Numerical Linear Algebra
Fundamentals of LA

Zahra Lakdawala

October 2, 2021

Outline

1 Matrix Multiplications

2 Range, Rank and Inverses

3 Inner products

4 Unitary matrices

5 Vector Norms

6 Matrix Norms

7 SVD

Matrix Multiplications

Definition

- Matrix-vector product $\mathbf{b} = \mathbf{A}\mathbf{x}$

$$b_i = \sum_{j=1}^{n} a_{ij} x_j$$

- All entries belong to $\mathbb{C}$, the field of complex numbers. The space of $m-$vectors is $\mathbb{C}^m$, and the space of $m \times n$ matrices is $\mathbb{C}^{m \times n}$.

Definition

- Matrix-vector product $\mathbf{b} = \mathbf{A}\mathbf{x}$

$$b_i = \sum_{j=1}^{n} a_{ij} x_j$$

- All entries belong to $\mathbb{C}$, the field of complex numbers. The space of $m-$vectors is $\mathbb{C}^m$, and the space of $m \times n$ matrices is $\mathbb{C}^{m \times n}$.
- The map $\mathbf{x} \to \mathbf{A}\mathbf{x}$ is linear, which means for any $\mathbf{x}, \mathbf{y} \in \mathbb{C}^n$ and any $\alpha \in \mathbb{C}$

$$\mathbf{A}(\mathbf{x} + \mathbf{y}) = \mathbf{A}\mathbf{x} + \mathbf{A}\mathbf{y}$$

$$\mathbf{A}(\alpha \mathbf{x}) = \alpha \mathbf{A}\mathbf{x}$$

Definition

- Matrix-vector product $\mathbf{b} = \mathbf{A}\mathbf{x}$

$$b_i = \sum_{j=1}^{n} a_{ij} x_j$$

- All entries belong to $\mathbb{C}$, the field of complex numbers. The space of $m-$vectors is $\mathbb{C}^m$, and the space of $m \times n$ matrices is $\mathbb{C}^{m \times n}$.
- The map $\mathbf{x} \to \mathbf{A}\mathbf{x}$ is linear, which means for any $\mathbf{x}, \mathbf{y} \in \mathbb{C}^n$ and any $\alpha \in \mathbb{C}$

$$\mathbf{A}(\mathbf{x} + \mathbf{y}) = \mathbf{A}\mathbf{x} + \mathbf{A}\mathbf{y}$$

$$\mathbf{A}(\alpha \mathbf{x}) = \alpha \mathbf{A}\mathbf{x}$$

- In lecture notes, I use **boldface UPPERCASE for matrices, and boldface lowercase letters for vectors.**

Pseudo code for Matrix-Vector Product

This should be straightforward to convert into real computer codes in any programming language.

Pseudo-code for $\mathbf{b} = \mathbf{Ax}$

for $i = 1$ **to** m **do**
    $b(i) = 0$;
    for $j = 1$ **to** n **do**
        $b(i) = b(i) + A(i,j) * x(j)$;
    end for
end for

Linear Combination

- Alternatively, matrix-vector product can be viewed as

$$\mathbf{b} = \mathbf{A}\mathbf{x} = \sum_{j=i}^{n} x_j \mathbf{a}_j$$

i.e **b** is a linear combination of column vectors of **A**

Linear Combination

- Alternatively, matrix-vector product can be viewed as

$$\mathbf{b} = \mathbf{A}\mathbf{x} = \sum_{j=i}^{n} x_j \mathbf{a}_j$$

  i.e $\mathbf{b}$ is a linear combination of column vectors of $\mathbf{A}$

- Two different views of matrix-vector products:
  1. $b_i = \sum_{j=1}^{n} a_{ij} x_j$: $\mathbf{A}$ acts on $\mathbf{x}$ to produce $\mathbf{b}$; scalar operations
  2. $\mathbf{b} = \sum_{j=i}^{n} x_j \mathbf{a}_j$: $\mathbf{x}$ acts on $\mathbf{A}$ to produce $\mathbf{b}$; vector operations

Linear Combination

- Alternatively, matrix-vector product can be viewed as

$$\mathbf{b} = \mathbf{A}\mathbf{x} = \sum_{j=i}^{n} x_j \mathbf{a}_j$$

  i.e $\mathbf{b}$ is a linear combination of column vectors of $\mathbf{A}$

- Two different views of matrix-vector products:
  1. $b_i = \sum_{j=1}^{n} a_{ij} x_j$: $\mathbf{A}$ acts on $\mathbf{x}$ to produce $\mathbf{b}$; scalar operations
  2. $\mathbf{b} = \sum_{j=i}^{n} x_j \mathbf{a}_j$: $\mathbf{x}$ acts on $\mathbf{A}$ to produce $\mathbf{b}$; vector operations
- If $\mathbf{A}$ is $m \times n$, $\mathbf{Ax}$ can be viewed as a mapping from $\mathbb{C}^n$ to $\mathbb{C}^m$

Takeaways

- Space travel: Matrix Vector multiplication takes you from one 'space' to another.
- Two ways to do the same thing
  - We like the product expressed as linear combinations
  - Let's convince ourselves: Algorithmic investigation
- Getting started with Python and numpy package/library.

Matrix-Matrix Multiplication

- If **A** is $l \times m$ and **C** is $m \times n$, then $\mathbf{B} = \mathbf{AC}$ is $l \times n$, with entries defined by

$$b_{ij} = \sum_{k=1}^{m} a_{ik} c_{kj}.$$

- Written in columns, we have

$$\mathbf{b}_j = \mathbf{A}\mathbf{c}_j = \sum_{k=1}^{m} c_{kj} \mathbf{a}_k.$$

- In other words, each column of **B** is a linear combination of the columns of **A**.

