計算科学における情報圧縮

Information Compression in Computational Science

2021.10.21

#3:Review of linear algebra

+ singular value decomposition

理学系研究科　大久保　毅

Graduate School of Science, **Tsuyoshi Okubo**

Outline

- Matrix and linear map (cont.)
- Eigenvalue problem and diagonalization
- Singular value decomposition (SVD)
- Generalized inverse matrix

Matrix and linear map

Matrix（行列）

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

M × N matrix

$$A = \begin{pmatrix} A_{11} & A_{12} & \cdots & A_{1,N} \\ A_{21} & A_{22} & \cdots & A_{2,N} \\ \vdots & \vdots & & \vdots \\ A_{M1} & A_{M2} & \cdots & A_{M,N} \end{pmatrix}$$

$$A_{ij} \in \mathbb{C}(\text{ or } \mathbb{R})$$

Product of matrices: $C = AB$

$$A : M \times K$$
$$B : K \times N$$
$$C : M \times N$$

$$C_{ij} = \sum_{k=1}^{K} A_{ik} B_{kj}$$

In general: $XY \neq YX$

*We also know addition, multiplication of scalar.

Identity matrix（単位行列）

Identity matrix:

N × N matrix
(Square matrix)

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

Product:

$$IA = A \qquad A : N \times M$$

$$BI = B \qquad B : K \times N$$

* 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:
(複素共役)

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

Adjoint:
(随伴)

or

Hermitian conjugate:
(エルミート共役)

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

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

("Dagger" is convention in physics)

Multiplication to coordinate vector

$$A : M \times N \qquad \vec{v} \in \mathbb{C}^N \qquad \vec{v}' \in \mathbb{C}^M$$

$$\begin{pmatrix} A_{11} & A_{12} & \cdots & A_{1,N} \\ A_{21} & A_{22} & \cdots & A_{2,N} \\ \vdots & \vdots & & \vdots \\ \vdots & \vdots & & \vdots \\ A_{M1} & A_{M2} & \cdots & A_{M,N} \end{pmatrix} \begin{pmatrix} v_1 \\ v_2 \\ \vdots \\ v_N \end{pmatrix} = \begin{pmatrix} v'_1 \\ v'_2 \\ \vdots \\ \vdots \\ v'_M \end{pmatrix}$$

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

M × N matrix ⟷ **Linear map:** $\mathbb{C}^N \to \mathbb{C}^M$

1 to 1 （線形写像）

General linear map

Map: $f : \mathbb{V} \to \mathbb{V}'$

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

Linear map:

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

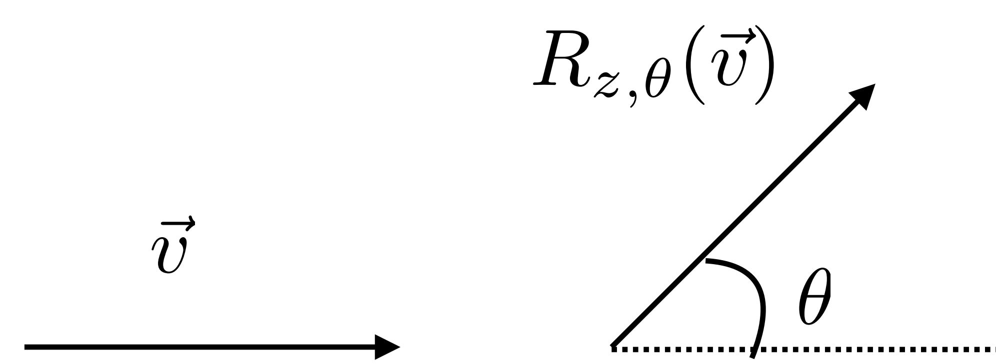
$$f(c\vec{x}) = cf(\vec{x})$$

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

Examples:

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

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



Hamiltonian operator

$$\mathcal{H} : \mathbb{V} \to \mathbb{V}$$

$$|\Psi\rangle \quad \Rightarrow \quad \mathcal{H}|\Psi\rangle$$

Matrix representation of linear map

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

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

Vector space	$\mathbb{V} : \dim \mathbb{V} = N$	$\Rightarrow$	$\mathbb{V}' : \dim \mathbb{V}' = M$
Basis	$\{\vec{e}_1, \vec{e}_2, \cdots, \vec{e}_N\}$		$\{\vec{e'}_1, \vec{e'}_2, \cdots, \vec{e'}_M\}$

Transformation of basis vectors:

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

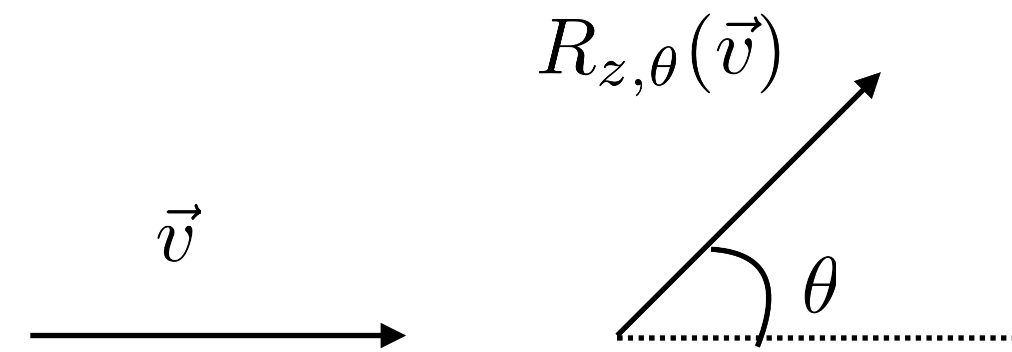
$\Rightarrow$ $f : \mathbb{V} \to \mathbb{V}'$ $\Leftrightarrow$
$$\begin{pmatrix} f_{11} & f_{12} & \cdots & f_{1,N} \\ f_{21} & f_{22} & \cdots & f_{2,N} \\ \vdots & \vdots & & \vdots \\ f_{M1} & f_{M2} & \cdots & f_{M,N} \end{pmatrix}$$

