2.1 Different Faces of Linear Systems

9/27

An $m \times n$ system of linear equations (or an $m \times n$ linear system) is a collection of m linear equations in n unknowns $x_1, \to x_m$:

$$\begin{cases} a_{1,1}x_1 + a_{1,2}x_2 + \cdots + a_{1,n}x_n = b_1 \\ a_{2,1}x_1 + a_{2,1}x_2 + \cdots + a_{2,n}x_n = b_2 \\ \vdots \qquad\qquad\qquad\qquad\qquad \vdots \\ a_{m,1}x_1 + a_{m,2}x_2 + \cdots + a_{m,n}x_n = b_m \end{cases}$$

Solving the system means finding all n-tuples of numbers x_1, x_n that solve all m equations simultaneously (or showing none exist).

- Using vector and matrix notation we can compress this system: let $\vec{x} = (x_1, \to x_n)^T$, $\vec{b} = (b_1, \dots, b_m)^T$, and let $\underbrace{A = (a_{i,j})}_{\text{coefficient matrix}} \in M_{m \times n}$. Then the linear system is equivalent to

$$A\vec{x} = \vec{b}$$

Solving this means finding all vectors $\vec{x}$ satisfying the above (or showing none exist).
If we know A is invertible and can compute A^{-1}, solving will be pretty easy.
But in fact, solving the system will teach us how to compute A^{-1} (if it exists).

- Denote the columns of A by $\vec{a}_1, \dots, \vec{a}_n$. So

$$\vec{a}_j = \begin{pmatrix} a_{1,j} \\ a_{2,j} \\ \vdots \\ a_{m,j} \end{pmatrix}$$

Then recall

$$\begin{aligned} A\vec{x} &= \vec{a}_1 x_1 + \vec{a}_2 x_2 + \cdots + \vec{a}_n x_n \\ &= x_1\vec{a}_1 + x_2\vec{a}_2 + \cdots + x_n\vec{a}_n. \longleftarrow \text{lin. comb. of } \vec{a}_1, \dots, \vec{a}_n \end{aligned}$$

So solving $A\vec{x} = \vec{b}$ means writing $\vec{b}$ as a lin. comb of $\vec{a}_1, \dots, \vec{a}_n$. So knowing whether or not the system $\vec{a}_1, \dots, \vec{a}_n$ is generating or linearly independent will be useful.

- These three viewpoints offer different perspectives on the same problem. They will lead to different insights and strategies for solving linear systems.

- Observe that in order to solve the system, we only need the information stored in A and $\vec{b}$. We store them together in an augmented matrix:

$$(A \mid \vec{b}) = \left(\begin{array}{cccc|c} a_{1,1} & a_{1,2} & \cdots & a_{1,n} & b_1 \\ \vdots & \vdots & & \vdots & \vdots \\ a_{m,1} & a_{m,2} & \cdots & a_{m,n} & b_m \end{array}\right)$$

2.2 Solution of a Linear System

EX Solve
$$\begin{cases} 2x_1 - x_2 = 1 \\ x_1 + x_2 = 2 \end{cases} \longrightarrow \begin{cases} 2x_1 - x_2 = 1 \\ 3x_1 = 3 \end{cases} \longrightarrow \begin{cases} 2x_1 - x_2 = 1 \\ x_1 = 1 \end{cases}$$

$$\longrightarrow \begin{cases} - x_2 = -1 \\ x_1 = 1 \end{cases} \longrightarrow \begin{cases} x_2 = 1 \\ x_1 = 1 \end{cases} \longrightarrow \begin{cases} x_1 = 1 \\ x_2 = 1 \end{cases}$$

Using just the augmented matrix:

$$\left(\begin{array}{cc|c} 2 & -1 & 1 \\ 1 & 1 & 2 \end{array}\right) \xrightarrow[R_2 \mapsto R_2 + R_1]{} \left(\begin{array}{cc|c} 2 & -1 & 1 \\ 3 & 0 & 3 \end{array}\right) \xrightarrow[R_2 \mapsto \frac{1}{3} R_2]{} \left(\begin{array}{cc|c} 2 & -1 & 1 \\ 1 & 0 & 1 \end{array}\right)$$

$$\xrightarrow{R_1 \mapsto R_1 - 2R_2} \left(\begin{array}{cc|c} 0 & -1 & -1 \\ 1 & 0 & 1 \end{array}\right) \xrightarrow{R_1 \mapsto -R_1} \left(\begin{array}{cc|c} 0 & 1 & 1 \\ 1 & 0 & 1 \end{array}\right) \xrightarrow{R_1 \leftrightarrow R_2} \left(\begin{array}{cc|c} 1 & 0 & 1 \\ 0 & 1 & 1 \end{array}\right) \Rightarrow \begin{array}{l} x_1 = 1 \\ x_2 = 1 \end{array} \ \square$$

The method we used here is called "Gauss-Jordan elimination" or more commonly "row reduction."

Row Operations: The following operations preserve the solutions to a linear system:

1. interchange two rows (row exchange)
2. multiply a row by a non-zero scalar (scaling)
3. replace a row by its sum with a scalar multiple of another row (row replacement)

- The reason why these operations preserve solutions is because they can be "undone" or inverted. In fact implementing each of these operations is equivalent to multiplying the coefficient matrix A by an invertible matrix on the left.

Elementary Matrices

These are matrices three types each corresponding to one of the row operations.

1. For
$$E = \begin{array}{c|ccccc|} & & j & & k & \\ & 1 \ddots 1 & & & & 0 \\ j & & 0 & & 1 & \\ & & & 1 \ddots 1 & & \\ k & & 1 & & 0 & \\ & 0 & & & & 1 \ddots 1 \end{array}$$
EA is the same as A, but with row j and row k interchanged
AE ——— 〃 ——— column j and column k ——〃——

Note that $E^2 = I_m$, so E is invertible with inverse E.

2. For
$$E = \begin{array}{c|ccc|} & & j & \\ & 1 \ddots 1 & & 0 \\ j & & a & \\ & 0 & & 1 \ddots 1 \end{array},$$
EA is the same as A but with row j multiplied by a.
AE ——— 〃 ——— column j ——〃——

Note that $E^{-1} = \begin{pmatrix} 1 \ddots 1 & & 0 \\ & a^{-1} & \\ 0 & & 1 \ddots 1 \end{pmatrix}$ exists so long as $a \neq 0$

3. For $E = \begin{array}{cccccc} & & j & & k & \\ & 1 & & & & 0 \\ j & & 1 & \cdots & 0 & \\ & & \vdots & \ddots & \vdots & \\ k & & a & \cdots & 1 & \\ & 0 & & & & 1 \end{array}$ EA is the same as A but with row k replaced by its sum with $a \cdot (\text{row}_j)$

AE ——— " ——— column k ——— " ——— $a\,(\text{column}_j)$

Note that $E^{-1} = \begin{array}{cccccc} & & j & & k & \\ & 1 & & & & 0 \\ j & & 1 & & 0 & \\ & & \vdots & \ddots & & \\ k & & -a & & 1 & \\ & 0 & & & & 1 \end{array}$ exists regardless of a.

Since all of these elementary matrices E are invertible,

$$A\vec{x} = \vec{b} \iff EA\vec{x} = E\vec{b}$$

$$\underset{\xleftarrow{\quad\quad} E^{-1}x}{Ex \xrightarrow{\quad\quad}}$$

Row Reduction Strategy

The general row reduction strategy is to iterate the following process:

1. Find the leftmost non-zero column of a matrix
2. Use row exchange to make sure the top entry of this column is non-zero. This entry is called the pivot
3. (Using scaling to make the pivot equal 1.)
4. Using row replacement to make all entries below the pivot equal to zero.

After applying this process to a matrix, we leave the pivot's row alone and apply the process to only the rows below (which will create a new pivot). Iterating until we run out of rows, we obtain a matrix in row echelon form (REF)

Ex

$$\begin{cases} 2x_1 + 4x_2 + 6x_3 = 2 \\ x_1 + x_2 - x_3 = 6 \\ 2x_1 + x_2 + 2x_3 = 1 \end{cases} \longrightarrow \left(\begin{array}{ccc|c} \overset{\text{pivot}}{\boxed{2}} & 4 & 6 & 2 \\ 1 & 1 & -1 & 6 \\ 2 & 1 & 2 & 1 \end{array}\right) \xrightarrow{R_1 \mapsto \frac{1}{2}R_1} \left(\begin{array}{ccc|c} 1 & 2 & 3 & 1 \\ 1 & 1 & -1 & 6 \\ 2 & 1 & 2 & 1 \end{array}\right)$$

$$\xrightarrow[R_3 \mapsto R_3 - 2R_1]{R2 \mapsto R_2 - R1} \left(\begin{array}{ccc|c} 1 & 2 & 3 & 1 \\ 0 & \underset{\text{new pivot}}{\boxed{-1}} & -4 & 5 \\ 0 & -3 & -4 & -1 \end{array}\right) \xrightarrow{R2 \mapsto -1R2} \left(\begin{array}{ccc|c} 1 & 2 & 3 & 1 \\ 0 & 1 & 4 & -5 \\ 0 & -3 & -4 & -1 \end{array}\right) \xrightarrow[R_3 \mapsto R_3 + 3R_2]{} \left(\begin{array}{ccc|c} 1 & 2 & 3 & 1 \\ 0 & 1 & 4 & -5 \\ 0 & 0 & \underset{\text{new pivot}}{\boxed{8}} & -16 \end{array}\right)$$

$$\xrightarrow{R_3 \mapsto \frac{1}{8}R_3} \left(\begin{array}{ccc|c} 1 & 2 & 3 & 1 \\ 0 & 1 & 4 & -5 \\ 0 & 0 & 1 & -2 \end{array}\right) \iff \begin{cases} x_1 + 2x_2 + 3x_3 = 1 \\ \quad\quad x_2 + 4x_3 = -5 \\ \quad\quad\quad\quad x_3 = -2 \end{cases}$$

