18.06

Zack Remscrim

Review for the Final Exam

Main Topics:

- Equation $Ax = b$
- Four Fundamental Subspaces
- SVD and other matrix factorizations
- Fundamental Theorem of Linear Algebra
- Determinant
- Eigenvalues and Eigenvectors

Four Fundamental Spaces

For a matrix A

Column space: $\mathrm{col}(A)$

Row space: $\mathrm{row}(A)$

Nullspace: $\mathrm{null}(A)$

Left Nullspace: $\mathrm{null}(A^{\top})$

What are these spaces?

Why do we study them?

How can we compute them?

Nullspace

A $m \times n$ matrix

Nullspace: $\text{null}(A) = \{x \in \mathbb{R}^n | Ax = 0\}$

A is a function

input: $x \in \mathbb{R}^n$

output: $y \in \mathbb{R}^m$

$y = Ax$

$\text{null}(A)$: $\{x \in \mathbb{R}^n | A(x) = 0\}$

Familiar example:

f is a function

input: $x \in \mathbb{R}$

output: $y \in \mathbb{R}$

$y = f(x)$

$\text{zeros}(f)$: $\{x \in \mathbb{R} | f(x) = 0\}$

Nullspace

A $m \times n$ matrix: $\text{null}(A) = \{x \in \mathbb{R}^n | Ax = 0\}$

Consider equation $Ax = 0$

This equation always has a solution

Set of all solutions is $\text{null}(A)$

Consider equation $Ax = b$

This equation may or may not have a solution

If has solution, let x_p denote some solution ($Ax_p = b$)

Set of all solutions is $x_p + \text{null}(A)$

$$\{x_p + x_n | x_n \in \text{null}(A)\}$$

Column Space

A $m \times n$ matrix

Column space: $\text{col}(A) = \{$ all linear combinations of columns $\}$

A is a function

input: $x \in \mathbb{R}^n$

output: $y \in \mathbb{R}^m$

$y = Ax$

$\text{col}(A)\colon \{y \in \mathbb{R}^m | Ax = y, x \in \mathbb{R}^n\}$

Familiar example:

f is a function

input: $x \in \mathbb{R}$

output: $y \in \mathbb{R}$

$y = f(x)$

$\text{range}(f)\colon \{y \in \mathbb{R} | f(x) = y, x \in \mathbb{R}\}$

Column Space

A $m \times n$ matrix: $\text{col}(A) = \{y \in \mathbb{R}^m | Ax = y, x \in \mathbb{R}^n\}$

Consider equation $Ax = b$

This equation may or may not have a solution

Has at least one solution exactly when $b \in \text{col}(A)$

Row Space and Left Nullspace

A $m \times n$ matrix

Row space: $\text{row}(A) = \text{col}(A^\top)$

Left Nullspace: $\text{null}(A^\top)$

$A^\top$ is a function

input: $y \in \mathbb{R}^m$

output: $x \in \mathbb{R}^n$

$x = A^\top y$

$\text{row}(A)$: $\{x \in \mathbb{R}^n | A^\top y = x, y \in \mathbb{R}^m\}$

$\text{null}(A^\top) = \{y \in \mathbb{R}^m | A^\top y = 0\}$

Careful: $A^\top$ is not (generally) the inverse of A

Singular Value Decomposition

A $m \times n$ matrix, of rank r

Write $A = \begin{bmatrix} U_1 & U_2 \end{bmatrix} \begin{bmatrix} \Sigma_r & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} V_1^\top \\ V_2^\top \end{bmatrix}$

Σ_r $r \times r$ diagonal matrix, with all positive singular values on diagonal

U_1 first r columns of U

V_1 first r columns of V

U_1, U_2, V_1, V_2 are all orthogonal matrices

Why Use the SVD?