Pseudo-Code for Matrix-Matrix Multiplication

Pseudo-code for $\mathbf{B} = \mathbf{AC}$

for $i = 1$ **to** l **do**
  for $j = 1$ **to** n **do**
    $B(i,j) = 0$;
    for $k = 1$ **to** m **do**
      $B(i,j) = B(i,j) + A(i,k) * C(k,j)$;
    end for
  end for
end for

Pseudo-Code for Matrix-Matrix Multiplication

Pseudo-code for $\mathbf{B} = \mathbf{AC}$

```
for i = 1 to l do
    for j = 1 to n do
        B(i,j) = 0;
        for k = 1 to m do
            B(i,j) = B(i,j) + A(i,k) * C(k,j);
        end for
    end for
end for
```

TODO: Write the pseudo code where **B** is expressed as a linear combination of columns of **A**

Rank-1 Matrices

- Full-rank matrices are important
- Another interesting space case is rank-1 matrices
- A matrix $\mathbf{A}$ is rank-1 if it can be written as $\mathbf{A} = \mathbf{u}\mathbf{v}^*$ where $\mathbf{u}$ and $\mathbf{v}$ are non zero vectors
- $\mathbf{u}\mathbf{v}^*$ is called the outer product of the two vectors, as opposed to the inner product $\mathbf{u}^*\mathbf{v}$

Perspective: Vector Space

A useful way in understanding matrix operations is to think in terms of vector spaces

- Vector space spanned by a set of vectors is composed of linear combinations of these vectors
  - It is closed under addition and scalar multiplication
  - $\mathbf{0}$ is always a member of a subspace
  - Space spanned by m-vectors is subspace of $\mathbb{C}^m$
- If S_1 and S_2 are two subspaces, then $S_1 \cap S_2$ is a subspace, so is $S_1 + S_2$, the space of sum of vectors from S_1 and S_2 .
  - Note that $S_1 + S_2$ is different from $S_1 \cup S_2$
- Two subspaces S_1 and S_2 of $\mathbb{C}^m$ are complementary subspaces of each other if $S_1 + S_2 = \mathbb{C}^m$ and $S_1 \cap S_2 = \{\mathbf{0}\}$.
  - In other words, $\dim(S_1) + \dim(S_2) = m$ and $S_1 \cap S_2 = \{\mathbf{0}\}$

Range, Rank and Inverses

Range

Definition

The range of a matrix $\mathbf{A}$, written as $\text{range}(\mathbf{A})$, is the set of vectors that can be expressed as $\mathbf{Ax}$ for some $\mathbf{x}$.

Range

Definition

The range of a matrix $\mathbf{A}$, written as range($\mathbf{A}$), is the set of vectors that can be expressed as $\mathbf{Ax}$ for some $\mathbf{x}$.

Theorem

range($\mathbf{A}$) is the space spanned by the columns of $\mathbf{A}$.

Therefore, the range of $\mathbf{A}$ is also called the column space of $\mathbf{A}$.

Range and Null Space

Definition

The null space of $\mathbf{A} \in \mathbb{C}^{m \times n}$, written as null($\mathbf{A}$), is the set of vectors $\mathbf{x}$ that satisfy $\mathbf{Ax} = \mathbf{0}$.

Entries of $\mathbf{x} \in \mathbf{null}(\mathbf{A})$ give coefficient of $\sum x_i \mathbf{a_i} = \mathbf{0}$. Note: The null space of $\mathbf{A}$ is in general **not** a complimenary subspace of range($\mathbf{A}$).

Rank

Definition

The column rank of a matrix is the dimension of its column space. The row rank is the dimension of the space spanned by its rows.

- Question: Can the column rank and row rank be different?

Rank

Definition

The column rank of a matrix is the dimension of its column space. The row rank is the dimension of the space spanned by its rows.

- Question: Can the column rank and row rank be different?
  - Answer: No
  - Simply referred as the rank of a matrix.

Rank

Definition

The column rank of a matrix is the dimension of its column space. The row rank is the dimension of the space spanned by its rows.

- Question: Can the column rank and row rank be different?
  - Answer: No
  - Simply referred as the rank of a matrix.
- Question: Given $\mathbf{A} \in \mathbb{C}^{m \times n}$, what is $\dim(\text{null}(\mathbf{A})) + \text{rank}(\mathbf{A})$ equal to?

Rank

Definition

The column rank of a matrix is the dimension of its column space. The row rank is the dimension of the space spanned by its rows.

- Question: Can the column rank and row rank be different?
  - Answer: No
  - Simply referred as the rank of a matrix.
- Question: Given $\mathbf{A} \in \mathbb{C}^{m \times n}$, what is $\dim(\text{null}(\mathbf{A})) + \text{rank}(\mathbf{A})$ equal to?
  - Answer: n

Full Rank

Definition

A matrix has *full* rank if it has the maximum possible rank, i.e., $\min\{m, n\}$

Full Rank

Definition

A matrix has *full* rank if it has the maximum possible rank, i.e., $\min\{m, n\}$

Theorem

A matrix $\mathbf{A} \in \mathbb{C}^{m \times n}$ with $m \geq n$ has full rank if and only if it maps no two distinct vectors to the same vector.

In other words, the linear mapping defined by $\mathbf{Ax}$ for $\mathbf{x} \in \mathbb{C}^n$ is one-to-one

Proof

($\Rightarrow$) Column vectors of $\mathbf{A}$ forms a basis of range($\mathbf{A}$), so every $\mathbf{b} \in$ range($\mathbf{A}$) has a unique linear expansion in terms of the columns of $\mathbf{A}$. ($\Leftarrow$) If $\mathbf{A}$ does not have full rank, then its column vectors are linear dependent, so its vectors do not have a unique linear combination

Inverse

Definition

A nonsingular or invertible matrix is a square matrix of full rank.

Inverse

Definition

A nonsingular or invertible matrix is a square matrix of full rank.

Definition

Given a nonsingular matrix $\mathbf{A}$, its inverse is written as $\mathbf{A}^{-1}$, and $\mathbf{A}\mathbf{A}^{-1} = \mathbf{A}^{-1}\mathbf{A} = \mathbf{I}$

- Note that $(\mathbf{AB})^{-1} = (\mathbf{B}^{-1}\mathbf{A})^{-1}$
- $(\mathbf{A}^{-1})^{*} = (\mathbf{A}^{*})^{-1}$, and we use $\mathbf{A}^{-*}$ as a shorthand for it

Inverse

Theorem

For $\mathbf{A} \in \mathbb{C}^{m \times m}$, the following conditions are equivalent:
(a) $\mathbf{A}$ has an inverse $\mathbf{A}^{-1}$
(b) $rank(\mathbf{A})$ is m
(c) $range(\mathbf{A})$ is $\mathbb{C}^m$
(d) $null(\mathbf{A})$ is $\{0\}$
(e) 0 is not an eigenvalue of $\mathbf{A}$
(f) 0 is not a singular value of $\mathbf{A}$
(g) $det(\mathbf{A}) \neq 0$

