Felix Breuer
G-04

Linear Algebra
Week 7

Recap: Solutions of SLE

1. Any homogenous system of linear equations (SLE) has at least one solution. TRUE $x=0$ always solves $Ax=0$
2. If $A \in \mathbb{R}^{n \times n}$ is invertible there is at most one nonzero solution $x \in \mathbb{R}^n$ to $Ax=0$ TRUE (there are 0 nonzero solutions)
3. For $A \in \mathbb{R}^{m \times n}$, $x \in \mathbb{R}^n$, $b \in \mathbb{R}^m$. $Ax=b$ has a solution if and only if $b \in C(A)$ TRUE

RREF, recap Gauss-Elimination

Applying Gaussian Elimination on $Ax=b$ to get $REF(A)=U$
i.e. multiplying with elimination matrices from left

- doesn't affect solution set of underlying SLE
  $\longrightarrow Ux=\tilde{b}$ has same solutions as $Ax=b$,
  specifically for $b=0$
  $\rightarrow$ we have to apply elimination matrices on b as well!
  (span of rows is preserved $\rightarrow$ this is an exercise this week)
- preserves linear dependence relations between columns

This is why we can compute $N(A)$, $C(A)$ the way we did last week

row echelon form (REF)

1. All zero rows at bottom
2. First nonzero entry is strictly to the right to first nonzero element of row above

(⇒ all entries below pivot zero)

reduced (RREF) if also

3. Each pivot is 1
4. All entries aside pivot in each column are zero

Some facts

Let $A \in \mathbb{R}^{n \times n}$:

- RREF(A) is unique!
- A is invertible $\Longleftrightarrow$ RREF(A) = I
- The R in the CR decomposition is RREF(A) without zero rows

Which of the following are in reduced row echelon form (RREF)?

$$\begin{bmatrix} 3 & 2 & 4 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix} \quad \begin{bmatrix} 0 & 1 & 2 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \quad \begin{bmatrix} 0 & 1 & 2 \\ 1 & 0 & 3 \\ 0 & 0 & 4 \end{bmatrix} \quad \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix}$$

Gauss Elimination vs. Gauss-Jordan Elimination

Gauss Elimination

- reduce to REF
- easier to compute, more commonly used

Gauss-Jordan Elimination

"can't simplify further with elimination"

- reduce to RREF
- allows us to practically read off solution
- requires more elimination steps

Example

we computed this in week 4

$$A = \begin{bmatrix} 1 & 2 & 3 \\ 1 & 3 & 2 \\ 2 & 1 & 1 \end{bmatrix}, \quad U = \begin{bmatrix} 1 & 2 & 3 \\ 0 & 1 & -1 \\ 0 & 0 & -8 \end{bmatrix} = \mathrm{REF}(A)$$

$\mathrm{RREF}(A) = ?$

$\longrightarrow$ eliminate as much as possible:

$$\begin{bmatrix} 1 & 2 & 3 \\ 0 & 1 & -1 \\ 0 & 0 & -8 \end{bmatrix} \sim \begin{bmatrix} 1 & 2 & 3 \\ 0 & 1 & -1 \\ 0 & 0 & 1 \end{bmatrix} \sim \begin{bmatrix} 1 & 2 & 3 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \sim \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

The CR decomposition, RREF(A) and N(A)

$$A = CR$$

C: (first) independent columns of A

R: information how to combine columns in C to get A

Example from week 3:

https://www.felixgbreuer.com/week3.pdf

Example

$$A = \begin{bmatrix} 3 & -3 & 1 & 8 & 0 \\ 2 & -2 & 0 & 4 & 0 \\ 4 & -4 & 0 & 8 & 1 \end{bmatrix} \qquad C = \begin{bmatrix} 3 & 1 & 0 \\ 2 & 0 & 0 \\ 4 & 0 & 1 \end{bmatrix}$$

$$R = \begin{bmatrix} 1 & -1 & 0 & 2 & 0 \\ 0 & 0 & 1 & 2 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix}$$

col 4 of A is 2x col 1 + 2x col 2 of C

Which of these are true?