1 to 1
(if we fix basis)

Examples of matrix

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

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



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

Hamiltonian operator

$$\mathcal{H} : \mathbb{V} \to \mathbb{V}$$

$$\mathcal{H} \to \begin{pmatrix} H_{0,0;0,0} & H_{0,0;0,1} & H_{0,0;1,0} & H_{0,0;1,1} \\ H_{0,1;0,0} & H_{0,1;0,1} & H_{0,1;1,0} & H_{0,1;1,1} \\ H_{1,0;0,0} & H_{1,0;0,1} & H_{1,0;1,0} & H_{1,0;1,1} \\ H_{1,1;0,0} & H_{1,1;0,1} & H_{1,1;1,0} & H_{1,1;1,1} \end{pmatrix}$$

Matrix element: $H_{\alpha,\beta;\alpha',\beta'} \equiv \langle \alpha\beta | \mathcal{H} | \alpha'\beta' \rangle$
（行列要素）

* In this notation, basis should be orthonormal.

Image of a map

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

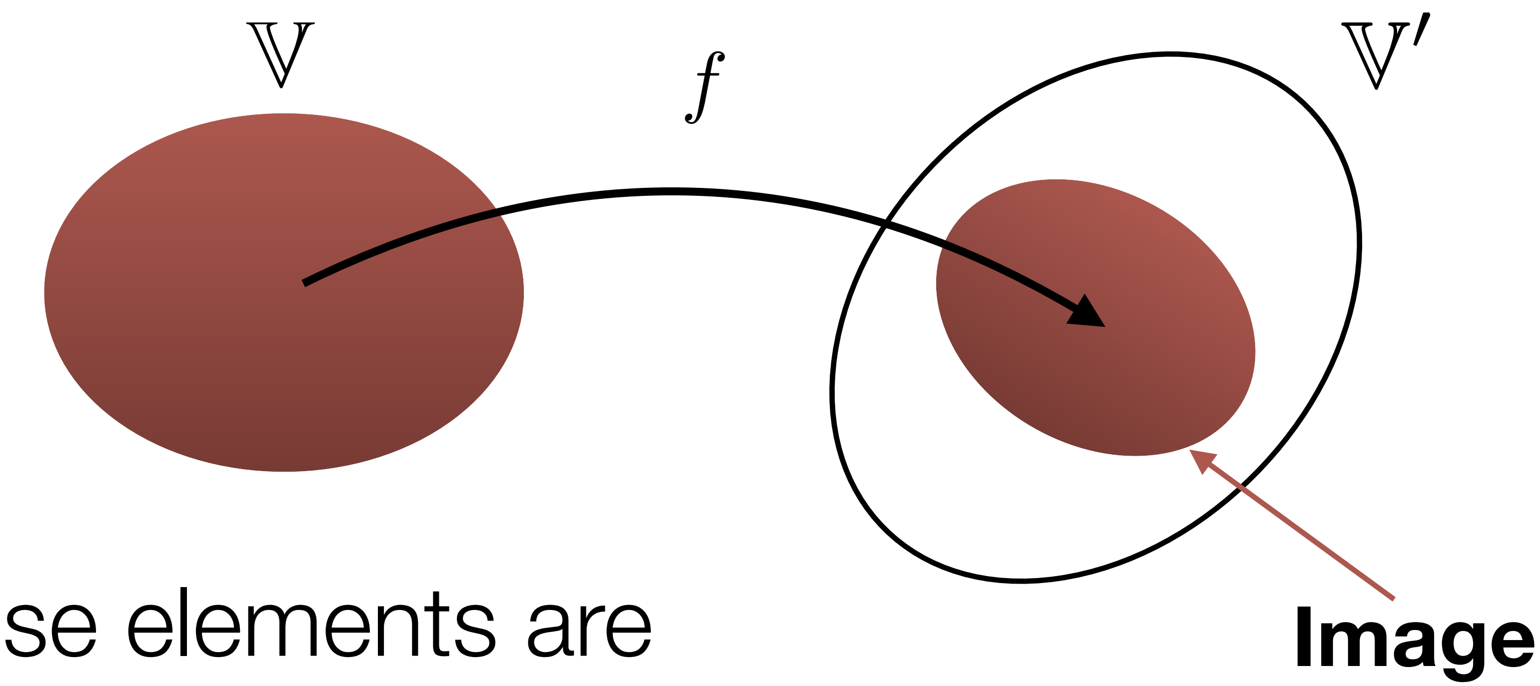


Image of f:
(像)

