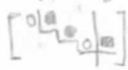
<u>CHAPTER 1</u> Skipped, just notation

<u>CHAPTER 2</u> Linear Systems

'A linear system has either:

① No solution  pivot in b'

② Exactly one unique solution  $\#_{pivot} = \#_{variable}$, no pivot in b'

③ ∞-many solution  $\#_{pivot} < \#_{variable}$, no pivot in b'

<u>Homogeneous</u> system when $\vec{b} = \vec{0}$

<u>Inhomogeneous</u> otherwise

Note that homogeneous $\Rightarrow \vec{b} = \vec{0} \Rightarrow \vec{b}'$ has no pivots $\Rightarrow \exists$ solution

<u>Gaussian Elimination</u> To find $\vec{x}$ satisfying $A\vec{x} = \vec{b}$,

① Write augmented matrix $[A|\vec{b}]$

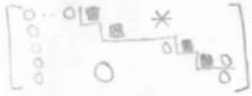
② Apply admissible row operations, i.e.

a. multiply row with scalar $\lambda \neq 0$ $(R_2' = 2R_2)$

b. add one row to another $(R_2' = R_1 + R_2)$

c. permute rows $(R_2 \leftrightarrow R_1)$

Until $[A|\vec{b}]$ is in reduced form (aka row-echelon),

 where pivots are <u>strictly</u> to the right of previous pivots

(Note that $\begin{bmatrix}0&0\\0&0\end{bmatrix}$ is reduced, though it has no pivots)

<u>Rank</u> $\operatorname{rank}(A) = \#_{pivots}$ in reduced A'

<u>"Rank-Nullity Theorem"</u> Consider the general solution to $A\vec{x} = \vec{0}$, $A_{m\times n}$.

$$\left.\begin{array}{c}\#_{\text{free parameters}} = n - \operatorname{rank}(A) \\ \text{(a variable is either fixed as a pivot or free)}\end{array}\right\} \text{revisited later}$$

<u>Square</u> $A_{n\times n}$ is a square matrix

<u>Non-singular</u> $A_{n\times n}$ and $(A\vec{x} = \vec{0} \Rightarrow \vec{x} = \vec{0})$
("invertible")

<u>THM</u> $A_{n\times n}$ non-singular

$\Leftrightarrow A \to A' =$ matrix without free parameters

$\Leftrightarrow \operatorname{rank}(A) = n$

$\Leftrightarrow A\vec{x} = \vec{b}$ has unique solution $\forall \vec{b}$

<u>Complex Numbers</u> ← ?? what is this topic doing here

$i^2 = -1$, standard algebraic rules apply

<u>Complex conjugate</u> of $z = a+ib$ is $\bar{z} = a - ib$

Note $z\bar{z} = a^2 + b^2$

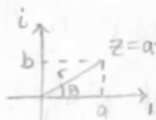
$\overline{z\cdot u} = \bar{z}\cdot\bar{u}$

$\overline{z+u} = \bar{z} + \bar{u}$

<u>Complex division</u> by multiplying $\frac{z}{u}$ with $\frac{\bar{u}}{\bar{u}}$ since $u\cdot\bar{u} \in \mathbb{R}$

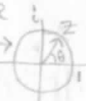
<u>Complex inverse</u> $\frac{1}{z} = z^{-1} = \frac{\bar{z}}{z\bar{z}} = \frac{\bar{z}}{a^2+b^2} = \frac{\bar{z}}{|z|^2}$

<u>Polar form</u> generally nicer to work with



$z = a+ib$ $r = \sqrt{a^2+b^2} = |z|$

$z = re^{i\theta} = r(\cos\theta + i\sin\theta)$

unit circle in $\mathbb{C}$ → 

$z = e^{i\theta}$, $|z| = 1$

Note that $re^{i\theta} r'e^{i\theta'} = rr'e^{i(\theta+\theta')}$

$$\frac{Re^{i\theta}}{re^{i\alpha}} = \frac{R}{r}e^{i(\theta-\alpha)}$$

$z = re^{i\theta} \Rightarrow z^n = r^n e^{in\theta}$

This lets us find the n n^{th} roots of 1 (roots of unity) as $e^{i\frac{2\pi}{n}k}$, $k = 1 \ldots n$

<u>THM</u> $(\mathbb{C}, +, \cdot)$ is a field with 0, 1 as the neutral elements of + and ·

(??! all fields are vector spaces, but not all vector spaces are fields – e.g. field requires multiplicative inverse, see math.stackexchange.com/a/969737. but wot)

CHAPTER 3 Computational Rules for Matrices, Vectors

	Associative	Commutative	Distributive
Matrix Multiplication	✓ $A(BC)=(AB)C$	✗ NOT IN GENERAL though sometimes	✓ $C(A+B)=CA+CB$
Matrix Addition	✓ $A+(B+C)=(A+B)+C$	✓ $A+B=B+A$	$\lambda(A+B)=\lambda A+\lambda B$

Can also view matrix operations graphically - see 3blue1brown

CHAPTER 4 Transpose, Inverse

$A_{m\times n}=\underbrace{[a_{ij}]}_{n \text{ columns}}\}m \text{ rows}$

Transpose $A^t_{n\times m}=\underbrace{[a^t_{ij}]}_{m\text{ columns}}\}n \text{ rows},\ a^t_{ij}=a_{ji}\quad \forall i,j$

Properties:

$(A^t)^t=A$

$(A+B)^t=A^t+B^t$

$(AB)^t=B^tA^t$

in general be able to prove properties by looking at a generic entry, e.g.

$(AB)^t_{ij}=(AB)_{ji}=\sum_k a_{jk}b_{ki}=\sum_k a^t_{kj}b^t_{ik}=\sum_k b^t_{ik}a^t_{kj}=(B^tA^t)_{ij}$

scalars, can be swapped

Inverse $A_{n\times n}$ is invertible iff $\exists B,C$ such that

$AB=I=CA$ (B= right inverse, C= left inverse)

Furthermore, $AB=I \Rightarrow (CA)B=CI \Rightarrow B=C$, i.e. left inverse = right inverse

$AB=I=AB' \Rightarrow AB=AB' \Rightarrow CAB=CAB' \Rightarrow B=B'$, i.e. inverse is unique

THM A invertible $\Rightarrow \exists! A^{-1}$ such that $A^{-1}A=AA^{-1}=I$

THM A has right inverse $\Rightarrow A$ has left inverse and $\therefore A$ invertible

(with this thm, only need to check right inverse in this class)

Properties:

$(AB)^{-1}=B^{-1}A^{-1}$ (consider $B^{-1}\underbrace{A^{-1}A}B$)

$(A^t)^{-1}=(A^{-1})^t$ (recall $B^tA^t=(AB)^t$, so $A^t(A^{-1})^t=(A^{-1}A)^t=I^t=I$)

$(A^{-1})^{-1}=A$

$A\neq B \Rightarrow A^{-1}\neq B^{-1}$ (because if $A^{-1}=B^{-1}$, then $\left.\begin{matrix}B^{-1}A=A^{-1}A=I\\ B^{-1}B=I\end{matrix}\right\}BB^{-1}A=BB^{-1}B \Rightarrow A=B$)

CHAPTER 5 Gauss-Jordan, computing inverses

Gauss-Jordan Algorithm $[A|I]\xrightarrow{\text{rref}}[I|A^{-1}]$

THM $A_{n\times n}$ non-singular $\Rightarrow A$ invertible

Special cases:

$D=\begin{bmatrix} d_1&0&0\\0&d_2&0\\0&0&d_n\end{bmatrix}$, $d_i\neq 0$, then $D^{-1}=\begin{bmatrix} 1/d_1&0&0\\0&1/d_2&0\\0&0&1/d_n\end{bmatrix}$

$A_{2\times 2}=\begin{bmatrix}a&b\\c&d\end{bmatrix}$, A regular $\Leftrightarrow ad-bc=0$ and $A^{-1}=\frac{1}{ad-bc}\begin{bmatrix}d&-b\\-c&a\end{bmatrix}$

we like knowing A^{-1} since $A\vec{x}=\vec{b} \Rightarrow \vec{x}=A^{-1}\vec{b}$

THM For a matrix $A_{n\times n}$,

A regular ($=\underset{\Leftrightarrow A\vec{x}=\vec{0}\Rightarrow\vec{x}=\vec{0}}{\text{non-singular}}=$ invertible)

$\Leftrightarrow \forall\vec{b}, \exists!\vec{x}$ such that $A\vec{x}=\vec{b}$

$\Leftrightarrow \exists! A^{-1}$ such that $AA^{-1}=A^{-1}A=I$

$\Leftrightarrow$ row-rank(A)=rank$(A)=n$

$\Leftrightarrow$ column-rank(A)= rank$(A^t)=n$

$\left.\begin{matrix}\text{rank}(A)\\ =\text{rank}(A^t)\end{matrix}\right.$

CHAPTER 6 LDU Factorization

Key observation: Gaussian elimination can be expressed as matrix multiplication

Scaling a row $[A] \xrightarrow{R_2' = \lambda R_2} [A'] \iff \begin{pmatrix} 1&0&0\\0&\lambda&0\\0&0&1 \end{pmatrix} A = A'$

Adding a row $[A] \xrightarrow{R_2' = 2R_1 + R_2} [A'] \iff \begin{pmatrix} 1&0&0\\2&1&0\\0&0&1 \end{pmatrix} A = A'$ ← recall that you can read $\begin{pmatrix} 1&0&0\\2&1&0\\0&0&1 \end{pmatrix} A$ as $\begin{pmatrix} R_1\\2R_1+R_2\\R_3 \end{pmatrix}$

Lower triangular $\begin{pmatrix} * & & 0\\ & * & \\ * & & * \end{pmatrix}$, normalized lower $\Delta = \begin{pmatrix} 1 & & 0\\ & 1 & \\ * & & 1 \end{pmatrix}$

Upper triangular $\begin{pmatrix} * & & *\\ & * & \\ 0 & & * \end{pmatrix}$, normalized upper $\Delta = \begin{pmatrix} 1 & & *\\ & 1 & \\ 0 & & 1 \end{pmatrix}$

Notice (and prove!) that if L_1, L_2 are lower Δ, then L_1^{-1} and $L_1 \cdot L_2$ are also lower Δ.

So if we can do Gaussian elimination without swapping rows, $A \longrightarrow A'$,

then $\underset{\substack{\uparrow \\ \text{row ops as normalized lower } \Delta}}{(L_k \cdots L_2 L_1)} A = \underbrace{A' = U}_{\text{upper triangular}}$

L_i^{-1} exists $\forall i$ since L_i of full rank

$\therefore A = (L_1^{-1} L_2^{-1} \cdots L_k^{-1}) A'$

$\Rightarrow A = \underset{\substack{\uparrow \\ \text{normalized}}}{L} U$ ← might not be normalized

Notice that we can only normalize U if A' is of full rank $\iff A$ is invertible.

By convention, LDU factorization has $A = \underbrace{L D U}_{\text{normalized}}$, otherwise trivially $\exists D = I \; \forall A$

But so if we can normalize U, then we can pull out the pivots as scaling matrix D

The above assumes we don't swap rows. If we need to, do it at the start with permutation matrix P, i.e.

$$PA = LU \quad (\text{or } PA = LDU \text{ if } A \text{ invertible})$$

THM $\forall A, \exists! P, L, U$ such that $PA = LU$

THM A non-singular $\iff PA = LDU$ exists

(permutation; norm lower Δ; diagonal with pivots; norm upper Δ)

THM LU factorization unique

e.g. $A = \begin{bmatrix} 0&5&7\\4&9&2\\8&1&6 \end{bmatrix} \xrightarrow[P = \begin{bmatrix} 0&1&0\\1&0&0\\0&0&1 \end{bmatrix}]{R_1 \leftrightarrow R_2} \begin{bmatrix} 4&9&2\\0&5&7\\8&1&6 \end{bmatrix} \xrightarrow[L_1 = \begin{pmatrix} 1&0&0\\0&1&0\\-2&0&1 \end{pmatrix}]{R_3 = -2R_1 + R_3} \begin{bmatrix} 4&9&2\\0&5&7\\0&-17&2 \end{bmatrix} \xrightarrow[L_2 = \begin{pmatrix} 1&0&0\\0&1&0\\0&\frac{17}{5}&1 \end{pmatrix}]{R_3' = \frac{17}{5}R_2 + R_3} \begin{bmatrix} 4&9&2\\0&5&7\\0&0&\frac{129}{5} \end{bmatrix}$

$\therefore \begin{bmatrix} 0&1&0\\1&0&0\\0&0&1 \end{bmatrix} A = \begin{bmatrix} 1&0&0\\0&1&0\\-2&0&1 \end{bmatrix}^{-1} \begin{bmatrix} 1&0&0\\0&1&0\\0&\frac{17}{5}&0 \end{bmatrix}^{-1} \begin{bmatrix} 4&9&2\\0&5&7\\0&0&\frac{129}{5} \end{bmatrix} = \begin{bmatrix} 1&0&0\\0&1&0\\2&-\frac{17}{5}&0 \end{bmatrix} \begin{bmatrix} 4&0&0\\0&5&0\\0&0&\frac{129}{5} \end{bmatrix} \begin{bmatrix} 1&\frac{9}{4}&\frac{1}{2}\\0&1&\frac{7}{5}\\0&0&1 \end{bmatrix}$

CHAPTER 7 Vector spaces, linear maps

Note: $\mathbb{R}^n$ = n-dimensional Euclidean space

$\vec{a} = (a_1, a_2, \dots, a_n)$ $a_i \in \mathbb{R}$ = n-tuple of real numbers

$\xrightarrow{\text{is associated}} \cong \begin{bmatrix} a_1 \\ \vdots \\ a_n \end{bmatrix}_{n\times 1}$ column vector

$\underline{a} = [a_1 a_2 \cdots a_n]_{1\times n}$ = row vector $= \vec{a}^t$

Vector space A set V is a vector space if its elements ("vectors") can be added and multiplied with scalars and still remain in V, i.e.

