Systems of Equations and Linear Transformations

We know that solving a system of equations

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1\ell}x_\ell &= b_1 \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2\ell}x_\ell &= b_2 \\ \vdots \qquad \vdots \qquad\qquad &\vdots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{m\ell}x_\ell &= b_m \end{aligned}$$

is equivalent to solving

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

with

$$A = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1\ell} \\ a_{21} & a_{22} & & \\ \vdots & & \ddots & \vdots \\ a_{m1} & & \cdots & a_{m\ell} \end{pmatrix} \qquad \vec{x} = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_\ell \end{pmatrix} \qquad \vec{b} = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{pmatrix}$$

Hence is equivalent to solving $L(\vec{x}) = \vec{b}$ for the linear transformation L associated to A.

Questions about Solutions to Systems of Equations

When we interpret the matrix A as the matrix for linear transformation $L : \mathbb{R}^\ell \to \mathbb{R}^m$ and think about the system $A\vec{x} = \vec{b}$ then:

- Asking is there a solution?

  is asking is $\vec{b} \in \mathrm{Im}(L)$.

- Asking what are the solutions when $\vec{b} = \vec{0}$, that is solutions to $A\vec{x} = \vec{0}$?

  is asking what is $\mathrm{Ker}(L)$.

- Writing the general form of the solution to $A\vec{x} = \vec{0}$?

  is asking to find a basis for $\mathrm{Ker}(L)$, That is writing the solutions in the form:

  $$\vec{x} = s_1\vec{k}_1 + \cdots + s_n\vec{k}_n$$

  for some basic solutions $\vec{k}_1, \ldots, \vec{k}_n$ which are a a basis for $\mathrm{Ker}(L)$.

- Finding a solution to $A\vec{x} = \vec{b}$

  is finding a vector $\vec{u}$ with $L(\vec{u}) = \vec{b}$

- Finding all the solutions to $A\vec{x} = \vec{b}$

  is like finding a single solution $\vec{u}$, then noticing the general form of the solution is

  $$\vec{u} + s_1\vec{k}_1 + \cdots + s_n\vec{k}_n$$

  where $s_1\vec{k}_1 + \cdots + s_n\vec{k}_n \in \mathrm{Ker}(L)$ is the general form of a solution to $A\vec{x} = \vec{0}$.

Observations - Injectivity and solutions to systems of equations.

Injective

- If L is injective then if there is a solution to $A\vec{x} = \vec{b}$ then it is unique!
- If $\mathrm{Ker}(L) = \{\vec{0}\}$ then if there is a solution to $A\vec{x} = \vec{b}$ then it is unique!
- If $\mathrm{Null}(L) = 0$ then if there is a solution to $A\vec{x} = \vec{b}$ then it is unique!

Not Injective

- If L is **not** injective then if there is a solution to $A\vec{x} = \vec{b}$ then there are infinitely many!
- If $\mathrm{Ker}(L) \neq \{\vec{0}\}$ then if there is a solution to $A\vec{x} = \vec{b}$ then there are infinitely many!
- If $\mathrm{Null}(L) > 0$ then if there is a solution to $A\vec{x} = \vec{b}$ then there are infinitely many!

Observations - Surjectivity/bijectivity and solutions to systems of equations.

Surjective

- If L is surjective then $A\vec{x} = \vec{b}$ always has at least one solution!
- If $\mathrm{Im}(L) = \mathbb{R}^m$ then $A\vec{x} = \vec{b}$ always has at least one solution!
- If $\mathrm{Rank}(L) = m$ then $A\vec{x} = \vec{b}$ always has at least one solution!

Bijective

- If L is bijective then there is always a unique solution to $A\vec{x} = \vec{b}$.
- If $\mathrm{Ker}(L) = \{\vec{0}\}$ and $\mathrm{Im}(L) = \mathbb{R}^m$ then there is always a unique solution to $A\vec{x} = \vec{b}$.
- If $\mathrm{Null}(L) = 0$ and $\mathrm{Rank}(L) = m$ then there is always a unique solution to $A\vec{x} = \vec{b}$.

So injective/surjective/bijective just distinguish between cases: no solutions, unique solution, infinitely many solutions

Reduced Row Echelon Form

Suppose A is a matrix, and E is an invertible matrix.

The most interesting case is if EA is in reduced row echelon form!

Lemma

We have $\mathrm{Ker}(A) = \mathrm{Ker}(EA)$ and hence $\mathrm{null}(A) = \mathrm{null}(EA)$.

Proof Idea The solutions to the two systems are the same implies the kernels are equal. If the vector spaces are the same, the ranks are the same.

Lemma

We have $\mathrm{Im}(EA) = E(\mathrm{Im}(A))$ and hence if $\vec{f}_1, \ldots, \vec{f}_\ell$ are a basis for $\mathrm{Im}(A)$ then $E\vec{f}_1, \ldots, E\vec{f}_\ell$ a basis for $\mathrm{Im}(\mathrm{EA})$ and hence $\mathrm{rank}(A) = \mathrm{rank}(EA)$.

Proof Idea For functions it is true that

$$x \in \mathrm{Im}(E \circ A) \iff x \in E(\mathrm{A})$$

so these define the same sets.

Now, because E is injective, it takes linearly independent to linearly independent, so

$$E\vec{f}_1, \ldots, E\vec{f}_\ell$$

are linearly independent, and as they also generate $E(\mathrm{Im}(\mathrm{A}))$ they are a basis.

Because $\mathrm{Im}(A)$ and $\mathrm{Im}(EA)$ have basis of the same size, they have the same dimension, and hence A and EA have the same rank.

Rank/Nullity and Reduced Row Echelon Form

The following are easy to check for EA, but by the previous slide, are true for A!.

- The nullity of L is exactly the number of **parameters** needed to describe the general form of a solution to $A\vec{x} = \vec{b}$ whenever a solution exists.
- The nullity of L is exactly the number of **non-pivot columns** when A is put into row reduced echelon form.
- The rank of L is the dimension of the subspace of $\mathbb{R}^m$ for which a solution will actually exists.
- The rank of L is exactly the number of non-zero **rows** when A is put into row reduced echelon form.
- The rank of L is thus exactly the number of **pivot columns** when A is put into row reduced echelon form.
- The rank of L plus the nullity of L must be equal to the total number of **columns** of A, that is ℓ.

Theorem[Rank-Nullity Theorem]
If $L : U \to V$ is a map of finite dimensional vector spaces then

$$\mathrm{Rank}(L) + \mathrm{Null(L)} = \mathrm{Dim}(\text{domain of L})$$

The above is our first *proof* of this, we shall see another later.

Pessimistic/Optimistic Perspective: Transformations vs Equations

Pessimistic

Everything we have just done, definitions of linear transformations, images, kernels, ranks, nullities has been to find a way to make what we were already doing in previous linear algebra courses harder and more confusing.

Optimistic

Everything we are doing now is just as easy as what we had been doing in previous courses.

Reality

Probably both, except that it is **much much** easier to prove things about images/kernels/ranks/nullities in an abstract way than it is to prove things about matricies in an explicit way.

$(AB)^t = B^t A^t$ is one of the easier things to prove about matricies, and it is already notationally heavy.

A bonus question on A3 gives an abstract interpretation of taking a transpose, in this abstract interpretation, the analog of $(AB)^t = B^t A^t$ is equivalent to

$$h \circ (L \circ M) = (h \circ L) \circ M$$

that is composition of functions is associative.

Sadly, explaining why an abstract interpretation is equivalent to a concrete one is often still hard.