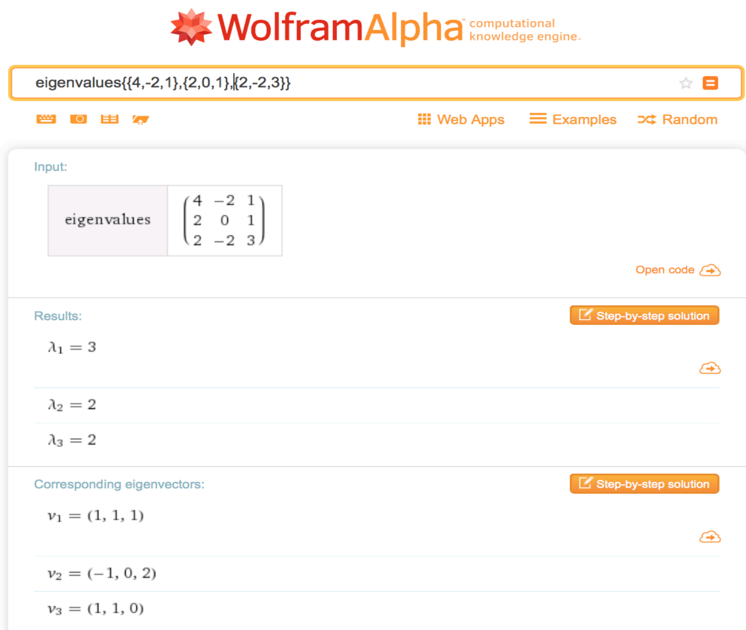
$$B := \begin{bmatrix} 4 & -2 & 1 \\ 2 & 0 & 1 \\ 2 & -2 & 3 \end{bmatrix}.$$

(i) Find the characteristic polynomial and factor it to find the eigenvalues.

(ii) for each eigenvalue, find bases for the corresponding eigenspaces.

(iii) Can you describe the transformation $T(\underline{\mathbf{x}}) = B\underline{\mathbf{x}}$ geometrically using the eigenbases? Does $det(B)$ have anything to do with the geometry of this transformation?

Your solution will be related to the output below:



In all of our examples so far, it turns out that by collecting bases from each eigenspace for the matrix $A_{n \times n}$, and putting them together, we get a basis for $\mathbb{R}^n$. This lets us understand the <u>geometry</u> of the transformation

$$T(\underline{\mathbf{x}}) = A\,\underline{\mathbf{x}}$$

almost as well as if A is a diagonal matrix. This is actually something that does not always happen for a matrix A. When it does happen, we say that A is <u>diagonalizable.</u> Here's an example of a matrix which is NOT diagonalizable:

<u>Exercise 5</u>: Find matrix eigenvalues and eigenspace basis for each eigenvalue, for

$$A = \begin{bmatrix} 3 & 2 \\ 0 & 3 \end{bmatrix}.$$

Explain why there is no basis of $\mathbb{R}^2$ consisting of eigenvectors of A.