We now use a method called back substitution. The last equation says $x_3 = -2$. Substituting this into the second equation gives:

$$x_2 + 4(-2) = -5$$
$$x_2 - 8 = -5$$
$$x_2 = 3$$

Finally, substituting $x_2 = 3$ and $x_3 = -2$ into the first equation gives:

$$x_1 + 2(3) + 3(-2) = 1$$
$$x_1 + 6 - 6 = 1$$
$$x_1 = 1.$$

So the solution to the system is $\begin{cases} x_1 = 1 \\ x_2 = 3 \\ x_3 = -2. \end{cases}$ or $\vec{x} = \begin{pmatrix} 1 \\ 3 \\ -2 \end{pmatrix}$

Instead of doing back sub., we could also keep working the matrix, and now eliminate any non-zero entries above the pivots:

$$\left(\begin{array}{ccc|c} 1 & 2 & 3 & 1 \\ 0 & 1 & 4 & -5 \\ 0 & 0 & 1 & -2 \end{array}\right) \xrightarrow[R_2 \mapsto R_2 - 4R_3]{R_1 \mapsto R_1 - 3R_3} \left(\begin{array}{ccc|c} 1 & 2 & 0 & 7 \\ 0 & 1 & 0 & 3 \\ 0 & 0 & 1 & -2 \end{array}\right) \xrightarrow{R_1 \mapsto R_1 - 2R_2} \left(\begin{array}{ccc|c} 1 & 0 & 0 & 1 \\ 0 & 1 & 0 & 3 \\ 0 & 0 & 1 & -2 \end{array}\right) \longleftrightarrow \begin{cases} x_1 = 1 \\ x_2 = 3 \\ x_3 = -2. \end{cases}$$

$$\underbrace{\begin{matrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{matrix}}_{I_3}$$

The matrix we obtain after doing this is in reduced row echelon form (RREF)

Def A matrix is in row echelon form if

1. Any rows of all zeros are at the bottom.
2. The leading non-zero entry of a row (i.e. the pivot) is to the right of a the leading non-zero entry in the row above it.

It is in reduced row echelon form if it further satisfies:

3. All pivots equal 1.
4. Each pivot is the only non-zero entry in its column.

Ex

$$A = \left(\begin{array}{ccccc|c} \boxed{2} & 1 & 2 & 0 & 3 & -1 \\ 0 & 0 & \boxed{1} & 4 & -2 & 0 \\ 0 & 0 & 0 & 0 & \boxed{1} & 3 \end{array}\right) \qquad B = \left(\begin{array}{ccccc|c} \boxed{1} & 2 & 0 & 0 & 0 & 3 \\ 0 & 0 & \boxed{1} & 4 & 0 & -2 \\ 0 & 0 & 0 & 0 & \boxed{1} & 1 \end{array}\right)$$

Both A and B are in REF, but only B is in RREF. B corresponds to the lin. system:

9/26

$$\begin{cases} x_1 + 2x_2 = 3 \\ x_3 + 4x_4 = -2 \\ x_5 = 1 \end{cases}$$

We see $x_5 = 1$, but for the other variables we can only say:

$$x_1 = 3 - 2x_2$$
$$x_3 = -2 - 4x_4$$

We call x_2 and x_4 free variables. These are the variables corresponding to any non-zero column without a pivot. In this case our solution would be:

$$\vec{x} = \begin{pmatrix} 3 - 2x_2 \\ x_2 \\ -2 - 4x_4 \\ x_4 \\ 1 \end{pmatrix} \quad x_2, x_3 \in \mathbb{F} \qquad \text{or} \quad \vec{x} = \begin{pmatrix} 3 \\ 0 \\ -2 \\ 0 \\ 1 \end{pmatrix} + x_2 \begin{pmatrix} -2 \\ 1 \\ 0 \\ 0 \\ 0 \end{pmatrix} + x_4 \begin{pmatrix} 0 \\ 0 \\ -4 \\ 1 \\ 0 \end{pmatrix} \quad x_2, x_4 \in \mathbb{F}.$$

$\square$

2.3 Analyzing the Pivots

Ex The matrix

$$\left(\begin{array}{ccc|c} 1 & 0 & 0 & 2 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 0 & 2 \end{array}\right)$$

is in RREF, but corresponds to a lin. system with no solutions. Indeed, the last row gives the equation $0 = 2$, which is always false. On the other hand

$$\left(\begin{array}{ccc|c} 1 & 0 & 0 & 2 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{array}\right)$$

has solutions $\vec{x} = \begin{pmatrix} 2 \\ 1 \\ x_3 \end{pmatrix}$ for $x_3 \in \mathbb{F}$. □

Def We say a linear system is inconsistent if it has no solutions. This happens if and only if the row echelon form of the augmented matrix has a row of the form:

$$(0\ 0 \cdots 0 \mid b)$$

with $b \neq 0$. We say the linear system is consistent if it has at least one solution.

- Note that being inconsistent is equivalent to there being a pivot in the last column of the augmented matrix.

Observations (About coefficient matrix)

① A solution (if it exists) is unique iff there are no free variables. Equivalently iff the row echelon form of the coefficient matrix has a pivot in every column

Pf Assuming a solution exists, free variables precisely determine whether or not it is unique. And they exist precisely when a column lacks a pivot. □

② $A\vec{x} = \vec{b}$ is consistent for all right sides $\vec{b}$ iff the row echelon form of the coefficient matrix has a pivot in every row

Pf (⇐): This means there cannot be a pivot in the last column of the augmented matrix. Hence the system is consistent.

(⇒): We use proof by contrapositive. Assume the row echelon form A_e of A does not have a pivot in every row. Then A_e has a row of all zeros at the bottom. Set $\vec{b}_e = (0, \dots, 0, 1)^T$, then

$$\ast \qquad A_e \vec{x} = \vec{b}_e$$

is inconsistent. Since A_e is the row echelon form of A, $A_e = EA$ for $E = E_1 \cdots E_d$ a product of elementary matrices. Therefore E is invertible and so $(\ast)$ is equivalent to $A\vec{x} = E^{-1}\vec{b}_e$, and in particular is inconsistent. □

③ $A\vec{x} = \vec{b}$ has a unique solution for any right side $\vec{b}$ if and only if the row echelon form of the coefficient matrix A has a pivot in every column and row.

Pf This follows from combining 1 and 2 above. □

Corollaries about linear independence

Prop Let $\vec{v}_1, \dots, \vec{v}_n \in \mathbb{F}^m$ be a system of vectors and define

$$A := (\vec{v}_1 \vec{v}_2 \cdots \vec{v}_n) \in M_{m \times n}$$

Then:

① The system $\vec{v}_1, \dots, \vec{v}_n$ is linearly independent if and only if the row echelon form of A has a pivot in every column.

② The system $\vec{v}_1, \dots, \vec{v}_n$ is generating for $\mathbb{F}^m$ if and only if the row echelon form of A has a pivot in every row.

③ The system $\vec{v}_1, \dots, \vec{v}_n$ is a basis for $\mathbb{F}^m$ if and only if the row echelon form of A has a pivot in every column and every row.

Proof (1): $\vec{v}_1, \dots, \vec{v}_n$ is lin. indep. iff the equation

$$x_1 \vec{v}_1 + \cdots + x_n \vec{v}_n = \vec{0}$$

has $\vec{x} = \vec{0}$ as its unique solution. This is equivalent to $A\vec{x} = \vec{0}$ having the unique solution $\vec{x} = \vec{0}$. By Observation 1 above, this is equivalent to the row echelon form of A having a pivot in every column.

(2): $\vec{v}_1, \dots, \vec{v}_m$ is generating iff the equation

$$x_1 \vec{v}_1 + \cdots + x \vec{v}_n = \vec{b}$$

$$A\vec{x} = \vec{b}$$

has a solution for every right hand side $\vec{b}$. By Observation 2 above, this is equivalent to the row echelon form of A having a pivot in every row.

(3): $\vec{v}_1, \dots, \vec{v}_n$ is a basis iff it is generating and lin. indep. So we simply apply 1 and 2. □

Prop Any linearly independent system of vectors in $\mathbb{F}^m$ cannot have more than m vectors in it.

Proof Let $\vec{v}_1, \dots, \vec{v}_n \in \mathbb{F}^m$ be a lin. indep. system. Then for

$$A := (\vec{v}_1 \cdots \vec{v}_n) \in M_{m \times n},$$

the row echelon form of A has a pivot in every column (by the previous Prop). Thus there are n pivots, but each row can have at most one pivot. Since there are only m rows, we must have $n \le m$ $\square$

Prop (A second proof) Any two bases in a vector space V have the same number of vectors in them.

Proof Suppose $\vec{v}_1, \dots, \vec{v}_n$ and $\vec{w}_1, \dots, \vec{w}_m$ are both bases for V. Without loss of generality (WLOG) we may assume $n \le m$. So it suffices to prove $n \ge m$. Define an isomorphism $A : \mathbb{F}^n \to V$ by

$$A\vec{e}_k = \vec{v}_k \quad \text{for } k = 1, \dots, n.$$

Then A^{-1} is also an isomorphism, and so by Exercise 1 on Homework 4,

$$A^{-1}(\vec{w}_1), \dots, A^{-1}(\vec{w}_m)$$

is a basis for $\mathbb{F}^n$. In particular, the system is lin. indep. and so the previous Prop. implies $m \le n$, as desired. $\square$

Prop Any generating system in $\mathbb{F}^m$ must have at least m vectors.

Proof Let $\vec{v}_1, \dots, \vec{v}_n$ be a generating system in $\mathbb{F}^m$, and define

$$A := (\vec{v}_1 \cdots \vec{v}_n) \in M_{m \times n}.$$

By part 2 of the first Proposition in this section, the row echelon form of A has a pivot in every row. So there are m pivots, and since each column can contain at most one pivot we must have $m \le n$. $\square$

Corollaries about invertible matrices

Prop A matrix A is invertible if and only if its row echelon form has a pivot in every column and every row.