Matrix Inverse Times a Vector

- When writing $\mathbf{x} = \mathbf{A}^{-1}\mathbf{b}$, it means $\mathbf{x}$ is the solution of $\mathbf{Ax} = \mathbf{b}$
- In other words, $\mathbf{A}^{-1}\mathbf{b}$ is a vector of coefficients of the expansion of $\mathbf{b}$ in the basis of columns of $\mathbf{A}$
- Multiplying $\mathbf{b}$ by $\mathbf{A}^{-1}$ is a change of basis operations from $\{\mathbf{a_1}, \mathbf{a_2}, \ldots, \mathbf{a_m}\}$ to $\{\mathbf{e_1}, \mathbf{e_2}, \ldots, \mathbf{e_m}\}$
- Multiplying $\mathbf{A}^{-1}\mathbf{b}$ by $\mathbf{A}^{-1}$ is a change of basis operations from $\{\mathbf{e_1}, \mathbf{e_2}, \ldots, \mathbf{e_m}\}$ to $\{\mathbf{a_1}, \mathbf{a_2}, \ldots, \mathbf{a_m}\}$

Inner products

Transpose and Adjoint

- Transpose of $\mathbf{A}$, denoted by $\mathbf{A^T}$, is the matrix $\mathbf{B}$ with $b_{ij} = a_{ji}$
- *Adjoint* or *Hermitian conjugate*, denoted by $\mathbf{A}^*$ or $\mathbf{A}^H$, is the matrix $\mathbf{B}$ with $b_{ij} = \bar{a}_{ji}$
- Note that, $(\mathbf{AB})^T = \mathbf{B^T A^T}$ and $(\mathbf{AB})^* = \mathbf{B}^* \mathbf{A}^*$
- A matrix $\mathbf{A}$ is symmetric if $\mathbf{A} = \mathbf{A^T}$ (i.e., $a_{ij} = a_{ji}$). It is *Hermitian* if $\mathbf{A} = \mathbf{A^T}$ (i.e., $a_{ij} = \bar{a}_{ji}$)
- For $\mathbf{A} \in \mathbb{R}^{m \times n}$, null($\mathbf{A}$) and range($\mathbf{A^T}$) are complementary subspaces. In addition, null($\mathbf{A}$) and range($\mathbf{A^T}$) are orthogonal to each other (to be explained later)
- For $\mathbf{A} \in \mathbb{C}^{m \times n}$, null($\mathbf{A}$) and range($\mathbf{A}^*$) are complementary subspaces

Inner Product

- Inner product (dot product) of two column vectors $\mathbf{u}$, $\mathbf{v} \in \mathbb{C}$ is $\mathbf{u}^*\mathbf{v}$
- In contrast, *outer* product of $\mathbf{u}$ and $\mathbf{v}$ is $\mathbf{u}\mathbf{v}^*$
- Note that *cross* product is different

Inner Product

- Inner product (dot product) of two column vectors $\mathbf{u}, \mathbf{v} \in \mathbb{C}$ is $\mathbf{u}^*\mathbf{v}$
- In contrast, *outer* product of $\mathbf{u}$ and $\mathbf{v}$ is $\mathbf{u}\mathbf{v}^*$
- Note that *cross* product is different

Different ways to see the inner product

1. Vector-vector multiplication: $\mathbf{u}^*\mathbf{v} = \sum_{i=1}^{m} \bar{u}_i v_i$
2. Euclidean length of $\mathbf{u}$ is the square root of the inner product of $\mathbf{u}$ with itself, i.e., $\sqrt{\mathbf{u}^*\mathbf{u}}$
3. Inner product of two unit vectors $\mathbf{u}$ and $\mathbf{v}$ is the cosine of the angle α between $\mathbf{u}$ and $\mathbf{v}$, i.e., $\cos \alpha = \frac{u^* v}{||u||||v||}$

Inner Product is bilinear

Inner product is *bilinear*, in the sense that it is linear in each vertex separately:

- $(u_1 + u_2)^* v = u_1^* v + u_2^* v$
- $u^*(v_1 + v_2) = u^* v_1 + u^* v_2$
- $(\alpha u)^*(\beta v) = \bar{\alpha}\beta u^* v$

Orthogonal Vectors

Definition

A pair of vectors are *orthogonal* if $\mathbf{x}^*\mathbf{y} = 0$.

In other words, the angle between them is 90 degrees

Orthogonal Vectors

Definition

A pair of vectors are *orthogonal* if $\mathbf{x}^*\mathbf{y} = 0$.

In other words, the angle between them is 90 degrees

Definition

Two sets of vectors X and Y are orthogonal if every $\mathbf{x} \in X$ is orthogonal to every $\mathbf{y} \in Y$.

Orthogonal Vectors

Definition

A pair of vectors are *orthogonal* if $\mathbf{x}^*\mathbf{y} = 0$.

In other words, the angle between them is 90 degrees

Definition

Two sets of vectors X and Y are orthogonal if every $\mathbf{x} \in X$ is orthogonal to every $\mathbf{y} \in Y$.

Definition

A set of nonzero vectors S is *orthogonal* if they are pairwise orthogonal. They are *orthonormal* if it is orthogonal and in addition each vector has unit Euclidean length.

Orthogonal Vectors

Theorem

The vectors in an orthogonal set S are linearly independent.

Proof

Prove by contradiction. If a vector can be expressed as linear combination of the other vectors in the set, then it is orthogonal to itself

Question: If the column vectors of an $m \times n$ matrix $\mathbf{A}$ are orthogonal, what is the rank of $\mathbf{A}$?

Orthogonal Vectors

Theorem

The vectors in an orthogonal set S are linearly independent.

Proof

Prove by contradiction. If a vector can be expressed as linear combination of the other vectors in the set, then it is orthogonal to itself

Question: If the column vectors of an $m \times n$ matrix $\mathbf{A}$ are orthogonal, what is the rank of $\mathbf{A}$?
Answer: $n = min\{m, n\}$. In other words, $\mathbf{A}$ has full rank

Components of Vector

- Given an orthonormal set $\{\mathbf{q}_1, \mathbf{q}_2, \cdots, \mathbf{q}_m\}$ forming a basis of $\mathbb{C}^m$, vector $\mathbf{v}$ can be decomposed into orthogonal components as $\mathbf{v} = \sum_{i=1}^{m} (\mathbf{q}_i^* \mathbf{v}) \mathbf{q}_i$

Components of Vector

- Given an orthonormal set $\{\mathbf{q}_1, \mathbf{q}_2, \cdots, \mathbf{q}_m\}$ forming a basis of $\mathbb{C}^m$, vector $\mathbf{v}$ can be decomposed into orthogonal components as $\mathbf{v} = \sum_{i=1}^{m} (\mathbf{q}_i^* \mathbf{v}) \mathbf{q}_i$
- Another way to express the condition is $\mathbf{v} = \sum_{i=1}^{m} (\mathbf{q}_i^* \mathbf{q}_i) \mathbf{v}$
- $\mathbf{q}_i \mathbf{q}_i^*$ is an *orthogonal projection matrix*. Note that it is NOT an orthogonal matrix.

