Singular Value Decomposition and Applications

Swanand Khare

Department of Mathematics, IIT Kharagpur

Motivation

- Simple question:
  - Did you pass the examination? Answer: yes/no
  - Can we do better than yes/no type answer?
  - Definitely!! But how?

- Ask smarter question!!
  - By how many marks, I failed/passed?
  - More refined version: grades EX,A,B,C,D,P,F

- Similar situation arises in many system and control questions, in particular in matrix theory
  - To a class of problems, the answer lies in SVD

Important Question

- A very useful concept in linear algebra is rank of a matrix.

- $A \in R^{n \times n}$, rank of A is the number of linearly independent rows or columns.

- Non-singularity is same as full rank case in square matrices.

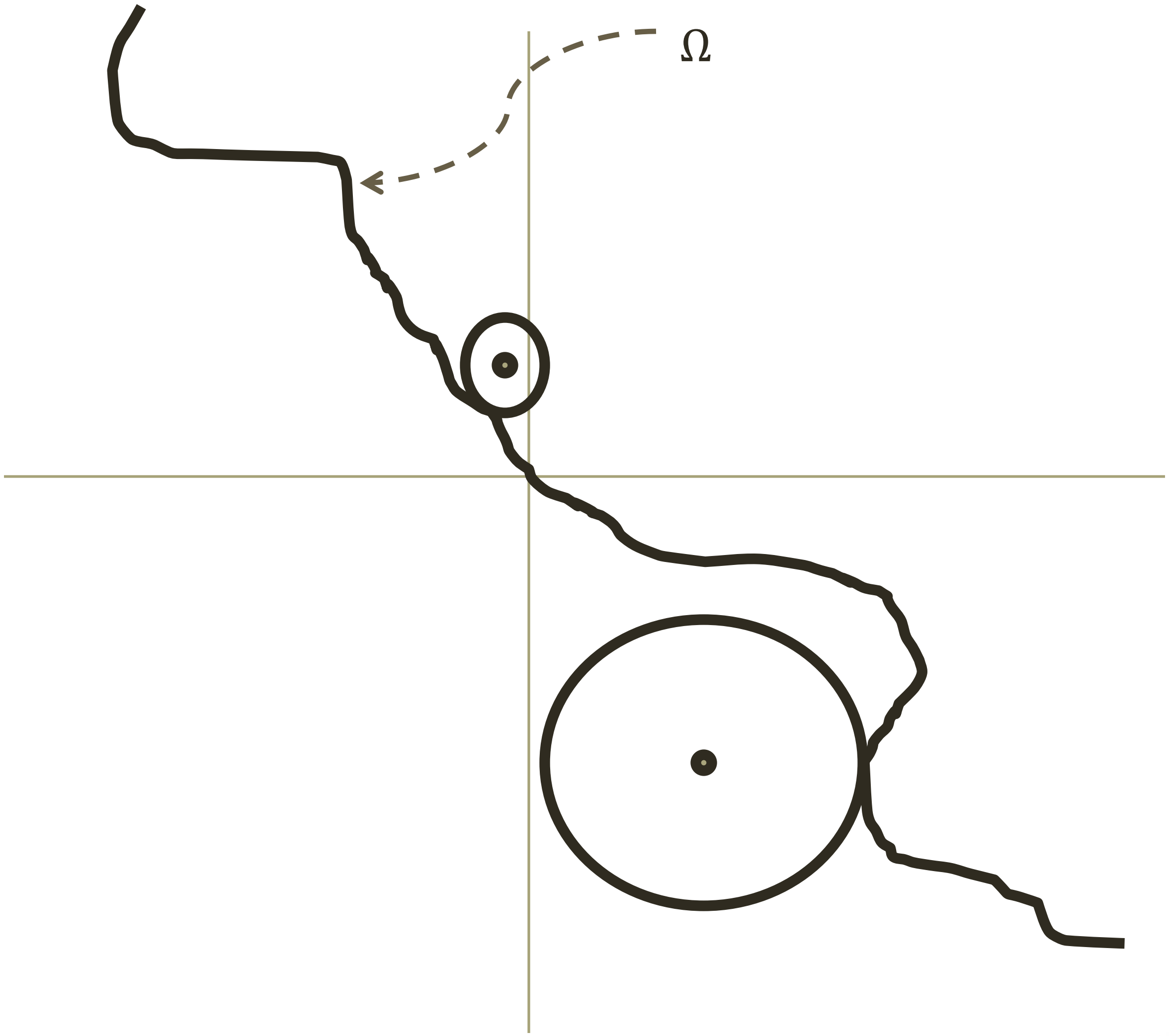
- $A = \begin{bmatrix} 1 & 1 \\ 1 & 1.01 \end{bmatrix}$; is A invertible ?