Proof Let $A \in M_{m \times n}$ and let $\vec{e}_1, \dots, \vec{e}_n$ be the std. basis for $\mathbb{F}^n$. Recall that $A\vec{e}_1, \dots, A\vec{e}_n$ are the columns of A. Now, using a pair of theorems from Section 1.6, we know A is invertible if and only if $A\vec{e}_1, \dots, A\vec{e}_n$ form a basis. By part 3 of the first proposition of this section, this happens if and only if the row echelon form of A has a pivot in every column and every row. $\square$

- Note that this gives us another proof the fact that only square matrices can be invertible.

Prop If a square matrix is either left or right invertible, then it is invertible.

Proof First suppose A is left invertible, with left inverse B. Then $\vec{x}=\vec{0}$ is the only solution to $A\vec{x}=\vec{0}$ since multiplying by B on the left gives:

$$\vec{x} = B(\vec{0}) = \vec{0}.$$

By Obs. 1, this means the row echelon form of A has a pivot in every row. But A is square and so is its row echelon form, so it also has a pivot in every column. By the previous Prop. A is invertible.

Next suppose A is right invertible with right inverse C. Then for any $\vec{b}$ the system $A\vec{x}=\vec{b}$ is consistent since $\vec{x} := C\vec{b}$ is a solution. Indeed

$$A\vec{x} = A(C\vec{b}) = (AC)\vec{b} = \vec{b}$$

So by Obs. 2, the row echelon form of A has a pivot in every column. Since it is a square matrix, it has a pivot in every row too Thus A is invertible by the previous Prop. □

10/1

2.4 Finding A^{-1} with Row Reduction

- In last section, we saw $A \in M_{n\times n}$ is invertible iff its row echelon form has a pivot in every column and every row. That means its <u>reduced</u> row echelon form must be I_n.

- The converse is also true: if the RREF of A is I_n, then $EA = I_n$ where $E = E_1 \cdots E_n$ is a product of elementary matrices. So A is left invertible and square, hence invertible by last section.

Thm $A \in M_{n\times n}$ is invertible iff its reduced row echelon form is I_n.

- Note that by the above argument, $A^{-1} = E$. So to find A^{-1} we just need to remember which row operations we did to go from A to I_n. The following trick lets us do this in a slick way:

1. Form an augmented $n \times 2n$ matrix $(A \mid I)$
2. Perform row operations on this augmented matrix until A becomes I: $(A \mid I) \xrightarrow{\text{row ops.}} (I \mid B)$
3. The resulting matrix B will be A^{-1}.
4. If A cannot be transformed in I via row operations, then A is <u>not</u> invertible.

Why does this work? Suppose $E_1, E_2, \dots, E_\ell$ are the elementary matrices corresponding to the row operations we do. Then if $E = E_\ell \cdots E_2 E_1$ we have $EA = I$ and $E = A^{-1}$. But then

$$E(A \mid I) = (EA \mid EI) = (I \mid E) = (I \mid A^{-1}).$$

Ex Let $A = \begin{pmatrix} 0 & 2 & -1 \\ 1 & 0 & -1 \\ 1 & -1 & 0 \end{pmatrix}$. We do not know yet if A is invertible, but this doesn't matter because the above algorithm will tell us!

$$(A \mid I) = \left(\begin{array}{ccc|ccc} 0 & 2 & -1 & 1 & 0 & 0 \\ 1 & 0 & -1 & 0 & 1 & 0 \\ 1 & -1 & 0 & 0 & 0 & 1 \end{array}\right) \xrightarrow{R2 \leftrightarrow R1} \left(\begin{array}{ccc|ccc} 1 & 0 & -1 & 0 & 1 & 0 \\ 0 & 2 & -1 & 1 & 0 & 0 \\ 1 & -1 & 0 & 0 & 0 & 1 \end{array}\right)$$

$$\xrightarrow{R3 \mapsto R3 - R1} \left(\begin{array}{ccc|ccc} 1 & 0 & -1 & 0 & 1 & 0 \\ 0 & 2 & -1 & 1 & 0 & 0 \\ 0 & -1 & 1 & 0 & -1 & 1 \end{array}\right) \xrightarrow{R2 \mapsto R2 + 2R3} \left(\begin{array}{ccc|ccc} 1 & 0 & -1 & 0 & 1 & 0 \\ 0 & 0 & 1 & 1 & -2 & 2 \\ 0 & -1 & 1 & 0 & -1 & 1 \end{array}\right)$$

$$\xrightarrow[R3 \mapsto R3 - R2]{R1 \mapsto R1 + R2} \left(\begin{array}{ccc|ccc} 1 & 0 & 0 & 1 & -1 & 2 \\ 0 & 0 & 1 & 1 & -2 & 2 \\ 0 & -1 & 0 & -1 & 1 & -1 \end{array}\right) \xrightarrow[R2 \leftrightarrow R3]{R3 \mapsto -R3} \left(\begin{array}{ccc|ccc} 1 & 0 & 0 & 1 & -1 & 2 \\ 0 & 1 & 0 & 1 & -1 & 1 \\ 0 & 0 & 1 & 1 & -2 & 2 \end{array}\right)$$

So evidently A is invertible with

$$A^{-1} = \begin{pmatrix} 1 & -1 & 2 \\ 1 & -1 & 1 \\ 1 & -2 & 2 \end{pmatrix}$$

Exercise: use matrix mult. to check $AA^{-1} = A^{-1}A = I$. □

- Our justification of the algorithm also yields the following:

Thm Any invertible matrix can be written as a product of elementary matrices.

- If you knew what a "group" is, this theorem says that the elementary matrices generate the group of invertible matrices.

2.5 Dimension and Finite-Dimensional Spaces

- Recall that we previously defined the dimension of a vector space V, $\dim(V)$, to be the number of vectors in a basis (hence any basis)
- By convention, if $V = \{\vec{0}\}$ is the vector space consisting of just a zero vector, then

$$\dim(\{\vec{0}\}) := 0.$$

If V does not have a finite basis, then

$$\dim(V) := \infty$$

Def We say a vector space V is <u>finite-dimensional</u> if $\dim(V)$ is a finite number. Otherwise, we say V is <u>infinite-dimensional</u>.

- Recall that in Section 1.2 we proved any finite generating system contains a (finite) basis. Since any finite basis is also a finite generating system, we obtain the following:

Prop A vector space is finite-dimensional if and only if it contains a finite generating system.

Thm If V is a finite-dimensional vector space V with $\dim(V) = n$, then $V \cong \mathbb{F}^n$.

Proof Since $\dim(V) = n$, there exists a basis for V of size n: $\vec{v}_1, \dots, \vec{v}_n$. Define $T: V \longrightarrow \mathbb{F}^n$ by

$$T(\vec{v}_k) = \vec{e}_k \qquad k = 1, \dots, n.$$

Since the basis $\vec{v}_1, \dots, \vec{v}_n$ is mapped to the basis $\vec{e}_1, \dots, \vec{e}_n$ for $\mathbb{F}^n$, A is an isomorphism. □

- This theorem allows us to carry our results from Section 2.3 about the sizes of certain systems in $\mathbb{F}^n$ to general (finite-dimensional) vector spaces.

Cor Let V be a finite-dimensional vector space.

① Any linearly independent system in V cannot have more than $\dim(V)$ vectors in it.

② Any generating system in V must have at least $\dim(V)$ vectors in it.

Proof Suppose $\dim(V) = n$ and let $T: V \longrightarrow \mathbb{F}^n$ be an isomorphism. If $\vec{v}_1, \dots, \vec{v}_m$ is a linearly independent (resp. generating) system in V, then by Exercise 1 from Homework 4 we know $T(\vec{v}_1), \dots, T(\vec{v}_m)$ is a linearly independent (resp. generating) system in $\mathbb{F}^n$. Our results from Section 2.3 then imply $m \leq n$ (resp. $m \geq n$). □

Rem In order to use this Corollary, must <u>already</u> know the dimension of V.

- We have previously seen that a generating system can be reduced to a basis. The following shows a linearly independent system can be increased to a basis.

Prop (Completion to a basis) Any linearly independent system in a finite-dimensional vector space can be completed to a basis. That is, if $\dim(V) = n$ and $\vec{v}_1, \dots, \vec{v}_r$ is a linearly independent system, then there exists $\vec{v}_{r+1}, \dots, \vec{v}_n$ s.t. $\vec{v}_1, \dots, \vec{v}_n$ is a basis for V.

Proof Let $\vec{v}_1, \dots, \vec{v}_r$ be a lin. indep. system in V. Then we necessarily have $r \le n = \dim(V)$. If the system is generating, then it is a basis (so $r = n$) and we are done. Otherwise there exists a vector in

$$V \setminus \operatorname{span}\{\vec{v}_1, \dots, \vec{v}_r\}.$$

Call any such vector $\vec{v}_{r+1}$. By Exercise 5 on Homework 2, the system $\vec{v}_1, \dots, \vec{v}_r, \vec{v}_{r+1}$ is linearly independent. If it is generating, then we are done. Otherwise we iterate the above procedure until we find a generating set.

The process will necessarily stop after finitely many steps. Namely, when the system reaches size n, because we cannot have a lin. indep system in V of size $n+1$ or higher. □

- Finally, this last proposition tells us how dimensions interact with subspaces. Recall that a subspace is always a vector space in its own right, so we can talk about the dimension of a subspace too

Prop Let V_0 be subspace of a finite-dimensional vector space V. Then $\dim(V_0) \le \dim(V)$. Moreover, if $\dim(V_0) = \dim(V)$ then $V_0 = V$.