Components of Vector

- Given an orthonormal set $\{\mathbf{q}_1, \mathbf{q}_2, \cdots, \mathbf{q}_m\}$ forming a basis of $\mathbb{C}^m$, vector $\mathbf{v}$ can be decomposed into orthogonal components as $\mathbf{v} = \sum_{i=1}^{m} (\mathbf{q}_i^* \mathbf{v}) \mathbf{q}_i$
- Another way to express the condition is $\mathbf{v} = \sum_{i=1}^{m} (\mathbf{q}_i^* \mathbf{q}_i) \mathbf{v}$
- $\mathbf{q}_i \mathbf{q}_i^*$ is an *orthogonal projection matrix*. Note that it is NOT an orthogonal matrix.
- More generally, given an orthonormal set $\{\mathbf{q}_1, \mathbf{q}_2, \cdots, \mathbf{q}_n\}$ with $n \leq m$, we have

$$\mathbf{v} = \mathbf{r} + \sum_{i=1}^{n} (\mathbf{q}_i^* \mathbf{v}) \mathbf{q}_i = \mathbf{r} + \sum_{i=1}^{n} (\mathbf{q}_i \mathbf{q}_i^*) \mathbf{v} \text{ and } \mathbf{r}^* \mathbf{q}_i = 0,\ 1 \leq i \leq n$$

- Let $\mathbf{Q}$ be composed of column vectors $\{\mathbf{q}_1, \mathbf{q}_2, \cdots, \mathbf{q}_n\}$. $\mathbf{Q}\mathbf{Q}^* = \sum_{i=1}^{n} (\mathbf{q}_i \mathbf{q}_i^*)$ is an orthogonal projection matrix.

Unitary matrices

Unitary Matrices

Definition

A matrix is unitary if $\mathbf{Q}^* = \mathbf{Q}^{-1}$,i.e. if $\mathbf{Q}^*\mathbf{Q} = \mathbf{Q}\mathbf{Q}^* = I$

- In the real case, we say the matrix is *orthogonal*. Its column vectors are *orthonormal*.
- In other words, $\mathbf{q}_i^*\mathbf{q}_j = \delta_{ij}$, the *Kronecker delta*

Unitary Matrices

Definition

A matrix is unitary if $\mathbf{Q}^* = \mathbf{Q}^{-1}$,i.e. if $\mathbf{Q}^*\mathbf{Q} = \mathbf{Q}\mathbf{Q}^* = I$

- In the real case, we say the matrix is *orthogonal*. Its column vectors are *orthonormal*.
- In other words, $\mathbf{q}_i^*\mathbf{q}_j = \delta_{ij}$, the *Kronecker delta*

Question: What is the geometric meaning of multiplication by a unitary matrix?

Unitary Matrices

Definition

A matrix is unitary if $\mathbf{Q}^* = \mathbf{Q}^{-1}$,i.e. if $\mathbf{Q}^*\mathbf{Q} = \mathbf{Q}\mathbf{Q}^* = I$

- In the real case, we say the matrix is *orthogonal*. Its column vectors are *orthonormal*.
- In other words, $\mathbf{q}_i^*\mathbf{q}_j = \delta_{ij}$, the *Kronecker delta*

Question: What is the geometric meaning of multiplication by a unitary matrix?
Answer: It preserves angles and Euclidean length. In the real case, multiplication by an orthogonal matrix $\mathbf{Q}$ is a rotation (if $\det(\mathbf{Q}) = 1$) or reflection (if $\det(\mathbf{Q}) = -1$).

Vector Norms

Definition of Norms

- Norm captures "size" of vector or "distance" between vectors
- There are many different measures for "sizes" but a norm must satisfy some requirements:

Definition

A norm is a function $||\cdot||$: $\mathbb{C}^m \to \mathbb{R}$ that assigns a real-valued length to each vector. It must satisfy the following conditions:

1. $\|\mathbf{x}\| \geq 0$, and $\|\mathbf{x}\| = 0$ only if $\mathbf{x} = 0$
2. $\|\mathbf{x}+\mathbf{y}\| \leq \|\mathbf{x}\| + \|\mathbf{y}\|$
3. $\|\alpha\mathbf{x}\| = |\alpha|\,\|\mathbf{x}\|$.

- An example is Euclidean length (i.e.$||\mathbf{x}|| = \sqrt{\sum_{i=1}^{m} |x_i|^2}$)

p-norms

- Euclidean length is a special case of p-norms, defined as

$$\|\mathbf{x}\|_p = \Big(\sum_{i=1}^{m} |x_i|^p\Big)^{\frac{1}{p}}$$

for $1 \leq p \leq \infty$

p-norms

- Euclidean length is a special case of p-norms, defined as

$$\|\mathbf{x}\|_p = (\sum_{i=1}^{m} |x_i|^p)^{\frac{1}{p}}$$

  for $1 \leq p \leq \infty$
- Euclidean norm is 2-norm $\|\mathbf{x}\|_2$ (i.e., $p = 2$)

p-norms

- Euclidean length is a special case of *p*-norms, defined as

$$\|\mathbf{x}\|_p = \left(\sum_{i=1}^{m} |x_i|^p\right)^{\frac{1}{p}}$$

  for $1 \leq p \leq \infty$
- Euclidean norm is 2-norm $\|\mathbf{x}\|_2$ (i.e., $p = 2$)
- 1-norm: $\|\mathbf{x}\|_1 = \sum_{i=1}^{m} |x_i|$

p-norms

- Euclidean length is a special case of p-norms, defined as

$$\|\mathbf{x}\|_p = (\sum_{i=1}^{m} |x_i|^p)^{\frac{1}{p}}$$

  for $1 \leq p \leq \infty$
- Euclidean norm is 2-norm $\|\mathbf{x}\|_2$ (i.e., $p = 2$)
- 1-norm: $\|\mathbf{x}\|_1 = \sum_{i=1}^{m} |x_i|$
- ∞-norm: $\|\mathbf{x}\|_\infty$. What is its value? Answer: $\|\mathbf{x}\||_\infty = \max_{1 \leq i \leq m} |x_i|$
- Why we require $p \geq 1$? What happens if $0 \leq p < 1$?