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

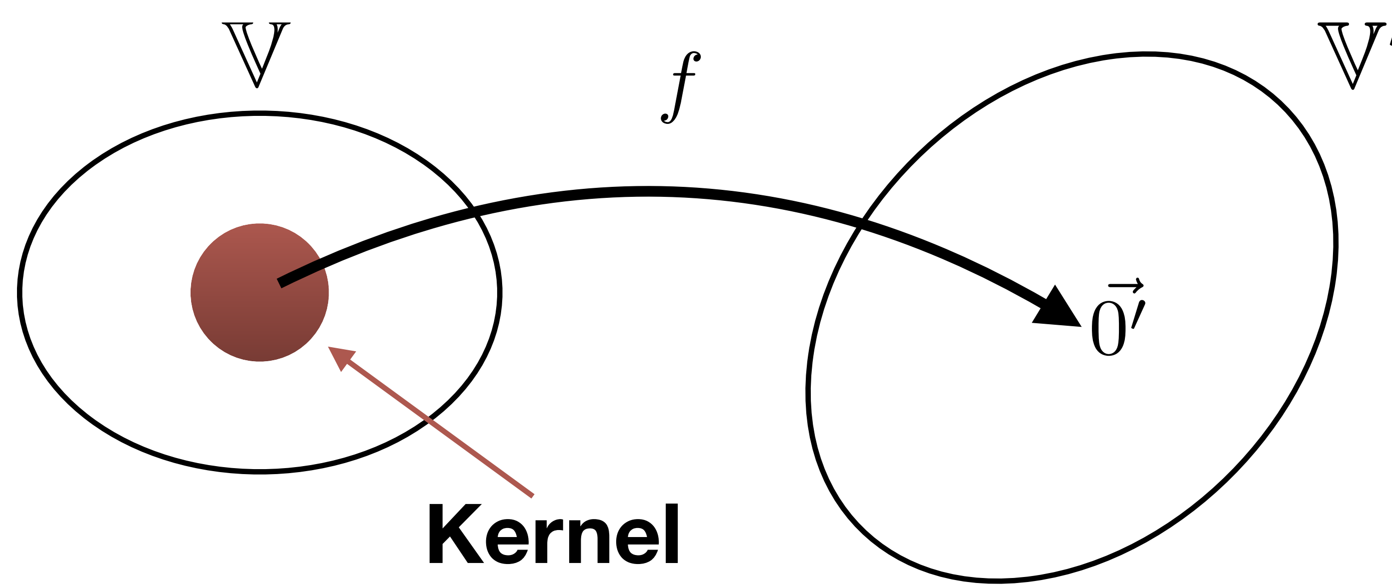
$$\mathrm{img}(f) = \{\vec{v}' | \vec{v} \in \mathbb{V}, \vec{v}' = f(\vec{v})\}$$

Kernel of a map

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

Kernel of f:

(核)



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

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

Theorem:

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

Rank of matrix

Rank （ランク or 階数）of a matrix *A*:

$$\mathrm{rank}(A) \equiv \dim(\mathrm{img}(A))$$

Rank is identical with

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

Maximum # of linearly independent row vectors （行ベクトル）in *A*

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

for a *N*×*M* matrix *A*.

$$\begin{pmatrix} A_{11} & A_{12} & \cdots & A_{1,N} \\ A_{21} & A_{22} & \cdots & A_{2,N} \\ \vdots & \vdots & & \vdots \\ \vdots & \vdots & & \vdots \\ A_{M1} & A_{M2} & \cdots & A_{M,N} \end{pmatrix}$$

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 $\Longleftrightarrow$ $\mathrm{rank}(A) = N$

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

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

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

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

N > *M*: Underdetermined problem （劣決定問題）

N < *M*: Overdetermined problem （優決定問題）

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

It is related to the topic "sparse modeling".
(Especially for underdetermined problems.)

Eigenvalue problems and diagonalization

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）

$\lambda \in \mathbb{C}$: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.

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.

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

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. $\Longleftrightarrow$ *A* has *N* linearly independent eigenvectors.

necessary and sufficient

$$\alpha_i = \lambda_i$$

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

$$(P^{-1})^t = (\vec{u}_1^*, \vec{u}_2^*, \cdots, \vec{u}_N^*)$$

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

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.

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

* The determinant of A is invariant under this transformation:

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

$$\det(A) = \prod_i^N \lambda_i$$

(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）： $P^t = P^{-1}, (P_{ij} \in \mathbb{R})$

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

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

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）： $A^t = A, \quad (A_{ij} \in \mathbb{R})$

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

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

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** (特異値分解) is an generalization of the diagonalization.

(We can also consider a decomposition of a tensor.)

Singular value decomposition

Diagonalization

Diagonalizaiton（対角化）：

$$A : N \times N$$

(Square matrix)

$$P^{-1}AP = \begin{pmatrix} \lambda_1 & & & \\ & \lambda_2 & & \\ & & \ddots & \\ & & & \lambda_N \end{pmatrix}$$

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

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

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

$$(P^{-1})^t = (\vec{u}_1^*, \vec{u}_2^*, \cdots, \vec{u}_N^*)$$

- Eigenvalue problems and diagonalizations are defined for a square matrix.
- Even if A is a square matrix, it may not be diagonalized.
  - Normal or Hermitian matrices are always diagonalized by a unitary matrix

Spectral decomposition

(For a normal matrix A,)

Spectral decomposition
(スペクトル分解)

$$A = U \begin{pmatrix} \lambda_1 & & & \\ & \lambda_2 & & \\ & & \ddots & \\ & & & \lambda_N \end{pmatrix} U^\dagger$$

$$= \sum_{i=1}^{N} \lambda_i \underline{\vec{u}_i \vec{u}_i^\dagger}$$

$$\left(= \sum_{i=1}^{N} \lambda_i |u_i\rangle\langle u_i| \right)$$

Note:

$$\vec{u}_i \vec{u}_i^\dagger = \begin{pmatrix} u_1 u_1^* & u_1 u_2^* & \cdots & u_1 u_N^* \\ u_2 u_1^* & u_2 u_2^* & \cdots & u_2 u_N^* \\ \vdots & \vdots & \cdots & \vdots \\ u_N u_1^* & u_N u_2^* & \cdots & u_N u_N^* \end{pmatrix}$$

Matrix decomposition into a sum of projectors onto its eigen subspaces.