- Are you comfortable answering this question?

Determinant and rank

- For $A \in R^{n \times n}$, A being invertible is equivalent to $\det(A) \neq 0$.
- Geometry of this set: $\Omega = \{A \in R^{2\times 2} \mid \det(A) = 0\}$
- If $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$, then $\det(A) = ad - bc$
- $\det(A) = 0$ means $ad - bc = 0$.
- In four dimensional space, this is an algebraic variety.
- The set Ω is a set of measure zero.
    - in the language of layman, this is a very *thin* set

Geometry of Ω a very simplistic approach



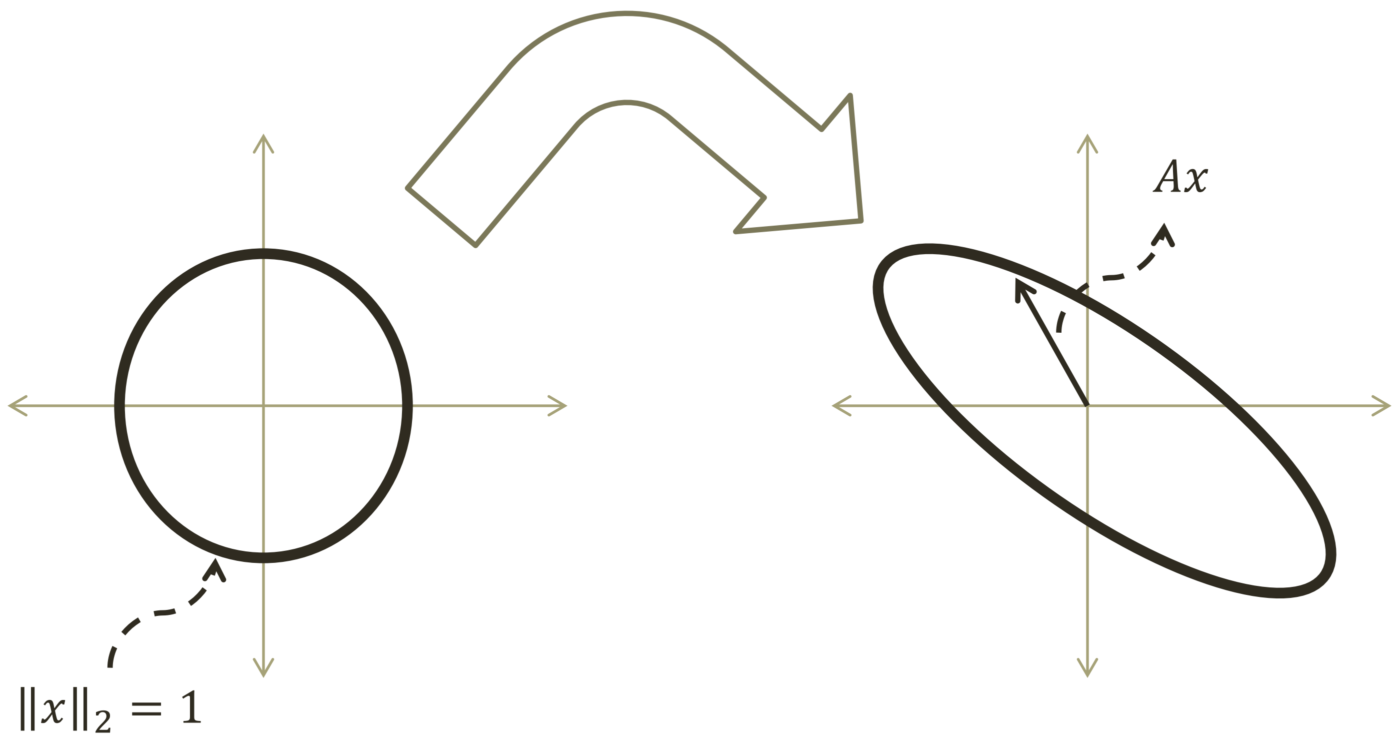
Concept of distance

- Norm of a vector
  - Length of a vector
  - Most popular 2 norm, spectral norm

