Hinatori8

2025 年 12 月 29 日

1. Fundamentals

1.1 Orthogonal Vectors and Matrices

Adjoint

hermitian conjugate もしくは adjoint はエルミート転置もしくは共役転置を意味する。

$$A = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \\ a_{31} & a_{32} \end{bmatrix} \qquad \Rightarrow \qquad A^* = \begin{bmatrix} \overline{a}_{11} & \overline{a}_{21} & \overline{a}_{31} \\ \overline{a}_{12} & \overline{a}_{22} & \overline{a}_{32} \end{bmatrix}$$

もし、$A = A^*$ であるなら行列 A は hermitian matrix（エルミート行列）であると言う。行列の要素が複素数であった場合は、転置と同時に共役な要素に変わることを意味するが、実数のみの行列であるならば、単純に行列が転置しただけと変わらない。実対称行列もエルミート行列の仲間である。今まで使ってきた A^T は実行列の中では transpose となるが、複素数まで拡大するとエルミート転置を意味していると考えられる。つまり、実行列に対する転置を意味する記号として、hermitian conjugate を意味する $*$ を使うことができる。

Inner Product

$x, y \in \mathbb{C}^m$ に対して、内積は

$$x^* y = \sum_{i=1}^{m} \overline{x}_i y_i$$

と定義される。今までは実数ベクトルに対しての内積を考えていたため、転置するだけであったが、複素数まで対応させた内積はエルミート転置を使う必要がある。前述の通り、実数であったために共役を考える必要がなかった。

Orthogonal Vectors

- A set of nonzero vectors S is **orthogonal**(直交) if its elements are pairwise orthogonal, i.e., if for $x, y \in S,\ x \neq y \Rightarrow x^* y = 0$.
- A set of nonzero vectors S is **orthonormal**(正規直交) if it is orthogonal and if for all $x \in S,\ \|x\| = 1$.

Unitary Matrices

$Q^* = Q^{-1},\ Q^*Q = I$ を満たす正方行列 Q を **unitary** という。また、Q が実行列であるなら、Orthogonal となる。つまり、実行列は大きく言えば unitary であり、同じく条件を満たす。

1.2 Norms

Vector Norms

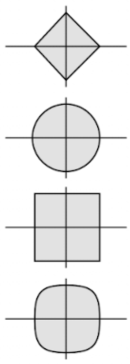
norm はそれぞれのベクトルに対して実数値の長さを割り当てる関数 $\|\cdot\| : \mathbb{C}^m \to \mathbb{R}$ である。そして、次の 3 つの条件を満たす必要がある。

$$\begin{aligned} &\|x\| \geq 0 \text{ for all } x \in \mathbb{C}^m,\ \text{and } \|x\| = 0 \iff x = 0, \\ &\|x + y\| \leq \|x\| + \|y\| \text{ for all } x, y \in \mathbb{C}^m, \\ &\|\alpha x\| = |\alpha|\|x\| \text{ for all } x \in \mathbb{C}^m,\ \alpha \in \mathbb{C}. \end{aligned} \tag{1.1}$$

Matrix Norms Induced by Vector Norms

■ **p-norm** The most important class of vector norms are the **p-norms** are defined below.The closed unit ball $\{x \in \mathbb{C}^m : \|x\|_p = 1\}$ corresponding to each norm is illustrated to the right for the case $m = 2$.

$$\begin{aligned} \|x\|_1 &= \sum_{i=1}^{m} |x_i|, \\ \|x\|_2 &= \left(\sum_{i=1}^{m} |x_i|^2 \right)^{1/2} = \sqrt{x^* x}, \\ \|x\|_\infty &= \max_{1 \leq i \leq m} |x_i|, \\ \|x\|_p &= \left(\sum_{i=1}^{m} |x_i|^p \right)^{1/p} \quad (1 \leq p < \infty). \end{aligned}$$