(actually there are 10 axioms? 13? that must be satisfied)

$\forall \vec{v}, \vec{w} \in V$, $\vec{v}+\vec{w} \in V$
$\lambda\vec{v} \in V \ \forall \lambda \in \mathbb{R}$

Subspace U is a subspace of the vector space V iff U is closed under addition and scalar multiplication, i.e.

① $\forall \vec{a}, \vec{b} \in U$, $\vec{a}+\vec{b} \in U$
② $\forall \vec{a} \in U, \lambda \in \mathbb{R}$, $\lambda\vec{a} \in U$

} not necessarily $\in V$

Note that U is a subspace $\Leftrightarrow$ U is a vector space

Corollary of conditions for being a vector space/subspace: closed under linear combinations

Some Subspaces

$\mathcal{C}(A)$ = column space = span$(\vec{a}_1, \dots \vec{a}_n)$ is a subspace of $\mathbb{R}^m$

$\mathcal{N}(A)$ = null-space = $\{\vec{x} \in \mathbb{R}^n \mid A\vec{x} = \vec{0}\}$ is a subspace of $\mathbb{R}^n$

CHAPTER 8 Linear Independence, basis, dimension

Independent $\vec{a}_1, \dots \vec{a}_k \in V$ are independent iff $\sum_i \lambda_i \vec{a}_i = \vec{0} \Rightarrow \vec{\lambda} = \vec{0}$ we call this linear combination trivial

Notice that this is like needing $\underbrace{\left[\begin{array}{c|c} \ddots & * \\ 0 & 0 \end{array}\right]}_{k}\Big\} n$, i.e. we need $k \le n$ for independence and so more than n vectors in $\mathbb{R}^n$ are dependent

Furthermore in a reduced matrix,
pivot rows are independent
pivot cols are independent
} prove!

And the ordering of independent vectors doesn't matter

Basis Let V be a vector space. $\{a_1 \dots a_k\}$ are a basis for V iff

1. $V = \text{span}(\vec{a}_1, \dots \vec{a}_k)$ ← enough to make everything
2. $\vec{a}_1, \dots \vec{a}_k$ are independent ← nothing extra, no redundant

} note that basis is not unique!

Coordinate vector $[\vec{v}]_A = \vec{\lambda}$ is the coordinate vector of V in basis A such that $\sum_i \lambda_i \vec{a}_i = \vec{v}$.

$[\vec{v}]_A$ is unique!

Dimension of V, dim(V) = number of vectors in any basis for V. If dim(V) finite, then V is finite dimensional.

THM If $A = \{\vec{a}_1 \dots \vec{a}_k\}$ and $B = \{\vec{b}_1 \dots \vec{b}_m\}$ are both bases for V, then $k = m \Leftrightarrow \dim(A) = \dim(B)$

CHAPTER 9 Row-space, column-space of reduced matrix A'

$\mathcal{R}(A) \leq \mathbb{R}^n$ with pivot rows of A' forming a basis for $\mathcal{R}(A') \Rightarrow \dim \mathcal{R}(A') = \text{rank}(A') = p$, $\{\underline{a}_1' \cdots \underline{a}_p'\}$ basis $\text{span}(\underline{a}_1' \cdots \underline{a}_m')$

$\mathcal{C}(A') \leq \mathbb{R}^m$ with pivot cols of A' forming a basis for $\mathcal{C}(A') \Rightarrow \dim \mathcal{C}(A') = \text{rank}(A') = p, \{\vec{a}_1' \cdots \vec{a}_p'\}$ basis $\text{span}(\vec{a}_1' \cdots \vec{a}_m')$

THM $\text{rank}(A) = \dim(\text{span}(\underline{a}_1 \cdots \underline{a}_m)) = \dim \mathcal{R}(A) = \dim \mathcal{C}(A')$

THM $A \to A'$, pivot rows $\underline{a}_1' \cdots \underline{a}_p'$ are basis for $\text{span}(a_1 \cdots a_m)$

whereas corresponding column vectors $\{\vec{a}_1 \cdots \vec{a}_p\}$ are basis for $\text{span}(\vec{a}_1 \cdots \vec{a}_m)$

Fundamental Subspaces

$\mathcal{R}(A) \leq \mathbb{R}^n$ $\quad \mathcal{C}(A) \leq \mathbb{R}^m$ $\quad \mathcal{N}(A) \leq \mathbb{R}^n$

$\mathcal{R}(A) = \mathcal{C}(A^t)$, $\mathcal{C}(A) = \mathcal{R}(A^t)$, $\mathcal{N}(A) = \{\vec{x} \mid A\vec{x} = \vec{0}\}$

THM

$\dim \mathcal{C}(A) = \text{rank}(A)$

$\dim \mathcal{N}(A) = n - \text{rank}(A)$

$\dim \mathcal{C}(A^t) = \dim \mathcal{R}(A) = \text{rank}(A)$

$\text{rank}(A) = \text{rank}(A^t)$

$\dim \mathcal{N}(A^t) = m - \text{rank}(A)$

$\mathcal{R}(A) = \mathcal{R}(A') \Leftarrow\Rightarrow \mathcal{C}(A) = \mathcal{C}(A')$

$\dim \mathcal{C}(A) = \dim \mathcal{C}(A') = \#\text{ pivots} = \text{rank}(A)$

NOTE $\mathcal{C}(A) \neq \mathcal{C}(A')$! Row operations preserve row space but not column space, e.g. $\begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}$

CHAPTER 10 Linear Transformations

$f: S \to T$ such that $\forall s \in S, \exists! f(s) \in T$

Image If $A \subseteq S$, $f(A) \subseteq T$ is the image of A

Preimage If $B \subseteq T$, $f^{-1}(B) = \{s \in S \mid \exists t \in B, f(s) = t\} \subseteq S$

$f^{-1}(f(A)) \supseteq A$ (try $f(x) = x^2$)

Injective f injective iff $\forall s, s' \in S$, $s \neq s' \Rightarrow f(s) \neq f(s')$

Surjective f surjective iff $f(S) = \text{Im}(f) = T$

Bijective If f is both injective and surjective, then f is bijective $\Leftarrow\Rightarrow \exists f^{-1}$

$f: V \to V'$ is a linear map iff $f(\sum_i \lambda_i \vec{v}_i) = \sum_i \lambda_i f(\vec{v}_i)$, i.e.

1. $f(\lambda \vec{v}) = \lambda f(\vec{v})$
2. $f(\vec{v}_1 + \vec{v}_2) = f(\vec{v}_1) + f(\vec{v}_2)$

Consequence: $f(\vec{0}) = \vec{0}$ (use 2.)

THM Let $\vec{v}_1 \cdots \vec{v}_k$ be a basis in V, and let $\vec{v}_i' = f(\vec{v}_i) \in V'$

Then $\vec{v}_i'$ are uniquely determined and $\text{Im}(f) = \text{span}(\vec{v}_1' \cdots \vec{v}_k')$, i.e. $\text{Im}(f) \leq V'$

Cont next page

cont. CHAPTER 10 Matrices as linear maps

$f : \mathbb{R}^n \to \mathbb{R}^m$ can be represented as $A = \underbrace{\left[\quad\right]}_{n}\}m$

$\vec{x} \mapsto A\vec{x}$

THM $A : \mathbb{R}^n \to \mathbb{R}^m$ is

injective $\iff \operatorname{rank}(A) = n$ (all input used uniquely)

surjective $\iff \operatorname{rank}(A) = m$ (all output reachable)

bijective $\iff A_{n \times n}$ and $\operatorname{rank}(A) = n$

THM $\dim \operatorname{Im}(A) = \dim \mathcal{C}(A) = n - \dim \mathcal{N}(A)$

COMPOSITION OF MAPS

$$\begin{tikzcd} V \arrow[r, "f"] \arrow[rr, bend right, "g \circ f"'] & V' \arrow[r, "g"] & V'' \end{tikzcd}$$

CHAPTER 11 Coordinate maps, basic 2D linear maps

Coordinate map $K_B = \begin{cases} V \xrightarrow{K_B} \mathbb{R}^n \\ \vec{v} \longmapsto [\vec{v}]_B = K_B(\vec{v}) \end{cases}$

THM K_B is linear and bijective

$\Rightarrow K_B^{-1}$ exists

$K_B^{-1}(\vec{x}) = \sum_i x_i \vec{b}_i$

$$\begin{tikzcd} V \arrow[r, "f"] \arrow[d, shift right, "K_B"'] & V' \arrow[d, "K'_{B'}"] \\ \mathbb{R}^n \arrow[u, shift right, "K_B^{-1}"'] \arrow[r, dashed, "\varphi"'] & \mathbb{R}^m \end{tikzcd}$$

The induced map φ is given by

$\varphi = \underbrace{K'_{B'} \circ f \circ K_B^{-1}}_{\substack{\text{all linear,} \\ \therefore\ \varphi \text{ also linear}}}$, $\varphi(\vec{x}) = F(\vec{x})$ where $F = \left[\cdots \left[\varphi(\vec{e}_k)\right] \cdots \right]$

Since $K_B^{-1}(\vec{e}_k) = \sum_j \underbrace{(\vec{e}_k)_j}_{= \begin{cases} 0 & j \neq k, \\ 1 & j = k \end{cases}} \vec{b}_j = \vec{b}_k$, $\varphi = K'_{B'} \circ f(\vec{b}_k) = \left[f(\vec{b}_k) \right]_{B'}$

$F = \left[\cdots \left[f(\vec{b}_k) \right]_{B'} \cdots \right]$

BASIC 2D LINEAR MAPS

Rotations (Q), Projections (P), Reflections (H)

$Q_\theta = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}$ } generally think about where $\begin{bmatrix} 1 \\ 0 \end{bmatrix}$ and $\begin{bmatrix} 0 \\ 1 \end{bmatrix}$ go.

recall $\begin{bmatrix} 1 \\ 0 \end{bmatrix} \mapsto$, $\begin{bmatrix} 0 \\ 1 \end{bmatrix} \mapsto$ } for projection and reflection, rotate problem, solve, rotate back.