See next page

Proof ✗ Let $\dim(V_0) = m$ and let $\vec{v}_1, \dots, \vec{v}_m$ be a basis for V_0. Then in particular they are linearly independent. Since $V_0 \subset V$, it follows that $\vec{v}_1, \dots, \vec{v}_m$ is (still) a linearly independent system in V. So by part 1 of the above Corollary:

$$\dim(V_0) = m \le \dim(V).$$

Now, suppose $\dim(V_0) = \dim(V)$ and let $\vec{v}_1, \dots, \vec{v}_m$ be a basis for V_0. We claim it is also a basis for V. Indeed, we saw above it is a lin. indep system in V, so it suffices to show it is a generating system in V. Suppose, towards a contradiction, that it is not. Then by Exercise 5 on Homework 2, we can find $\vec{v}_{m+1} \in V$ so that $\vec{v}_1, \dots, \vec{v}_m, \vec{v}_{m+1}$ is lin. indep in V. But $\dim(V) = m < m+1$, so this contradicts part 2 of the above Corollary. Thus it must be that $\vec{v}_1, \dots, \vec{v}_m$ is generating in V and therefore a basis in V.

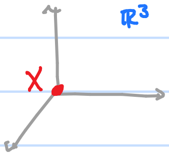
Now, let $\vec{v} \in V$ be arbitrary. Then there exist scalars $\alpha_1, \dots, \alpha_m$ such that

$$\vec{v} = \sum_{k=1}^{m} \alpha_k \vec{v}_k$$

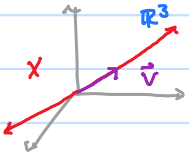
Since $\vec{v}_1, \dots, \vec{v}_m \in V_0$ and V_0 is a subspace, we have $\vec{v} \in V_0$. Since $\vec{v} \in V$ was arbitrary, it follows that $V \subset V_0$. We also obviously have $V_0 \subset V$, and so $V_0 = V$. □

Ex Let $X \subset \mathbb{R}^3$ be a subspace. Since $\dim(\mathbb{R}^3)=3$, we must have $\dim(X) \in \{0,1,2,3\}$. Let's examine each of these cases.

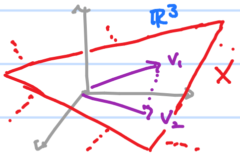
- $\dim(X)=0$: this implies $X=\{\vec{0}\}$



- $\dim(X)=1$: Let $\vec{v}$ be a basis for X. Then for any other $\vec{w} \in X$ we have $\vec{w}=\alpha\vec{v}$ for some scalar. Thus X is the line parallel to $\vec{v}$.
  Note that this line must go through the origin.



- $\dim(X)=2$: Let $\vec{v}_1,\vec{v}_2$ be a basis for X. Then for any other $\vec{w} \in X$ we have $\vec{w}=\alpha_1\vec{v}_1+\alpha_2\vec{v}_2$ for some scalars α_1,α_2.
  Thus X is the plane containing $\vec{v}_1,\vec{v}_2$
  Note that this plane must go through the origin.



- $\dim(X)=3$: So $\dim(X)=\dim(\mathbb{R}^3)$, which implies $X=\mathbb{R}^3$.

10/8

* **Lemma** Any subspace of a finite-dimensional vector space is also finite-dimensional.

Proof Let V be a fin. dim'l vector space, and $V_0 \subset V$ be a subspace.
We must first argue that V_0 is finite-dimensional, so that $n<\infty$
If $V_0=\{\vec{0}\}$, then $\dim(V_0)=0<\infty$, and the result follows immediately.
Next, assume $V_0 \neq \{\vec{0}\}$. Then $\exists\, \vec{v}_1 \in V_0$ s.t. $\vec{v}_1 \neq \vec{0}$. Now, the system v_1 is lin. indep. (since $\vec{v}_1 \neq \vec{0}$). If it is generating for V_0, then $\dim(V_0)=1<\infty$ and we're done. Otherwise, by Exercise 5 on Homework 2 $\exists\, v_2 \in V_0$ s.t. v_1,v_2 is lin. indep. Iterating this argument, we will either obtain a basis for V_0, or obtain a linearly independent system $v_1,\dots,v_n$ where $n=\dim(V)$. It must then be the case that $v_1,\dots,v_n$ is generating for V_0, because otherwise we will obtain a lin. indep. system of size $n+1>\dim(V)$, a contradiction.
So in either case we obtain a finite basis for V_0. ☐

2.6 General Solution of Linear System

Def A homogeneous linear system of equations has the following form

$$A\vec{x} = \vec{0}$$

That is, the right-hand side equals the zero vector. If $A\vec{x} = \vec{b}$ is an arbitrary linear system, we say $A\vec{x}$ is the associated homogeneous linear system.

- Note that a homogeneous lin. sys. is always consistent, since $\vec{x} = \vec{0}$ is a solution.

Thm (General solution of a linear system) Let $\vec{x}_1$ be any vector solving a linear system $A\vec{x} = \vec{b}$. Let H be the set of all solutions of the homogeneous linear system $A\vec{x} = \vec{0}$. Then the set

$$X := \{ \vec{x} = \vec{x}_1 + \vec{x}_h : \vec{x}_h \in H \}$$

is the set of all solutions to $A\vec{x} = \vec{b}$.

Proof First observe that for $\vec{x} = \vec{x}_1 + \vec{x}_h$ with $\vec{x}_h \in H$ we have

$$A\vec{x} = A(\vec{x}_1 + \vec{x}_h) = A\vec{x}_1 + A\vec{x}_h = \vec{b} + \vec{0} = \vec{b}.$$

So every element in the above set is a solution of $A\vec{x} = \vec{b}$.

Now let $\vec{x}_2$ be any solution of $A\vec{x} = \vec{b}$. Observe

$$A(\vec{x}_2 - \vec{x}_1) = A\vec{x}_2 - A\vec{x}_1 = \vec{b} - \vec{b} = \vec{0},$$

so $\vec{x}_2 - \vec{x}_1 \in H$. But then

$$\vec{x}_2 = \vec{x}_2 - \vec{x}_1 + \vec{x}_1 = \vec{x}_1 + (\vec{x}_2 - \vec{x}_1) \in X.$$

Since $\vec{x}_2$ was an arbitrary solution of $A\vec{x} = \vec{b}$, we see that all solutions are in X. □

- This Theorem says:

$$\boxed{\begin{matrix}\text{General solution}\\ \text{of } A\vec{x} = \vec{b}\end{matrix}} = \boxed{\begin{matrix}\text{Particular Solution}\\ \text{of } A\vec{x} = \vec{b}\end{matrix}} + \boxed{\begin{matrix}\text{General Solution}\\ \text{of } A\vec{x} = \vec{0}\end{matrix}}$$

This same theorem makes an appearance in Differential Equations. It is really the same theorem because we did not require $\vec{x}$ or $\vec{b}$ to live in a fin. dim. vector space.

- Observe that $H = \text{Null}(A)$ and so is a subspace.

Ex Someone claims

$$\vec{x} = \begin{pmatrix} 3 \\ 1 \\ 0 \\ 2 \\ 0 \end{pmatrix} + s\begin{pmatrix} -2 \\ 1 \\ 1 \\ 0 \\ 0 \end{pmatrix} + t\begin{pmatrix} 0 \\ 0 \\ 1 \\ 2 \\ 1 \end{pmatrix} \qquad s, t \in \mathbb{F}$$

is the solution set to

$$\begin{pmatrix} 2 & 3 & 1 & 4 & -9 \\ 1 & 1 & 1 & 1 & -3 \\ 1 & 1 & 1 & 2 & -5 \\ 2 & 2 & 2 & 3 & -8 \end{pmatrix} \vec{x} = \begin{pmatrix} 17 \\ 6 \\ 8 \\ 14 \end{pmatrix}$$

How do we (quickly) check this?

First verify that $\begin{pmatrix}3\\1\\0\\2\\0\end{pmatrix}$ is a particular solution.

Next, check that $\begin{pmatrix}-2\\1\\1\\0\\0\end{pmatrix}$ and $\begin{pmatrix}0\\0\\1\\2\\1\end{pmatrix}$ are solutions of the associated homogeneous lin. sys.

Since $H = \{\vec{x}_h : A\vec{x}_h = \vec{0}\}$ is a subspace, for all $s,t \in \mathbb{F}$ we have

$$s\begin{pmatrix}-2\\1\\1\\0\\0\end{pmatrix} + t\begin{pmatrix}0\\0\\1\\2\\1\end{pmatrix} \in H.$$

The theorem then tells us that the claimed set of solutions are all indeed solutions. But has this accounted for all of the solutions?

If we knew $\dim(H)$, then we could quickly determine this. Indeed, $\begin{pmatrix}-2\\1\\1\\0\\0\end{pmatrix}, \begin{pmatrix}0\\0\\1\\2\\1\end{pmatrix}$ is readily seen to be linearly independent. So if $\dim(H) = 2$, then they must form a basis for H. Otherwise $\dim(H) > 2$ and

$$\operatorname{span}\left\{\begin{pmatrix}-2\\1\\1\\0\\0\end{pmatrix}, \begin{pmatrix}0\\0\\1\\2\\1\end{pmatrix}\right\} \subsetneq H$$

So we're missing solutions.

In the next section, we will find a method for quickly determining $\dim(H)$ $\square$

2.7 The fundamental subspaces of a matrix, and rank

- Let $A \in M_{m\times n}$, which we think of as a lin. trans $A: \mathbb{F}^n \to \mathbb{F}^m$. Recall that we can associate two subspaces to A:

$$Ker(A) = Null(A) = \{\vec{x}\in\mathbb{F}^n : A\vec{x}=\vec{0}\}$$

$$Ran(A) = \{\vec{b}\in \mathbb{F}^m : \vec{b}=A\vec{x} \text{ for some } \vec{x}\in\mathbb{F}^n\}$$

So $Ker(A)$ is all solutions to the homogeneous lin system $A\vec{x}=\vec{0}$, while $Ran(A)$ is all the choices of $\vec{b}$ so that $A\vec{x}=\vec{b}$ is consistent.