- ☐ $C(A) = C(R)$
- ☐ $C(A) = C(C)$
- ☐ $R(C) = R(R)$
- ☐ $R(A) = R(R)$
- ☐ the columns of C form a basis of $C(A)$
- ☐ the rows of C form a basis of $R(A)$
- ☐ R with zero rows removed equals RREF(A)

(see end of next page for solutions)

columns of C span columns of A
rows of R span rows of A

Why? -> row/column view

$$\underbrace{\begin{bmatrix} - & a_1 & - \\ & \vdots & \\ - & a_m & - \end{bmatrix}}_{=A} \underbrace{\begin{bmatrix} | & & | \\ b_1 & \cdots & b_n \\ | & & | \end{bmatrix}}_{=B} = \begin{bmatrix} | & & | \\ Ab_1 & \cdots & Ab_n \\ | & & | \end{bmatrix} = \begin{bmatrix} - & a_1 B & - \\ & \vdots & \\ - & a_m B & - \end{bmatrix}$$

What's the connection to $C(A)$, $N(A)$, basis?

⊛ C's columns form a basis of $C(A)$

R's rows form a basis of $R(A)$

solutions to the true/false questions above: ✗ ✓ ✗ ✓   ✓ ✗ ✓

The following is a derivation showing that choosing $R = RREF(A)$ without zero rows yields a factorization $A = CR$ where the properties ⊛ hold.

We can reduce A to $rref(A)$ by elementary row operations:

$$\underbrace{E}_{\substack{\text{elementary}\\\text{row ops}}} A = rref(A) = \begin{bmatrix} R_{r \times n} \\ 0_{m-r \times n} \end{bmatrix} \begin{array}{l} \leftarrow \text{rref (A) without zero rows} \\ \leftarrow \text{zero rows at bottom in rref(A)} \end{array}$$

$E \in \mathbb{R}^{m \times m}$ is the product of these elimination matrices

$E = E_k \cdots \underbrace{E_1}_{\substack{\text{first step}\\\text{in Gauss-Jordan}\\\text{elimination}}}$ where all E_i are invertible.

Hence E^{-1} exists:

$$A = E^{-1} E A = E^{-1} \begin{bmatrix} R_{r \times n} \\ 0_{m-r \times n} \end{bmatrix} = \Big[\underbrace{E_1^{-1}}_{\substack{\text{first}\\ r \text{ columns}}} \; \underbrace{E_2^{-1}}_{\substack{\text{last}\\ n-r\\ \text{columns}}} \Big] \begin{bmatrix} R_{r \times n} \\ 0_{m-r \times n} \end{bmatrix} =$$

$$= E_1^{-1} R + E_2^{-1} 0 \quad = E_1^{-1} R_{r \times n}$$

The rows of R span $R(A)$ as elementary row ops don't change span of rows (and we only removed 0 rows). In $RREF(A)$ the nonzero rows are linearly independent, hence R's rows are a basis for $R(A)$.

With the definition of matrix multiplication ("column/row-view") we find that the columns of A are linear combinations of the columns of E_1^{-1}. As no row of R is zero and the columns of E_1^{-1} are linearly independent (it is invertible) $C(A) = C(E_1^{-1})$ and the columns of E_1^{-1} form a basis of $C(A)$.

A general solution to $Ax=b$

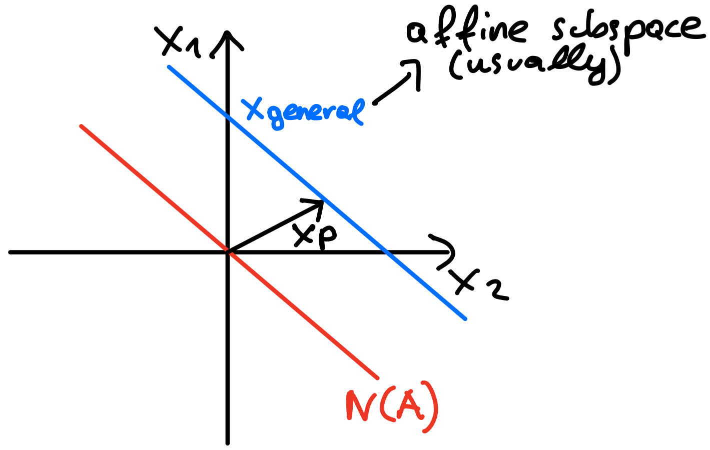
The general solution set $x_{general} = \{x \in \mathbb{R}^n \mid Ax=b\}$ of $Ax=b$ for any $A \in \mathbb{R}^{m \times n}$ can be expressed as

$$x_{general} = \{\underbrace{x_p}_{\substack{\text{one particular}\\\text{solution of } Ax=b\\(Ax_p=b)}} + x_H \mid x_H \in \underbrace{N(A)}_{\substack{\text{general homogenous}\\\text{solution set of } Ax=0}}\}$$

