

---

# Dynamic Mode Decomposition

岡田崇

2026年6月14日

# 目次

---

このノートの目的	1
1 因果関係とは	1
2 Dynamic Mode Decomposition	2
2.1 基本的な考え方	2
2.2 データ行列	2
2.3 DMD は最小二乗問題である	3
2.4 準備: Moore–Penrose Pseudoinverse(疑似逆行列)	4
2.5 導出 1: $X$ が行フルランクの時	5
2.6 SVD から見た最小ノルム解	6
2.7 DMD モードと固有値の意味	8
3 実務的には	9
3.1 Bootstrapping	9
3.2 Regularization	9
4 発展的な内容	10
4.1 Kalman Filter	10
4.2 Koopman Operator	11
4.3 Sparse Identification of Nonlinear Dynamics, SINDy	12
5 まとめ	13
A Singular Value Decomposition と Moore–Penrose 疑似逆行列	14
A.1 SVD とは何か	14
A.2 なぜ SVD が存在するのか	15
A.3 Moore–Penrose Inverse	17
A.4 フルランクの場合の明示公式	17
A.5 [補足] コンパクト SVD から通常の SVD へ	20

参考文献 ..... 21

## このノートの目的

---

このノートでは、Dynamic Mode Decomposition、略して DMD を説明する。概要は以下の通りである。

1. 因果関係と同時相関の違いを確認する。
2. DMD を「線形の力学系を仮定した最小二乗問題」として理解する。
3. 実務で重要になる Bootstrapping と Regularization を整理する。
4. Kalman Filter、Koopman Operator、SINDy との関係を概観する。

## 1 因果関係とは

---

データ分析で最初に注意したいのは、相関があることと、因果関係があることは違うという点である。

たとえば、2つの時系列  $x_t$  と  $y_t$  があり、同じ時刻  $t$  において強く相関していたとする。

$$\text{corr}(x_t, y_t) \text{ が大きい}$$

しかし、これだけでは次のどれが起きているのか分からない。

- ▶  $x$  が  $y$  の原因になっている。
- ▶  $y$  が  $x$  の原因になっている。
- ▶ 第三の変数  $z$  が  $x$  と  $y$  の両方を動かしている。
- ▶ 季節性、トレンド、周期性などによって、たまたま同じように動いている。
- ▶ 因果効果はあるが、同時刻ではなく遅れて現れている。

### 注意

同時相関だけでは因果関係は分からない。

$x_t$  と  $y_t$  が同時に動いていても、どちらがどちらを動かしたのか、あるいは別の変数が両方を動かしたのかは判断できない。

特に時系列データでは、同時刻の相関だけを見ると危険である。原因は結果より前に起きる必要が

あるため、少なくとも時間順序を考える必要がある。

因果を議論するには、少なくとも次の視点が重要になる。

1. **時間順序** 原因は結果より前に起きる必要がある。
2. **交絡の考慮** 第三の変数が両方を動かしていないかを確認する必要がある。
3. **介入または仮定** 「もし  $x$  を外から変えたら  $y$  は変わるか」という視点が重要になる。

DMD が推定するのは

$$x_{t+1} \approx Ax_t$$

という形の線形時間発展モデルである。隠れた交絡変数などがある場合には、やはり DMD の  $A$  をそのまま因果関係として解釈するのは危険であるが、DMD は「時間方向の予測構造」を推定する手法であり、同時相関よりは因果に近い情報を含む。

## 2 Dynamic Mode Decomposition

### 2.1 基本的な考え方

DMD は、時系列データから 1 ステップ先の時間発展を表す線形写像を推定する方法である。基本的な仮定は

#### 基本式

$$x_{t+1} \approx Ax_t$$

である。ここで、 $x_t \in \mathbb{R}^n$  は時刻  $t$  における状態ベクトル、 $A \in \mathbb{R}^{n \times n}$  は状態を 1 ステップ先に進める行列である。