CHAPTER 12 Change of basis, coordinates

$$\begin{tikzcd} & V \arrow[dl, "K_\theta"'] \arrow[dr, "K_N"] & \\ \mathbb{R}^k \arrow[rr, dashed, "\varphi = \varphi_C"'] & & \mathbb{R}^k \end{tikzcd}$$

$\varphi = \underbrace{K_N \circ K_\theta^{-1}}_{\text{bijective, linear}} \Rightarrow \exists!\, C$ such that $\varphi(\vec{x}) = C\vec{x}$, i.e $[\vec{v}]_N = C[\vec{v}]_\theta$, $C = C_{N \leftarrow \theta}$

$C_{C \leftarrow A} = C_{C \leftarrow B} C_{B \leftarrow A} = \left[\cdots [\vec{a}_k]_C \cdots \right]$, $C_{C \leftarrow A}^{-1} = C_{A \leftarrow C}$ (one is usually easier to find)

$F_{N' \leftarrow N} = C_{N' \leftarrow B'} F_{B' \leftarrow B} C_{B \leftarrow N}$ (to apply a function)

NO CHAPTER 13

CHAPTER 14 Inner product, orthogonality

Length of vector $\vec{x} \in \mathbb{R}^n = \|\vec{x}\| = \left(\sum_k^n x_k^2\right)^{1/2}$

Orthogonal First consider $\mathbb{R}^2$. $\vec{a}, \vec{b}$ orthogonal

$\iff \vec{a} \perp \vec{b}$, Pythagoras

$\iff \|\vec{a}\|^2 + \|\vec{b}\|^2 = \|\vec{c}\|^2$

$\iff \sum_k a_k^2 + \sum_k b_k^2 = \sum_k (a_k + b_k)^2 = \sum_k (a_k^2 + b_k^2 + 2a_k b_k)$

$\iff \sum_k a_k b_k = 0$

Inner Product $\langle \vec{a}, \vec{b} \rangle = \langle \vec{a} | \vec{b} \rangle = \sum_k a_k b_k$

$\vec{0}$ is orthogonal to everything by this definition!

So by definition, $\vec{a} \perp \vec{b} \iff \langle \vec{a}, \vec{b} \rangle = 0$

$\langle \vec{a}, \vec{b} \rangle = \sum_k a_k b_k = \langle \vec{b}, \vec{a} \rangle$ (symmetric)

$\langle \vec{a}, \vec{b} \rangle = \vec{a}^t \cdot \vec{b} = \vec{b}^t \cdot \vec{a}$

$\left.\begin{array}{l} \langle \vec{a}+\vec{c}, \vec{b} \rangle = \langle \vec{a}, \vec{b} \rangle + \langle \vec{c}, \vec{b} \rangle \\ \langle \lambda \vec{a}, \vec{b} \rangle = \lambda \langle \vec{a}, \vec{b} \rangle \end{array}\right\}$ $\langle \lambda\vec{a} + \vec{c}, \vec{b} \rangle = \lambda \langle \vec{a}, \vec{b} \rangle + \langle \vec{c}, \vec{b} \rangle$, i.e. $\langle \cdot, \cdot \rangle$ is linear in both arguments (bilinear)

(Mutually) Orthogonal $\vec{v}_1, \ldots \vec{v}_k \in \mathbb{R}^n$ are mutually orthogonal iff they are pairwise orthogonal ($\forall i \neq j$, $\vec{v}_i \perp \vec{v}_j$)

THM If $\vec{v}_1, \ldots \vec{v}_k \in \mathbb{R}^n$, $\vec{v}_i \neq \vec{0}$, are orthogonal then they are independent

Orthogonal Subspaces $V, W \leq \mathbb{R}^n$ are subspaces, they are orthogonal iff $\forall \vec{v} \in V, \forall \vec{w} \in W$, $\vec{v} \perp \vec{w}$ ($\iff \langle \vec{w}, \vec{v} \rangle = 0$)

Orthogonal Complement $V^\perp = \{x \in \mathbb{R}^n \mid \vec{x} \perp \vec{v} \text{ for all } \vec{v} \in V\}$

$\iff \langle \vec{x}, \vec{v} \rangle = 0 \iff A\vec{x} = \vec{0} \iff \vec{x} \in N(A)$

$V \cap V^\perp = \{\vec{0}\}$

THM Let $\{\vec{v}_1, \ldots \vec{v}_k\}$ be a basis for V and $\{\vec{u}_1, \ldots \vec{u}_k\}$ be a basis for $V^\perp$.

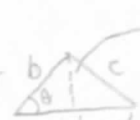
Then $\{\vec{v}_1, \ldots \vec{v}_k, \vec{u}_1, \ldots \vec{u}_k\}$ is a basis for $\mathbb{R}^n$, and

1. $\dim V^\perp = n - \dim V$
2. $\vec{a} \in \mathbb{R}^n \Rightarrow \exists! \vec{v} \in V, \vec{u} \in V^\perp$ such that $\vec{a} = \vec{v} + \vec{u}$ (i.e. $\mathbb{R}^n = \mathrm{span}(V, V^\perp)$)
3. $(V^\perp)^\perp = V$

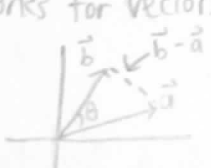
THM $\mathbb{R}^n \geq \mathcal{R}(A) \perp N(A) \leq \mathbb{R}^n$

$\mathcal{R}(A)^\perp = N(A)$

CHAPTER 15 Projections

Cosine Rule  $c^2 = a^2 + b^2 - 2ab\cos\theta$ (proof: consider )

This works for vectors, so for $\vec{a}, \vec{b} \in \mathbb{R}^n$



$\|\vec{b} - \vec{a}\|^2 = \|\vec{a}\|^2 + \|\vec{b}\|^2 - 2\|\vec{a}\|\|\vec{b}\|\cos\theta$

$\iff \sum_i (b_i - a_i)^2 = \sum_i a_i^2 + \sum_i b_i^2 - 2\|\vec{a}\|\|\vec{b}\|\cos\theta$

$\iff \sum a_i b_i = \underline{\langle \vec{a} | \vec{b} \rangle = \|\vec{a}\|\|\vec{b}\|\cos\theta}$!

(corollary: Cauchy-Schwartz Inequality $|\langle \vec{a}, \vec{b} \rangle| \leq \|\vec{a}\| \cdot \|\vec{b}\|$ since $|\cos\theta| \leq 1$)

Projection onto a line

Generally, $V = \operatorname{span}(\vec{a})$ and $\vec{p}$ = orthogonal projection of $\vec{x}$ onto V,

then
$$\vec{p} = \vec{a} \cdot \frac{1}{\|\vec{a}\|} \|\vec{x}\| \cos\theta$$
$$= \vec{a} \frac{\|\vec{a}\| \|\vec{x}\|}{\|\vec{a}\|^2} \cos\theta$$
$$= \vec{a} \frac{\langle \vec{a} | \vec{x} \rangle}{\langle \vec{a} | \vec{a} \rangle}$$
$$= \vec{a} \frac{\vec{a}^t \vec{x}}{\vec{a}^t \vec{a}}$$
$$= P\vec{x}, \text{ where } P = \frac{\vec{a}\,\vec{a}^t}{\vec{a}^t \vec{a}}$$

And generally P is symmetric,

and $P^2 = P$

Projection onto a subspace

P is the orthogonal projection onto V such that

$\forall \vec{b} \in \mathbb{R}^n, \vec{p} \in V$

$\forall \vec{b} \in \mathbb{R}^n, (\vec{b} - \vec{p}) \perp V$

Since $\vec{b} - \vec{p} \perp V$, and $\vec{p} \in V \Rightarrow \vec{p} = \sum \hat{x}_i \vec{a}_i = A\hat{x}$

Then $\vec{b} - A\hat{x} \perp \vec{a}_i$

$A^t(\vec{b} - A\hat{x}) = \vec{0}$

$A^t A \hat{x} = A^t \vec{b}$

Consider $N(A^tA)$, i.e. $A^tA\vec{x} = \vec{0} \Rightarrow \langle \vec{x} | A^tA\vec{x} \rangle = 0$ since $A^tA\vec{x} = \vec{0}$

$\Rightarrow \vec{x}^t A^t A\vec{x} = \vec{0} \Rightarrow (A\vec{x})^t (A\vec{x}) = \vec{0} \Rightarrow \langle A\vec{x}, A\vec{x} \rangle = 0 \Rightarrow A\vec{x} = \vec{0}$

but since $N(A) = \{\vec{0}\}$, $\vec{x} = \vec{0}$

assumed $A = \underbrace{[\vec{a}_1 \cdots \vec{a}_k]}_{\text{basis}} \}n,\ n > k$

$\therefore N(A^tA) = \{\vec{0}\} \Rightarrow \exists (A^tA)^{-1}$

$\therefore \hat{x} = (A^tA)^{-1} A^t \vec{b}$

$\therefore \vec{p} = A\hat{x} = A(A^tA)^{-1} A^t \vec{b}$

$\Rightarrow P_V^\perp = A(A^tA)^{-1}A^t \wedge P\vec{b} = \vec{p}$

Fundamentally, what we did: we can now project orthogonally onto lines and subspaces, and this projection is the closest.

(cont next Application: Approximate Solutions to Overdetermined Systems)

cont/ CHAPTER 15 Application: Approximate Solutions to Overdetermined Systems

Suppose $A\vec{x}=\vec{b}$ no solution, i.e. $A\vec{b} \notin \mathcal{C}(A)$ (recall $\begin{pmatrix}\vec{a}_1 & \vec{a}_2 & \vec{a}_3\end{pmatrix}\begin{pmatrix}x\\y\\z\end{pmatrix}=x\begin{bmatrix}\vec{a}_1\end{bmatrix}+y\begin{bmatrix}\vec{a}_2\end{bmatrix}+z\begin{bmatrix}\vec{a}_3\end{bmatrix}$)

But we can find the $\perp$ projection of $\vec{b}$ on $\mathcal{C}(A)$,

and this projection is the closest!

Since the projection $\vec{p} \in V=\mathcal{C}(A)$ and $\hat{x}=\vec{b}-\vec{p}$ is the straight line closest distance

And we know $(A^tA)\hat{x}=A^t\vec{b}$

Least Square Fitting

Similarly suppose we have a linear function $f(t)=C+Dt$

Then given many data points (t,b),

$$\underbrace{\begin{bmatrix}1 & t_1\\1 & t_2\\ & \vdots\\1 & t_n\end{bmatrix}}_{A}\begin{bmatrix}C\\D\end{bmatrix}=\underbrace{\begin{bmatrix}b_1\\b_2\\ \vdots\\b_n\end{bmatrix}}_{b}$$

we can approximately reconstruct C and D

$$A^tA\vec{x}=A^t\vec{b} \iff \underbrace{\begin{bmatrix}1 & 1 & 1 & 1 & \cdots & 1\\t_1 & t_2 & t_3 & t_4 & \cdots & t_n\end{bmatrix}\begin{bmatrix}1 & t_1\\1 & t_2\\ \vdots & \vdots\\1 & t_n\end{bmatrix}}_{\begin{bmatrix}n & \Sigma t_i\\ \Sigma t_i & \Sigma t_i^2\end{bmatrix}}\vec{x}=\underbrace{\begin{bmatrix}1 & 1 & \cdots & 1\\t_1 & t_2 & \cdots & t_n\end{bmatrix}\begin{bmatrix}b_1\\b_2\\ \vdots\\b_n\end{bmatrix}}$$

$$\begin{bmatrix}n & \Sigma t_i\\ \Sigma t_i & \Sigma t_i^2\end{bmatrix}\begin{bmatrix}\hat{C}\\ \hat{D}\end{bmatrix}=\begin{bmatrix}\Sigma b_i\\ \Sigma t_i b_i\end{bmatrix}$$

Characterization of Projections

THM ① P is a projection iff $P^2=P \wedge P^t=P$

② Any P satisfying ① is a projection onto $\mathcal{C}(P)$

CHAPTER 16 Orthogonal Matrices, Gram-Schmidt

(ON) Orthonormal vectors $\vec{q}_1 \cdots \vec{q}_k \in \mathbb{R}^n$ satisfy

① Orthogonal
② Unit length/normalized

$\}\ \forall i,j: \langle \vec{q}_i, \vec{q}_j \rangle = \begin{cases}0 & i\neq j\\1 & i=j\end{cases}$

Orthogonal matrices are square and have ON columns

Consequence: Q orthogonal $\iff Q^{-1}=Q^t \iff$ ON columns $\iff$ ON rows

Properties:

Angle-preserving $\measuredangle(\vec{a},\vec{b})=\measuredangle(Q\vec{a},Q\vec{b})$
Length-preserving $\|\vec{a}\|=\|Q\vec{a}\|$

$\}$ in general, $\langle Q\vec{a} | Q\vec{b}\rangle = \langle \vec{a} | \vec{b}\rangle$

(ONB) Orthonormal basis formed by $\vec{q}_1 \cdots \vec{q}_k$ in the subspace $V \in \mathbb{R}^n$ iff

① $\vec{q}_1 \cdots \vec{q}_k$ are ON
② span$(q_1, \cdots q_k)=V$

We like ONB because $[\vec{b}]_Q=\begin{bmatrix}\langle \vec{q}_1 | \vec{b}\rangle\\ \vdots\\ \langle \vec{q}_n | \vec{b}\rangle\end{bmatrix}$, i.e. $\vec{b}=\Sigma\langle \vec{q}_i, \vec{b}\rangle \vec{q}_i$

(alternatively, $\vec{x}=Q^{-1}\vec{b}=Q^t\vec{b}) \Leftarrow$

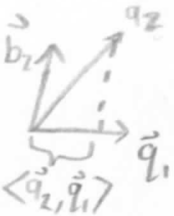
cont. CHAPTER 16 Gram-Schmidt

Suppose we have a basis $\vec{a}_1 \cdots \vec{a}_k$ in $V \subseteq \mathbb{R}^n$.