Weighted p-norms

- A generalization of p-norm is weighted p-norm, which assigns different weights (priorities) to different components.
  - It is anisotropic instead of isotropic
- Algebraically, $\|\mathbf{x}\|_{\mathbf{W}} = \|\mathbf{W}\mathbf{x}\|$, where $\mathbf{W}$ is diagonal matrix with $i-$th diagonal entry $w_i \neq 0$ being weight for ith component
- In other words,

$$\|\mathbf{x}\|_{\mathbf{W}} = (\sum_{i=1}^{m} |w_i x_i|^p)^{\frac{1}{p}}$$

- What happens if we allow $w_i = 0$?

Weighted p-norms

- A generalization of p-norm is weighted p-norm, which assigns different weights (priorities) to different components.
  - It is anisotropic instead of isotropic
- Algebraically, $\|\mathbf{x}\|_{\mathbf{W}} = \|\mathbf{W}\mathbf{x}\|$, where $\mathbf{W}$ is diagonal matrix with $i-$th diagonal entry $w_i \neq 0$ being weight for ith component
- In other words,

$$\|\mathbf{x}\|_{\mathbf{W}} = \Big(\sum_{i=1}^{m} |w_i x_i|^p\Big)^{\frac{1}{p}}$$

- What happens if we allow $w_i = 0$?
- Can we further generalize it to allow $\mathbf{W}$ being arbitrary matrix?
- No. But we can allow $\mathbf{W}$ to be arbitrary nonsingular matrix.

Matrix Norms

Matrix Norms Induced by Vector Norms

- Viewing $m \times n$ matrix as mn-vectors is not always useful, as operations involving $m \times n$ matrices do not behave this way
- *Induced matrix norms* capture such behavior

Definition

Given vector norms $||\cdot||_{(n)}$ and $||\cdot||_{(m)}$ on domain and range of $\mathbf{A} \in \mathbb{C}^{m\times n}$, respectively, the induced matrix norm $||\mathbf{A}||_{(m,n)}$ is the smallest number $\mathbb{C} \in \mathbb{R}$ for which the following inequality holds for all $\mathbf{x} \in \mathbb{C}^n$:

$$\|\mathbf{Ax}\|_{(m)} \leq C\,\|x\|_{(n)}\,.$$

Matrix Norms Induced by Vector Norms

- Viewing $m \times n$ matrix as mn-vectors is not always useful, as operations involving $m \times n$ matrices do not behave this way
- *Induced matrix norms* capture such behavior

Definition

Given vector norms $||\cdot||_{(n)}$ and $||\cdot||_{(m)}$ on domain and range of $\mathbf{A} \in \mathbb{C}^{m\times n}$, respectively, the induced matrix norm $||\mathbf{A}||_{(m,n)}$ is the smallest number $\mathbb{C} \in \mathbb{R}$ for which the following inequality holds for all $\mathbf{x} \in \mathbb{C}^n$:

$$\|\mathbf{Ax}\|_{(m)} \leq C \|x\|_{(n)}.$$

- In other words, it is supremum of ratio $||\mathbf{Ax}||_{(m)}/||\mathbf{x}||_{(n)}$ for all nonzero vectors $\mathbf{x} \in \mathbb{C}^n$
- Maximum factor by which $\mathbf{A}$ can "stretch" $\mathbf{x} \in \mathbb{C}^n$

$$\|\mathbf{A}\|_{(m,n)} = \sup_{\mathbf{x}\in\mathbb{C}^n, \mathbf{x}\neq 0} \|\mathbf{Ax}\|_{(m)} / \|\mathbf{x}\|_{(n)} = \sup_{\mathbf{x}\in\mathbb{C}^n, \|\mathbf{x}\|_{(n)}=1} \|\mathbf{Ax}\|_{(m)}$$

Matrix Norms Induced by Vector Norms

- Viewing $m \times n$ matrix as mn-vectors is not always useful, as operations involving $m \times n$ matrices do not behave this way
- *Induced matrix norms* capture such behavior

Definition

Given vector norms $||\cdot||_{(n)}$ and $||\cdot||_{(m)}$ on domain and range of $\mathbf{A} \in \mathbb{C}^{m\times n}$, respectively, the induced matrix norm $||\mathbf{A}||_{(m,n)}$ is the smallest number $\mathbb{C} \in \mathbb{R}$ for which the following inequality holds for all $\mathbf{x} \in \mathbb{C}^n$:

$$\|\mathbf{Ax}\|_{(m)} \leq C \, \|x\|_{(n)} .$$

- In other words, it is supremum of ratio $||\mathbf{Ax}||_{(m)}/||\mathbf{x}||_{(n)}$ for all nonzero vectors $\mathbf{x} \in \mathbb{C}^n$
- Maximum factor by which $\mathbf{A}$ can "stretch" $\mathbf{x} \in \mathbb{C}^n$

$$\|\mathbf{A}\|_{(m,n)} = \sup_{\mathbf{x}\in\mathbb{C}^n, \mathbf{x}\neq 0} \|\mathbf{Ax}\|_{(m)} / \|\mathbf{x}\|_{(n)} = \sup_{\mathbf{x}\in\mathbb{C}^n, \|\mathbf{x}\|_{(n)}=1} \|\mathbf{Ax}\|_{(m)}$$

- Is vector norm consistent with matrix norm of $m \times 1$-matrix?

1-norm

- By definition

$$||\mathbf{A}||_1 = \sup_{\mathbf{x}\in\mathbb{C}^n, ||\mathbf{x}||_1=1} ||\mathbf{A}\mathbf{x}||_1$$

1-norm

- By definition

$$||\mathbf{A}||_1 = \sup_{\mathbf{x}\in\mathbb{C}^n, ||\mathbf{x}||_1=1} ||\mathbf{Ax}||_1$$

- What is it equal to?

