計算科学における情報圧縮

Information Compression in Computational Science **2017.10.5**

#3:情報圧縮の数理1 (線形代数の復習)

Review of linear algebra

理学系研究科 物理学専攻 大久保 毅 Department of Physics, **Tsuyoshi Okubo**

Outline

- Vector space- Abstract vectors-
 - Geometric vectors
 - General vector space (with inner product)
 - Basis and relation to coordinate vector space
- Matrix and linear map
 - Relation between matrices and linear maps
 - Important properties and operations for matrices
- Eigenvalue problem and diagonalization

Vector space -Abstract vectors-

Geometric vector

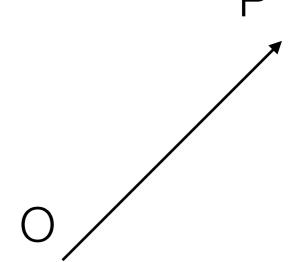
Geometric vector: Arrow on the plane (or the space),

which has "Direction" and "Length"

$$\vec{v} \equiv \overrightarrow{OP}$$

We can express a vector by its component:

$$\vec{v} = \begin{pmatrix} v_x \\ v_y \\ v_z \end{pmatrix} = \begin{pmatrix} x_p - x_o \\ y_p - y_o \\ z_p - z_o \end{pmatrix}$$



Properties of vector

Properties of addition:

$$\vec{a} + \vec{b} = \vec{b} + \vec{a}$$

$$(\vec{a} + \vec{b}) + \vec{c} = \vec{a} + (\vec{b} + \vec{c})$$

$$\vec{a} + \vec{0} = \vec{a}$$

$$\vec{a} + (-\vec{a}) = \vec{0}$$

Commutative property (交換法則)

Associative property (結合法則)

zero vector

inverse vector



Multiplication of scaler $c \in \mathbf{R}$ (実数):

$$c(\vec{a} + \vec{b}) = c\vec{b} + c\vec{a}$$
$$(c+d)\vec{a} = c\vec{a} + d\vec{a}$$
$$(cd)\vec{a} = c(d\vec{a})$$

Distributive property (分配法則)

Inner product of vector

Inner product:

$$(\vec{a}, \vec{b}) \equiv \vec{a} \cdot \vec{b}$$
$$= a_x b_x + a_y b_y + a_z b_z$$

Properties:

$$(\vec{a}, \vec{a}) \ge 0$$

 $(\vec{a}, \vec{b}) = (\vec{b}, \vec{a})$
 $(\vec{a} + \vec{b}, \vec{c}) = (\vec{a}, \vec{c}) + (\vec{b}, \vec{c})$
 $(c\vec{a}, \vec{b}) = c(\vec{a}, \vec{b})$ $c \in \mathbf{R}$

Norm (length):

$$\|\vec{a}\| \equiv \sqrt{(\vec{a}, \vec{a})}$$

Example:

$$\vec{a} = \begin{pmatrix} a_x \\ a_y \\ a_z \end{pmatrix}, \vec{b} = \begin{pmatrix} b_x \\ b_y \\ b_z \end{pmatrix}$$

Vector space (linear space)

Vector space V: generalization of geometric vector

Set of elements (vectors) satisfying following axioms (公理)

Properties of addition:

$$\vec{a} + \vec{b} = \vec{b} + \vec{a}$$

$$(\vec{a} + \vec{b}) + \vec{c} = \vec{a} + (\vec{b} + \vec{c})$$

$$\vec{a} + \vec{0} = \vec{a}$$

$$\vec{a} + (-\vec{a}) = \vec{0}$$

Multiplication of scaler $\,c\,$:

$$c(\vec{a} + \vec{b}) = c\vec{b} + c\vec{a}$$

$$(c + d)\vec{a} = c\vec{a} + d\vec{a}$$

$$(cd)\vec{a} = c (d\vec{a})$$

Commutative property (交換法則)

Associative property(結合法則)