- Let us associate two more subspaces to A:

Def Let $A\in M_{m\times n}$ have columns $\vec{v}_1,\dots,\vec{v}_n\in\mathbb{F}^m$ and rows $\vec{w}_1^T,\dots,\vec{w}_m^T\in(\mathbb{F}^m)^T$:

$$A=(\vec{v}_1\cdots\vec{v}_n) = \begin{pmatrix}\vec{w}_1^T\\ \vdots\\ \vec{w}_m^T\end{pmatrix}$$

The column space of A, sometimes denoted $Col(A)$, is the subspace $\operatorname{span}\{\vec{v}_1,\dots,\vec{v}_n\}$.
The row space of A is the subspace $\operatorname{span}\{\vec{w}_1,\dots,\vec{w}_n\}$.

- Recall that $A\vec{x}=\vec{b}$ has a solution $\vec{x}=(x_1,\dots,x_n)^T$ iff

$$\vec{b} = x_1\vec{v}_1+\cdots+x_n\vec{v}_n \in Col(A).$$

That is $\vec{b}\in Ran(A)$ iff $\vec{b}\in Col(A)$. Thus $Ran(A) = Col(A)$.

- Observe that the rowspace of A is the column space of A^T. So by the above, we know the row space of $A = Ran(A^T)$.

Def For $A\in M_{m\times n}$, its four fundamental subspaces are $Ran(A), Ran(A^T), Ker(A), Ker(A^T)$.

- We will see how these subspaces are related to each other.

Def Let $T: V\to W$ be a linear transformation. The rank of T is

$$rank(T) := dim(Ran(T)).$$

Computing fundamental subspaces and rank

Thm Let $A\in M_{m\times n}$ with RREF B. Suppose $A=(\vec{v}_1\cdots\vec{v}_n)$ and $B=\begin{pmatrix}\vec{w}_1^T\\ \vdots\\ \vec{w}_n^T\end{pmatrix}$.

① If B has pivots in columns $j_1,\dots,j_r$ then $\vec{v}_{j_1},\dots,\vec{v}_{j_r}$ is a basis for $Ran(A)$.

② If B has pivots in rows $i_1,\dots,i_r$ then $\vec{w}_{i_1},\dots,\vec{w}_{i_r}$ is a basis for $Ran(A^T)$.

In particular, if the row echelon form of A has r pivots, then $rank(A)=r$.

$\boxed{\text{Ex}}$ Recall the following matrix from Midterm 1:

$$A = \begin{pmatrix} 1 & -1 & -2 & 0 & 2 \\ 3 & -3 & 1 & 5 & -4 \\ -1 & 1 & 4 & 2 & -3 \end{pmatrix}$$

Then

$$\operatorname{Ran}(A) = \operatorname{span}\left\{ \begin{pmatrix} 1 \\ 3 \\ -1 \end{pmatrix}, \begin{pmatrix} -1 \\ -3 \\ 1 \end{pmatrix}, \begin{pmatrix} -2 \\ 1 \\ 4 \end{pmatrix}, \begin{pmatrix} 0 \\ 5 \\ 2 \end{pmatrix}, \begin{pmatrix} 2 \\ -4 \\ -3 \end{pmatrix} \right\}$$

but these vectors do not form a basis for Ran(A) (they are lin. dep.) However, the RREF of A is:

$$B = \begin{pmatrix} \textcircled{1} & -1 & 0 & 2 & 0 \\ 0 & 0 & \textcircled{1} & 1 & 0 \\ 0 & 0 & 0 & 0 & \textcircled{1} \end{pmatrix}$$

Since columns 1, 3, and 5 contain pivots, the theorem implies $\begin{pmatrix} 1 \\ 3 \\ -1 \end{pmatrix}, \begin{pmatrix} -2 \\ 1 \\ 4 \end{pmatrix}, \begin{pmatrix} 2 \\ -4 \\ -3 \end{pmatrix}$ forms a basis for Ran(A) (Exercise: check this directly). Thus

$$\operatorname{rank}(A) = \dim(\operatorname{Ran}(A)) = 3$$

Since Ran(A) is a subspace of $\mathbb{F}^3$, which also has dimension 3, it must be that $\operatorname{Ran}(A) = \mathbb{F}^3$. Consequently, $A\vec{x} = \vec{b}$ is consistent for every $\vec{b} \in \mathbb{F}^3$ (but we already knew this since the RREF of A has a pivot in every row).

This same computation also tells us that

$$\begin{pmatrix} 1 \\ -1 \\ 0 \\ 2 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 1 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 1 \end{pmatrix}$$

(the transposes of the rows of B) form a basis for $\operatorname{Ran}(A^T)$, the row space of A. Hence $\dim(\operatorname{Ran}(A^T)) = 3$, which is strictly less than the dimension of $\mathbb{F}^5$. So

$$\operatorname{Ran}(A^T) \subsetneq \mathbb{F}^5$$

Finally, we can also use this computation to find a basis for Ker(A). Indeed, this first requires solving the system $A\vec{x} = \vec{0}$:

$$(A \mid \vec{0}) \xrightarrow{\text{row ops.}} \left(\begin{array}{ccccc|c} \textcircled{1} & -1 & 0 & 2 & 0 & 0 \\ 0 & 0 & \textcircled{1} & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & \textcircled{1} & 0 \end{array}\right) \longrightarrow \begin{cases} x_1 - x_2 + 2x_4 = 0 \\ x_3 + x_4 = 0 \\ x_5 = 0 \end{cases} \longrightarrow \begin{cases} x_1 = x_2 - 2x_4 \\ x_3 = -x_4 \\ x_5 = 0 \end{cases}$$

So a solution to $A\vec{x} = \vec{0}$ has the form

$$\vec{x} = \begin{pmatrix} x_2 - 2x_4 \\ x_2 \\ -x_4 \\ x_4 \\ 0 \end{pmatrix} = x_2 \begin{pmatrix} 1 \\ 1 \\ 0 \\ 0 \\ 0 \end{pmatrix} + x_4 \begin{pmatrix} -2 \\ 0 \\ -1 \\ 1 \\ 0 \end{pmatrix} \qquad x_2, x_4 \in \mathbb{F}.$$

Thus $(1,1,0,0,0)^T$, $(-2,0,-1,1,0)^T$ are generating for Ker(A). They are also lin indep: the first vector has 1 in the x_2 position and 0 in the x_4 position, while the second has the opposite. $\square$

- This procedure for computing ker(A) via the RREF of A will always yield a basis by the same argument.
- To compute $\ker(A^T)$, it is necessary to find the RREF of A^T. So a separate computation is required.

10/10

Proof (of Thm) ①: Let B be the RREF of A, and suppose the pivots of B are located in columns $j_1, j_2, \ldots, j_r$. Then these columns of B are exactly the standard basis vectors $\vec{e}_1, \vec{e}_2, \ldots, \vec{e}_r$, which are lin. indep. Also, any other column in B has all of its non-zero entries in rows $1, 2, \ldots, r$ (by def. of RREF, since otherwise we've missed a pivot.) Hence $\vec{e}_1, \vec{e}_2, \ldots, \vec{e}_r$ is a also generating for $\mathrm{Ran}(B)$. Therefore $\vec{e}_1, \ldots, \vec{e}_r$ forms a basis for $\mathrm{Ran}(B)$.

Now, recall that there exists an invertible matrix E s.t. $A = EB$ (namely, E is a product of elementary matrices). It follows that columns $j_1, j_2, \ldots, j_r$ of A are precisely E times the corresponding columns of B: $\vec{v}_{j_1} := E\vec{e}_1$, $\vec{v}_{j_2} := E\vec{e}_2, \ldots, \vec{v}_{j_r} := E\vec{e}_r$. Since E is invertible and $\vec{e}_1, \ldots, \vec{e}_r$ are lin. indep., $\vec{v}_{j_1}, \ldots, \vec{v}_{j_r}$ are lin. indep. It remains to show $\vec{v}_{j_1}, \ldots, \vec{v}_{j_r}$ are generating for $\mathrm{Ran}(A)$. Let $\vec{v} \in \mathrm{Ran}(A)$, then $\exists\, \vec{x}$ s.t. $\vec{v} = A\vec{x}$. Consequently

$$\vec{v} = EB\vec{x}$$
$$E^{-1}\vec{v} = B\vec{x}$$

So $E^{-1}\vec{v} \in \mathrm{Ran}(B)$, which means $\exists$ scalars $\alpha_1, \ldots, \alpha_r$ s.t.

$$\begin{aligned} E^{-1}\vec{v} &= \alpha_1\vec{e}_1 + \cdots + \alpha_r\vec{e}_r \\ \Rightarrow \quad \vec{v} &= E(\alpha_1\vec{e}_1 + \cdots + \alpha_r\vec{e}_r) \\ &= \alpha_1 E\vec{e}_1 + \cdots + \alpha_r E\vec{e}_r \\ &= \alpha_1\vec{v}_{j_1} + \cdots + \alpha_r\vec{v}_{j_r}. \end{aligned}$$

So $\vec{v}_{j_1}, \ldots, \vec{v}_{j_r}$ is generating and therefore a basis for $\mathrm{Ran}(A)$.