1-norm

- By definition

$$||\mathbf{A}||_1 = \sup_{\mathbf{x}\in\mathbb{C}^n, ||\mathbf{x}||_1=1} ||\mathbf{Ax}||_1$$

- What is it equal to?
  - Maximum of 1-norm of column vectors of **A**
  - "maximum column sum" of **A** is oversimplified in the textbook
- To show it, note that for $\mathbf{x} \in \mathbb{C}^n$ and $||\mathbf{x}||_1 = 1$

$$||\mathbf{Ax}||_1 = ||\sum_{j=1}^{n} x_j \mathbf{a}_j||_1 \leq \max_{1\leq j\leq n} ||a_j||_1 ||\mathbf{x}||_1$$

- Let $k = \arg\max_{1\leq j\leq n} ||\mathbf{a}_j||_1$, then $||\mathbf{A}e_k||_1 = ||\mathbf{a}_k||_1$, so $\max_{1\leq j\leq n} ||a_j||_1$ is tight upper bound

$\infty-$norm

- By definition
$$||\mathbf{A}||_\infty = \sup_{\mathbf{x}\in\mathbb{C}^n, ||\mathbf{x}||_\infty=1} ||\mathbf{A}\mathbf{x}||_\infty$$
- What is $||\mathbf{A}||_\infty$ equal to?

$\infty-$norm

- By definition

$$||\mathbf{A}||_\infty = \sup_{\mathbf{x}\in\mathbb{C}^n, ||\mathbf{x}||_\infty=1} ||\mathbf{Ax}||_\infty$$

- What is $||\mathbf{A}||_\infty$ equal to?
  - Maximum of 1-norm of column vectors of $\mathbf{A}^\mathbf{T}$
- To show it, note that for$\mathbf{x} \in \mathbb{C}^n$ and $||\mathbf{x}||_\infty = 1$

$$||\mathbf{Ax}||_\infty = \max_{1\leq i\leq m}|\mathbf{a}_i^* x| \leq \max_{1\leq i\leq m}||\mathbf{a}_i^*||_1||\mathbf{x}||_\infty$$

where $\mathbf{a}_i^*$ denotes the $i-$th row vector of $\mathbf{A}$

- Furthermore, $\max_{1\leq i\leq m}||\mathbf{a}_i^*||_1$ is a tight bound.
  - Which vector can we choose to reach the bound?

2-norm

- What is 2-norm of a matrix?

2-norm

- What is 2-norm of a matrix?
- Answer: Its largest singular value.
- We will talk more about singular-value decomposition

2-norm

- What is 2-norm of a matrix?
- Answer: Its largest singular value.
- We will talk more about singular-value decomposition
- What is 2-norm of a diagonal matrix?

Cauchy-Schwarz and Holder Inequalities

- Holder inequality: Let p and q satisfy $1/p + 1/q = 1$ with $1 \leq p,\ q \leq \infty$, then

$$|\mathbf{x}^* \mathbf{y}| \leq ||\mathbf{x}||_p ||\mathbf{y}||_q$$

- Cauchy-Schwarz inequality

$$|\mathbf{x}^* \mathbf{y}| \leq ||\mathbf{x}||_2 ||\mathbf{y}||_2$$

- Cauchy-Schwarz inequality is a special case of Holder inequality

Cauchy-Schwarz and Holder Inequalities

- Holder inequality: Let p and q satisfy $1/p + 1/q = 1$ with $1 \leq p,\ q \leq \infty$, then

$$|\mathbf{x}^* \mathbf{y}| \leq ||\mathbf{x}||_p ||\mathbf{y}||_q$$

- Cauchy-Schwarz inequality

$$|\mathbf{x}^* \mathbf{y}| \leq ||\mathbf{x}||_2 ||\mathbf{y}||_2$$

- Cauchy-Schwarz inequality is a special case of Holder inequality
- Example: What is 2-norm of rank-one matrix? Hint: Use Cauchy-Schwarz inequality.

Bounding Matrix-Matrix Multiplication

- Let **A** be an $l \times m$ matrix and **B** an $m \times n$ matrix, then for $\mathbf{x} \in \mathbb{C}^n$

$$||\mathbf{AB}||_{(l,n)} \leq ||\mathbf{A}||_{(l,m)}||\mathbf{B}||_{(m,n)}$$

- To show it, note

$$||\mathbf{ABx}||_{(l)} \leq ||\mathbf{A}||_{(l,m)}||\mathbf{Bx}||_{(m)} \leq ||\mathbf{A}||_{(l,m)}||\mathbf{B}||_{(m,n)}||\mathbf{x}||_{(n)},$$

Bounding Matrix-Matrix Multiplication

- Let $\mathbf{A}$ be an $l \times m$ matrix and $\mathbf{B}$ an $m \times n$ matrix, then for $\mathbf{x} \in \mathbb{C}^n$

$$||\mathbf{AB}||_{(l,n)} \leq ||\mathbf{A}||_{(l,m)}||\mathbf{B}||_{(m,n)}$$

- To show it, note

$$||\mathbf{ABx}||_{(l)} \leq ||\mathbf{A}||_{(l,m)}||\mathbf{Bx}||_{(m)} \leq ||\mathbf{A}||_{(l,m)}||\mathbf{B}||_{(m,n)}||\mathbf{x}||_{(n)},$$

- In general, this inequality is not an equality
- In particular, $||\mathbf{A}^n|| \leq ||\mathbf{A}||^n$ but $||\mathbf{A}^n|| \neq ||\mathbf{A}||^n$ in general for $n \geq 2$

General Matrix Norms

One can view $m \times n$ matrices as mn-dimensional vectors and obtain *general matrix norms*, which satisfy (for $\mathbf{A}$, $\mathbf{B} \in \mathbb{C}^{m \times n}$)

1. $||\mathbf{A}|| \geq 0$, and $||\mathbf{A}|| = 0$ only if $\mathbf{A} = 0$
2. $||\mathbf{A} + \mathbf{B}|| \leq ||\mathbf{A}|| + ||\mathbf{B}||$
3. $||\alpha \mathbf{A}|| = |\alpha| ||\mathbf{A}||$