To construct ONB,

$\vec{b}_1 = \vec{a}_1$, $\vec{q}_1 = \frac{\vec{b}_1}{\|\vec{b}_1\|}$

$\vec{b}_2 = \vec{a}_2 - \langle \vec{a}_2 | \vec{q}_1 \rangle \vec{q}_1$, $\vec{q}_2 = \frac{\vec{b}_2}{\|\vec{b}_2\|}$



we find the $\perp$ projection of $\vec{a}_2$ onto $\vec{q}_1$, then normalize it

$$\vec{b}_3 = \vec{a}_3 - \underbrace{\langle \vec{a}_3 | \vec{q}_1 \rangle \vec{q}_1}_{\substack{\text{component of} \\ \vec{a}_3 \text{ in } \vec{q}_1 \\ \text{direction}}} - \langle \vec{a}_3 | \vec{q}_2 \rangle \vec{q}_2, \quad \vec{q}_3 = \frac{\vec{b}_3}{\|\vec{b}_3\|}$$

So at every step we take the old basis vector, make it $\perp$ by removing components in already chosen directions, then normalize it

A = QR factorization

Reverse the GS process expressing $\vec{a}_j$ in terms of $\vec{q}_k$,

we had $\vec{b}_j = \vec{a}_j - \sum_{\ell<j} \langle \vec{a}_j, \vec{q}_\ell \rangle \vec{q}_\ell$ and $\vec{q}_j = \frac{\vec{b}_j}{\|\vec{b}_j\|}$

$$\therefore \vec{a}_j = \sum_{\ell<j} \langle \vec{a}_j, \vec{q}_\ell \rangle \vec{q}_\ell + \vec{b}_j$$

$$\vec{b}_j = \vec{q}_j \underbrace{\|\vec{b}_j\|}_{\langle \vec{q}_j, \vec{b}_j \rangle = \langle \vec{q}_j, \vec{a}_j \rangle}$$

$$\Rightarrow \vec{a}_j = \sum_{\ell \leq j} \langle \vec{a}_j | \vec{q}_\ell \rangle \vec{q}_\ell$$

THM If $A_{n \times k}$, $k \leq n$, has independent columns then

$$A = QR$$

$$n\{\underbrace{[A]}_{k} = \underbrace{[Q]}_{k} \underbrace{[{}_0\nabla]}_{k}\}k = \begin{bmatrix} \vec{q}_1 & \vec{q}_2 \cdots \vec{q}_k \end{bmatrix} \begin{bmatrix} \langle \vec{q}_1 | \vec{a}_1 \rangle & \langle \vec{q}_1 | \vec{a}_2 \rangle & \cdots & \langle \vec{q}_1 | \vec{a}_k \rangle \\ 0 & \langle \vec{q}_2 | \vec{a}_2 \rangle & \cdots & \langle \vec{q}_2 | \vec{a}_k \rangle \\ & & \ddots & \vdots \\ & 0 & & \langle \vec{q}_k | \vec{a}_k \rangle \end{bmatrix}$$

ON cols (Q); upper Δ invertible (R)

THM If $A_{n \times n}$, $Q_{k \times k}$ is orthogonal

CHAPTER 17 DETERMINANTS

Let $n \geq 1$, then $\exists!$ function determinant (det) satisfying:

$\det : A_{n\times n} \mapsto \underset{\in \mathbb{R}}{\det(A)}$ or $|A|$

① $\det(I) = 1$

② sign(det) changes when two rows are exchanged (det is alternating)

③ det is a linear function of the rows, $\det\begin{bmatrix} [a+\lambda b] \\ [c_2] \\ [c_n] \end{bmatrix} = \det\begin{pmatrix} [a] \\ [c_2] \\ [c_n] \end{pmatrix} + \lambda \det\begin{pmatrix} [b] \\ [c_2] \\ [c_n] \end{pmatrix}$

$\Rightarrow$ det (diagonal matrix) = product of diagonals, $\det\begin{pmatrix} a & & 0 \\ & b & \\ 0 & & c \end{pmatrix} = a \det\begin{pmatrix} 1 & & \\ & b & \\ & & c \end{pmatrix} \cdots = abc \det(I) = abc$

$\Rightarrow$ det (matrix with 0 row) = 0

$\Rightarrow$ det (matrix with dependent rows) = 0 (since switching rows to A', we need $\det(A') = -\det(A)$)

THM A regular $\Leftrightarrow \det(A) \neq 0$

THM $\det(AB) = \det(A)\det(B)$ ⟵ prove this! hint: $\det(ULPA) = \det(D)$

$\Rightarrow \det(A^{-1}) = \frac{1}{\det(A)}$

THM $\det(A) = \det(A^t)$

(Leibniz formula (in class, notes 17A)

$$\det(A) = \sum_{\sigma \in S_n} \operatorname{sgn}(\sigma) \prod_{i=1}^{n} a_{i,\sigma(i)}$$

i.e. det is sum of <u>all</u> permutations)

CHAPTER 18 COMPUTING DETERMINANTS

① "Easiest" to compute $A = PLU$

Then $\det(A) = \underset{(-1)^{\#\text{swaps}}}{\det(P)} \; \underset{\substack{1 \text{ since} \\ \text{norm} \\ \text{lower } \triangle}}{\det(L)} \; \underset{\substack{\text{product of} \\ \text{diagonals}}}{\det(U)}$

② Otherwise by cofactors,

$$\det(A) = \sum_j a_{kj} \underbrace{(-1)^{k+j} |A|^{kj}}_{\text{cofactor } C_{kj}} \quad \forall k \leq n$$

, where $|A|^{kj}$ = subdeterminant = minor = det(matrix A with row k and column j deleted)

This method is nice for sparse matrices since 0's make life easier.

APPLICATIONS

① Checking regularity of A

② Formula for A^{-1} if A is regular,

$(A^{-1})_{ij} = \frac{C_{ji}}{\det(A)} = \frac{(-1)^{i+j}|A|^{ji}}{\det(A)}$, let $C := [c_{ij}]$ be the cofactor matrix, then

$$A^{-1} = \frac{1}{\det(A)} C^t$$

③ Cramer's formula

For A regular, $A\vec{x} = \vec{b}$,

$$x_j = \frac{1}{\det(A)} \det(\vec{a}_1 \, \vec{a}_2 \cdots \vec{a}_{j-1} \, \underset{j^{th} \text{ column}}{\vec{b}} \, \vec{a}_{j+1} \cdots \vec{a}_n)$$

④ Volume of ~~paralleletetete~~ parallelepiped B

If B spanned by $\perp$ $\vec{a}_1 \cdots \vec{a}_n$ in $\mathbb{R}^n$, $\det(A) = \operatorname{vol}(B)$

MAGIC: This works for arbitrary $\vec{a}_1 \cdots \vec{a}_n$

CHAPTER 17A Permutations, Constructing det

Let A be a finite set $A = \{a_1, \dots a_n\}$

Permutation σ is a bijection $A \to A$, (note that we can let $A = \mathbb{N}$),

$$\sigma : A \to A$$
$$a_i \mapsto \sigma(a_i)$$

Transposition τ if only two elements swapped, all else map to themselves

Symmetric Group S_n is the collection of all permutations of $\mathbb{N}_n$

Lemma $(S_n, *)$ is a group, $(\sigma * \sigma')(k) = \sigma \circ \sigma'(k) = \sigma(\sigma'(k))$,

neutral element $= id : \mathbb{N}_n \to \mathbb{N}_n,\ k \mapsto k$,

$|S_n| = n!$

Note that S_n is not commutative

We can also have cycle decompositions, e.g. $\sigma = \begin{pmatrix} 1&2&3&4&5&6 \\ 3&2&5&6&1&4 \end{pmatrix} = (1,3,5)(2)(4,6) = (1,3,5)(4,6)$ ← cycle notation, omits fixed points

THM $\forall \sigma \in S_n$, σ is the product of disjoint cyclic permutations (and when disjoint, they commute).
Furthermore the cyclic decomposition is unique (up to cycle ordering)

Lemma We have the collapse rule, if $a_1, \dots a_k, a_{k+1} \cdots a_n, c \in \mathbb{N}_n$ distinct then

$$(a_1 \cdots a_k, c)(c, a_{k+1}, \dots a_n) = (a_1 \cdots a_k\, c\, a_{k+1} \cdots a_n)$$

Lemma Every permutation is a product of transpositions (NOT unique)

Sign $\gamma = (a_1, \dots a_k) \in S_n$, $sign(\gamma) = (-1)^{k-1}$

THM $sgn(\sigma\sigma') = sgn(\sigma)\, sgn(\sigma')$

Corollary $\sigma = \tau_1 \tau_2 \cdots \tau_k = \tau_1' \tau_2' \cdots \tau_\ell' \Rightarrow k, \ell$ have same parity

See notes for constructing det, essentially properties of permutations + Leibniz det definition satisfies det properties

CHAPTER 18A

Group $(G, \cdot)$

① closed under addition
$\forall a, b \in G,\ a \cdot b \in G$

② associativity
$\forall a, b, c \in G,\ a \cdot (b \cdot c) = (a \cdot b) \cdot c$

③ identity
$\exists e, \forall a \in G,\ a \cdot e = e \cdot a = a$

④ inverse
$\forall a \in G, \exists a^{-1} \in G,\ a \cdot a^{-1} = a^{-1} \cdot a = e$

Abelian Group

⑤ commutativity
$\forall a, b \in G,\ a \cdot b = b \cdot a$

Ring $(R, +, \cdot)$

1-5 on $+$, with identity denoted 0,

⑥ associativity of $\cdot$
$a \cdot (b \cdot c) = (a \cdot b) \cdot c$

⑦ distributivity
↳ left $a \cdot (b + c) = a \cdot b + a \cdot c$
↳ right $(a + b) \cdot c = a \cdot c + b \cdot c$

⑧* exists identity on $\cdot$ denoted 1 (only some textbooks, most don't)

Field $(F, +, \cdot)$

⑧ Commutative on $\cdot$

⑨ $\exists$ identity on $\cdot$

⑩ $\exists$ inverse on $\cdot$

Let F be a field $\mathbb{R}$ or $\mathbb{C}$, then the set of polynomials in x with coefficients in F is
$F[x] := \{ \sum_{k=0}^{n} a_k x^k \mid n \in \mathbb{N}, a_k \in F \}$ which is a ring (no multiplicative inverse)

DEF 0 polynomial $= 0[x] = 0 + 0x + 0x^2 + \dots = 0$

Lemma Given $f, g \in F[x]$, $z \in F$,
① $(f+g)(z) = f(z) + g(z)$
② $(f \cdot g)(z) = f(z) \cdot g(z)$
③ $(f \cdot g) \equiv 0 \Rightarrow f \equiv 0$ or $g \equiv 0$

Degree $\deg(f) =$ largest index of k, $a_k \neq 0$
$\deg(0) = -\infty$

THM Division with remainder
$f, g \in F[x], g \neq 0 \Rightarrow \exists!\, Q, R \in F[x]$
$f = Q \cdot g + R,\ \deg(R) < \deg(g)$

Root $z \in F$ is root of $f \in F[x] \Leftrightarrow f(z) = 0$

Lemma z root $\Rightarrow (x - z)$ divides f w/o remainder

DEF F algebraically closed $\Leftrightarrow$ every non-constant polynomial has a root

FACT Every field has essentially unique smallest extension to be algebraically closed

Lemma $p(z) = 0 \Rightarrow p(\bar{z}) = 0$, i.e. roots come in complex conj pairs

Lemma $(x - z)(x - \bar{z}) = x^2 + cx + d$, $c, d \in \mathbb{R}$

THM Fundamental Thm of Algebra for real polynomials
all non-constant $p(x)$,
$p(x) = c(x - x_1) \dots (x - x_k) q_1(x) \cdots q_\ell(x)$
$c, x_i \in \mathbb{R}$
$q_i(x) = x^2 + b_i x + c_i = (x - z_i)(x - \bar{z}_i)$, $b_i, c_i \in \mathbb{R}$, $z_i \notin \mathbb{R}$
$\deg(p) = k + 2\ell$