Projector:

$$P^2 = P$$

Singular value decomposition (SVD)

Singular value decomposition
(特異値分解)

$$A : M \times N$$

$$A_{ij} \in \boldsymbol{C}$$

$$A = \underline{U}\Sigma\underline{V}^\dagger$$

$U : M \times M$
Unitary

$V : N \times N$
Unitary

$$\Sigma = \begin{pmatrix} \Sigma_{r\times r} & 0_{r\times(N-r)} \\ 0_{(M-r)\times r} & 0_{(M-r)\times N-r} \end{pmatrix}$$

$$\Sigma_{r\times r} = \begin{pmatrix} \sigma_1 & & & \\ & \sigma_2 & & \\ & & \ddots & \\ & & & \sigma_r \end{pmatrix}$$

Diagonal matrix with
non-negative real elements

$$\sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_r > 0$$

Singular values

Properties of SVD 1

1. Any matrices can be decomposed as SVD: $A = U\Sigma V^\dagger$

$A : M \times N$ ➡ $A^\dagger A : N \times N$

* $A^\dagger A$ is a Hermitian matrix.

$$(A^\dagger A)^\dagger = A^\dagger A$$

➡

It can be diagonalized by
a unitary matrix V .

$$V^\dagger (A^\dagger A) V = \mathrm{diag}\{\lambda_1, \lambda_2, \cdots, \lambda_N\}$$

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

$\vec{v}_i$: eigenvector

* $A^\dagger A$ is a positive semi-definite matrix.
(半正定値、準正定値)

$$\vec{x}^* \cdot (A^\dagger A \vec{x}) = \|A\vec{x}\|^2 \geq 0$$

⬌

Its eigenvalues are
non-negative

$$\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_N \geq 0$$

Properties of SVD 1

1. Any matrices can be decomposed as SVD: $A = U\Sigma V^\dagger$

$$V^\dagger (A^\dagger A) V = \mathrm{diag}\{\lambda_1, \lambda_2, \cdots, \lambda_N\}$$
$$V = (\vec{v}_1, \vec{v}_2, \cdots, \vec{v}_N)$$

$$\Longleftrightarrow$$

$$(A\vec{v}_i)^* \cdot (A\vec{v}_j) = \lambda_i \delta_{ij}$$
$$(\|A\vec{v}_i\|^2 = \lambda_i)$$

Suppose $\lambda_1 \geq \lambda_2 \geq \cdots \underline{\geq \lambda_r > 0} = \lambda_{r+1} = \cdots = \lambda_N$

(There are r positive eigenvalues.)

Make new orthonormal basis $U = (\vec{u}_1, \vec{u}_2, \cdots, \vec{u}_M)$ in $\mathbb{C}^M$

For $(i = 1, 2, \ldots, r)$ $\quad \sigma_i = \sqrt{\lambda_i}\,, \vec{u}_i = \dfrac{1}{\sigma_i} A\vec{v}$

For $(i = r+1, \ldots, M)$ Any orthonormal basis orthogonal to $\vec{u}_i \quad (i = 1, 2, \ldots, r)$

$$\vec{u}_i^* \cdot (A\vec{v}_j) = \sigma_i \delta_{ij} \quad (i = 1, \ldots, M; j = 1, \ldots, N)$$

(For simplicity, we set $\sigma_i = 0$ for $i > r$.)

Properties of SVD 1

1. Any matrices can be decomposed as SVD: $A = U\Sigma V^\dagger$

We can perform same "proof" by using $AA^\dagger$.

$U = (\vec{u}_1, \vec{u}_2, \cdots, \vec{u}_M)$ is the unitary matrix which diagonalize $AA^\dagger$ as

$$U^\dagger (AA^\dagger) U = \mathrm{diag}\{\lambda_1, \lambda_2, \ldots, \lambda_r, \underbrace{0, \ldots, 0}_{M-r}\}$$

In summary,

- A matrix A can be decomposed as SVD: $A = U\Sigma V^\dagger$
- Singular values are related to the eigenvalues of $A^\dagger A$ and $AA^\dagger$ as

$$\sigma_i = \sqrt{\lambda_i}\ .$$

- V and U are eigenvectors of $A^\dagger A$ and $AA^\dagger$,respectively.