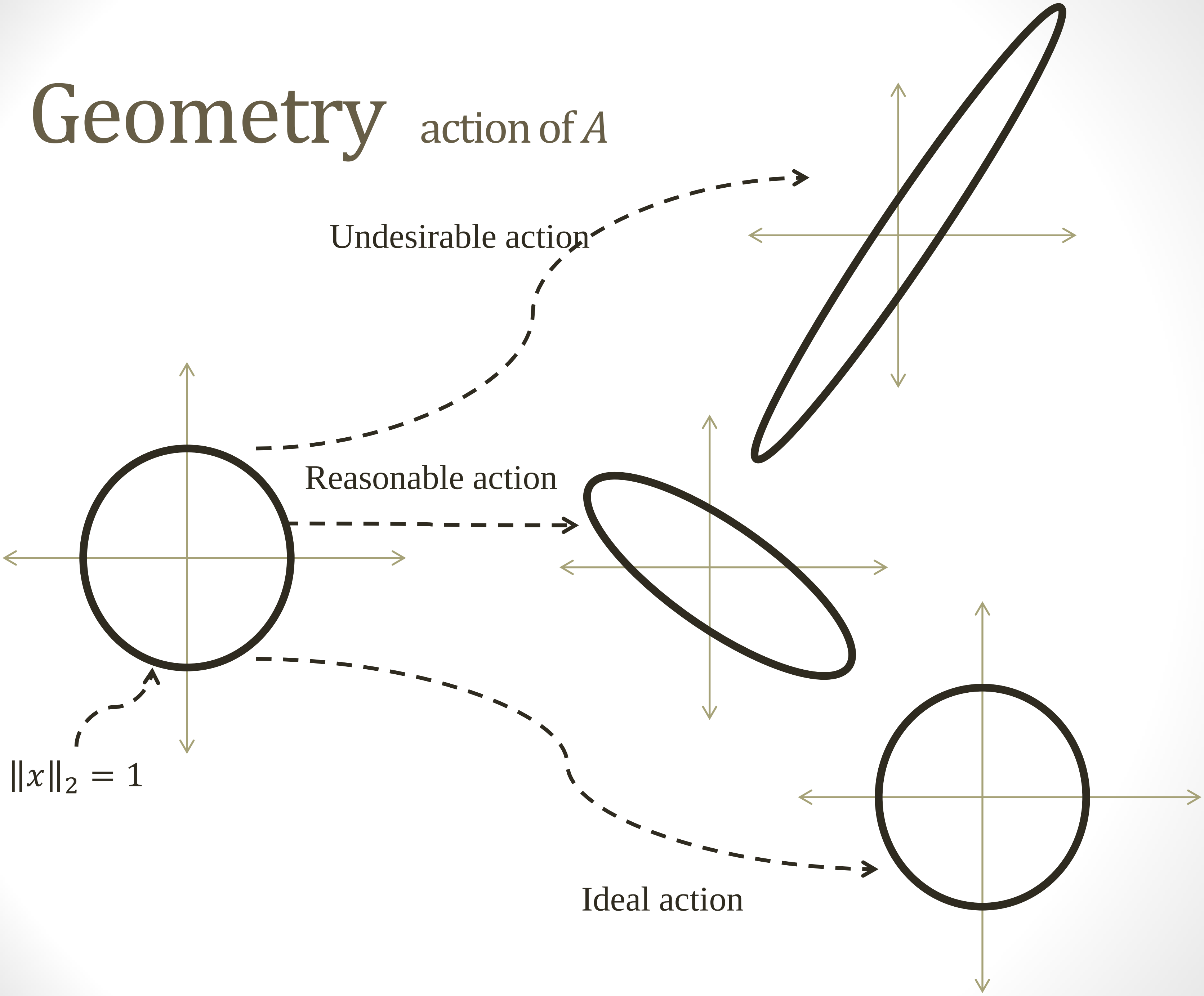
- Norm of a matrix
  - A matrix $A \in R^{n\times m}$ can be viewed as a map from R^m to R^n
  - Induced norm of A is defined as

$$\|A\|_2 = \max_{\|x\|_2=1} \|Ax\|_2$$

Geometry action of A



Geometry action of A



Ideal matrices Orthogonal Matrices

- Orthogonal matrices (in general isometry) $Q^T Q = QQ^T = I$
  - Map unit circle to unit circle

- Example 1: (Rotator)
  - $A = \begin{bmatrix} \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{bmatrix}$

- Example 2: (Reflector)
  - $A = \begin{bmatrix} 0 & -1 \\ -1 & 0 \end{bmatrix}$

- Can you think of any other matrix which maps circles to circles?