$A \in \mathbb{C}^{m \times n}$,respectively, the induced matrix norm $\|A\|_{mn}$ is the smallest number C for which the following inequality holds for all $x \in \mathbb{C}^n$:

$$\|Ax\|_m \leq C\|x\|_n$$

In other words, $\|A\|_{mn}$is the supremum(上限) of the ratio $\|Ax\|_m / \|x\|_n$ over vectors $x \neq 0$ in $\mathbb{C}^n$– the maximum factor by which A can "stretch" a vector x. 上限とは最大値のようなものである。たとえば、$x < 5$ は 5 は含まれないが、5 より小さい値であればどんなに 5 に近くても良い。つまり、5 は上限である。比率であり、上限であるから A が作用したことによって、x がどれだけ伸びうるかを表しており、誤差に対して重要な役割を果たす。

$$\|A\|_{mn} = \sup_{\substack{x \in \mathbb{C}^n \\ x \neq 0}} \frac{\|Ax\|_m}{\|x\|_n} = \sup_{\substack{x \in \mathbb{C}^n \\ \|x\|_n = 1}} \|Ax\|_m \tag{1.2}$$

行列ノルムは A によって誤差の最大何倍ベクトルを増幅させるかを表している。行列ノルムはを基準とし、どの p-norm を使うかで行列ノルムの値が変わる。

■ **1-norm** 1 ノルムは以下のように定義される。

$$\|x\|_1 = \sum_{j=1}^{n} |x_j|$$

行列ではないので、要素の絶対値の和となる。Ax は行単位でシグマを用いると、

$$Ax = \begin{bmatrix} \sum_j a_{1j}x_j \\ \sum_j a_{2j}x_j \\ \cdots \\ \sum_j a_{mj}x_j \end{bmatrix}$$

$$\|Ax\|_1 = \sum_{i=1}^{m} \left| \sum_{j=1}^{n} a_{ij}x_j \right|$$

これは絶対値の和であるから、三角不等式より、

$$\sum_{i=1}^{m} \left| \sum_{j=1}^{n} a_{ij}x_j \right| \leq \sum_{i=1}^{m} \sum_{j=1}^{n} |a_{ij}x_j| = \sum_{i=1}^{m} \sum_{j=1}^{n} |a_{ij}||x_j| = \sum_{j=1}^{n} \left(\sum_{i=1}^{m} |a_{ij}| \right) |x_j|$$

$\|x\|_1 = \sum_{j=1}^{n} |x_j|$ であるから、

$$\sup_{\substack{x \in \mathbb{C}^n \\ x \neq 0}} \frac{\|Ax\|_1}{\|x\|_1} = \max_{1 \leq j \leq n} \sum_{i=1}^{m} |a_{ij}| = \max_{1 \leq j \leq n} \|a_j\|_1 \tag{1.3}$$

つまり行列 A の 1-norm は**各列ベクトルの要素和における最大値**となる。

■ **∞-norm** ∞ ノルムは以下のように定義される。

$$\|x\|_\infty = \max_{1 \leq j \leq m} |x_j|$$

これは要素の絶対値の中での最大値となる。同様にして、行列 Ax では

$$\|Ax\|_\infty = \max_{1 \leq i \leq m} \left| \sum_{j=1}^{n} a_{ij}x_j \right| \leq \max_{1 \leq i \leq m} \left(\sum_{j=1}^{n} |a_{ij}| \right) |x_j| = \max_{1 \leq i \leq m} \sum_{j=1}^{n} |a_{ij}| \|x\|_\infty$$

$$\sup_{\substack{x \in \mathbb{C}^n \\ x \neq 0}} \frac{\|Ax\|_\infty}{\|x\|_\infty} = \max_{1 \leq i \leq m} \sum_{j=1}^{n} |a_{ij}| = \max_{1 \leq i \leq m} \|a_i^*\|_1 \tag{1.4}$$

つまり行列 A の 1-norm は**各行ベクトルの要素和における最大値**となる。

■ **Cauchy–Schwarz and Hölder Inequalities** Computing matrix p-norm with $p \neq 1, \infty$ is more difficult, and to approach this problem, we note that inner products can be bounded using p-norms. Let p and q satisfy $1/p + 1/q = 1$, with $1 \leq p, q \leq \infty$. Then the **Hölder inequality** sates that, for any vectors x and y,

$$|x^* y| \leq \|x\|_p \|y\|_q.$$

The Cauchy–Schwarz inequality is the special case of Hölder inequality with $p = q = 2$:

$$|x^* y| \leq \|x\|_2 \|y\|_2.$$

■ **The 2-Norm of a Row Vector** Consider a matrix A containing a single row. The matrix can be written as $A = x^*$, where a is a column vector. The Cauchy–Schwarz inequality allows us to obtain the induced matrix 2-norm. For any x, we have $\|Ax\|_2 = |a^* x| \leq \|a\|_2 \|x\|_2$.