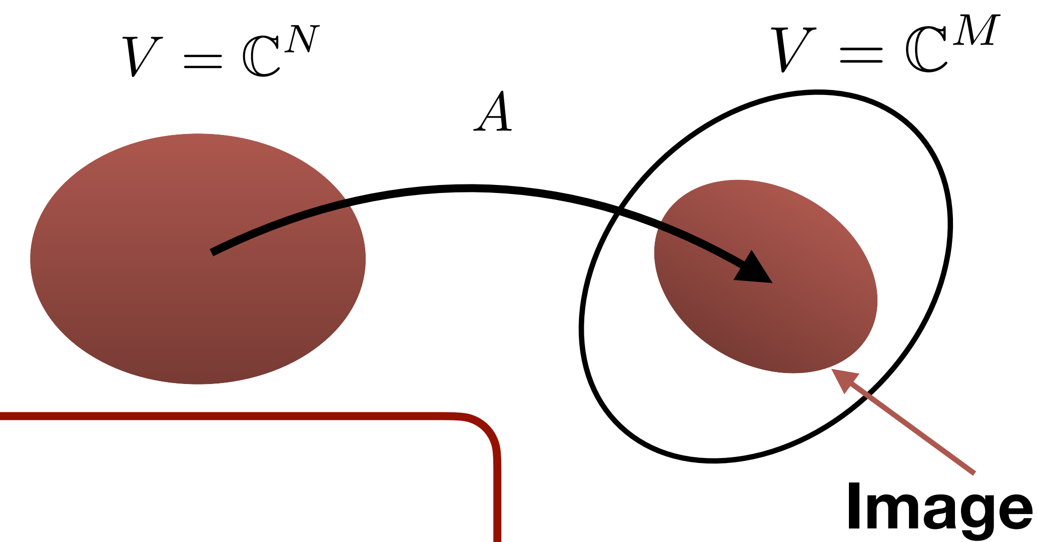
Properties of SVD 2

$$A = U\Sigma V^\dagger$$

2. # of positive singular values is identical with the rank.

$$A : M \times N \quad \Rightarrow \quad A : \mathbb{C}^N \to \mathbb{C}^M$$

$$\mathrm{rank}(A) \equiv \dim(\mathrm{img}(A))$$



Remember

The orthonormal basis $\{\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_N\}$ satisfies

$$(A\vec{v}_i)^* \cdot (A\vec{v}_j) = \lambda_i \delta_{ij}$$

Here, $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_r > 0 = \lambda_{r+1} = \cdots = \lambda_N$ and $\sigma_i = \sqrt{\lambda_i}$

$$\forall \vec{x} \in \mathbb{C}^N, \ \vec{x} = \sum_{i=1}^{N} C_i \vec{v}_i \quad \Rightarrow \quad A\vec{x} = \sum_{i=1}^{N} C_i (A\vec{v}_i) = \sum_{i=1}^{r} C_i (A\vec{v}_i)$$

$$\Rightarrow \quad \mathrm{img}(A) = \mathrm{Span}\{A\vec{v}_1, A\vec{v}_2, \ldots, A\vec{v}_r\}$$

$$\Rightarrow \quad \dim(\mathrm{img}(A)) = r$$ **= # of positive singular values**

Properties of SVD 3

$$A = U\Sigma V^\dagger$$

3. Singular vectors

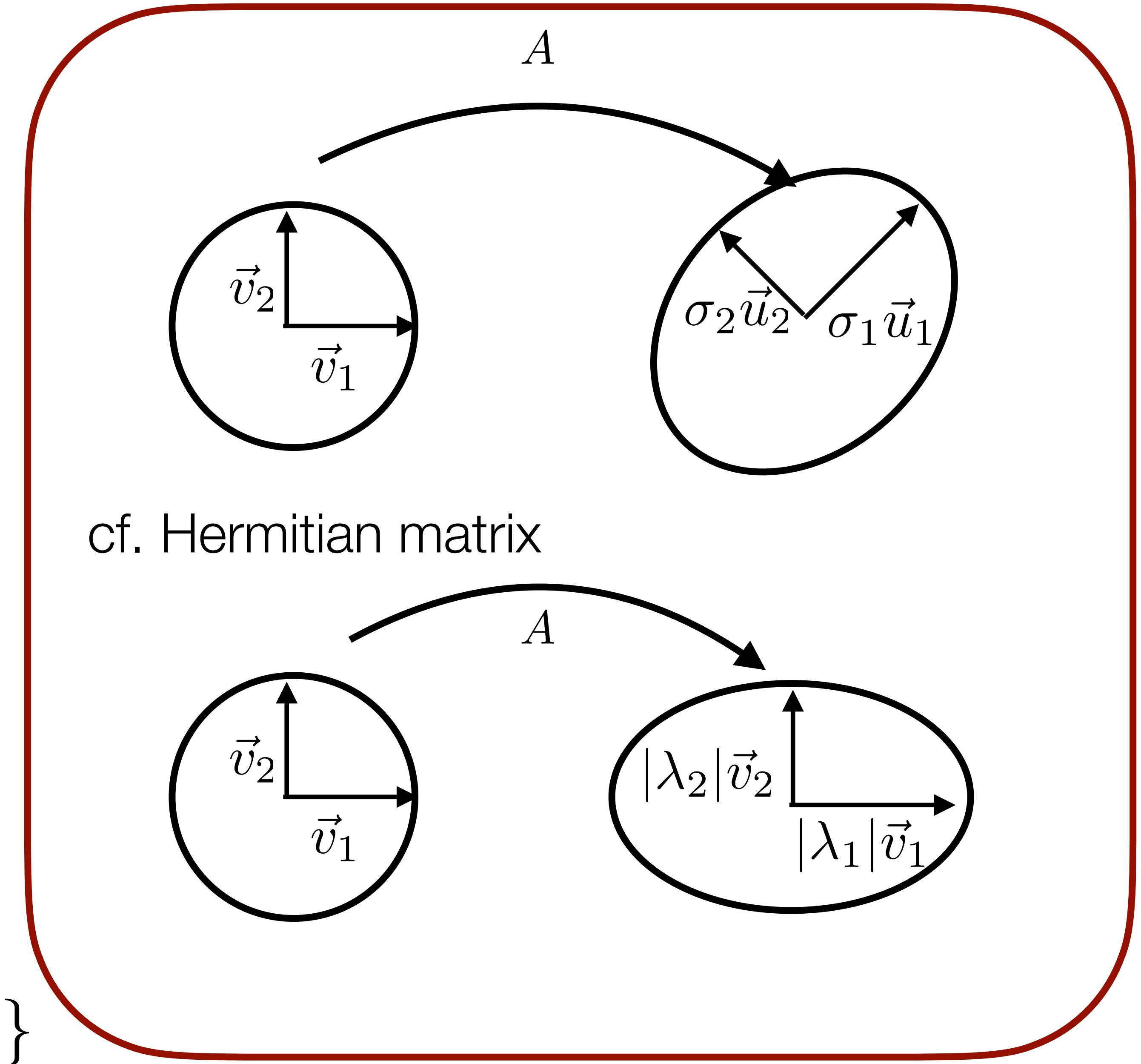
$A : M \times N \qquad U = (\vec{u}_1, \vec{u}_2, \cdots, \vec{u}_M)\,,\; V = (\vec{v}_1, \vec{v}_2, \cdots, \vec{v}_N)$