SVD theorem

- A given matrix $A \in R^{m \times n}$ can be decomposed as $A = U\Sigma V^T$
  - $U \in R^{m \times m}$ orthogonal
  - $V \in R^{n \times n}$ orthogonal
  - $\Sigma \in R^{n \times m}$ diagonal

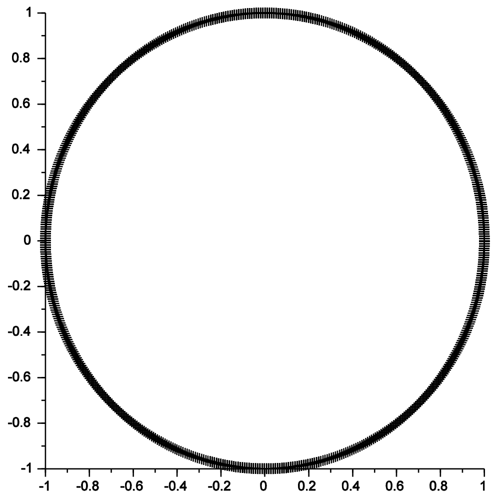
- Columns of U: Left singular vectors
- Columns of V: Right singular vectors
- Diagonal entries of Σ: Singular values (always non-increasing and non-negative)

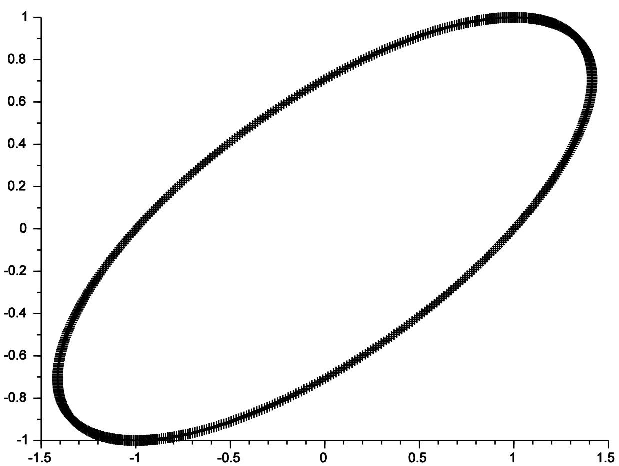
Properties of SVD

- $AV = U\Sigma$
- $A[v_1 \ \ v_2 \cdots \ \ v_n] = [u_1\, u_2 \cdots u_m]\begin{bmatrix} \sigma_1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \sigma_m \end{bmatrix}$
- If $r$singular values are non-zero, then
$$Av_1 = \sigma_1 u_1, \dots, Av_r = \sigma_r u_r$$
- $rank\ (A) = r$
- Nullspace: $\mathrm{N}(A) = span\{v_{r+1}, \dots, v_n\}$
- Range space: $\mathrm{R}(A) = span\{u_1, \dots, u_r\}$

Revisiting geometry

- $A = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}$
- SVD of A is $A = U\Sigma V^T$ where
  - $U = \begin{bmatrix} 0.85 & -0.52 \\ 0.52 & 0.85 \end{bmatrix}$, $V = \begin{bmatrix} 0.52 & -0.85 \\ 0.85 & 0.52 \end{bmatrix}$
  - $\Sigma = \begin{bmatrix} 1.61 & 0 \\ 0 & 0.61 \end{bmatrix}$