CHAPTER 19 Eigenvalues, eigenvectors, diagonalizable

Let $f: V \to V$ linear,

DEF If $\vec{x} \in V$, $\vec{x} \neq \vec{0}$, $f(x) = \lambda \vec{x}$ for some $\lambda \in F$, then $\vec{x}$: eigenvector, λ: eigenvalue

① $(f(\vec{x}) = A\vec{x} \wedge A\vec{x} = \lambda\vec{x}) \Rightarrow \vec{x}, \lambda$ are eigenvector and eigenvalue of A

② $\vec{x}$ eigenvector $\Rightarrow c\vec{x}$ eigenvector

③ $\mathrm{Ker}(f) \neq \{\vec{0}\} \Rightarrow \lambda = 0$ eigenvalue, all $\vec{y} \in \mathrm{Ker}(f)$ eigenvector

DEF $E_\lambda = \{\vec{x} \in V \mid f(\vec{x}) = \lambda \vec{x}\} \leq V$ is the eigenspace to λ

① λ eigenvalue $\Rightarrow \dim(E_\lambda) \geq 1$

② $E_\lambda \setminus \{\vec{0}\}$ is the collection of all eigenvectors to λ

DEF ③ $\dim(E_\lambda) = \mathrm{geom}(\lambda) =$ geometric multiplicity of λ

DEF $p_A(\lambda) = \det(A - \lambda I_n) =$ characteristic polynomial of A and of f, where A is the matrix map representing f in basis B

Finding eigenvectors, eigenvalues

① Start $f: V \to V$. Pick any basis, obtain $A = \left[f(b_k)\right]_B$

② Solve $p_A(\lambda)$ for λ – eigenvals

③ For each λ, find $\mathrm{Ker}(A - \lambda I)$ – eigenvecs

Remarks

① $A \sim A'$ (i.e. $A' = CAC^{-1}$), then $p_A(\lambda) = p_{A'}(\lambda)$ ← good to prove. hint: $I = CC^{-1} = CIC^{-1}$

② D diagonal $\Leftrightarrow$ Evalues $(d_{ii})_{i=1,\dots,n}$, evectors $\vec{e}_1, \dots \vec{e}_n$

③ T triangular, same as above

④ not every linear map has evals/evecs, e.g. rotation

↳ to have eigenvalues and eigenvectors is to say the function is just stretching along some vectors

DEF ⑤ If $(\lambda - \lambda_i)^k$ is a factor of $p_A(\lambda)$, then the algebraic multiplicity of $\lambda_i = k$.

Note $\mathrm{geom}(\lambda)$ is between 1 and k.

⑥ A invertible, has eval λ evec $\vec{x} \Rightarrow A^{-1}$ has eval $\frac{1}{\lambda}$, evec $\vec{x}$

Lemma $\lambda_1, \dots \lambda_k$ distinct evals $\wedge$ $\vec{x}_1, \dots \vec{x}_k$ corresponding evecs $\Rightarrow \vec{x}_1 \dots \vec{x}_k$ indep

Diagonalizable

DEF A diagonalizable $\Leftrightarrow A \sim D \Leftrightarrow \exists T, A = TDT^{-1}$

THM A diagonalizable $\Leftrightarrow A$ has n independent eigenvectors

in which case $D = \Lambda =$ eval matrix

$T =$ evec matrix

on time evolution matrix

$A_{n\times n} \Rightarrow e^A = I_n + A + \frac{1}{2!}A^2 + \dots$ converges, $e^A_{n \times n}$

A diagonal $\Rightarrow e^A = \begin{bmatrix} e^{d_1} & & 0 \\ & \ddots & \\ 0 & & e^{d_n} \end{bmatrix} = Se^{\Lambda}S^{-1}$

① $e^{tA}e^{sA} = e^{(t+s)A}$

② $e^{0A} = I_n$

③ $\frac{d}{dt} e^{tA} = Ae^{tA}$

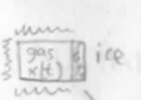
BUT $e^{A+B} \neq e^A e^B$ in general

CHAPTER 20 Eigenvalue Applications

① A diagonalizable $\Rightarrow A^k$ is too $\Rightarrow$ easy to compute $A^k = S\Lambda^k S^{-1}$

② Closed form for Fibonacci numbers by considering $u_{k+1} = Au_k = \begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} F_{k+1} \\ F_k \end{bmatrix}$

③ Linear Differential Equations



$\Delta x(t) = -\gamma(x(t) - T_0)$

If $T_0 = 0$, then

$x'(t) = -\gamma x(t)$ with $x(0) = x_0$

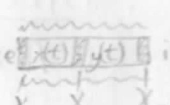
$\Rightarrow$ general solution $x(t) = ce^{-\gamma t}$

$\Rightarrow$ to find unique solution,

$x(0) = ce^{-\gamma 0} = x_0$

$\Rightarrow c = x_0$

$\Rightarrow x(t) = x_0 e^{-\gamma t}$



$x(t), y(t)$ – temperature of gas container 1, 2

$\gamma_1, \gamma_2, \gamma_3$ – thermal conductivities of walls

Model as diff eq $(x(t), y(t)) = \begin{bmatrix} x(t) \\ y(t) \end{bmatrix} = \vec{x}(t)$

Note $x'(t) = -\gamma_1(x(t) - 0) - \gamma_3(x(t) - y(t))$

$y'(t) = -\gamma_2(y(t) - 0) - \gamma_3(y(t) - x(t))$

$\Rightarrow \vec{x}'(t) = \begin{bmatrix} -\gamma_1 - \gamma_3 & \gamma_3 \\ \gamma_3 & -\gamma_2 - \gamma_3 \end{bmatrix} \vec{x}(t)$, set initial condition $\begin{bmatrix} x(0) \\ y(0) \end{bmatrix} = \begin{bmatrix} 30 \\ 10 \end{bmatrix} = \vec{x}_0$

Suppose $\exists T$ invertible, to decouple let $\vec{z}(t) = T\vec{x}(t)$, $A = \begin{bmatrix} -\gamma_1 - \gamma_3 & \gamma_3 \\ \gamma_3 & -\gamma_2 - \gamma_3 \end{bmatrix}$

Then $\vec{z}'(t) = \frac{d}{dt}(T\vec{x}(t)) = T\vec{x}'(t) = TAT^{-1}\vec{z}(t)$. If $A = S\Lambda S^{-1}$, then pick $T = S^{-1} \Rightarrow \vec{z}'(t) = \Lambda \vec{z}(t)$

$\vec{z}(t) = e^{t\Lambda}\vec{z}(0) \Rightarrow \vec{x}(t) = S\vec{z}(t) = Se^{t\Lambda}\vec{z}(0) = Se^{t\Lambda}S^{-1}\vec{x}(0)$. Note time evolution matrix $e^{tA} = Se^{t\Lambda}S^{-1}$

CHAPTER 21 Schur's Lemma

$\ell: V \to V$, $\dim(V) = n$, then

ℓ is diagonalizable $\iff \exists$ basis B, $L = \ell_B$ diagonal

THM ℓ is diagonalizable $\iff \exists$ n independent eigenvectors of ℓ (idea: use the eigenvectors as a basis, then $\ell_B = \Lambda$ since $[\ell(\vec{s}_k)]_B = \begin{bmatrix} 0 \\ \lambda_k \\ 0 \end{bmatrix}$, $\vec{s}_k$ evec to λ_k)

ℓ is triagonalizable $\iff \exists$ basis b, $L = \ell_B$ triagonal $\left(T = \begin{bmatrix} * & * & * \\ 0 & * & * \\ & & * \end{bmatrix}\right)$

Schur's Lemma Let F be algebraically closed, then every $\ell: V \to V$ is triagonalizable

Proof by induction: pick $\vec{b}_1$ eigenvalue to λ, then $L = \left[[\ell(\vec{b}_1)]_B [\ell(\vec{b}_2)]_B \cdots [\ell(\vec{b}_n)]_B \right] = \left[\begin{bmatrix} \lambda \\ 0 \\ \vdots \\ 0 \end{bmatrix} [*][*] \cdots [*] \right] = \begin{bmatrix} [\lambda] & [\quad * \quad] \\ [0] & [\quad * \quad] \end{bmatrix}$ triagonal by induction

Note that by nature of the proof we have freedom to rearrange λ

THM Every real symmetric matrix is diagonalizable with an ON set of eigenvectors (proof deferred)

CHAPTER 22 Hermitian, Unitary, Spectral Theorem

In $\mathbb{C}^n$, $\langle \vec{z}, \vec{u} \rangle = \sum_k \bar{z}_k u_k$ ← not symmetric. But note that if $\vec{z}, \vec{u}$ real then $\langle \vec{z}, \vec{u} \rangle_{\mathbb{C}^n} = \langle \vec{z}, \vec{u} \rangle_{\mathbb{R}^n}$

① $\langle \alpha\vec{z} + \vec{u} | \vec{v} \rangle = \bar{\alpha} \langle \vec{z} | \vec{v} \rangle + \langle \vec{u} | \vec{v} \rangle$

② $\langle \vec{z} | \alpha\vec{u} + \vec{v} \rangle = \alpha \langle \vec{z} | \vec{u} \rangle + \langle \vec{z} | \vec{v} \rangle$

③ $\langle \vec{z} | \vec{z} \rangle = \|\vec{z}\|^2 \in \mathbb{R}^+ \cup \{0\}$

④ $\|\vec{z}\| = 0 \iff \vec{z} = \vec{0}$

⑤ $\langle \vec{z} | \vec{u} \rangle = \overline{\langle \vec{u} | \vec{z} \rangle}$

If A is a complex matrix,

$\bar{A} = [\bar{a}_{ij}]$ conjugate

$A^* = \bar{A}^t = \overline{A^t}$ adjoint = conjugate + transpose

① $(AB)^* = B^* A^*$

② $(A^*)^* = A$

③ $\langle \vec{x} | A\vec{y} \rangle = \langle A^* \vec{x} | \vec{y} \rangle$

Self-Adjoint/Hermitian $A^* = A$

① $\Rightarrow a_{kk} \in \mathbb{R}$

② $\Rightarrow$ if A only has real entries, $\underset{\text{self-adjoint}}{A = A^*} \iff \underset{\text{symmetric}}{A = A^t}$

THM $A_{n \times n}$ Hermitian, then

① λ eigenvalue $\Rightarrow \lambda \in \mathbb{R}$ and $\exists n$ real eigenvalues (prove by considering $\langle \vec{s}, A\vec{s} \rangle$, or $\langle \vec{x}, A\vec{x} \rangle$)

② $\lambda_1 \neq \lambda_2 \Rightarrow \vec{s}_1 \perp \vec{s}_2$ (again consider $\langle A\vec{s}_1, \vec{s}_2 \rangle$)

③ $\exists n$ independent (complex) eigenvectors, i.e. A diagonalizable

NOTE: A real + symmetric $\Rightarrow$ diagonalizable over $\mathbb{R}$

cont. next column

THM ($\mathbb{R}$) $L_{n \times n}$ symmetric $\Rightarrow \exists Q$, $L = Q \Lambda Q^{-1}$

Proof by induction like Schur's Lemma, but since L symmetric we have

$$\begin{bmatrix} [\lambda] & [\ * \ ] \\ [0] & [\ * \ ] \end{bmatrix} \xrightarrow[(B^{-1}LB)^t]{\text{symmetric}} \begin{bmatrix} [\lambda] & [\ 0 \ ] \\ [0] & [\ * \ ] \end{bmatrix} \cdots \begin{bmatrix} [\lambda] & \\ & [\Lambda] \end{bmatrix} = \Lambda$$

and remember we want orthogonal Q so normalize $\vec{b}_1 \ldots \vec{b}_k$, e.g. $\vec{b}_1 = \frac{\vec{s}_1}{\|\vec{s}_1\|}$

Unitary if $A_{n \times n}$ has ON columns, complex

① $\Rightarrow$ ON columns $\Rightarrow$ ON rows

② $\Rightarrow$ A unitary $\iff A^{-1} = A^*$

$\mathbb{R}^n$		$\mathbb{C}^n$
$\langle \vec{x} \| \vec{y} \rangle = \vec{x}^t \vec{y}$	----	$\langle \vec{x} \| \vec{y} \rangle = \bar{\vec{x}}^t \vec{y}$
orthogonal	----	unitary
symmetric	----	Hermitian

SPECTRAL THEOREM

A self-adjoint $\Rightarrow \exists U$ unitary, $A = \underset{\text{eigenvector}}{U} \underset{\text{eigenvalue}}{\Lambda} U^{-1}$