For $i = 1, 2, \ldots, r$

$$A\vec{v}_i = \sigma_i \vec{u}_i \;,\; A^\dagger \vec{u}_i = \sigma_i \vec{v}_i$$

$\vec{v}_i$: right singular vector

$\vec{u}_i$: left singular vector



Relation to image and kernel:

$$\mathrm{img}(A) = \mathrm{Span}\{\vec{u}_1, \vec{u}_2, \ldots, \vec{u}_r\}$$
$$\ker(A) = \mathrm{Span}\{\vec{v}_{r+1}, \vec{v}_{r+2}, \ldots, \vec{v}_N\}$$

$$\mathrm{img}(A^\dagger) = \mathrm{Span}\{\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_r\}$$
$$\ker(A^\dagger) = \mathrm{Span}\{\vec{u}_{r+1}, \vec{u}_{r+2}, \ldots, \vec{u}_M\}$$

Properties of SVD 4 (optional) $A = U\Sigma V^\dagger$

4. Min-max theorem (Courant-Fischer theorem)

$A : N \times N$, Hermitian matrix

Suppose its eigenvalues are $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_N$.

$$\lambda_k = \min_{\boldsymbol{S};\dim(\boldsymbol{S})\leq k-1} \max_{\vec{x}\in\boldsymbol{S}^\perp;\|\vec{x}\|=1} \vec{x}^* \cdot A\vec{x}$$

$$\boldsymbol{S}^\perp = \{\vec{x} : \vec{x}^* \cdot \vec{y} = 0, \vec{y} \in \boldsymbol{S}\}$$

Orthonormal complement（直交補空間）

We can prove this by considering vector subspace spanned by eigenvectors. (see references)

Intuitive examples:

Maximum appears for the eigenvector.

$$A\vec{u}_i = \lambda_i \vec{u}_i$$

$$\lambda_1 = \max_{\vec{x}\in\boldsymbol{C}^N;\|\vec{x}\|=1} \vec{x}^* \cdot (A\vec{x}) \Rightarrow \vec{x} = \vec{u}_1$$

$$\lambda_2 = \max_{\vec{x}\in\boldsymbol{C}^N;\vec{x}\perp\vec{u}_1,\|\vec{x}\|=1} \vec{x}^* \cdot (A\vec{x}) = \min_{\boldsymbol{S};\dim(\boldsymbol{S})\leq 1} \max_{\vec{x}\in\boldsymbol{S}^\perp;\|\vec{x}\|=1} \vec{x}^* \cdot (A\vec{x}) \Rightarrow \vec{x} = \vec{u}_2$$

Properties of SVD 4 (optional) $A = U\Sigma V^\dagger$

4. Min-max theorem (Courant-Fischer theorem)

$A : M \times N$

Suppose its singular values are $\sigma_1 \geq \sigma_2 \geq \cdots$

$$\sigma_k = \min_{\boldsymbol{S};\dim(\boldsymbol{S})\leq k-1} \max_{\vec{x}\in\boldsymbol{S}^\perp;\|\vec{x}\|=1} \|A\vec{x}\|$$

By setting k=1,

$$\sigma_1 = \max_{\vec{x}\in\boldsymbol{C}^N,\|\vec{x}\|=1} \|A\vec{x}\|$$

which means

$$\|A\vec{x}\| \leq \sigma_1 \|\vec{x}\|$$

for $\vec{x} \in \boldsymbol{C}^N$

We can easily prove this by using

$A^\dagger A$: Hermitian

$$A^\dagger A \vec{v}_i = \lambda_i$$

$$\sigma_i = \sqrt{\lambda_i}$$

Properties of SVD 5 (optional) $A = U\Sigma V^\dagger$

5. Singular values for multiplication and addition

$\sigma_i(A)$: singular value of matrix A

(for $i > \mathrm{rank}(A)$, we set $\sigma_i = 0$)

*Following properties can be proven by using min-max theorem.