Existence of unique zero vector

Existence of unique inverse vector

 $c \in \mathbf{R}$: Real vector space

 $c \in \mathbf{C}$: Complex vector space

Inner product space (metric vector space)

Inner product space:

(計量空間)

Vector space + definition of inner product

Inner product: (\vec{a}, \vec{b})

Axiom:

$$(\vec{a}, \vec{a}) \ge 0$$

$$(\vec{a}, \vec{b}) = (\vec{b}, \vec{a})^*$$

$$(\vec{a} + \vec{b}, \vec{c}) = (\vec{a}, \vec{c}) + (\vec{b}, \vec{c})$$

$$(c\vec{a}, \vec{b}) = c(\vec{a}, \vec{b})$$

*If a norm defined from the inner product is "complete" (完備),
that space is called **Hilbert space**.

Examples of vector spaces

(1) Coordinate space(数ベクトル空間) $m{R}^n, m{C}^n$

Vector:

$$ec{v} = egin{pmatrix} v_1 \ v_2 \ dots \ v_n \end{pmatrix} \quad v_i \in m{R} ext{ or } m{C}$$

Inner product:

$$(\vec{a}, \vec{b}) \equiv \vec{a} \cdot \vec{b}^*$$

(2) Wave vectors in quantum physics

Vector:

 $|\Psi
angle$

Inner product:

$$(|a\rangle, |b\rangle) = \langle b|a\rangle$$

Linearly independent or dependent

(線形独立) — (線形従属) —

Linear combination:

$$\vec{x} = c_1 \vec{v}_1 + c_2 \vec{v}_2 + \cdots c_k \vec{v}_k$$

 $\vec{v}_i \in \mathbf{V} \quad c_i \in \mathbf{R} \text{ or } \mathbf{C}$

A set $\{\vec{v}_1, \vec{v}_2, \cdots \vec{v}_k\}$ is linearly independent when

 $\vec{x} = \vec{0}$ is satisfied if and only if $c_1 = c_2 = \cdots = c_k = 0$



A set $\{\vec{v}_1, \vec{v}_2, \cdots \vec{v}_k\}$ is linearly dependent when

it is not linearly independent.

Basis of vector space

(基底)

A set $\{\vec{e}_1,\vec{e}_2,\cdots\vec{e}_n\}$ $(\vec{e}_i\in V)$ is a basis (基底) of V when

 $\{ \vec{e}_1, \vec{e}_2, \cdots \vec{e}_n \}$ is linearly independent.

Any vectors in V are represented by its linear combination.



 \vec{e}_i : basis vector

of basis vectors (n) is called **dimension** (次元) of V.

$$n = \dim V$$

Relation (map) to coordinate vector space

By using a basis $\{\vec{e}_1,\vec{e}_2,\cdots\vec{e}_n\}$, $\vec{v}\in V$ is uniquely represented as $\vec{v}=v_1\vec{e}_1+v_2\vec{e}_2+\cdots v_n\vec{e}_n$ (* From linear independency)



We can represent \vec{v} as a coordinate vector

$$ec{v}
ightharpoonup \left(egin{array}{c} v_1 \\ v_2 \\ \cdots \\ v_n \end{array}
ight) \in oldsymbol{C}^n (ext{ or } oldsymbol{R}^n)$$

By selecting a basis, we obtain a "concrete" coordinate vector for an "abstract" vector

Orthonormal basis (正規直交基底)

When a vector space has an inner product,

$$\vec{a}, \vec{b}$$
 is orthogonal (直交) if $(\vec{a}, \vec{b}) = 0$.

Orthonormal basis

A basis $\{\vec{e}_1, \vec{e}_2, \cdots \vec{e}_n\}$ is an orthonormal basis when

$$\|\vec{e}_i\| = 1$$
 $(i = 1, 2, ..., n)$
 $(\vec{e}_i, \vec{e}_j) = 0$ $(i \neq j; i, j = 1, 2, ..., n)$

*A basis can be transformed into an orthonormal basis.

cf. Gram-Schmidt orthonormalization

Vector subspace (linear subspace)

Vector subspace:

A subset W of a vector space V is a vector subspace of V when W satisfies the same axioms of vector space with V.