If A has (compact) SVD $A = U_1 \Sigma_r V_1^{\top}$, then $\mathrm{col}(A) = \mathrm{col}(U_1)$
Equation $Ax = b$ Has at least one solution exactly when $b \in \mathrm{col}(A) = \mathrm{col}(U_1)$

Columns of U_1 span $\mathrm{col}(A)$
Columns of A span $\mathrm{col}(A)$

Columns of U_1 are linearly independent
Columns of A may or may not be linearly independent

Columns of U_1 form a basis of $\mathrm{col}(A)$
Columns of A may or may not form a basis of $\mathrm{col}(A)$

Easier to test $b \in \mathrm{col}(U_1)$ than $b \in \mathrm{col}(A)$

Projection onto col(U): U orthogonal matrix

Given:

vector $b \in \mathbb{R}^m$

orthogonal $m \times n$ matrix U

Want:

vector $\tilde{b} \in$ col(U)

$\tilde{b}$ as close to b as possible

How:

$\tilde{b} = UU^\top b$

$UU^\top$ is "projection matrix" transforms b to $\tilde{b}$

Fundamental Theorem of Linear Algebra

A $m \times n$ matrix, of rank r, where $A = \begin{bmatrix} U_1 & U_2 \end{bmatrix} \begin{bmatrix} \Sigma_r & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} V_1^\top \\ V_2^\top \end{bmatrix}$

$\text{col}(A) = \text{col}(U_1)$
$\text{row}(A) = \text{col}(V_1)$
$\text{null}(A) = \text{col}(V_2)$
$\text{null}(A^\top) = \text{col}(U_2)$

Tells us "everything" about A

General Solution to $Ax = b$

A $m \times n$ matrix, of rank r

When does $Ax = b$ have at least one solution?

Exactly when $b \in \mathrm{col}(A)$

Exactly when $b \in \mathrm{col}(U_1)$

Exactly when $U_1 {U_1}^\top b = b$

General Solution to $Ax = b$

A $m \times n$ matrix, of rank r

If $Ax = b$ has at least one solution, when does it have only one solution?

Exactly when null$(A) = \{0\}$ (only contains the zero vector)