Multiplication: $A : M \times L, B : L \times N$

$$\sigma_k(AB) \leq \sigma_1(A)\sigma_k(B) \qquad (k = 1, 2, \dots)$$

$$(\sigma_k(AB) \leq \sigma_k(A)\sigma_1(B))$$

$$\Rightarrow \quad \mathrm{rank}(AB) \leq \min(\mathrm{rank}(A), \mathrm{rank}(B))$$

Addition: $A, B : M \times N$

$$\sigma_{k+j-1}(A+B) \leq \sigma_k(A) + \sigma_j(B) \qquad (k, j = 1, 2, \dots)$$

$$(\sigma_{k+j-1}(A+B) \leq \sigma_j(A) + \sigma_k(B))$$

$\Rightarrow$ If $\mathrm{rank}(B) \leq r$,

$$\sigma_{k+r}(A+B) \leq \sigma_k(A)$$

Libraries for SVD

There are **LAPACK** routines for SVD.

*Linear Algebra PACKage

At *netlib.org* (reference implementations)

+

A lot of vender implementations

- Intel MKL
- Apple Accelerate Framework
- Fujitsu SSLII
- ...

DGESDD, ZGESDD

DGESVD, ZGESVD

(For dense matrices)

numpy and **scipy** modules in python have routines for SVD.

numpy.linalg.svd

scipy.linalg.svd

(For dense matrices)

scipy.sparse.linalg.svds

(For sparse matrices or calculation of partial singular values)

Computation cost

For a $M \times N$ matrix ($M \leqq N$):

Full SVD: $O(NM^2)$

Partial SVD: $O(NMk)$

k : # of singular values to be calculated

Generalized inverse matrix

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 $\Longleftrightarrow$ $\mathrm{rank}(A) = N$

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

Generalized inverse matrix

Generalized inverse matrix（一般化逆行列）：

For $A : M \times N$, a matrix $A^{-} : N \times M$ satisfying

$$AA^{-}A = A$$

is called generalized inverse matrix.

Properties:

- Generalized inverse matrix is not unique.
  - At least one generalized matrix exists.

- If A is a regular matrix, $A^{-} = A^{-1}$

A^{-} is a generalization of inverse matrix.

Moore-Penrose pseudo inverse

Moore-Penrose pseudo inverse matrix（擬似逆行列）：

For $A : M \times N$, a matrix $A^+ : N \times M$ satisfying

(1) $AA^+A = A$ (2) $A^+AA^+ = A^+$

(3) $(AA^+)^\dagger = AA^+$ (4) $(A^+A)^\dagger = A^+A$

is called (Moore-Penrose) pseudo inverse matrix.

Relation to SVD

- Pseudo inverse is unique and calculated from SVD.

$$A = U\Sigma V^\dagger = U \begin{pmatrix} \Sigma_{r\times r} & 0_{r\times(N-r)} \\ 0_{(M-r)\times r} & 0_{(M-r)\times N-r} \end{pmatrix} V^\dagger$$

$$\Rightarrow \quad A^+ = V \begin{pmatrix} \Sigma^{-1}_{r\times r} & 0_{r\times(M-r)} \\ 0_{(N-r)\times r} & 0_{(N-r)\times M-r} \end{pmatrix} U^\dagger$$

$$A^+A = V \begin{pmatrix} \Sigma^{-1}_{r\times r} & 0_{r\times(M-r)} \\ 0_{(N-r)\times r} & 0_{(N-r)\times M-r} \end{pmatrix} U^\dagger U \begin{pmatrix} \Sigma_{r\times r} & 0_{r\times(N-r)} \\ 0_{(M-r)\times r} & 0_{(M-r)\times N-r} \end{pmatrix} V^\dagger$$

$$= \sum_{i=1}^{r} \vec{v}_i \vec{v}_i^\dagger \left(= \sum_{i=1}^{r} |v_i\rangle\langle v_i|\right)$$

A^+A is a projector onto img($A^\dagger$).
(AA^+ is a projector onto img(A).)

Simultaneous linear equation

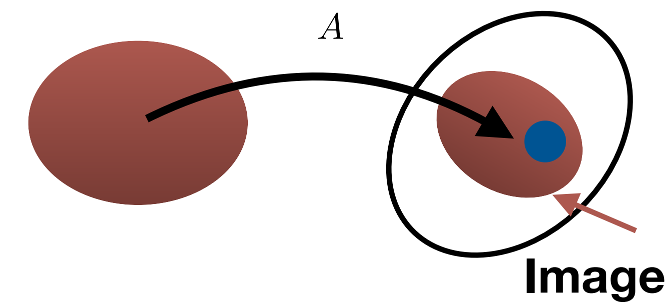
Simultaneous linear equation（連立一次方程式）

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

$\vec{b}$:

Two situations:

(1) There are solutions. ⟺ $\vec{b} \in \mathrm{img}(A)$



(i) There is the unique solution.

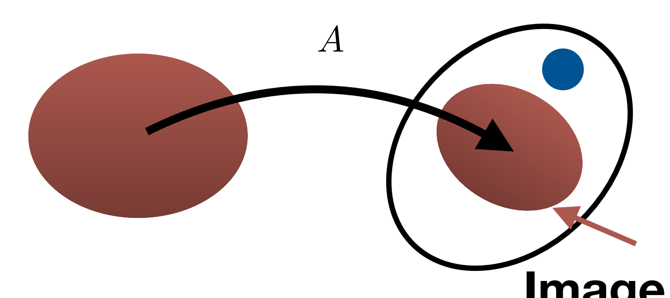
$$\mathrm{rank}(A) = N$$

(ii) There are infinite solutions (underdetermined).

$$\mathrm{rank}(A) < N$$

(We can add any vector $A\vec{y} = \vec{0}$.)