Frobenius Norm

- One useful norm is Frobenius norm (a.k.a. Hilbert-Schmidt norm)

$$\|\mathbf{A}\|_F = \sqrt{\sum_{i=1}^{m}\sum_{j=1}^{n}|a_{ij}|^2} = \sqrt{\sum_{j=1}^{n}\|\mathbf{a}_j\|_2^2}$$

  i.e., 2-norm of mn-vector

- Furthermore,

$$\|\mathbf{A}\|_F = \sqrt{tr(\mathbf{A}^T\mathbf{A})}$$

  where $tr(\mathbf{B})$ denotes trace of B, the sum of its diagonal entries

Frobenius Norm

- One useful norm is Frobenius norm (a.k.a. Hilbert-Schmidt norm)

$$\|\mathbf{A}\|_F = \sqrt{\sum_{i=1}^{m}\sum_{j=1}^{n}|a_{ij}|^2} = \sqrt{\sum_{j=1}^{n}\|\mathbf{a}_j\|_2^2}$$

  i.e., 2-norm of mn-vector
- Furthermore,

$$\|\mathbf{A}\|_F = \sqrt{tr(\mathbf{A}^\mathsf{T}\mathbf{A})}$$

  where $tr(\mathbf{B})$ denotes trace of B, the sum of its diagonal entries
- Note that

$$\|\mathbf{AB}\|_F \leq \|\mathbf{A}\|_F \|\mathbf{B}\|_F$$

  because

$$\|\mathbf{AB}\|_F^2 = \sum_{i=1}^{n}\sum_{j=1}^{m}|\mathbf{a}_i^*\mathbf{b}_j|^2 \leq \sum_{i=1}^{n}\sum_{j=1}^{m}(\|\mathbf{a}_i^*\|_2\|\mathbf{b}_j\|_2)^2 = \|\mathbf{A}\|_F^2\|\mathbf{B}\|_F^2$$

Invariance under Unitary Multiplication

Theorem

For any $\mathbf{A} \in \mathbb{C}^{m \times n}$ and unitary $\mathbf{Q} \in \mathbb{C}^{m \times m}$, we have

$$\|\mathbf{QA}\|_2 = \|\mathbf{A}\|_2 \text{ and } \|\mathbf{QA}\|_F = \|A\|_F$$

In other words, 2-norm and Frobenius norms are invariant under unitary multiplication. Proof for 2-norm: $\|\mathbf{Qy}\|_2 = \|\mathbf{y}\|_2$ for $\mathbf{y} \in \mathbb{C}^m$ and therefore $\|\mathbf{QAx}\|_2 = \|\mathbf{Ax}\|_2$ for $\mathbf{x} \in \mathbb{C}^n$. It then follows from definition of 2-norm.

SVD

Geometric Observation

- The image of unit sphere under any $m \times n$ matrix is a *hyperellipse*
- Give a unit sphere $\mathbf{S}$ in $\mathbb{R}^n$, let $\mathbf{AS}$ denote the shape after transformation
- SVD is

$$\mathbf{A} = \mathbf{U\Sigma V}^*$$

  where $\mathbf{U} \in \mathbb{C}^{m \times m}$ and $\mathbf{V} \in \mathbb{C}^{n \times n}$ is unitary and $\mathbf{\Sigma} \in \mathbb{R}^{m \times n}$ is diagonal
- Singular values are diagonal entries of $\mathbf{\Sigma}$, correspond to the principal semiaxes, with entries $\sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_n \geq 0$.
- Left singular vectors of $\mathbf{A}$ are column vectors of $\mathbf{U}$ and are oriented in the directions of the principal semiaxes of $\mathbf{AS}$
- Right singular vectors of $\mathbf{A}$ are column vectors of $\mathbf{V}$ and are the preimages of the principal semiaxes of $\mathbf{AS}$
- $\mathbf{A}\mathbf{v}_j = \sigma_j \mathbf{u}_j$ for $1 \leq j \leq n$

Two Different Types of SVD

- Full SVD: $\mathbf{U} \in \mathbb{C}^{m \times m}$, $\mathbf{\Sigma} \in \mathbb{R}^{m \times n}$, and $\mathbf{V} \in \mathbb{C}^{n \times n}$ is

$$\mathbf{A} = \mathbf{U\Sigma V}^*$$

- Reduced SVD: $\hat{\mathbf{U}} \in \mathbb{C}^{m \times n}$, $\hat{\mathbf{\Sigma}} \in \mathbb{R}^{n \times n}$ (assume $m \geq n$)

$$\mathbf{A} = \hat{\mathbf{U}}\hat{\mathbf{\Sigma}}\mathbf{V}^*$$

- Furthermore, notice that

$$\mathbf{A} = \sum_{i=1}^{min\{m,n\}} \sigma_i \mathbf{u}_i \mathbf{v}_i^*$$

so we can keep only entries of $\mathbf{U}$ and $\mathbf{V}$ corresponding to nonzero σ_i .

Existence of SVD

Theorem

(Existence) Every matrix $\mathbf{A} \in \mathbb{C}^{m\times n}$ has an SVD

Proof: Let $\sigma = \|\mathbf{A}\|_2$. There exists $\mathbf{v}_1 \in \mathbb{C}^n$ with $\|\mathbf{v}_1\|_2 = 1$ and $\|\mathbf{A}\mathbf{v}_1\|_2 = \sigma_1$. Let $\mathbf{U}_1$ and $\mathbf{V}_1$ be unitary matrices whose first columns are $\mathbf{u}_1 = \frac{\mathbf{A}\mathbf{v}_1}{\sigma_1}$ (or any unit-length vector if $\sigma_1 = 0$) and $\mathbf{v}_1$, respectively. Note that

$$\mathbf{U}_1^*\mathbf{A}\mathbf{V}_1 = \mathbf{S} = \begin{bmatrix} \sigma_1 & \omega^* \\ \mathbf{0} & \mathbf{B} \end{bmatrix}$$

Furthermore, $\omega = 0$ because $\|\mathbf{S}\|_2 = \sigma_1$, and

$$\left\| \begin{bmatrix} \sigma_1 & \omega^* \\ \mathbf{0} & \mathbf{B} \end{bmatrix} \begin{bmatrix} \sigma_1 \\ \omega \end{bmatrix} \right\| \geq \sigma_1^2 + \omega^*\omega = \sqrt{\sigma_1^2 + \omega^*\omega} \left\| \begin{bmatrix} \sigma_1 \\ \omega \end{bmatrix} \right\|_2,$$

implying that $\omega_1 \geq \sqrt{\sigma_1^2 + \omega^*\omega}$ and $\omega = 0$

Existence of SVD Cont'd

We then prove by induction using (1). If $m = 1$ or $n = 1$, then $\mathbf{B}$ is empty and we have $\mathbf{A} = \mathbf{U}_1 \mathbf{S} \mathbf{V}_1^*$. Otherwise, suppose $\mathbf{B} = \mathbf{U}_2 \boldsymbol{\Sigma}_2 \mathbf{V}_2^*$, and then

$$\mathbf{A} = \underbrace{\mathbf{U}_1 \begin{bmatrix} 1 & \mathbf{0}^* \\ \mathbf{0} & \mathbf{U}_2 \end{bmatrix}}_{\mathbf{U}} \underbrace{\begin{bmatrix} \sigma_1 & \mathbf{0}^* \\ \mathbf{0} & \boldsymbol{\Sigma}_2 \end{bmatrix}}_{\boldsymbol{\Sigma}}, \underbrace{\begin{bmatrix} 1 & \mathbf{0}^* \\ \mathbf{0} & \mathbf{V}_2^* \end{bmatrix} \mathbf{V}_1^*}_{\mathbf{V}^*}$$

where $\mathbf{U}$ and $\mathbf{V}$ are unitary.