Exactly when col$(V_2) = \{0\}$

Exactly when $r = n$

General Solution to $Ax = b$

A $m \times n$ matrix, of rank r

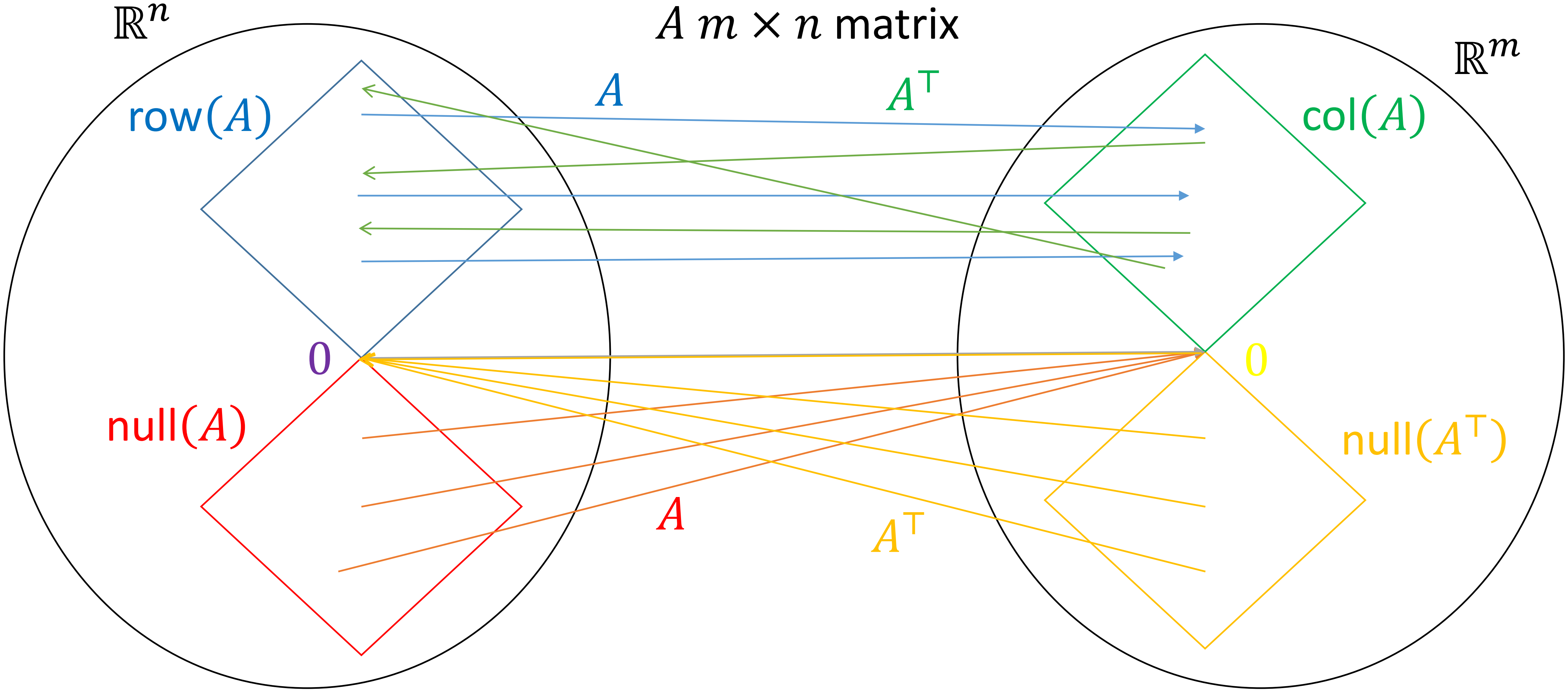
If $Ax = b$ has at least one solution, what are all of the solutions?

$x_p = V_1 \Sigma_r^{-1} U_1^\top b$ is a solution ($A = U_1 \Sigma_r V_1^\top$)

$x_p + \text{null}(A)$ is the set of all solutions

$x_p + \text{col}(V_2)$ is the set of all solutions

Four Fundamental Spaces



Determinant: Geometric Meaning in $\mathbb{R}^2$

Consider $A = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}$, 2×2 real matrix

and $e_1 = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$, $e_2 = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$ in $\mathbb{R}^2$