$$\|A\|_2 = \sup_{x \neq 0} \frac{\|Ax\|_2}{\|x\|_2} = \|a\|_2.$$

■ **The 2-Norm of an Outer Product** if $A = uv^*$, we can bounded $\|Ax\|_2$ as follows:

$$\|Ax\|_2 = \|uv^* x\|_2 = \|u\|_2 |v^* x| \leq \|u\|_2 \|v\|_2 \|x\|_2$$

Therefore, $\|A\|_2 = \|u\|_2 \|v\|_2$. Again, this inequality is equality: consider the case $x = v$.

General Matrix Norms

行列も式 (1.1) に従う。

$$\begin{aligned} &\|A\| \leq 0, \text{and} \|A\| = 0 \text{only if} A = 0, \\ &\|A + B\| \leq \|A\| + \|B\|, \\ &\|\alpha A\| = |\alpha| \|A\|. \end{aligned} \tag{1.5}$$

The most important matrix norm which is **not induced** by a vector norm is the **Frobenius norm** or **Frobenius norm**, defined by

$$\|A\|_F = \left(\sum_{i=1}^{m} \sum_{j=1}^{n} |a_{ij}|^2 \right)^{1/2}.$$

2-norm とトレースを用いても表すことができる。

$$\|A\|_F = \left(\sum_{j=1}^{n} \|a_j\|_2^2 \right)^{1/2} = \sqrt{\mathrm{tr}(A^* A)} = \sqrt{\mathrm{tr}(AA^*)}$$

Listing 1.1: Frobenius norm の検証コード

```
import numpy as np

A = np.array([[1, 2], [3, 4]])
sum1 = np.sqrt(np.sum(A**2))
sum2 = np.sqrt(np.trace(A.T@A))
sum3 = np.sqrt(np.trace(A@A.T))
print(sum1) # 5.477225575051661
print(sum2) # 5.477225575051661
print(sum3) # 5.477225575051661
```

また、Unitary 行列 Q に対して、Frobenius norm は不変である。

$$\|QA\|_F = \|A\|_F, \quad \|AQ\|_F = \|A\|_F$$

1.3 The Singular Value Decomposition

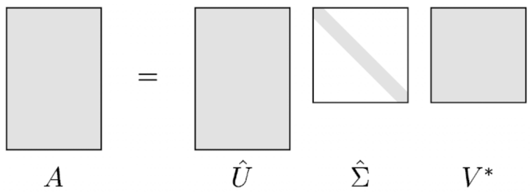
Reduced SVD

A は full rank n で V は unitay であるため、$V^{-1} = V^*$、$\hat{U}$ は $m \times n$ の直交行列、$\hat{\Sigma}$ は正方行列とする。

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

This factorization of A is called a **reduced singular value decomposition**, or **reduced SVD**, of A. Schematically, it looks likes this:

Reduced SVD $(m \geq n)$

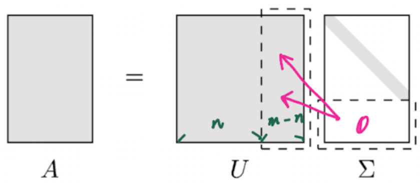


Full SVD

$\hat{U}$ is not unitary. However, by adjoining an additional $m - n$ orthonormal columns to $\hat{U}$ can be extended to a unitary matrix U.

If $\hat{U}$ is replaced by U, then $\hat{\Sigma}$ will have to change too. the last $m - n$ columns of U should be multiplied by zero.

Full SVD $(m \geq n)$



U と V はユニタリー行列となるから、

$$A = U\Sigma V^* \tag{1.6}$$

より、

$$A^{-1} = V\Sigma^{-1}U*$$

と逆行列演算が高速で安定する利点がある。

Having described the full SVD, we can now discard the simplifying assumption that A has full rank. if A is rank-deficient, the factorization (1.6) is still appropriate. 行列 A が横長でも式 (1.6) は成り立つ。

1.4 More on the SVD

全部大事