②: Let B be the RREF of A with pivots in rows $i_1, i_2, \ldots, i_r$. Let $\vec{\omega}_{i_1}, \vec{\omega}_{i_2}, \ldots, \vec{\omega}_{i_r}$ be the transposes of the corresponding rows of B. Suppose

$$\alpha_1\vec{\omega}_{i_1} + \alpha_2\vec{\omega}_{i_2} + \cdots + \alpha_r\vec{\omega}_{i_r} = \vec{0}$$

for some scalars $\alpha_1, \ldots, \alpha_r$. The first non-zero entry of $\alpha_1\vec{\omega}_{i_1}$ is then α_1, while the same entry in the remaining vectors $\alpha_2\vec{\omega}_{i_2}, \ldots, \alpha_r\vec{\omega}_{i_r}$ is zero. Thus we must have $\alpha_1 = 0$ and so

$$\alpha_2\vec{\omega}_{i_2} + \cdots + \alpha_r\vec{\omega}_{i_r} = \vec{0}$$

Iterating this argument yields $\alpha_2 = 0, \ldots, \alpha_r = 0$. Thus $\vec{\omega}_{i_1}, \ldots, \vec{\omega}_{i_r}$ are lin. indep. They are also generating for $\mathrm{Ran}(B^T)$, the row space of B, because they are precisely the non-zero rows of B. Thus $\vec{\omega}_{i_1}, \ldots, \vec{\omega}_{i_r}$ form a basis for $\mathrm{Ran}(B^T)$. Recall $A = EB$, so

$$\mathrm{Ran}(A^T) = \mathrm{Ran}(B^T E^T) = B^T(\mathrm{Ran}(E^T)) \underset{\substack{\uparrow \\ \text{since } E \text{ is invertible}}}{=} B^T(\mathbb{F}^m) \underset{\substack{\nwarrow \\ \text{def. of Range}}}{=} \mathrm{Ran}(B^T).$$

Therefore $\vec{\omega}_{i_1}, \ldots, \vec{\omega}_{i_r}$ is also a basis for $\mathrm{Ran}(A^T)$, since its the same space. $\square$

The Rank Theorem

As we have seen above, for $A \in M_{m \times n}$ $\mathrm{rank}(A) = \dim(\mathrm{Ran}(A))$ is given by the number of pivots in the RREF of A. This also gives $\dim(\mathrm{Ran}(A^T)) = \mathrm{rank}(A^T)$. This yields the following theorem:

Thm (The Rank Theorem) For any $A \in M_{m\times n}$, $\mathrm{rank}(A) = \mathrm{rank}(A^T)$.

Def Let $T: V \to W$ be a linear transformation. The nullity of T is

$$\mathrm{nullity}(T) := \dim(\mathrm{Ker}(T)).$$

Thm (Rank-nullity theorem) Let $T: V \to W$ be a linear transformation with V finite-dimensional. Then

$$\mathrm{rank}(T) + \mathrm{nullity}(T) = \dim(V).$$

Proof Let $n = \dim(V)$. Since Ker(T) is a subspace of V, it is finite-dimensional. Let $r = \dim(\mathrm{Ker}(T)) = \mathrm{nullity}(T)$, then we know $r \leq n$. Let $\vec{v}_1, \dots, \vec{v}_r$ be a basis for Ker(T). By the Completion to a basis theorem, we can find $\vec{v}_{r+1}, \dots, \vec{v}_n \in V$ so that $\vec{v}_1, \dots, \vec{v}_r, \vec{v}_{r+1}, \dots, \vec{v}_n$ forms a basis for V. We claim $T(\vec{v}_{r+1}), \dots, T(\vec{v}_n)$ is a basis for Ran(T).

Indeed, let $\vec{w} \in \mathrm{Ran}(T)$. Then $\exists \vec{v} \in V$ s.t. $T(\vec{v}) = \vec{w}$. Since $\vec{v}_1, \dots, \vec{v}_n$ is a basis for V, there are scalars $\alpha_1, \dots, \alpha_n$ s.t. $\vec{v} = \alpha_1\vec{v}_1 + \dots + \alpha_n\vec{v}_r$. Applying T to each side gives:

$$\begin{aligned}\vec{w} &= T(\alpha_1\vec{v}_1 + \dots + \alpha_r\vec{v}_r + \alpha_{r+1}\vec{v}_{r+1} + \dots + \alpha_n\vec{v}_n) \\ &= \alpha_1 T(\vec{v}_1) + \dots + \alpha_r T(\vec{v}_r) + \alpha_{r+1}T(\vec{v}_{r+1}) + \dots + \alpha_n T(\vec{v}_n) \\ &= \vec{0} + \dots + \vec{0} + \alpha_{r+1}T(\vec{v}_{r+1}) + \dots + \alpha_n T(\vec{v}_n) \\ &= \alpha_{r+1}T(\vec{v}_{r+1}) + \dots + \alpha_n T(\vec{v}_n).\end{aligned}$$

Since $\vec{w} \in \mathrm{Ran}(T)$ was arbitrary, this shows $T(\vec{v}_{r+1}), \dots, T(\vec{v}_n)$ is generating for Ran(T). Now, towards showing they are lin. indep. suppose there are scalars $\beta_{r+1}, \dots, \beta_n$ s.t.

$$\sum_{k=r+1}^{n} \beta_k T(\vec{v}_k) = \vec{0}$$

$$T\Big(\sum_{k=r+1}^{n} \beta_k \vec{v}_k\Big) = \vec{0}.$$

Thus $\vec{v}_0 := \sum_{k=r+1}^{n} \beta_k \vec{v}_k \in \mathrm{Ker}(T)$. Recall that $\vec{v}_1, \dots, \vec{v}_r$ is a basis for Ker(T), so there exists scalars $\gamma_1, \dots, \gamma_r$ s.t.

$$v_0 = \gamma_1\vec{v}_1 + \dots + \gamma_r\vec{v}_r \Rightarrow \vec{0} = \gamma_1\vec{v}_1 + \dots + \gamma_r\vec{v}_r + (-\beta_{r+1})\vec{v}_{r+1} + \dots + (-\beta_n)\vec{v}_n$$

Since the system $\vec{v}_1, \dots, \vec{v}_n$ is lin. indep., all scalars must equal zero. In particular, $\beta_{r+1} = 0, \dots, \beta_n = 0$, and so $T(\vec{v}_{r+1}), \dots, T(\vec{v}_n)$ is lin. ind., hence a basis for Ran(T). Thus

$$\mathrm{rank}(T) = \dim(\mathrm{Ran}(T)) = n - r = \dim(V) - \mathrm{nullity}(T).$$

□

- Note that this implies Ran(T) is finite-dimensional so long as the domain of T is finite-dimensional.

- Applying this theorem to matrices (and doing a little arithmetic) yields the following:

Cor For $A \in M_{m \times n}$

$$\operatorname{rank}(A) + \operatorname{nullity}(A) = \operatorname{rank}(A^T) + \operatorname{nullity}(A) = n$$
$$\operatorname{rank}(A^T) + \operatorname{nullity}(A^T) = \operatorname{rank}(A) + \operatorname{nullity}(A^T) = m$$

In particular,

$$\operatorname{nullity}(A) - \operatorname{nullity}(A^T) = n - m.$$

Ex Recall the example from the previous section, where we wanted to verify the claim that

$$\vec{x} = \begin{pmatrix} 3 \\ 1 \\ 0 \\ 2 \\ 0 \end{pmatrix} + s\begin{pmatrix} -2 \\ 1 \\ 1 \\ 0 \\ 0 \end{pmatrix} + t\begin{pmatrix} 0 \\ 0 \\ 1 \\ 2 \\ 1 \end{pmatrix} \qquad s, t \in \mathbb{F}$$

is the solution set to

$$\begin{pmatrix} 2 & 3 & 1 & 4 & -9 \\ 1 & 1 & 1 & 1 & -3 \\ 1 & 1 & 1 & 2 & -5 \\ 2 & 2 & 2 & 3 & -8 \end{pmatrix} \vec{x} = \begin{pmatrix} 17 \\ 6 \\ 8 \\ 14 \end{pmatrix}$$

We reduced this down to determining the dimension of the solution set to the associated homogeneous lin. sys. Row operations on the above coefficient matrix yields this REF:

$$\begin{pmatrix} 1 & 1 & 1 & 1 & -3 \\ 0 & 1 & -1 & 2 & -3 \\ 0 & 0 & 0 & 1 & -2 \\ 0 & 0 & 0 & 1 & -2 \end{pmatrix}$$

So there are 3 pivots $\Rightarrow \operatorname{rank}(A) = 3$. Hence $\operatorname{nullity}(A) = 5 - 3 = 2$. Thus the claimed set is indeed <u>all</u> of the solutions. $\square$

Thm Let $A \in M_{m \times n}$. Then

$$A\vec{x} = \vec{b}$$

is consistent for all $\vec{b} \in \mathbb{F}^m$ if and only if

$$A^T\vec{x} = \vec{0}$$

has a unique solution.

Proof Homework 6 $\square$

Completion to a basis algorithm

We know given any lin. indep. system $\vec{v}_1, \dots, \vec{v}_r$ in a fin. dim. vector space V, we can find vectors $\vec{v}_{r+1}, \dots, \vec{v}_n$ s.t. $\vec{v}_1, \dots, \vec{v}_n$ is a basis for V. However the proof did not give us a practical way of finding $\vec{v}_{r+1}, \dots, \vec{v}_n$ (unless we know $\operatorname{span}\{\vec{v}_1, \dots, \vec{v}_r\}$ very well). However, there is an algorithm for doing this when $V = \mathbb{F}^n$:

Algorithm: Given a lin. indep. system $\vec{v}_1, \ldots, \vec{v}_r$ in $\mathbb{F}^n$

① Form $A \in M_{r \times n}$ by $A = \begin{pmatrix} \vec{v}_1^{\,T} \\ \vdots \\ \vec{v}_r^{\,T} \end{pmatrix}$

② Compute the RREF (or even just the REF) of A.

③ Determine the columns $j_1, j_2, \ldots, j_{n-r}$ without pivots.

Then $\vec{v}_1, \ldots, \vec{v}_r, \vec{e}_{j_1}, \vec{e}_{j_2}, \ldots, \vec{e}_{j_{n-r}}$ will be a basis for $\mathbb{F}^n$.