Example

$$A = \begin{bmatrix} 1 & 1 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \overset{b}{\overset{\shortparallel}{\begin{bmatrix} 5 \\ 0 \end{bmatrix}}}, \quad x_p = \begin{bmatrix} 3 \\ 2 \end{bmatrix}$$

$$N(A) = \text{span}\left(\left\{\begin{bmatrix} 1 \\ -1 \end{bmatrix}\right\}\right)$$



Let $A \in \mathbb{R}^{m \times n}$, $x \in \mathbb{R}^n$, $b \in \mathbb{R}^m$, $Ax=b$:
We now consider x_p such that $Ax_p = b$, $x_H \in N(A)$.
First, we confirm that in fact $A(x_p + x_H) = b$:

$$A(x_p + x_H) \overset{\text{distr.}}{=} Ax_p + Ax_H \overset{\text{def } N(A)}{=} b + 0 = b$$

Now we show that any x that solves $Ax=b$ can be described as $x = x_p + x_H'$ for some $x_H' \in N(A)$:
We have $Ax=b$ and $Ax_p = b$.
Hence $Ax - Ax_p \overset{\text{distr.}}{=} A(x-x_p) = b-b = 0$. Let $x_H' = x - x_p$.
It directly follows that $x_H' \in N(A)$ and $x = x_p + x_H'$.

Geometric interpretation of matrices

Very useful to see them as functions that transform space: linear transformations (also often called linear map, linear functions)

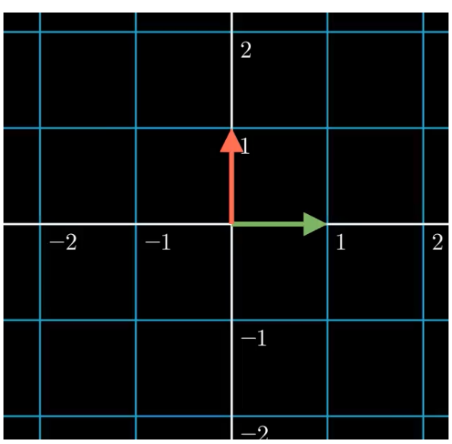
Let U, V be vector spaces over some field $\mathbb{F}$:
$f: X \rightarrow Y$ is a linear map if:
1. $f(u+v) = f(u) + f(v)$ for any $u, v \in X$
2. $f(\alpha v) = \alpha f(v)$ for any $\alpha \in \mathbb{F}$, $v \in X$

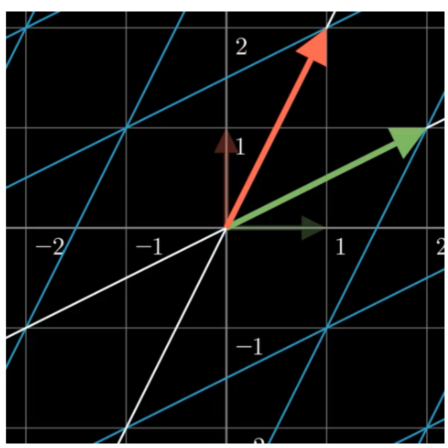
Update Nov 24: This is being covered in the lecture right now!

Some facts

- Any matrix is a linear map and we can express any linear map as a matrix if we fix a basis
- If we know how a linear map transforms each of our basis vectors, we know what the linear map does to any vector

Example (see felixgbreuer.com/week_7 for animations and code)





Recommendation: https://www.3blue1brown.com/lessons/linear-transformations

The four fundamental subspaces

Let $U, W \subseteq \mathbb{R}^n$ be subspaces of $\mathbb{R}^n$:

$U \perp W$ if for any $v \in U$, $w \in W$: $v \cdot w = 0$

(U is orthogonal to W)

$$U^\perp = \{v \in \mathbb{R}^n \mid v \cdot u = 0 \text{ for all } u \in U\}$$

is the orthogonal complement of U, the set of vectors that are orthogonal to all vectors in U

The following holds:

- $U^\perp$ is a subspace of $\mathbb{R}^n$
- $(U^\perp)^\perp = U$
- $\mathbb{R}^n = U^\perp \oplus U$
- $\dim \mathbb{R}^n = \dim U^\perp + \dim U$ (this holds for any direct sum)