The following conditions are necessary and sufficient.

$$\vec{a}, \vec{b} \in \mathbf{W}$$
 $\vec{a} + \vec{b} \in \mathbf{W}$ $\vec{a} \in \mathbf{W}, c \in \mathbf{C}$ $c\vec{a} \in \mathbf{W}$

(In the case of complex vector space)

Spanned vector subspace

Spanned subspace:

For a subset $\, {m S} \,$ of a vector space ${m V} \,$, a set of linear combinations

$$\{c_1\vec{s}_1 + c_2\vec{s}_2 \cdots + c_k\vec{s}_k | c_i \in C, \vec{s}_i \in S\}$$

becomes a vector subspace of $oldsymbol{V}$.

We often use

$$\operatorname{Span}\{\vec{s}_1, \vec{s}_2, \cdots, \vec{s}_k\}$$

to represents a vector subspace spanned by a set of vectors

$$\{\vec{s}_1,\vec{s}_2,\cdots,\vec{s}_k\}$$

(This representation may appear in Krylov subspace method)

Matrix and linear map

Matrix (行列)

Matrix: "Table" of (complex) numbers in a rectangular form

$$M \times N$$
 matrix

$$A=egin{pmatrix} A_{11} & A_{12} & \cdots & A_{1,N} \ A_{21} & A_{22} & \cdots & A_{2,N} \ dots & dots & dots \ A_{M1} & A_{M2} & \cdots & A_{M,N} \end{pmatrix}$$
 trices: $C=AB$

Product of matrices: C = AB

$$C_{ij} = \sum_{l=1}^{K} A_{ik} B_{kj} \qquad B: K \times N \\ C: M \times N$$

In general: $XY \neq YX$

*We also know addition, multiplication of scalar.

 $A: M \times K$

Identity matrix (単位行列)

Identity matrix:

$$I = \begin{pmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 \end{pmatrix}$$

Product:

$$IA = A$$

$$BI = B$$

$$B:K\times N$$

* Element of the identity matrix: $I_{ij} = \delta_{ij}$ (Kronecker delta)

$$\delta_{ij} = \begin{cases} 1 & (i=j) \\ 0 & (i \neq j) \end{cases}$$

Transpose, complex conjugate and adjoint

Transpose: (転置)

$$A^t \qquad (A^t)_{ij} = A_{ji}$$

Complex conjugate: A^* $(A^*)_{ij} = A^*_{ij}$ (複素共役)

$$A^* \qquad (A^*)_{ij} = A^*_{ij}$$

Adjoint: (随伴)

$$A^{\dagger} = (A^t)^* = (A^*)^t$$

or

$$(A^{\dagger})_{ij} = A^*_{ji}$$

Hermitian conjugate:

(エルミート共役)

("Dagger" is convention in physics)

Multiplication to coordinate vector

M × N matrix transforms a N-dimensional coordinate vector to a M-dimensional coordinate vector.

M imes N matrix igodiagned Linear map: $oldsymbol{C}^N o oldsymbol{C}^M$ 1 to 1 (線形写像)

General linear map

Map:
$$f: oldsymbol{V} o oldsymbol{V}'$$

$$f(\vec{v}) = \vec{v}' \qquad (\vec{v} \in oldsymbol{V}, \vec{v}' \in oldsymbol{V}')$$

Linear map:

$$f(\vec{x} + \vec{y}) = f(\vec{x}) + f(\vec{y})$$

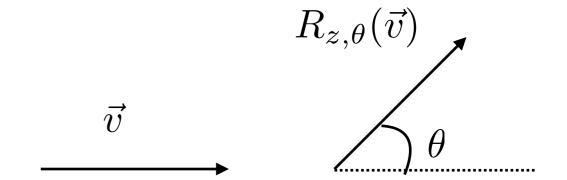
$$f(c\vec{x}) = cf(\vec{x})$$

$$(\vec{x}, \vec{y} \in \mathbf{V}, c \in \mathbf{C})$$

Examples:

Rotation (e.g. θ rotation around z-axis)

$$R_{z,\theta}: \mathbf{C}^3 \to \mathbf{C}^3$$



Hamiltonian operator

$$\mathcal{H}:oldsymbol{V} ooldsymbol{V}$$



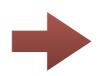
Matrix representation of linear map

By using a basis, we can represent a linear map in a matrix.

$$f: \mathbf{V} \to \mathbf{V}'$$

Vector space $V : \dim V = N$

Basis
$$\{ \vec{e}_1, \vec{e}_2, \cdots, \vec{e}_N \}$$

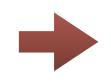


 $V': \dim V' = M$ $\{\vec{e'}_1, \vec{e'}_2, \cdots, \vec{e'}_M\}$

$$\{\vec{e'}_1, \vec{e'}_2, \cdots, \vec{e'}_M\}$$

Transformation of basis vectors:

$$f(\vec{e}_j) = f_{1j}\vec{e'}_1 + f_{2j}\vec{e'}_2 + \dots + f_{Mj}\vec{e'}_M$$



$$f: oldsymbol{V} o oldsymbol{V}'$$
 1 to 1 (if we fix basis

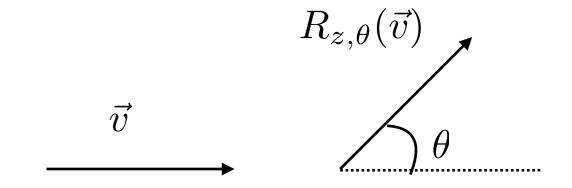
$$f: \mathbf{V} \to \mathbf{V}' \qquad \qquad \begin{cases} f_{11} & f_{12} & \cdots & f_{1,N} \\ f_{21} & f_{22} & \cdots & f_{2,N} \\ \vdots & \vdots & & \vdots \\ f_{M1} & f_{M2} & \cdots & f_{M,N} \end{cases}$$

Examples of matrix

Rotation (e.g. θ rotation around z-axis)

$$R_{z,\theta}: \mathbf{C}^3 \to \mathbf{C}^3$$

$$R_{z,\theta} = \begin{pmatrix} \cos \theta & -\sin \theta & 0 \\ \sin \theta & \cos \theta & 0 \\ 0 & 0 & 1 \end{pmatrix}$$



Hamiltonian operator

$$\mathcal{H}:oldsymbol{V} ooldsymbol{V}$$

$$\mathcal{H}: oldsymbol{V}
ightarrow oldsymbol{V} \ \mathcal{H}: oldsymbol{V}
ightarrow oldsymbol{V} \ \mathcal{H}: oldsymbol{V}
ightarrow egin{pmatrix} H_{0,0;0,0} & H_{0,0;0,1} & H_{0,0;1,0} & H_{0,0;1,1} \ H_{0,1;0,0} & H_{0,1;0,1} & H_{0,1;1,0} & H_{0,1;1,1} \ H_{1,0;0,0} & H_{1,0;0,1} & H_{1,0;1,0} & H_{1,0;1,1} \ H_{1,1;0,0} & H_{1,1;0,1} & H_{1,1;1,0} & H_{1,1;1,1} \ \end{pmatrix}$$

$$H_{\alpha,\beta;\alpha',\beta'} \equiv \langle \alpha\beta | \mathcal{H} | \alpha'\beta' \rangle$$