Uniqueness of SVD

Theorem

(Uniqueness) The singular values $\{\sigma_j\}$ are uniquely determined. If $\mathbf{A}$ is square and the σ_j are distinct, the left and right singular vectors are uniquely determined **up to complex signs** (i.e., complex scalar factors of absolute value 1).

Geometric argument: If the lengths of semiaxes of a hyperellipse are distinct, then the semiaxes themselves are determined by the geometry up to signs.

Uniqueness of SVD Cont'd

Algebraic argument: Based on 2-norm and prove by induction. Consider the case where the σ_j are distinct. The 2-norm is unique, so is σ_1 . If $\mathbf{v}_1$ is not unique up to sign, then the orthonormal bases of these vectors are right singular vectors of $\mathbf{A}$, implying that σ_1 is not a simple singular value.

Once σ_1 , $\mathbf{u}_1$, and $\mathbf{v}_1$ are determined, the remainder of SVD is determined by the space orthogonal to $\mathbf{v}_1$. Because $\mathbf{v}_1$ is unique up to sign, the orthogonal subspace is uniquely defined. Then prove by induction.

Uniqueness of SVD Cont'd

Algebraic argument: Based on 2-norm and prove by induction. Consider the case where the σ_j are distinct. The 2-norm is unique, so is σ_1 . If $\mathbf{v}_1$ is not unique up to sign, then the orthonormal bases of these vectors are right singular vectors of $\mathbf{A}$, implying that σ_1 is not a simple singular value.

Once σ_1 , $\mathbf{u}_1$, and $\mathbf{v}_1$ are determined, the remainder of SVD is determined by the space orthogonal to $\mathbf{v}_1$. Because $\mathbf{v}_1$ is unique up to sign, the orthogonal subspace is uniquely defined. Then prove by induction.

- Question: What if we change the sign of a singular vector?

Uniqueness of SVD Cont'd

Algebraic argument: Based on 2-norm and prove by induction. Consider the case where the σ_j are distinct. The 2-norm is unique, so is σ_1 . If $\mathbf{v}_1$ is not unique up to sign, then the orthonormal bases of these vectors are right singular vectors of $\mathbf{A}$, implying that σ_1 is not a simple singular value.

Once σ_1 , $\mathbf{u}_1$, and $\mathbf{v}_1$ are determined, the remainder of SVD is determined by the space orthogonal to $\mathbf{v}_1$. Because $\mathbf{v}_1$ is unique up to sign, the orthogonal subspace is uniquely defined. Then prove by induction.

- Question: What if we change the sign of a singular vector?
- Question: What if σ_i is not distinct?

SVD vs Eigenvalue Decomposition

- Eigenvalue decomposition of nondefective matrix $\mathbf{A}$ is $\mathbf{A} = \mathbf{X}\mathbf{\Lambda}\mathbf{X}^{-1}$

Differences

- Not every matrix has eigenvalue decomposition, but every matrix has singular value decomposition
- Eigenvalues may not always be real numbers, but singular values are always non-negative real numbers
- Eigenvectors are not always orthogonal to each other (orthogonal for symmetric matrices), but left (or right) singular vectors are orthogonal to each other

SVD vs Eigenvalue Decomposition

Similarities

- Singular values of $\mathbf{A}$ are square roots of eigenvalues of $\mathbf{A}\mathbf{A}^*$ and $\mathbf{A}^*\mathbf{A}$, and their eigenvectors are left and right singular vectors, respectively
- Singular values of hermitian matrices are absolute values of eigenvalues, and eigenvectors are singular vectors (up to complex signs)
- This relationship can be used to compute singular values by hand

Matrix Properties via SVD

- Let r be number of nonzero singular values of $\mathbf{A} \in \mathbb{C}^{m \times n}$
  - rank($\mathbf{A}$) is r
  - range($\mathbf{A}$) $= < \mathbf{u}_1, \mathbf{u}_2, ..., \mathbf{u}_r >$
  - null($\mathbf{A}$) $= < \mathbf{u}_{r+1}, \mathbf{u}_{r+2}, ..., \mathbf{u}_n >$
- 2-norm and Frobenius norm
  - $\|\mathbf{A}\|_2 = \sigma_1$ and $\|\mathbf{A}\|_F = \sqrt{\sum_i \sigma_i^2}$
- Determinant of matrix
  - For $\mathbf{A} \in \mathbb{C}^{m \times m}$, $|det(\mathbf{A})| = \Pi_{i=1}^{m} \sigma_i$

Matrix Properties via SVD

- Let r be number of nonzero singular values of $\mathbf{A} \in \mathbb{C}^{m\times n}$
  - rank($\mathbf{A}$) is r
  - range($\mathbf{A}$) $= < \mathbf{u}_1, \mathbf{u}_2, ..., \mathbf{u}_r >$
  - null($\mathbf{A}$) $= < \mathbf{u}_{r+1}, \mathbf{u}_{r+2}, ..., \mathbf{u}_n >$
- 2-norm and Frobenius norm
  - $\|\mathbf{A}\|_2 = \sigma_1$ and $\|\mathbf{A}\|_F = \sqrt{\sum_i \sigma_i^2}$
- Determinant of matrix
  - For $\mathbf{A} \in \mathbb{C}^{m\times m}$, $|det(\mathbf{A})| = \Pi_{i=1}^m \sigma_i$
- However, SVD may not be the most efficient way in solving problems
- Algorithms for SVD are similar to those for eigenvalue decomposition and we will discuss them later in the semester