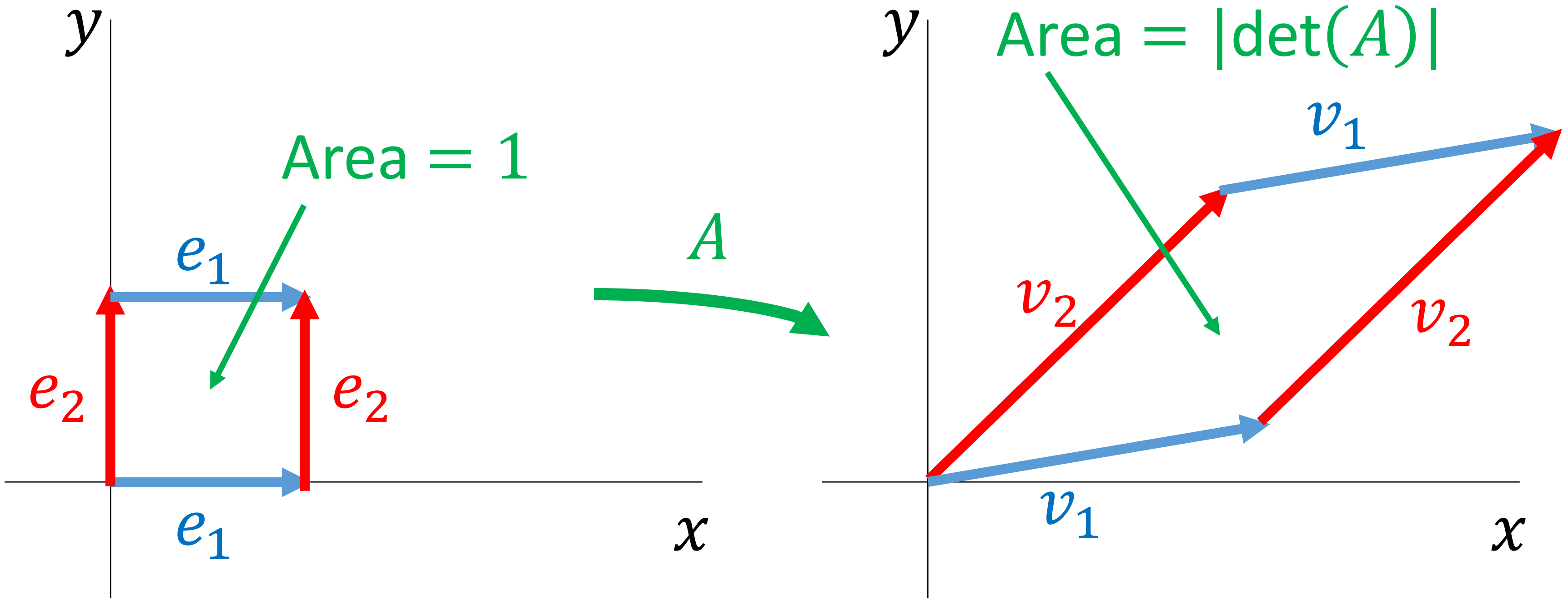
A transforms e_1, e_2 to

$$v_1 = Ae_1 = \begin{pmatrix} a_{11} \\ a_{21} \end{pmatrix}$$

$$v_2 = Ae_2 = \begin{pmatrix} a_{12} \\ a_{22} \end{pmatrix}$$

1st and 2nd col of A

A transforms square of area 1 to parallelogram of area $|\det(A)|$



Determinant: Geometric Meaning in $\mathbb{R}^n$

Consider A, $n \times n$ real matrix
and $e_1, e_2, \dots, e_n \in \mathbb{R}^n$, where e_i: i^{th} entry 1, all other entries 0

A transforms $e_1, e_2, \dots, e_n$ to

$v_1, v_2, \dots, v_n$, where $v_i = Ae_i$

v_i is i^{th} col of A

$e_1, e_2, \dots, e_n$ specify hypercube in $\mathbb{R}^n$ (n-dimensional version of cube)
$v_1, v_2, \dots, v_n$ specify parallelotope in $\mathbb{R}^n$ (n-dimensional version of parallelogram)

A transforms hypercube $e_1, e_2, \dots, e_n$ of volume 1
to parallelotope $v_1, v_2, \dots, v_n$ of volume $|\det(A)|$

Determinant: Geometric Meaning in $\mathbb{R}^n$

If $\det(A)$ is signed volume of parallelotope, det must satisfy certain properties

1) $\det(I) = 1$

   $v_1, v_2, \dots, v_n$ is $e_1, e_2, \dots, e_n$

2) If any column of A is all zeros, $\det(A) = 0$

   if i^{th} col of A is all zeros, $v_i = 0$, so volume is zero

3) Determinant is linear in each column separately

   $\text{Vol}(v_1, \dots, v_i, cv_{i+1}, v_{i+2}, \dots, v_n) = c\ \text{Vol}(v_1, \dots, v_i, v_{i+1}, v_{i+2}, \dots, v_n)$

   $\text{Vol}(v_1, \dots, v_i, v_{i+1} + v'_{i+1}, v_{i+2}, \dots, v_n)$

   $= \text{Vol}(v_1, \dots, v_i, v_{i+1}, v_{i+2}, \dots, v_n) + \text{Vol}(v_1, \dots, v_i, v'_{i+1}, v_{i+2}, \dots, v_n)$