Ex Complete $\vec{v}_1 = \begin{pmatrix} 1 \\ 2 \\ -2 \\ -1 \end{pmatrix}, \vec{v}_2 = \begin{pmatrix} 2 \\ 4 \\ -2 \\ 0 \end{pmatrix}$ to a basis for $\mathbb{F}^4$.

$$\begin{pmatrix} 1 & 2 & -2 & -1 \\ 2 & 4 & -2 & 0 \end{pmatrix} \xrightarrow[R2 \mapsto R2 - 2 \cdot R1]{} \begin{pmatrix} 1 & 2 & -2 & -1 \\ 0 & \underline{0} & 2 & \underline{2} \end{pmatrix}$$

no pivots

So $\vec{v}_1, \vec{v}_2, \begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \end{pmatrix}$ forms a basis for $\mathbb{F}^4$.

- Why does the algorithm work?

First note

$$\#\text{ of pivots in } A = \dim(\mathrm{Ran}(A^T)) = \dim(\mathrm{span}\{\vec{v}_1, \ldots, \vec{v}_r\}) = r$$

since $\vec{v}_1, \ldots, \vec{v}_r$ are lin indep. Now, let $\tilde{B} \in M_{n \times n}$ be the matrix obtained from the REF of A after inserting $\vec{e}_{j_1}^{\,T}, \vec{e}_{j_2}^{\,T}, \ldots, \vec{e}_{j_{n-r}}^{\,T}$ as rows $j_1, j_2, \ldots, j_{n-r}$ respectively. Then $\tilde{B}$ has a pivot in every column (and hence every row), and therefore is invertible. Let $\tilde{A}$ be the matrix obtained from A by the same procedure as above. Now, the same row ops that turn A into B will turn $\tilde{A}$ into $\tilde{B}$ (after accounting for the shifting of rows). Thus

$$\tilde{B} = E\tilde{A}$$

where E is a product of elementary matrices. But then $\tilde{A} = E^{-1}\tilde{B}$ and so is invertible as a product of invertible matrices. Thus $\tilde{A}^T$ is also invertible, which means its columns $(\vec{v}_1, \ldots, \vec{v}_r, \vec{e}_{j_1}, \ldots, \vec{e}_{j_{n-r}})$ form a basis for $\mathbb{F}^n$.

10/15

2.8 Matrix representations for arbitrary linear transformations

So far, we have seen how to represent linear transformations $T:\mathbb{F}^n \to \mathbb{F}^m$ as matrices. Now we will see how to do this for lin. trans. $T:V\to W$ between arbitrary (finite-dimensional) vector spaces. The key will be to pick bases for V and W, which allows us to identify $V\cong\mathbb{F}^n$ and $W\cong\mathbb{F}^m$.

Coordinate Vectors

Let V be a vector space with basis $\mathcal{B}:=\{\vec{v}_1,\vec{v}_2,\dots,\vec{v}_n\}$. For $\vec{v}\in V$ let $x_1,\dots,x_n$ be the unique scalars such that

$$\vec{v}=x_1\vec{v}_1+\dots+x_n\vec{v}_n.$$

Recall that $x_1,\dots,x_n$ are called the coordinates of $\vec{v}$ in the basis $\mathcal{B}$.
Denote

$$[\vec{v}]_{\mathcal{B}}:=\begin{pmatrix}x_1\\x_2\\\vdots\\x_n\end{pmatrix}\in\mathbb{F}^n$$

We call $[\vec{v}]_{\mathcal{B}}$ the coordinate vector of $\vec{v}$ relative to the basis $\mathcal{B}$.
Observe that

$$V\ni\vec{v}\longmapsto[\vec{v}]_{\mathcal{B}}\in\mathbb{F}^n$$

defines an isomorphism $V\cong\mathbb{F}^n$. In particular

$$[\vec{v}_1]_{\mathcal{B}}=\vec{e}_1,\quad[\vec{v}_2]_{\mathcal{B}}=\vec{e}_2,\quad\dots,\quad[\vec{v}_n]_{\mathcal{B}}=\vec{e}_n.$$

Rem The difficulty of computing $[\vec{v}]_{\mathcal{B}}$ for $\vec{v}\in V$ depends on how "nice" the basis $\mathcal{B}$ is. For example, if $V=P_2$ and $\mathcal{B}=\{1,x,x^2\}$ then the computation is easy:

$$[a_2x^2+a_1x+a_0]_{\mathcal{B}}=[a_0\cdot 1+a_1x+a_2x^2]_{\mathcal{B}}=\begin{pmatrix}a_0\\a_1\\a_2\end{pmatrix}$$

However, if $\mathcal{B}=\{x+1,2x^2-x+1,x^2+2\}$ then computing $[\cdot]_{\mathcal{B}}$ is hard because it's not clear how to write $a_2x^2+a_1x+a_0$ as a linear combination of these basis vectors. We will soon produce a change of coordinates formula to help us with this.

Matrix of a linear transformation

Let $T:V\to W$ be a linear trans. and let $\mathcal{A}=\{\vec{v}_1,\dots,\vec{v}_n\}$ and $\mathcal{B}=\{\vec{w}_1,\dots,\vec{w}_m\}$ be bases for V and W, respectively.

Def The matrix representation of T with respect to $\mathcal{A}$ and $\mathcal{B}$ is the matrix $[T]_{\mathcal{A}}^{\mathcal{B}}\in M_{m\times n}$ that satisfies

$$[T]_{\mathcal{A}}^{\mathcal{B}}[\vec{v}]_{\mathcal{A}}=[T(\vec{v})]_{\mathcal{B}}\qquad\text{for all }\vec{v}\in V.$$

Ex Consider $T: \mathbb{P}_2 \to \mathbb{P}_1$, $T(p(x)) = p'(x)$. Let $A = \{1, x, x^2\}$ and $B = \{1, x\}$. Then for $p(x) = a_2x^2 + a_1x + a_0$ we have:

$$[p(x)]_A = \begin{pmatrix} a_0 \\ a_1 \\ a_2 \end{pmatrix} \qquad [T(p(x))]_B = [2a_2x + a_1] = \begin{pmatrix} a_1 \\ 2a_2 \end{pmatrix}$$

So $[T]_A^B \in M_{2\times3}$ should satisfy: $$[T]_A^B \begin{pmatrix} a_0 \\ a_1 \\ a_2 \end{pmatrix} = \begin{pmatrix} a_1 \\ 2a_2 \end{pmatrix}$$

We can determine the columns of $[T]_A^B$ by choosing a_0, a_1, a_2 to give us the std basis vectors:

$$[T]_A^B \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}, \quad [T]_A^B \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \quad [T]_A^B \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} = \begin{pmatrix} 0 \\ 2 \end{pmatrix}$$

So

$$[T]_A^B = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 2 \end{pmatrix}.$$

□

- Note that if $A = \{\vec v_1, \dots, \vec v_n\}$, then $[\vec v_j]_A = \vec e_j$ and so

$$[T(\vec v_j)]_B = [T]_A^B [\vec v_j]_A = \underbrace{[T]_A^B \vec e_j}_{j\text{th column of } [T]_A^B}$$

So the columns of $[T]_A^B$ are given by $[T(\vec v_1)]_B, \dots, [T(\vec v_n)]_B$.

Prop Let $S: U \to V$, $T: V \to W$ be lin. trans. and let A, B, and C be bases for U, V, and W respectively. Then

$$[T \circ S]_A^C = [T]_B^C [S]_A^B$$

Proof Let $\vec u \in U$. Then

$$[T \circ S]_A^C [\vec u]_A = [T \circ S(\vec u)]_C = [T(S(\vec u))]_C$$

and

$$[T]_B^C [S]_A^B [\vec u]_A = [T]_B^C [S(\vec u)]_B = [T(S(\vec u))]_C$$

Since $\vec u \in U$ was arbitrary, we must have $[T \circ S]_A^C = [T]_B^C [S]_A^B$. □

- If $T: V \to W$ is invertible, then

$$[T^{-1}]_B^A [T]_A^B = [T^{-1} \circ T]_A^A = [I_V]_A^A = I_n$$

and similarly $[T]_A^B [T^{-1}]_B^A = I_n$, so $[T^{-1}]_B^A = ([T]_A^B)^{-1}$ (as matrices).

- Matrix representations give us the following diagram:

$$\begin{tikzcd} V \arrow[r, "T"] \arrow[d, swap, "\vec v \mapsto {[\vec v]_A}"] \arrow[dr, phantom, "\circlearrowleft"] & W \arrow[d, "\vec w \mapsto {[\vec w]_B}"] \\ \mathbb{F}^n \arrow[r, "{[T]_A^B}"] & \mathbb{F}^m \end{tikzcd}$$

Change of coordinate matrix

- Let A and B both be bases for V. Then $I = I_V$ satisfies for all $\vec{v} \in V$

$$[I]^B_A [\vec{v}]_A = [I(\vec{v})]_B = [\vec{v}]_B.$$

That is, to get the coordinate vector $[\vec{v}]_B$, just multiply the coordinate vector $[\vec{v}]_A$ by $[I]^B_A$. We call $[I]^B_A$ the change of coordinates matrix from A to B.

- Observe that since I is its own inverse we have:

$$[I]^A_B = [I^{-1}]^A_B = ([I]^B_A)^{-1}$$

So if we already know how to change coordinates from A to B (via $[I]^B_A$), then computing the inverse matrix will tell us how to compute the reverse.