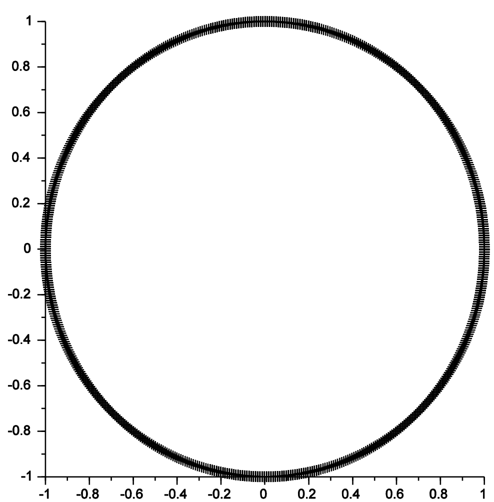
Revisiting geometry

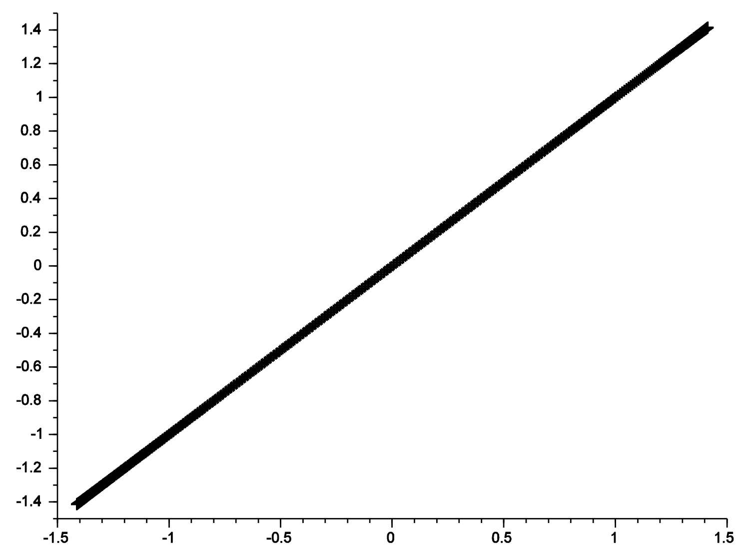
- v_1 is the direction in domain which gets maximum magnification!!
- u_1 is the direction of maximum magnification in range space. (A unit vector in the direction of major axis)
- σ_1 is the amount of maximum magnification.
- In fact, $\sigma_1 = \|A\|_2$

- v_2 is the direction in domain which gets minimum magnification!!
- u_2 is the direction of minimum magnification in range space. (A unit vector in the direction of minor axis)
- σ_2 is the amount of minimum magnification.
- Ratio of σ_1 and σ_2: Condition number of the matrix.

SVD rank deficient case

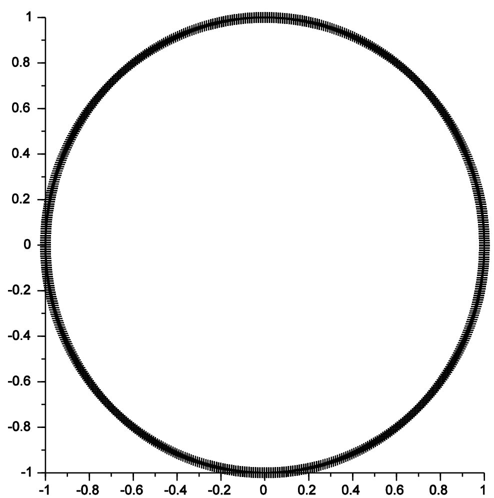
- $A = \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}$
- $U = \begin{bmatrix} -0.70 & -0.70 \\ -0.70 & 0.70 \end{bmatrix}, V = \begin{bmatrix} -0.70 & -0.70 \\ -0.70 & 0.70 \end{bmatrix}$
- $\Sigma = \begin{bmatrix} 2 & 0 \\ 0 & 0 \end{bmatrix}$