Proof is analogous

CHAPTER 23 Quadratic Forms, Positive (Semi) Definite Matrices

Quadratic Form associated with the matrix $A=[a_{ij}]$ is

$$q_A(\vec{x}) = q_A(x_1 \cdots x_n) = \sum_{i,j} a_{ij} x_i x_j, \quad a_{ij} \in \mathbb{R}$$

① q is a polynomial in $x_1 \cdots x_n$ containing only second order terms

② $q_A(\vec{x}) = \langle \vec{x}, A\vec{x} \rangle = [x_1 \cdots x_n][A]\begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}$

THM $\exists!$ symmetric A, $q_A(\vec{x}) = \langle \vec{x}, A\vec{x} \rangle \ \forall \vec{x}$

quadratic forms have a one-to-one correspondence with symmetric matrices

DEF Symmetric matrix A is

① Pos (neg) definite iff all eigenvalues >0 (<0)

② Pos (neg) semidefinite iff zero is an eigenvalue and the others are all >0 (<0)

③ Indefinite iff $\exists$ strictly positive and negative eigenvalues

THM ① A pos (neg) definite $\Leftrightarrow q_A(\vec{x}) > 0$ $(<)$ $\forall \vec{x} \neq \vec{0}$

② A indefinite $\Leftrightarrow \exists \vec{x}, \vec{y},\ q_A(\vec{x}) < 0 < q_A(\vec{y})$

NOTE If A symmetric and all upper-left subdeterminants $\left[\begin{bmatrix} \cdot \end{bmatrix} \cdots \right] > 0$, then A pos definite

APPLICATION: Principal Axes

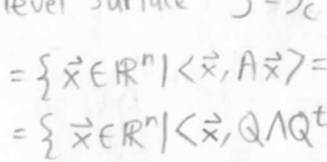
The level surface $S = S_c = \{\vec{x} \in \mathbb{R}^n \mid q(\vec{x}) = c\}$ can be better understood by considering

$$\begin{aligned} S &= \{\vec{x} \in \mathbb{R}^n \mid \langle \vec{x}, A\vec{x} \rangle = 1\} \\ &= \{\vec{x} \in \mathbb{R}^n \mid \langle \vec{x}, Q\Lambda Q^t \vec{x} \rangle = 1\} \\ &= \{\vec{x} \in \mathbb{R}^n \mid \langle Q^t \vec{x}, \Lambda Q^t \vec{x} \rangle = 1\}, \text{ let } Q^t\vec{x} = \vec{y} \Rightarrow \vec{x} = Q\vec{y}, \text{ then} \\ &= \{Q\vec{y} \mid \langle \vec{y}, \Lambda \vec{y} \rangle = 1\} \\ &= \underset{\text{rotation matrix}}{Q} \underbrace{\left\{\vec{y} \mid \left\langle \vec{y}, \begin{bmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{bmatrix} \vec{y} \right\rangle = 1 \right\}}_{=: S'\left(\frac{1}{\sqrt{\lambda_1}} \cdots \frac{1}{\sqrt{\lambda_n}}\right)} \overset{\lambda_1, \lambda_2}{=} Q\left\{\vec{y} \mid q(\vec{y}) = \lambda_1 x^2 + \lambda_2 y^2 = \left(\frac{x}{1/\sqrt{\lambda_1}}\right)^2 + \left(\frac{y}{1/\sqrt{\lambda_2}}\right)^2 = 1\right\} \text{ ellipse} \end{aligned}$$

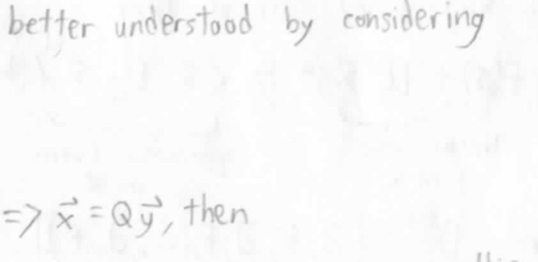
$=: S'\left(\frac{1}{\sqrt{\lambda_1}} \cdots \frac{1}{\sqrt{\lambda_n}}\right)$, a quadric with principal axes $\vec{e}_1 \cdots \vec{e}_n$

e.g. $q(\vec{x}) = 5x^2 + 8xy + 5y^2 = \left\langle \vec{x}, \begin{bmatrix} 5 & 4 \\ 4 & 5 \end{bmatrix} \vec{x} \right\rangle$, $\begin{bmatrix} 5 & 4 \\ 4 & 5 \end{bmatrix} = \overset{Q}{\begin{bmatrix} 1/\sqrt{2} & 1/\sqrt{2} \\ -1/\sqrt{2} & 1/\sqrt{2} \end{bmatrix}} \overset{\wedge}{\begin{bmatrix} 1 & 0 \\ 0 & 9 \end{bmatrix}} Q^t$

Then $S_1 = \{\vec{x} \in \mathbb{R}^2 \mid q(\vec{x}) = 1\} = Q\,S'(1, 1/3)$

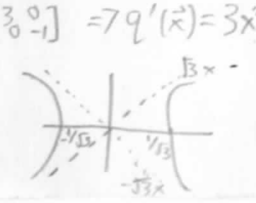


now use Q as new axes $\longrightarrow$



e.g. $q(\vec{x}) = x^2 + y^2 + 4xy$, $A = \begin{bmatrix} 1 & 2 \\ 2 & 1 \end{bmatrix} = \begin{bmatrix} 1/\sqrt{2} & -1/\sqrt{2} \\ 1/\sqrt{2} & 1/\sqrt{2} \end{bmatrix} \begin{bmatrix} 3 & 0 \\ 0 & -1 \end{bmatrix} Q^t$

$\begin{bmatrix} 3 & 0 \\ 0 & -1 \end{bmatrix} \Rightarrow q'(\vec{x}) = 3x^2 - y^2 = 1$

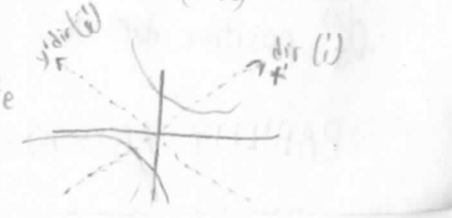


$y = 0,\ x = \pm\frac{1}{\sqrt{3}}$

$|x| < \frac{1}{\sqrt{3}},\ x = 0,\ y$ no sol^{n}

$|x| \geq \frac{1}{\sqrt{3}},\ y = \pm\sqrt{3x^2 - 1} = \pm\sqrt{3} \cdot \sqrt{x^2 - \frac{1}{\sqrt{3}}} \overset{x \to \infty}{\simeq} \pm\sqrt{3}\,x$

then rotate $\longrightarrow$



CHAPTER 24 More Quadratic Form Applications

APPLICATION: Local minima/maxima of real multivariable functions

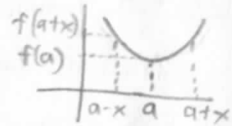
For smooth function f, recall Taylor series expansion

$$f(a+x) \underset{\text{for small }|x|}{\cong} f(a) + f'(a)x + \frac{1}{2!}f''(a)x^2 + \underset{\approx 0}{\text{error}}$$

Then f has a local minimum at a iff

① $f'(a) = 0$
② $f''(a) > 0$

} since then $f(a+x) - f(a) \cong \frac{1}{2!}f''(a)x^2 > 0 \ \forall$ nearby x, i.e. everywhere near a is higher



Recall the existence of partial derivatives, where ∂_i fixes every variable but x_i

e.g. $f(x_1, x_2, x_3) = x_1^2 x_2 + \sin(x_1 + x_3)$

$\partial_1 f(\vec{x}) = x_2 \, 2x_1 + \cos(x_1 + x_3)$

$\partial_2 f(\vec{x}) = x_1^2$

THM $\partial_{ij} f(\vec{x}) = \partial_i(\partial_j f(\vec{x})) = \partial_j(\partial_i(f(\vec{x}))) = \partial_{ji} f(\vec{x})$, i.e. order of partial derivatives doesn't matter

We can then generalize the Taylor series expansion to multivariable functions

$f = f(x_1, \dots x_n) = f(\vec{x})$ smooth, then

$$f(\vec{a}+\vec{x}) = f(\vec{a}) + \underbrace{D^{(1)}_{\vec{a}}}_{\substack{\text{linear map}\\ \text{in } \vec{x}}}\vec{x} + \frac{1}{2!}\underbrace{\langle \vec{x}, D^{(2)}_{\vec{a}}\vec{x}\rangle}_{\substack{\text{quadratic form}\\ \text{in } \vec{x}}} + \underset{\approx 0}{\text{error}}$$

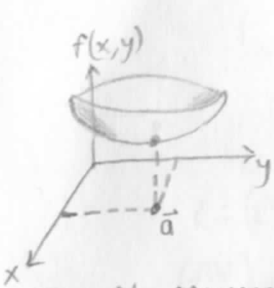
where $D^{(1)}_{\vec{a}} = (\partial_1 f, \partial_2 f, \dots, \partial_n f)\big|_{\vec{a}}$ is the first derivative of f

and $D^{(2)}_{\vec{a}} = \left[\partial_{ij} f(\vec{a})\right]$ is the second derivative of f (officially called the Hessian matrix of f I think)

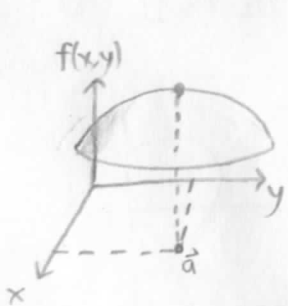
Then f has a local minima at $\vec{a}$ iff

① $D^{(1)}_{\vec{a}} = (0, 0, \dots, 0)$

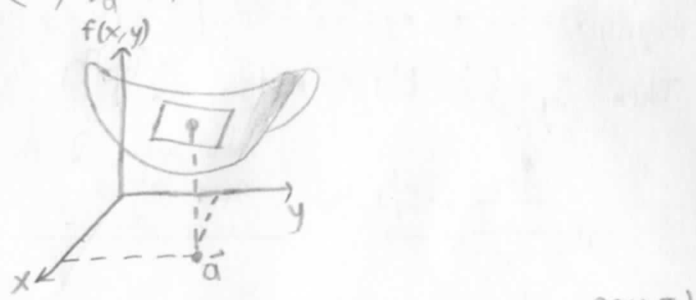
② $D^{(2)}_{\vec{a}}$ is a positive definite matrix, i.e. $\forall x$, $\langle \vec{x}, D^{(2)}_{\vec{a}} \vec{x}\rangle > 0$



LOCAL MINIMUM
$D^{(2)}_{\vec{a}}$ positive definite



LOCAL MAXIMUM
$D^{(2)}_{\vec{a}}$ negative definite



FLAT TANGENT PLANE (SADDLE POINT)
NO MIN OR MAX, no analogue in 1D
$D^{(2)}_{\vec{a}}$ is indefinite

PROBABILITY APPLICATION TO COVARIANCE MATRICES NOT INCLUDED

Jordan Normal Form: Prelude

Recall the purpose of diagonalization. Given $f: V \to V$ represented as $A_{n\times n}$ in some basis, we find $A = S\Lambda S^{-1}$ because it lets us understand a complicated object (the linear map f) as just a stretching along some vectors. More precisely, to apply f to a vector $\vec{x}$ is to stretch $\vec{x}$ in the direction of each eigenvector in S by its corresponding eigenvalue in Λ. (i.e. $A\vec{x} = \lambda\vec{x} \iff (A-\lambda I)\vec{x} = \vec{0}$)

So diagonalization allows us to understand linear maps as stretching, and more importantly tells us which vectors and how long to stretch by.

Unfortunately, not every linear map can be diagonalized, in particular, given $A_{n\times n}$:

A is diagonalizable $\iff$ A has n linearly independent eigenvectors (previous chapter)

Attempt 1

Consider for example $f(x,y,z) = \begin{pmatrix} 2x+2y+2z \\ 2y+2z \\ 2z \end{pmatrix}$ which can be written as $A = \begin{bmatrix} 2 & 2 & 2 \\ 0 & 2 & 2 \\ 0 & 0 & 2 \end{bmatrix}$

A has eigenvector $\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}$ to $\lambda = 2$ (note $\mathrm{geom}_A(2) = 1 < \mathrm{alg}_A(2) = 3$)

We need two more eigenvectors for A to be diagonalizable, but we can't find them since $2x+2y+2z$ and $2y+2z$ are more than simple stretching – they combine everything into one dimension while stretching too. So we cannot get a nice diagonal matrix, but

<u>IDEA</u> what if we combine the dimensions beforehand?

we want $\begin{pmatrix} 2x+2y+2z \\ 2y+2z \\ 2z \end{pmatrix}$ i.e. $\begin{bmatrix} 2 & 2 & 2 \\ 0 & 2 & 2 \\ 0 & 0 & 2 \end{bmatrix}$ in a simplified form.