(2) There is no solution. ⟺ $\vec{b} \notin \mathrm{img}(A)$
(overdetermined)



Pseudo inverse and simultaneous linear equation

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

(1) There are solutions. $\Leftrightarrow$ $\vec{b} \in \mathrm{img}(A)$

- A vector defined by the pseudo inverse as

$$\vec{x}' \equiv A^+\vec{b}$$

is one of the solutions.

Because $\vec{b} \in \mathrm{img}(A)$,there exists $\vec{v} : A\vec{v} = \vec{b}$.

$$\Rightarrow \quad A\vec{x}' = AA^+\vec{b} = AA^+A\vec{v} = A\vec{v} = \vec{b}$$

- $\vec{x}'$ has the smallest norm $||\vec{x}'||$ among the solutions.

$$||\vec{x}|| \underline{\geq} ||A^+A\vec{x}|| = ||A^+\vec{b}|| = ||\vec{x}'||$$

$\because A^+A$ is a projector.

The pseudo inverse gives us the smallest norm solution.

Pseudo inverse and simultaneous linear equation

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

(2) There is no solution. $\Leftrightarrow$ $\vec{b} \notin \mathrm{img}(A)$

- A vector defined by the pseudo inverse as

$$\vec{x}' \equiv A^+\vec{b}$$

minimizes the "distance" $||\vec{b} - A\vec{x}||$.

$$\vec{y} = A\vec{c} \in \mathrm{img}(A), \vec{c} \in \mathbb{C}^N$$

$$||\vec{y} - \vec{b}||^2 = ||\underbrace{\vec{y} - AA^+\vec{b}}_{\in\, \mathrm{img}(A)} - \underbrace{(I - AA^+)\vec{b}}_{\in\, \mathrm{img}(A)^\perp}||^2$$

$$= ||\vec{y} - AA^+\vec{b}||^2 + ||\vec{b} - AA^+\vec{b}||^2$$

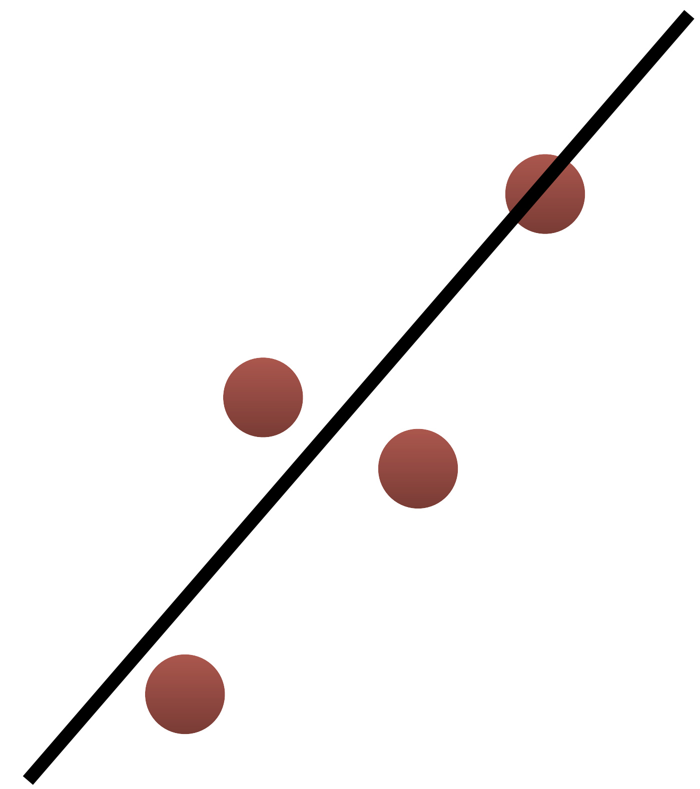
$$\geq ||\vec{b} - AA^+\vec{b}||^2 = ||\vec{b} - A\vec{x}'||^2$$

The pseudo inverse gives us approximate "least square solution".

Example of Least square solution problem

Fitting of a line to data points

$$y = ax + b$$



Data poins:

$$(x_1, y_1), (x_2, y_2), (x_3, y_3), (x_4, y_4)$$

$$\begin{pmatrix} x_1 & 1 \\ x_2 & 1 \\ x_3 & 1 \\ x_4 & 1 \end{pmatrix} \begin{pmatrix} a \\ b \end{pmatrix} = \begin{pmatrix} y_1 \\ y_2 \\ y_3 \\ y_4 \end{pmatrix}$$

$$A\vec{x} = \vec{b}$$

Least square fitting（最小二乗法）

$$\begin{pmatrix} a \\ b \end{pmatrix} = A^{+} \begin{pmatrix} y_1 \\ y_2 \\ y_3 \\ y_4 \end{pmatrix}$$

Next week

1st: Huge data in modern physics (Today)

2nd: Information compression in modern physics

(+review of linear algebra)

3rd: Review of linear algebra (+ singular value decomposition)

4th: Singular value decomposition and low rank approximation

5th: Basics of sparse modeling

6th: Basics of Krylov subspace methods

7th: Information compression in materials science

8th: Accelerating data analysis: Application of sparse modeling

9th: Data compression: Appication of Krylov subspace method

10th: Entanglement of information and matrix product states

11th: Application of MPS to eigenvalue problems

12th: Tensor network representation

13th: Information compression by tensor network renormalization