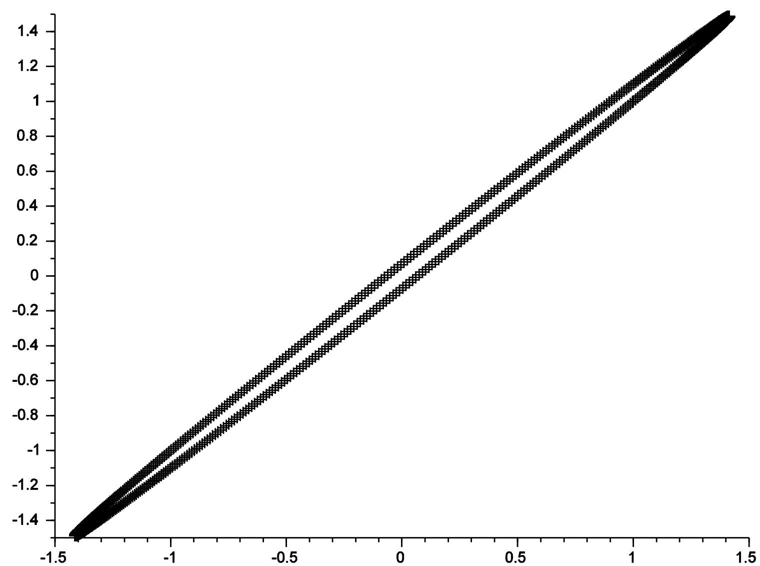




SVD nearness to singularity case

- $A = \begin{bmatrix} 1 & 1 \\ 1 & 1.01 \end{bmatrix}$
- $U = \begin{bmatrix} -0.705 & -0.708 \\ -0.708 & 0.705 \end{bmatrix}, V = \begin{bmatrix} -0.705 & -0.708 \\ -0.708 & 0.705 \end{bmatrix}$
- $\Sigma = \begin{bmatrix} 2.005 & 0 \\ 0 & 0.004 \end{bmatrix}$





Tolerance, Numerical Rank Application 1

- Number of non-zero singular values is equal to rank
- What is zero? Is zero really zero?
- Tolerance
- Numerical rank = number of non-zero singular values greater than the tolerance!!
  - Very useful in computation of nullspace and range space in when tolerance is specified.

Another form of SVD

- $A = U\Sigma V^T$

- For r non-zero singular values, A can be written as sum of r rank one matrices as
$$A = \sigma_1 u_1 v_1^T + \sigma_2 u_2 v_2^T + \cdots + \sigma_r u_r v_r^T$$

- Note that $u_i v_i^T$ is always a rank one matrix.

- Useful in stating Low Rank Approximation of a matrix.

Low Rank Approximation Application 2

- Given a matrix $A \in R^{n\times n}$ with rank n, find the nearest matrix A_{Low} such that rank of A_{Low} is $n-1$.

- $A_{Low} = \text{argmin}_{\text{B}} \|A - B\|_2$ where $rank\ (B) = n - 1$.

- Solution (Eckart-Young-Mirsky Theorem)
  - If $A = \sigma_1 u_1 v_1^T + \sigma_2 u_2 v_2^T + \cdots + \sigma_n u_n v_n^T$ is the svd of A, then A_{Low} is given by
    $\text{A}_{\text{Low}} = \sigma_1 u_1 v_1^T + \sigma_2 u_2 v_2^T + \cdots + \sigma_{n-1} u_{n-1} v_{n-1}^T$

  - $\|A - A_{Low}\|_2 = \sigma_n$

General case

- Given a matrix $A \in R^{m \times n}$ with rank r, find the nearest matrix A_{Low} such that rank of A_{Low} is k, where $k < r$.
- $A_{Low} = \text{argmin}_{\text{B}} \|A - B\|_2$ where $rank\ (B) = k$.
- Solution (Eckart-Young-Mirsky Theorem)
  - If $A = \sigma_1 u_1 v_1^T + \sigma_2 u_2 v_2^T + \cdots + \sigma_r u_r v_r^T$ is the svd of A, then A_{Low} is given by $\text{A}_{\text{Low}} = \sigma_1 u_1 v_1^T + \sigma_2 u_2 v_2^T + \cdots + \sigma_k u_k v_k^T$
  - $\|A - A_{Low}\|_2 = \sigma_{k+1}$

Structured LRA

- Along with the low rank constraint, there is a constraint of structure.

- Given a matrix $A \in R^{m \times n}$ with rank r and some specific structure, find the nearest matrix A_{Low} such that rank of A_{Low} is $k,$ where $k < r$ and the structure of A_{Low} is same as that of A.

- More challenging problem than SVD.

- A lot of applications of SLRA.
  - Hankel: model order reduction
  - Sylvester/Toeplitz: radius of controllability

Computation of SVD

- SVD of $A \in R^{m \times n}$ is equivalent to eigenvalue eigenvector decomposition of matrices $A^T A$ and AA^T.
- Computation of SVD is done as eigenvalue eigenvector decomposition.
- Iterative procedure.
- Golub-Reinch algorithm to compute SVD.

Books

- Matrix Computations
  - Golub and Van Loan
- Fundamentals of matrix computations
  - Watkins
- Numerical linear algebra and applications
  - Datta

Question???

THANK YOU!!!