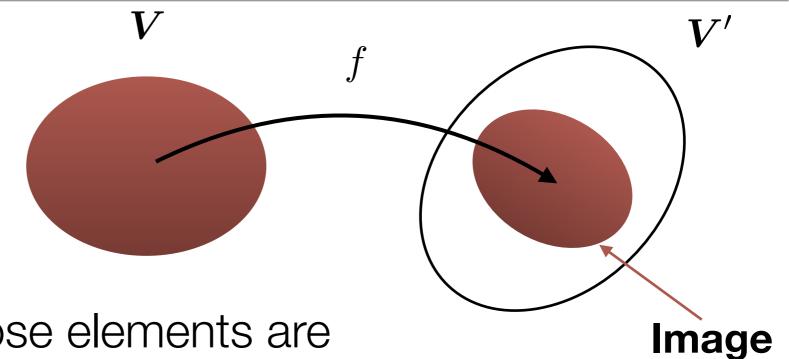
* In this notation, basis should be orthonormal.

Image of a map

$$f: \mathbf{V} \to \mathbf{V}'$$

Image of f:

(像)



Vector subspace whose elements are mapped from ${m V}$ by f .

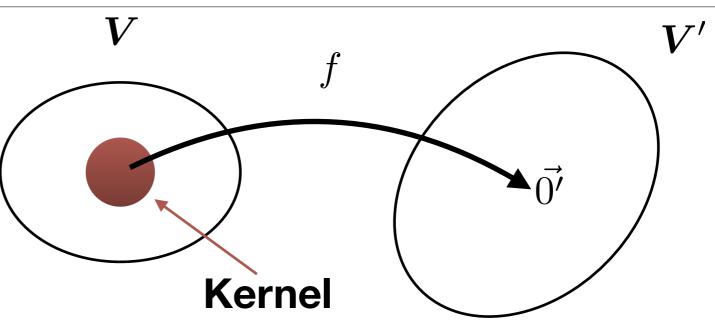
$$img(f) = {\vec{v}' | \vec{v} \in V, \vec{v}' = f(\vec{v})}$$

Kernel of a map

$$f: \mathbf{V} o \mathbf{V}'$$

Kernel of f:

(核)



Vector subspace whose elements are mapped into zero vector by f .

$$\ker(f) = \{\vec{v} | \vec{v} \in V, f(\vec{v}) = \vec{0'}\}$$

Theorem:

$$\dim(V) = \dim(\ker(f)) + \dim(\operatorname{img}(f))$$

Rank of matrix

Rank (ランク or 階数)of a matrix A:

$$rank(A) \equiv \dim(img(A))$$

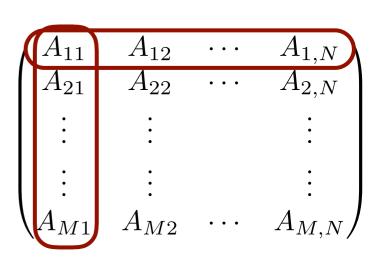
Rank is identical with

Maximum # of linearly independent column vectors (列ベクトル) in *A*Maximum # of linearly independent law vectors (行ベクトル) in *A*



$$\operatorname{rank}(A) \leq \min(M,N)$$

for a $N \times M$ matrix A.



Regular matrix and its inverse matrix

A square matrix A is a **regular matrix** (正則) if a matrix X satisfying

$$AX = XA = I$$

exists. The matrix X is called inverse matrix (逆行列) of A and it is written as $X = A^{-1}$

Properties:

 A^{-1} is unique.

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

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

A is a regular matrix $\operatorname{rank}(A) = N$



Can we consider an "inverse matrix" of a non-regular matrix (including a rectangular matrix)?



Simultaneous linear equation

Simultaneous linear equation (連立一次方程式)

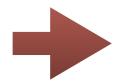
can be represented by a matrix and a vector as

$$A\vec{x} = \vec{b}$$
 $A: M \times N, \vec{x} \in \mathbb{C}^N, \vec{b} \in \mathbb{C}^M$

If A is a square matrix (N=M), and it has a inverse matrix (rank(A) = N), we can solve the equation as

$$\vec{x} = A^{-1}\vec{b}$$

How can we find a "solution" when A does not have a "inverse"?