Eigenvalues and Eigenvectors: Definition

Square matrix A

If $Av = \lambda v$, $v \neq 0$ a vector and λ a scalar

Then v is eigenvector of A with eigenvalue λ

v is a very special vector

Applying A to v only changes the magnitude of v, not direction

Applying A to v produces λv

λv is scaled version of v

Eigenvalues and Eigenvectors

Square matrix A,

- Vector $v \neq 0$ is an eigenvector with eigenvalue λ if $Av = \lambda v$
- Eigenvalues are the roots of $p_A(\lambda) = \det(\lambda I - A)$
- What about eigenvectors?
- If λ is an eigenvalue,
  - $v \neq 0$ is eigenvector with with eigenvalue λ if $(\lambda I - A)v = 0$
  - All $v \in$ null$(\lambda I - A)$, except $v = 0$
- Eigenvectors not unique,
  - Each eigenvalue λ has infinitely many corresponding eigenvectors
  - Example: If v eigenvector with eigenvalue λ, so is $2v, \frac{3}{2}v, -7v$, ...
  - Generally just write down a collection of lin. indep. eigenvectors

Why are Eigenvectors Useful?

Example: A a 2×2 matrix,

That has two lin. indep. eigenvectors v_1, v_2

With corresponding eigenvalues λ_1, λ_2

v_1, v_2 form a basis of $\mathbb{R}^2$ (eigenbasis)

Can write any $x \in \mathbb{R}^2$ as $c_1 v_1 + c_2 v_2$, for some $c_1, c_2 \in \mathbb{R}$

$$\begin{aligned} Ax &= A(c_1 v_1 + c_2 v_2) \\ &= c_1 A v_1 + c_2 A v_2 \\ &= c_1 \lambda_1 v_1 + c_2 \lambda_2 v_2 \end{aligned}$$

Each component scales by corresponding eigenvalue

Why are Eigenvectors Useful?

Example: A an $n \times n$ matrix,

That has n lin. indep. eigenvectors $v_1, \dots, v_n$

With corresponding eigenvalues $\lambda_1, \dots, \lambda_n$

$v_1, \dots, v_n$ form a basis of $\mathbb{R}^n$ (eigenbasis)

Can write any $x \in \mathbb{R}^n$ as $c_1 v_1 + \cdots + c_n v_n$, for some $c_i \in \mathbb{R}$

$$\begin{aligned} Ax &= A(c_1 v_1 + \cdots + c_n v_n) \\ &= c_1 A v_1 + \cdots + c_n A v_n \\ &= c_1 \lambda_1 v_1 + \cdots + c_n \lambda_n v_n \end{aligned}$$

Each component scales by corresponding eigenvalue

Much easier to compute in an eigenbasis, A behaves like diagonal matrix

Diagonalization

Consider A an $n \times n$ matrix,

That has n lin. indep. eigenvectors $v_1, \ldots, v_n$

With corresponding eigenvalues $\lambda_1, \ldots, \lambda_n$

Let X be $n \times n$ matrix with eigenvectors $v_1, \ldots, v_n$ in the columns

X invertible because it is a square matrix with lin. Indep. cols

Let Λ be $n \times n$ diagonal matrix with eigenvalues $\lambda_1, \ldots, \lambda_n$ along diagonal

Then $AX = X\Lambda$,

So $A = X\Lambda X^{-1}$

and $\Lambda = X^{-1}AX$

Diagonalization

If we can write $A = X\Lambda X^{-1}$, where Λ is diagonal,

we say we have diagonalized A

because A is "similar" to a diagonal matrix

We say matrices B and C are similar

if there is an invertible matrix T, such that $B = TCT^{-1}$

Similar matrices represent the same function, in different bases

Sometimes one basis is easier to work in than another

Diagonal matrices are especially easy to work with