Let $A \in \mathbb{R}^{m \times n}$ with $\operatorname{rank}(A) = r$:

$$C(A) = N(A^T)^\perp$$

$$\underbrace{C(A^T)}_{= R(A)} = N(A)^\perp$$

Now applying the properties listed above gives us:

$$\mathbb{R}^m = C(A) \oplus N(A^T)$$

$$\mathbb{R}^n = C(A^T) \oplus N(A)$$

$\dim C(A) = r$ $\quad$ $\dim N(A^T) = m - r$

$\dim C(A^T) = r$ $\quad$ $\dim N(A) = n - r$

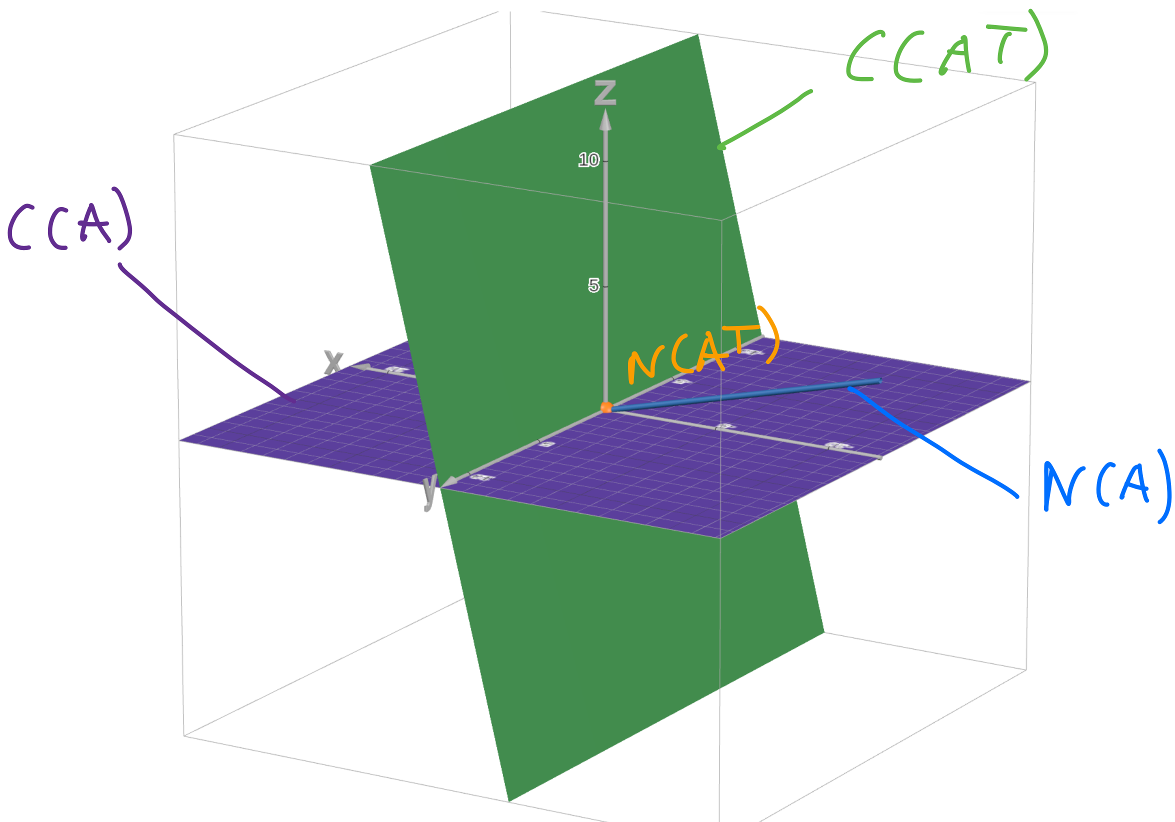
We can count the pivots in RREF

Example $A = \begin{bmatrix} 1 & 2 & 4 \\ 2 & 5 & 8 \end{bmatrix}$ ← we will discuss this (or a similar example) next week

subspace	basis	dimension
$C(A)$	$\left\{\begin{bmatrix}1\\2\end{bmatrix}, \begin{bmatrix}2\\5\end{bmatrix}\right\} \subseteq \mathbb{R}^2$	2
$N(A)$	$\left\{\begin{bmatrix}-4\\0\\1\end{bmatrix}\right\} \subseteq \mathbb{R}^3$	1
$C(A^T)$	$\left\{\begin{bmatrix}1\\2\\4\end{bmatrix}, \begin{bmatrix}2\\5\\8\end{bmatrix}\right\} \subseteq \mathbb{R}^3$	2
$N(A^T)$	$\{\ \} \subseteq \mathbb{R}^2$	0

$m = \#\text{rows} = 2 = \dim C(A) + \dim N(A^T) = 2 + 0$

$n = \#\text{columns} = 3 = \dim C(A^T) + \dim N(A) = 2 + 1$



https://www.desmos.com/3d/d8fb99479d

References:
Last years course for some definitions
Sergey Treil, Linear Algebra Done Wrong, https://www.math.brown.edu/streil/papers/LADW/LADW_2021_01-11.pdf