It is probably related to the topic "sparse modeling".

Determinant of matrix

For a square matrix A its **determinant**(行列式) is defined as

$$\det A = |A| = \sum_{\sigma} \operatorname{sgn}(\sigma) A_{1\sigma(1)} A_{2\sigma(2)} \cdots A_{N\sigma(N)}$$
$$= \sum_{\sigma} \operatorname{sgn}(\sigma) A_{\sigma(1)1} A_{\sigma(2)2} \cdots A_{\sigma(N)N}$$

 σ : permutation(置換) of $\{1,2,...,N\}$

$$\sigma = \begin{cases}
1 & \text{even permutation} & (偶置換) \\
-1 & \text{odd permutation} & (奇置換)
\end{cases}$$

Examples:

$$\begin{vmatrix} a & b \\ c & d \end{vmatrix} = ab - cd \qquad \begin{vmatrix} a & b & c \\ d & e & f \\ g & h & i \end{vmatrix} = aei + bfg + cdh$$

$$-afg - bdi - ceg$$

Determinant and inverse matrix

By using the determinant of A, we can represent its inverse matrix:

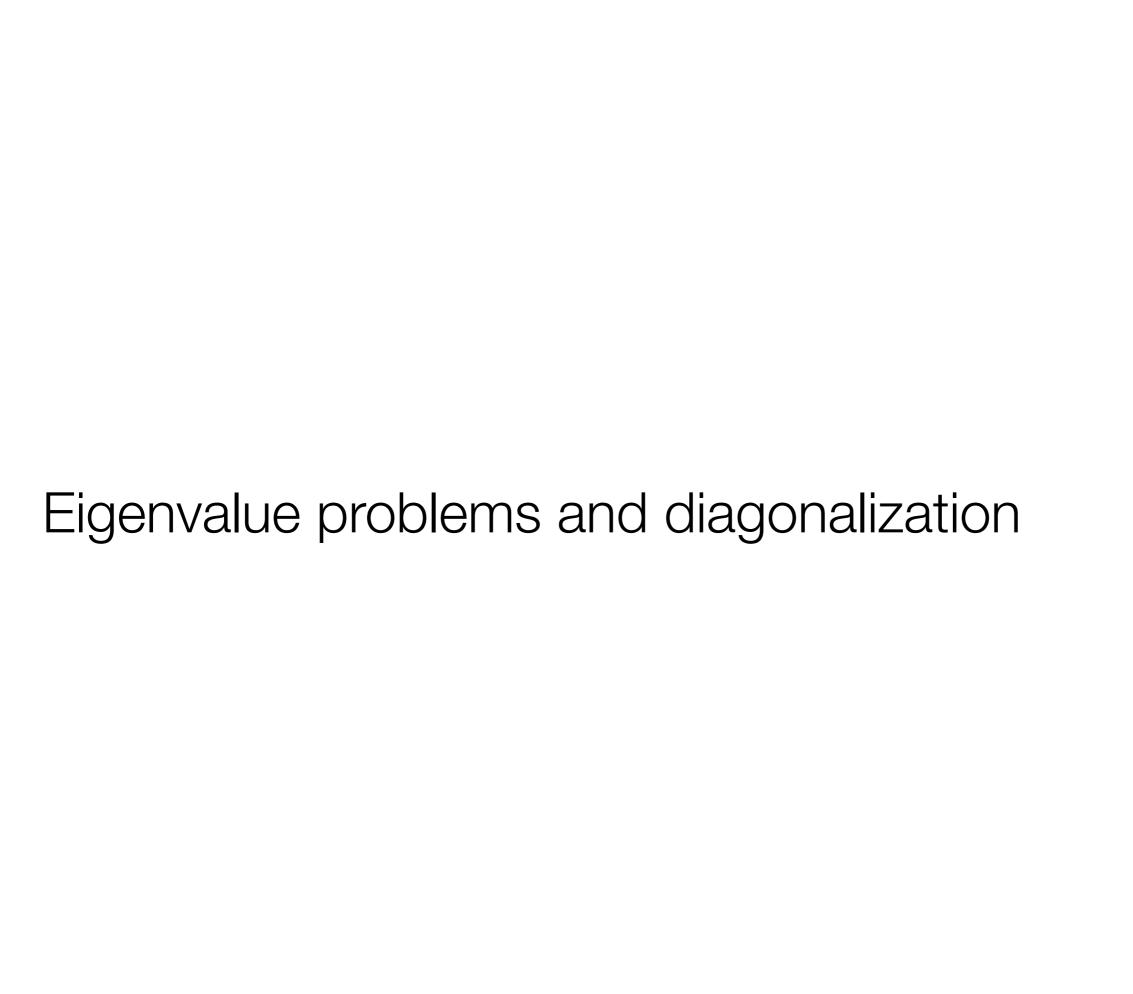
$$A^{-1} = \frac{1}{\det(A)} \begin{pmatrix} \tilde{A}_{11} & \tilde{A}_{21} & \cdots & \tilde{A}_{N1} \\ \tilde{A}_{12} & \tilde{A}_{22} & \cdots & \tilde{A}_{N2} \\ \vdots & \vdots & & \vdots \\ \tilde{A}_{1N} & \tilde{A}_{2N} & \cdots & \tilde{A}_{NN} \end{pmatrix}$$
 can see that
$$\tilde{A}_{ij} : \text{cofactor } (\text{余因子})$$