If we send $\begin{pmatrix} x \\ y \\ z \end{pmatrix} \xrightarrow[\text{change of basis!}]{\text{via } C = \left(\begin{smallmatrix} 1 & 0 & 0 \\ 0 & 2 & 2 \\ 0 & 0 & 4 \end{smallmatrix}\right)} \begin{pmatrix} x \\ 2y+2z \\ 4z \end{pmatrix}$, then we can write it as a matrix with

line above diagonal i.e. superdiagonal

$\begin{bmatrix} 2 & 1 & 0 \\ 0 & 2 & 1 \\ 0 & 0 & 2 \end{bmatrix} =: J$, which is fairly close to a diagonal matrix: apart from 1's on the superdiagonal and the eigenvalues on the diagonal, everything else is 0.

Of course, now JC represents $\begin{pmatrix} 2(x)+1(2y+2z) \\ 2(2y+2z)+1(4z) \\ 2(4z) \end{pmatrix} = \begin{pmatrix} 2x+2y+2z \\ 4y+8z \\ 8z \end{pmatrix}$ instead of our original linear map f, but multiplying by $C^{-1} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1/2 & -1/4 \\ 0 & 0 & 1/4 \end{bmatrix}$ fixes that and returns $\begin{pmatrix} 2x+2y+2z \\ 2y+2z \\ 2z \end{pmatrix}$.

Letting $S = C^{-1}$, we therefore were able to simplify $A = SJS^{-1}$.

We can (somewhat) view
$$A = \begin{bmatrix} 2 & 2 & 2 \\ 0 & 2 & 2 \\ 0 & 0 & 2 \end{bmatrix} = \begin{array}{c} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1/2 & -1/4 \\ 0 & 0 & 1/4 \end{bmatrix} \\ \begin{array}{ccc} \vec{x}_1 & \vec{x}_2 & \vec{x}_3 \end{array} \end{array} \begin{bmatrix} 2 & 1 & 0 \\ 0 & 2 & 1 \\ 0 & 0 & 2 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1/2 & -1/4 \\ 0 & 0 & 1/4 \end{bmatrix}^{-1} = SJS^{-1}$$

as stretching along $\vec{x}_1, \vec{x}_2$ and $\vec{x}_3$ by $\lambda = 2$, up to some error term generated by the other stretching vectors.

ATTEMPT 2

We like eigenvectors because they tell us what inputs remain "stable" (up to some stretching), i.e.

$A\vec{x} = \lambda\vec{x}$, equivalently $(A-\lambda I)\vec{x} = \vec{0}$

If we manage to find n linearly independent eigenvectors, then A is completely described by stretching along said vectors (A is diagonalizable).

Often, however, we cannot - it is not enough to consider the action of A on its input $\vec{x}$. Instead, we need to also consider the action of A on $A\vec{x}, A^2\vec{x}, \ldots$, i.e. we want to find all inputs that are eventually stable, and it turns out this is enough to describe A.

A generalized eigenvector of rank m and corresponding to λ satisfies $(A-\lambda I)^m \vec{x}_m = \vec{0}$ and $(A-\lambda I)^{m-1}\vec{x}_m = \vec{0}$, i.e. it takes m applications of A to become stable (ordinary eigenvectors are of rank 1).

It turns out that $A_{n\times n}$ $\Rightarrow$ there are always n linearly independent generalized eigenvectors
$\Rightarrow$ n linearly independent vector classes are stable under A eventually

e.g. $A = \begin{pmatrix} 2&2&2\\0&2&2\\0&0&2 \end{pmatrix} = \begin{pmatrix} 1&0&0\\0&1/2&-1/4\\0&0&1/4 \end{pmatrix}\begin{pmatrix} 2&1&0\\0&2&1\\0&0&2 \end{pmatrix}\begin{pmatrix} 1&0&0\\0&1/2&-1/4\\0&0&1/4 \end{pmatrix}^{-1} = SJS^{-1}$

$\lambda = 2$ $\Rightarrow (A-\lambda I)\begin{pmatrix}1\\0\\0\end{pmatrix} = \begin{pmatrix} 0&2&2\\0&0&2\\0&0&0 \end{pmatrix}\begin{pmatrix}1\\0\\0\end{pmatrix} = \begin{pmatrix}0\\0\\0\end{pmatrix}$ and $(A-\lambda I)^0\begin{pmatrix}1\\0\\0\end{pmatrix} = \begin{pmatrix}1\\0\\0\end{pmatrix} \neq \vec{0}$

$\Rightarrow (A-\lambda I)^2\begin{pmatrix}0\\1/2\\0\end{pmatrix} = \begin{pmatrix} 0&0&4\\0&0&0\\0&0&0 \end{pmatrix}\begin{pmatrix}0\\1/2\\0\end{pmatrix} = \begin{pmatrix}0\\0\\0\end{pmatrix}$ and $(A-\lambda I)^1\begin{pmatrix}0\\1/2\\0\end{pmatrix} = \begin{pmatrix} 0&2&2\\0&0&2\\0&0&0 \end{pmatrix}\begin{pmatrix}0\\1/2\\0\end{pmatrix} = \begin{pmatrix}1\\0\\0\end{pmatrix} \neq \vec{0}$

$\Rightarrow (A-\lambda I)^3\begin{pmatrix}0\\-1/4\\1/4\end{pmatrix} = \begin{pmatrix} 0&0&0\\0&0&0\\0&0&0 \end{pmatrix}\begin{pmatrix}0\\-1/4\\1/4\end{pmatrix} = \begin{pmatrix}0\\0\\0\end{pmatrix}$ and $(A-\lambda I)^2\begin{pmatrix}0\\-1/4\\1/4\end{pmatrix} = \begin{pmatrix} 0&0&4\\0&0&0\\0&0&0 \end{pmatrix}\begin{pmatrix}0\\-1/4\\1/4\end{pmatrix} = \begin{pmatrix}1\\0\\0\end{pmatrix} \neq \vec{0}$

The span of the generalized eigenvectors to λ form the generalized eigenspace of λ.

Essentially, eigenvectors reduce maps to scalar multiplication, and generalized eigenvectors reduce maps to something like scalar multiplication. (more on that later)

More importantly, generalized eigenspaces are invariant subspaces - in some sense we are splitting n-dimensional Euclidean space into subspaces that could be interesting to work with/have features of their own. (Maybe.)

CHAPTER 25 JORDAN NORMAL FORM (MATRICES)

JORDAN THM In $\mathbb{C}$, given $A_{n\times n}$, $\exists B$ invertible such that $A = BJB^{-1}$, where J is the Jordan form of A, a matrix consisting of Jordan blocks

$$J = \begin{bmatrix} \overbrace{[J_1]}^{m_1} & & & & 0 \\ & \overbrace{[J_2]}^{m_2} & & & \\ & & \ddots & & \\ & & & \overbrace{[J_k]}^{m_k} & \\ 0 & & & & \overbrace{[J_s]}^{m_s} \end{bmatrix}$$

s-blocks in total, $m_1 + m_2 + \dots + m_s = n$,

and $J_k = \begin{bmatrix} \lambda_k & 1 & & 0 \\ & \lambda_k & 1 & \\ & & \lambda_k & \ddots & \\ & & & \ddots & 1 \\ 0 & & & & \lambda_k \end{bmatrix}$ (1's on superdiagonal)

A and J have (not necessarily distinct) eigenvalues $\lambda_1 \dots \lambda_s$ with algebraic multiplicity $m_1 \dots m_s$

J's eigenvectors are $\vec{e}_1, \vec{e}_{1+m_1}, \vec{e}_{1+m_1+m_2}, \dots \vec{e}_{1+m_1+\dots+m_{s-1}}$ (corresponding to start of block), and so A's eigenvectors are the corresponding columns of $\underset{\vec{b}_1}{\underset{\parallel}{B\vec{e}_1}}, \underset{\vec{b}_{1+m_1}}{\underset{\parallel}{B\vec{e}_{1+m_1}}}, \dots$

Note: A diagonalizable $\Leftrightarrow$ all J-blocks size 1

NILPOTENT $T_{n\times n}$ is nilpotent iff $T^k = [0]$ for some k

$\Leftrightarrow$ all of T's eigenvalues are 0, since otherwise $T\vec{x} = \lambda\vec{x} \Rightarrow \underbrace{T^k\vec{x}}_{0} = \underbrace{\lambda^k\vec{x}}_{\text{must be } 0}$

THM (JNF) If T is nilpotent, $\exists C$ invertible such that

$$J = \begin{bmatrix} 0 & \Delta & & 0 \\ & 0 & \Delta & \\ & & \ddots & \ddots \\ 0 & & & 0 \end{bmatrix} = C^{-1}TC, \quad \Delta = 0 \text{ or } 1$$

NOTE If T has real entries, we can choose C to have real entries too

CHAPTER 26 JORDAN NORMAL FORM (MAPS)

NILPOTENT (map version) $f: V \to V$ ($\mathbb{R}$ or $\mathbb{C}$) is nilpotent $\Leftrightarrow \exists k \in \mathbb{N}$, $f^k = \underbrace{f \circ f \circ f \dots \circ f}_{k \text{ times}} \equiv 0$.

If f is nilpotent, $\vec{x} \in V$, then

① $T = \{\vec{x}, f(\vec{x}), f^2(\vec{x}), \dots \underbrace{f^{k-1}(\vec{x})}_{\neq \vec{0}}\}$ is the trajectory of $\vec{x}$

② $k = |T| = \min\{j \mid f^j(\vec{x}) = \vec{0}\}$ is the index of $\vec{x}$.

③ A trajectory is full $\Leftrightarrow \vec{x} \notin \mathrm{Im}(f)$, i.e. you can't extend backwards

Let $I^j = \mathrm{Im}\, f^j$, $K^j = \ker f^j$, k = index of f, then

① $V \underset{\neq}{\geq} I^1 \underset{\neq}{\geq} I^2 \underset{\neq}{\geq} I^3 \dots I^{k-1} \underset{\neq}{\geq} I^k = \{\vec{0}\}$

② $\{\vec{0}\} \underset{\neq}{\leq} K^1 \underset{\neq}{\leq} K^2 \underset{\neq}{\leq} \dots \quad K^{k-1} \underset{\neq}{\leq} K^k = V$

LEMMA If $\tau_1 \dots \tau_t$ are trajectories with independent last elements, their union is also independent

Γ $t=1$: Consider $\left(\sum_{i=0}^{k-1} \lambda_i f^i(\vec{x}) = \vec{0}\right)(f^{k-1})$

$\Rightarrow \sum_{i=0}^{k-1} \lambda_i \underset{\uparrow\, 0 \text{ for } i \neq 0}{f^{i+k-1}(\vec{x})} = \vec{0} \quad \Rightarrow \lambda_0 = 0,$

similarly $\lambda_0, \dots \lambda_{k-1} = 0$

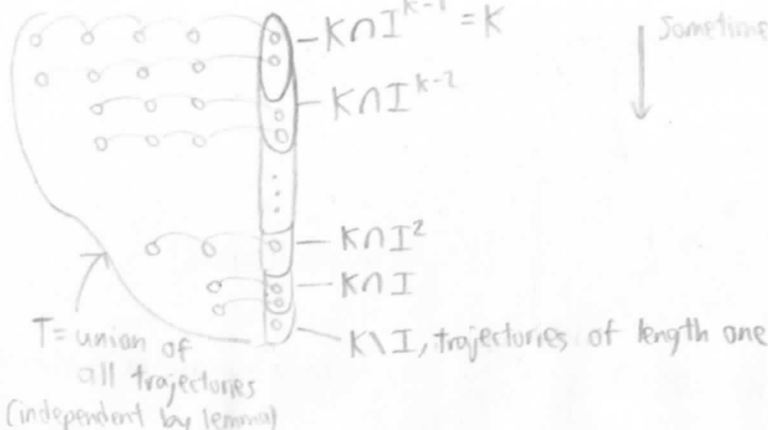
$\Rightarrow$ independent

$t=2$: $\vec{x}, f(\vec{x}), f^2(\vec{x}) \dots f^{k-1}(\vec{x}) = \tau_1$
$\quad\;\; \vec{y}, f(\vec{y}), \dots f^{k-2}(\vec{y}) = \tau_2$

$\left.\begin{array}{l}\vec{a} = f^{k-1}(\vec{x}) \\ \vec{b} = f^{k-2}(\vec{y})\end{array}\right\}$ independent by assumption

But since $(f^{k-2})^{-1}(\vec{a}) = f(\vec{x})$ and $(f^{k-2})^{-1}(\vec{b}) = \vec{y}$, then $f(\vec{x}), \vec{y}$ are also independent

CONSTRUCTING A BASIS FROM TRAJECTORIES



Extend basis from the top to bottom.
Sometimes $K \cap I^j = K \cap I^{j-1}$, i.e. no Jordan block of size $j-1$, no new basis vector

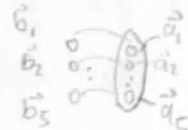
CLAIM $\langle T \rangle = \text{span}(T) = V$

Consider $K^i \leq \langle T \cap K^i \rangle$

$i=1$: $K^1 = \text{Ker} f \leq \langle T \rangle$ by construction

$i=2$: Let $\vec{v} \in K^2 \setminus K^1$ (so $\vec{v}$ has index 2).