Ex Consider the following bases for $\mathbb{P}_2$:

$$S = \{1, x, x^2\} \qquad \mathcal{B} = \{x+1, 2x^2-x+1, x^2+2\}$$

Let's compute $[x^2+x+1]_B$. First note

$$[x^2+x+1]_S = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}.$$

Then

$$[x^2+x+1]_B = [I]^B_S \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}.$$

So we must compute $[I]^B_S$. Recall that its columns are given by

$$[I(1)]_B = [1]_B, \quad [I(x)]_B = [x]_B, \quad [I(x^2)]_B = [x^2]_B$$

but computing these isn't any easier than the original problem. Instead, lets try and compute $[I]^S_B$. Its columns are:

$$[I(x+1)]_S = [x+1]_S = \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}$$

$$[I(2x^2-x+1)]_S = [2x^2-x+1]_S = \begin{pmatrix} 1 \\ -1 \\ 2 \end{pmatrix} \implies [I]^S_B = \begin{pmatrix} 1 & 1 & 2 \\ 1 & -1 & 0 \\ 0 & 2 & 1 \end{pmatrix}$$

$$[I(x^2+2)]_S = [x^2+2]_S = \begin{pmatrix} 2 \\ 0 \\ 1 \end{pmatrix}$$

To find $[I]^B_S$, we then compute the inverse:

$$([I]^S_B \mid I) = \left(\begin{array}{ccc|ccc} 1 & 1 & 2 & 1 & 0 & 0 \\ 1 & -1 & 0 & 0 & 1 & 0 \\ 0 & 2 & 1 & 0 & 0 & 1 \end{array}\right) \xrightarrow{R2-R1} \left(\begin{array}{ccc|ccc} 1 & 1 & 2 & 1 & 0 & 0 \\ 0 & -2 & -2 & -1 & 1 & 0 \\ 0 & 2 & 1 & 0 & 0 & 1 \end{array}\right) \xrightarrow[R3+R2]{-\frac{1}{2}R2} \left(\begin{array}{ccc|ccc} 1 & 1 & 2 & 1 & 0 & 0 \\ 0 & 1 & 1 & 1/2 & -\frac{1}{2} & 0 \\ 0 & 0 & -1 & -1 & 1 & 1 \end{array}\right)$$

$$\xrightarrow[-R3]{\substack{R1+2R3 \\ R2+R3}} \left(\begin{array}{ccc|ccc} 1 & 1 & 0 & -1 & 2 & 2 \\ 0 & 1 & 0 & -1/2 & 1/2 & 1 \\ 0 & 0 & 1 & 1 & -1 & -1 \end{array}\right) \xrightarrow{R1-R2} \left(\begin{array}{ccc|ccc} 1 & 0 & 0 & -1/2 & 3/2 & 1 \\ 0 & 1 & 0 & -1/2 & 1/2 & 1 \\ 0 & 0 & 1 & 1 & -1 & -1 \end{array}\right)$$

$$\underbrace{\begin{matrix} -1/2 & 3/2 & 1 \\ -1/2 & 1/2 & 1 \\ 1 & -1 & -1 \end{matrix}}_{[I]^B_S}$$

So

$$[x^2+x+1]_B = \begin{pmatrix} -1/2 & 3/2 & 1 \\ -1/2 & 1/2 & 1 \\ 1 & -1 & -1 \end{pmatrix} \cdot \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} = \begin{pmatrix} 2 \\ 1 \\ -1 \end{pmatrix}$$

Let's verify this:

$$2(x+1) + 1 \cdot (2x^2-x+1) - 1 \cdot (x^2+2) = 2x+2+2x^2-x+1-x^2-2 = x^2+x+1 \checkmark$$

□

- In the previous example, it really helped that one of the two bases was the standard basis. But one can always ensure that a sufficiently nice basis S appears. Indeed let A and B be two arbitrary bases for V, and let S be a "nice" basis (i.e. one we can compute $[\cdot]_S$ easily for). Then

$$[I]^B_A = [I \circ I]^B_A = [I]^B_S \, [I]^S_A$$

Then $[I]^S_B$ and $[I]^S_A$ will be easy to compute and you just need to invert $\nearrow$ and multiply by $\nwarrow$.

Ex Let $S = \{1, x, x^2\}$ and $B = \{x+1, 2x^2-x+1, x^2+2\}$, as before. Now consider a third basis:

$$A = \{x^2+x, \; x^2+2x+1, \; 2x^2+x+2\}$$

Then

$$[I]^S_A = \begin{pmatrix} 0 & 1 & 2 \\ 1 & 2 & 1 \\ 1 & 1 & 2 \end{pmatrix}$$

So using our previous computation we have:

$$[I]^B_A = [I]^B_S \cdot [I]^S_A = \begin{pmatrix} -1/2 & 3/2 & 1 \\ -1/2 & 1/2 & 1 \\ 1 & -1 & -1 \end{pmatrix} \cdot \begin{pmatrix} 0 & 1 & 2 \\ 1 & 2 & 1 \\ 1 & 1 & 2 \end{pmatrix} = \begin{pmatrix} 5/2 & 7/2 & 5/2 \\ 3/2 & 3/2 & 3/2 \\ -2 & -2 & -1 \end{pmatrix}$$

We can verify this on the basis A:

$$\tfrac{5}{2}(x+1) + \tfrac{3}{2}(2x^2-x+1) - 2(x^2+2) = x^2+x \checkmark$$
$$\tfrac{7}{2}(x+1) + \tfrac{3}{2}(2x^2-x+1) - 2(x^2+2) = x^2+2x+1 \checkmark$$
$$\tfrac{5}{2}(x+1) + \tfrac{3}{2}(2x^2-x+1) - 1(x^2+2) = 2x^2+x+2 \checkmark$$

10/17

$\square$

Changing coordinates of a matrix representation

- Let $T: V \to W$ be a linear trans., and suppose we know $[T]^B_A$ for bases A for V and B for W. Let $\tilde{A}, \tilde{B}$ be new bases for V and W, respectively. How do we compute $[T]^{\tilde{B}}_{\tilde{A}}$ from $[T]^B_A$? Observe

$$T = I_W \circ T \circ I_V$$

So

$$[T]^{\tilde{B}}_{\tilde{A}} = [I_W \circ T \circ I_V]^{\tilde{B}}_{\tilde{A}} = [I_W]^{\tilde{B}}_B \, [T]^B_A \, [I_V]^A_{\tilde{A}}$$

Thus we simply multiply $[T]^B_A$ by change of coordinate matrices.

- Now, consider the special case $T: V \to V$. Let A and B be bases for V. Then we have

$$[T]^B_B = [I_V]^B_A \, [T]^A_A \, [I_V]^A_B$$

and recall

$$[I_V]^B_A = \left([I_V]^A_B\right)^{-1}$$

So denoting $Q := [I]^A_B$, we have

$$[T]^B_B = Q^{-1}[T]^A_A Q$$

This special relationship between $[T]^B_B$ and $[T]^A_A$ is an example of a more general relationship between matrices:

Def For $A, B \in M_{n\times n}$, we say A is similar to B if there exists an invertible $Q \in M_{n \times n}$ such that

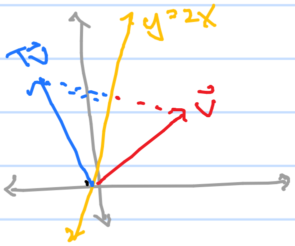
$$A = Q^{-1}BQ$$

- Observe that

$$A = Q^{-1}BQ \iff QAQ^{-1} = B$$

so if A is similar to B, then B is similar to A, and therefore we can just say A and B are similar.

Ex Define $T: \mathbb{R}^2 \to \mathbb{R}^2$ by letting $T(\vec{v})$ be the reflection of $\vec{v}$ over the line $y = 2x$. We encountered this example back in Section 1.5, where we showed:



$$* \quad [T] = \underbrace{\frac{1}{5}\begin{pmatrix}1 & -2\\ 2 & 1\end{pmatrix}}_{R_\theta}\begin{pmatrix}1 & 0\\ 0 & -1\end{pmatrix}\underbrace{\begin{pmatrix}1 & 2\\ -2 & 1\end{pmatrix}}_{R_{-\theta}} = \frac{1}{5}\begin{pmatrix}-3 & 4\\ 4 & 3\end{pmatrix}$$

where θ is the angle between $y = 2x$ and $y = 0$. Let

$$S = \left\{\begin{pmatrix}1\\0\end{pmatrix}, \begin{pmatrix}0\\1\end{pmatrix}\right\}$$

Then $[T] = [T]^S_S$. Consider

$$R_\theta\begin{pmatrix}1\\0\end{pmatrix} = \begin{pmatrix}1/5\\2/5\end{pmatrix} =: \vec{b}_1 \quad \text{and} \quad R_\theta\begin{pmatrix}0\\1\end{pmatrix} = \begin{pmatrix}-2/5\\1/5\end{pmatrix} =: \vec{b}_2$$

Then $B = \{\vec{b}_1, \vec{b}_2\}$ is another basis s.t. $R_\theta = [I]^S_B$ and $R_{-\theta} = [I]^B_S$. Indeed

$$\vec{b}_i = [\vec{b}_i]_S = [I]^S_B[\vec{b}_i]_B = [I]^S_B \vec{e}_i$$

So the columns of $[I]^S_B$ are $\vec{b}_1, \vec{b}_2$, which means it is R_θ. Then

$$R_{-\theta} = R_\theta^{-1} = \left([I]^S_B\right)^{-1} = [I]^B_S.$$

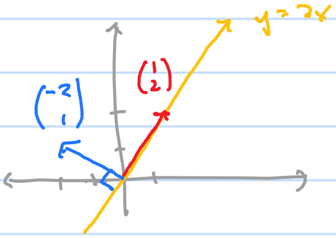
So (*) is then equivalent to:

$$[T]_S^S = [I]_B^S \underbrace{\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}}_{[T]_B^B.} [I]_S^B$$

Indeed, the columns of $[T]_B^B$ are:

$$[T(\vec{b}_1)]_B = [\vec{b}_1]_B = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$$

$$[T(\vec{b}_2)]_B = [-\vec{b}_2]_B = \begin{pmatrix} 0 \\ -1 \end{pmatrix}$$



So rephrasing our original method for finding $[T]_S^S$, what we did was find a basis $\mathcal{B}$ which was more compatible with T, so that $[T]_B^B$ is easy to compute. Then we simply used change of coordinates to find $[T]_S^S$. □