We can see that

$$det(A) = 0$$
 A-1 diverges

Indeed,

A is a regular matrix. $\det(A) \neq 0$ necessary and sufficient



Right and left eigenvectors

In general, left eigenvectors can be different from the right eigenvectors.

$$A\vec{v} = \lambda \vec{v}$$
$$(\vec{u}^*)^t A = \lambda (\vec{u}^*)^t$$

 \vec{v} : Right eigenvector

 $(\vec{u}^*)^t$:Left eigenvector

Properties:

Set of eigenvalues are identical between the right and the left eigenvectors.

A left eigenvector and a right eigenvector are orthogonal when they correspond to different eigenvalues.

$$\vec{u}_i^* \cdot \vec{v}_j = 0 \qquad (\lambda_i \neq \lambda_j)$$

Eigenvalue and Eigenvector

For a square matrix A

$$A\vec{v} = \lambda \vec{v}$$

 $\vec{v} \neq \vec{0}$:eigenvector (固有ベクトル)

 $\lambda \in C$:eigenvalue (固有値)

Properties:

If \vec{v} is an eigenvector, $c\vec{v}$ is also an eigenvector.

Eigenspace (固有空間):

The set of eigenvectors corresponds an eigenvalue λ .

Eigenvectors corresponding to different eigenvalues are linearly independent.

Diagonalization

Diagonalizaiton(対角化):

$$A: N \times N$$

$$P^{-1}AP = \begin{pmatrix} \alpha_1 & & \\ & \alpha_2 & \\ & & \ddots & \\ & & & \alpha_N \end{pmatrix}$$

A can be diagonalized.



A has N linearly independent eigenvectors.

$$\alpha_{i} = \lambda_{i}$$

$$P = (\vec{v}_{1}, \vec{v}_{2}, \cdots, \vec{v}_{N})$$

$$(P^{-1})^{t} = (\vec{u}_{1}^{*}, \vec{u}_{2}^{*}, \cdots, \vec{u}_{N}^{*})$$

Normalization: $\vec{u}_i^* \cdot \vec{v} = 1$

Meaning of diagonalization

General transform using a regular matrix: $P^{-1}AP$

It is a transform of the basis:

$$\{\vec{e}_1,\vec{e}_2,\cdots,\vec{e}_N\} \rightarrow \{P\vec{e}_1,P\vec{e}_2,\cdots,P\vec{e}_N\}$$