Suppose $T \cap K \cap I$ ($\in$ basis in $K \cap I$) has $\vec{a}_1, \dots \vec{q}_k$



So $K \cap I = f(\vec{v}) = \sum_1^5 \lambda_i \vec{a}_i$, let $\vec{u} = \sum_1^5 \lambda_i \vec{b}_i$, then

$f(\vec{v} - \vec{u}) = f(\vec{v}) - \sum_1^5 \lambda_i f(b_i) = \vec{0}$

i.e. $\vec{v} - \vec{u} \in K$

$\Rightarrow \vec{v} = \underbrace{\vec{u}}_{\substack{\in (K^2 \setminus K) \cap T \\ \text{i.e. index } 2}} + \underbrace{\vec{v} - \vec{u}}_{\substack{\in \langle K \cap T \rangle \\ \text{i.e. index } 1}} \in \langle T \cap K^2 \rangle$

$i=k$: analogous to $i=2$, split into vectors of index k and less than k ┘

① Finding basis for $K \cap I^j$

$$\left.\begin{array}{l} \vec{y} \in \text{Im } A^j \Rightarrow \vec{y} = A^j \vec{x} \text{ for some } \vec{x} \in \mathbb{R}^n \\ \vec{y} \in \text{Ker } A \Rightarrow A\vec{y} = \vec{0} \end{array}\right\} \left.\begin{array}{l} A(A^j \vec{x}) = \vec{0} \\ \Leftrightarrow A^{j+1} \vec{x} = \vec{0} \end{array}\right\} \Rightarrow K \cap I^j = \{A^j \vec{x} \mid \vec{x} \in \text{Ker } A^{j+1}\}$$

② Extending basis $K \cap I^j \to K \cap I^{j-1}$

Suppose $U \leq V \leq \mathbb{R}^n$ (or $\mathbb{C}^n$) and we have basis $\vec{u}_1, \dots \vec{u}_s$ for U (triangular by Schur)

Now we want to extend to $V = \langle \vec{v}_1, \dots \vec{v}_k \rangle$.

Set $A = \begin{bmatrix} \vec{u}_1 \\ \vec{u}_2 \\ \vdots \\ \vec{u}_s \\ \hline \vec{v}_1 \\ \vec{v}_2 \\ \vdots \\ \vec{v}_k \end{bmatrix} \begin{array}{l} \left.\vphantom{\begin{array}{c} a \\ a \\ a \\ a \end{array}}\right\} \text{fixed, don't modify} \\ \left.\vphantom{\begin{array}{c} a \\ a \\ a \\ a \end{array}}\right\} \text{Gaussian elimination until at most 1 pivot in every column} \end{array}$ $\Rightarrow$ remaining vectors are independent and with $\vec{u}_1 \dots \vec{u}_s$ will span V

This general procedure works but computing $\cap$ sucks.

For small matrices,
① find longest chain $\{\vec{x}, T\vec{x}, \dots T^{k-1}\vec{x}\}$ by raising T until $T^k = [0]$
② solve $T^{k-1}\vec{x} = \vec{0}$ to get $\vec{x}$
③ you have $\vec{x}$, you have $T \Rightarrow$ you have the chain

Recall that in $\mathbb{C}$, Schur's Lemma tells us $A_{n\times n} = CTC^{-1}$,

$$T = \begin{bmatrix} \boxed{\begin{matrix}\lambda_1 & * \\ 0 & \ddots\ \lambda_1\end{matrix}} & & & * \\ & \boxed{\begin{matrix}\lambda_2 & * \\ 0 & \ddots\ \lambda_2\end{matrix}} & & \\ 0 & & \ddots & \\ & & & \boxed{\begin{matrix}\lambda_k & * \\ 0 & \ddots\ \lambda_k\end{matrix}} \end{bmatrix}$$

eigenvalue blocks

GOAL We want to zero out the top right $*$, i.e. we want to find B, $T' = BTB^{-1}$ such that

$$T' = \begin{bmatrix} \boxed{\begin{matrix}\lambda_1 & * \\ 0 & \ddots\ \lambda_1\end{matrix}} & & 0 \\ & \ddots & \\ 0 & & \boxed{\begin{matrix}\lambda_k & * \\ 0 & \ddots\ \lambda_k\end{matrix}} \end{bmatrix}$$

NOTE The vectors $\vec{b}_{i_1} \dots \vec{b}_{i,k_i}$ corresponding to the i^{th} block $\begin{bmatrix}\lambda_i & * \\ 0 & \ddots\ \lambda_i\end{bmatrix}$ together span an invariant subspace of the original map ℓ, i.e. $U_i = \mathrm{span}(b_{i_1} \dots b_{i k_i})$, $\ell(U_i) \subseteq U_i$

We have a procedure to clean out the $*$'s one at a time

Consider $B(j,k,\alpha) = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ & 1 & 0 & \alpha & & \\ & & 1 & 0 & & \\ & & & 1 & 0 & \\ & 0 & & & \ddots & \\ & & & & & 1 \end{bmatrix} = B_\alpha$ — j^{th} row, k^{th} column

Then $B_\alpha^{-1} C = \begin{bmatrix} [\ \vec{c}_1\ ] \\ [\ \vec{c}_2\ ] \\ \vdots \\ [\vec{c}_j - \alpha c_k] \\ \vdots \\ [\ \vec{c}_n\ ] \end{bmatrix}$ and $CB = \begin{bmatrix} \begin{bmatrix}\vec{c}_1\end{bmatrix} \begin{bmatrix}\vec{c}_2\end{bmatrix} \cdots \begin{bmatrix}\vec{c}_k + \alpha \vec{c}_j\end{bmatrix} \cdots \begin{bmatrix}\vec{c}_n\end{bmatrix} \end{bmatrix}$

and since T is triangular, the only nonzero changes are above and to the right of α

Furthermore so if we want to zero out t_{jk}, pick $\alpha = \dfrac{t_{jk}}{t_{kk} - t_{jj}} = \dfrac{\text{point to be 0'd}}{\text{diagonal below} - \text{diagonal to left}}$

So now we have $T' = \begin{bmatrix} \boxed{\begin{matrix}\lambda_1 & * \\ 0 & \ddots\ \lambda_1\end{matrix}} & & 0 \\ & \ddots & \\ 0 & & \boxed{\begin{matrix}\lambda_k & * \\ 0 & \ddots\ \lambda_k\end{matrix}} \end{bmatrix}$.

Each eigenvalue-block $\lambda_j = \begin{bmatrix}\lambda_j & \ddots & * \\ 0 & & \lambda_j\end{bmatrix} = \begin{bmatrix}0 & & * \\ & 0 & \\ 0 & & 0\end{bmatrix} + \lambda_j I$

$= BJB^{-1} + \lambda_j BIB^{-1}$

$= B(J + \lambda_j)B^{-1}$

$=$ tada! $\begin{bmatrix}\lambda_j & \Delta & 0 \\ & \lambda_j & \\ 0 & & \ddots\ \lambda_j\end{bmatrix}$ $\Delta = 0$ or 1

So we have JNF for arbitrary $n\times n$ matrices, since

$$\underset{[*]}{A_{n\times n}} \xrightarrow{\text{Schur}} T_{n\times n} \xrightarrow{\text{cleaning}} T'_{n\times n} \xrightarrow{\text{JNF for nilpotent}} J$$

$$\begin{bmatrix} \boxed{\begin{matrix}\lambda_1 & * \\ & \ddots\ \lambda_1\end{matrix}} & * \\ 0 & \boxed{\begin{matrix}\lambda_n & * \\ & \ddots\ \lambda_n\end{matrix}} \end{bmatrix} \quad \begin{bmatrix} \boxed{\begin{matrix}\lambda_1 & * \\ 0 & \lambda_1\end{matrix}} & 0 \\ 0 & \boxed{\begin{matrix}\lambda_k & * \\ 0 & \lambda_k\end{matrix}} \end{bmatrix} \quad \begin{bmatrix}\lambda_1 & \Delta & & 0 \\ & \ddots\ \lambda_1 & & \\ & & \lambda_k & \\ 0 & & & \ddots\ \lambda_k\end{bmatrix} \quad \Delta = 0 \text{ or } 1$$

CHAPTER 28 SINGULAR VALUE DECOMPOSITION

$A_{m\times n}$ in $\mathbb{R}$, then

<u>Lemma</u> A^tA is symmetric, positive definite

$\ulcorner (A^tA)^t = A^tA \Rightarrow$ symmetric, since real
$\Rightarrow A^tA = Q\Lambda Q^{-1} \Rightarrow$ eigenvectors of unit length

$\lambda = \lambda\|s\|^2 = \langle \vec{s}|\lambda\vec{s}\rangle = \langle\vec{s}|A^tA\vec{s}\rangle = \langle A\vec{s}|A\vec{s}\rangle = \|A\vec{s}\|^2 \geq 0 \lrcorner$

<u>DEF</u> A^tA has eigenvalues $\lambda_1 \geq \lambda_2 \geq \cdots \lambda_n \geq 0$, then

$\sigma_k = \sqrt{\lambda_k}$ are the <u>singular values</u> of A

<u>THM</u> $\exists\, Q, S$ orthogonal,

$$\Sigma_{m\times n} = \left[\begin{array}{c|c} \boxed{\begin{matrix}\sigma_1 & & \\ & \ddots & \\ & & \sigma_r\end{matrix}} & 0 \\ \hline 0 & 0 \end{array}\right] = Q^tAS$$

singular values >0

$\ulcorner$ Already have $A = Q\Lambda Q^{-1}$, but pick $\vec{q}_i = \frac{1}{\sigma_i}A\vec{s}_i$. (eigenvector: $\vec{s}_i$)

$\vec{q}_1 \cdots \vec{q}_r$ are ON since for $i\neq j$, $\langle\vec{s}_i|\vec{s}_j\rangle = 0$

and $\lambda_i = \|A\vec{s}_i\|^2 \Rightarrow \sigma_i = \|A\vec{s}_i\| \Rightarrow 1 = \|\frac{1}{\sigma_i}A\vec{s}_i\| = \|\vec{q}_i\|$

Also if $r<m$, we can extend to basis in $\mathbb{R}^m$ with Gram Schmidt

Then $\Sigma = C_{Q\leftarrow E}\, A\, C_{E\leftarrow S} = Q^tAS$

$$= \left[[A\vec{s}_1]_Q \cdots [A\vec{s}_r]_Q \cdots [A\vec{s}_n]_Q \right]$$

$$= \left[[\sigma_1\vec{q}_1]_Q \cdots [\sigma_r\vec{q}_r]_Q\, [\vec{0}]_Q \cdots [\vec{0}]_Q \right]$$

$$= \left[\begin{array}{c|c} \boxed{\begin{matrix}\sigma_1 & & 0 \\ & \sigma_2 \ddots & \\ 0 & & \sigma_r\end{matrix}} & 0 \\ \hline 0 & 0 \end{array}\right] \lrcorner$$

We can write $A = Q\Sigma S^t$ in <u>outer product form</u>,

$$A = \vec{q}_1\sigma_1\vec{s}_1^{\,t} + \vec{q}_2\sigma_2\vec{s}_2^{\,t} + \cdots + \vec{q}_r\sigma_r\vec{s}_r^{\,t}$$

listed in order of contribution to A

$\Rightarrow$ application to compression, we can drop some of the last terms without much loss. Like a Taylor expansion

FIN