Diagonalization:

By using eigenvectors as a basis, we can obtain a simple linear map represented by a diagonal matrix.

$$A \to P^{-1}AP$$

* The determinant of A is invariant under this transformation:

$$\det(P^{-1}AP) = \det(P^{-1})\det(A)\det(P^{-1}) = \det(A)\det(P^{-1}P) = \det(A)$$



$$\det(A) = \prod_{i=1}^{N} \lambda_i$$

(This relation is true even if A cannot be diagonalized)

Unitary matrix

Unitary matrix (ユニタリ行列):
$$U^\dagger = U^{-1}$$

Real Orthogonal matrix(実直交行列): $P^t = P^{-1}, (P_{ij} \in \mathbf{R})$

When we consider a unitary matrix as a set of vectors:

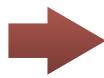
$$U = (\vec{v}_1, \vec{v}_2, \cdots, \vec{v}_N)$$

it is a orthonormal basis: $\vec{v}_i^* \cdot \vec{v}_j = \delta_{i,j}$

The linear map represented by a unitary matrix (unitary transformation) does not change

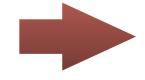
- the norm of a vector $||U\vec{v}|| = ||\vec{v}||$
- "distance" between two vectors

$$||U\vec{v}_1 - U\vec{v}_2|| = ||\vec{v}_1 - \vec{v}_2||$$



Normal matrix

Normal matrix(正規行列): $A^{\dagger}A = AA^{\dagger}$



We can always diagonalize it by a unitary matrix

$$U^{\dagger} = U^{-1}$$

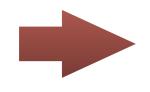
as
$$U^\dagger A U = egin{pmatrix} \lambda_1 & & & & \ & \lambda_2 & & \ & & \ddots & \ & & & \lambda_N \end{pmatrix} \qquad \lambda_i \in m{C}$$

Its eigenvalues could be complex. (even if A is a real matrix)

Hermitian matrix and its eigenvalue

Hermitian matrix(エルミート行列): $A^\dagger = A$

Real symmetric matrix(実対称行列): $A^t = A, \quad (A_{ij} \in \mathbf{R})$



It is a special normal matrix. $A^{\dagger}A = AA^{\dagger} = AA$ Its eigenvalues are real.

We can always diagonalize it by a unitary matrix

$$U^{\dagger}AU = \begin{pmatrix} \lambda_1 & & & \\ & \lambda_2 & & \\ & & \ddots & \\ & & & \lambda_N \end{pmatrix} \qquad \lambda_i \in \boldsymbol{R}$$

Hermitian (or real symmetric) matrices often appear in physics.

Generalization of diagonalization

- Eigenvalue problems and diagonalizations are defined for a square matrix.
- Even if A is a square matrix, it may not be diagonalized.



- Is it possible to transform all square matrixes into diagonal forms by generalizing the diagonalization?
- Is it possible to generalize it to a rectangular matrices?

Yes. The **singular value decomposition** (特異值分解) is an generalization of the diagonalization.

Next week

第1回: 現代物理学における巨大なデータ

第2回: 情報圧縮と繰り込み

第3回: 情報圧縮の数理1 (線形代数の復習)

第4回: 情報圧縮の数理2 (特異値分解と低ランク近似)

(Singular value decomposition and low rank approximation)

第5回: 情報圧縮の数理3 (スパース・モデリングの基礎)

第6回: 情報圧縮の数理4 (クリロフ部分空間法の基礎)

第7回: 物質科学における情報圧縮

第8回: スパース・モデリングの物質科学への応用

第9回: クリロフ部分空間法の物質科学への応用

第10回: 行列積表現の基礎

第11回: 行列積表現の応用

第12回: テンソルネットワーク表現への発展

第13回: テンソルネットワーク繰り込みと低ランク近似の応用