たとえば、気温、圧力、速度場、売上、センサー値などをまとめたものを  $x_t$  と考える。DMD では、それらの変化をできるだけ一つの線形写像  $A$  で説明する。

### 2.2 データ行列

時系列データとして

$$x_0, x_1, \dots, x_N$$

が得られているとする。DMD では、現在側のスナップショットと次時刻側のスナップショットを

$$X = \begin{bmatrix} | & | & & | \\ x_0 & x_1 & \cdots & x_{N-1} \\ | & | & & | \end{bmatrix}, \quad X' = \begin{bmatrix} | & | & & | \\ x_1 & x_2 & \cdots & x_N \\ | & | & & | \end{bmatrix}$$

と並べる。ここで

$$X, X' \in \mathbb{R}^{n \times N}$$

である。

各時刻で

$$x_{k+1} \approx Ax_k$$

が成り立つと仮定すると、すべての関係式はまとめて

$$X' \approx AX$$

と書ける。したがって、DMD の目標は、この近似を最もよく満たす行列  $A$  を求めることである。

### 2.3 DMD は最小二乗問題である

$A$  に特別な制約を課さない場合、DMD は次の最小二乗問題として定式化される。

#### DMD の推定問題

$$\min_A \|X' - AX\|_F^2$$

ここで

$$\|M\|_F^2 = \sum_i \sum_k M_{ik}^2$$

は Frobenius norm であり、行列の全成分の二乗和を表す。

この問題の標準的な解は

#### Pseudoinverse による解

$$A = X'X^+$$

である。ここで  $X^+$  は  $X$  の Moore–Penrose 疑似逆行列である。

もし理想的に

$$X' = AX$$

が完全に成り立ち、さらに  $X$  が正方行列で逆行列を持つなら、

$$A = X'X^{-1}$$

と書ける。しかし実際には、 $X \in \mathbb{R}^{n \times N}$  は正方行列とは限らない。変数の数  $n$  とデータ数  $N$  は一般に一致せず、またノイズやデータの重複によってランクが落ちることもある。そのため、通常の逆行列の代わりに疑似逆行列  $X^+$  が現れる。

## 2.4 準備: Moore–Penrose Pseudoinverse(疑似逆行列)

Moore–Penrose inverse と singular value decomposition についてまとめておく (より詳細な議論は Appendix を参照)。

行列  $X \in \mathbb{R}^{n \times N}$  に対して、Moore–Penrose 疑似逆行列  $X^+ \in \mathbb{R}^{N \times n}$  は、次の 4 条件を満たす唯一の行列として定義される。

$$\begin{aligned} XX^+X &= X, & X^+XX^+ &= X^+, \\ (XX^+)^{\top} &= XX^+, & (X^+X)^{\top} &= X^+X. \end{aligned}$$

これらを Penrose conditions と呼ぶ。

特に、 $X$  が正方かつ正則なら

$$X^+ = X^{-1}$$

である。つまり、Moore–Penrose 疑似逆行列は、通常の逆行列を正方・正則でない行列にも拡張したものである。

実際の計算では、Singular Value Decomposition(SVD) を用いて定義するのが便利である。 $X$  のランクを  $r$  とする。普通の SVD では、ゼロ特異値に対応する成分も形式的に含めて

$$X = U\Sigma V^{\top}$$

と書く。しかし、ゼロ特異値に対応する部分は、積  $U\Sigma V^{\top}$  の中で結局ゼロをかけるだけなので、 $X$  の再構成には寄与しない。そこで、ゼロでない特異値だけを残して

$$X = U_r \Sigma_r V_r^{\top}$$

と書く。<sup>1)</sup>

---

1) この形を

bfseries コンパクト SVD という。ゼロ部分を省いただけなので、これは近似ではなく厳密な等式である。

## 2.5 導出 1: $X$ が行フルランクの時

目的関数を

$$J(A) = \frac{1}{2} \|X' - AX\|_F^2$$

と置く。成分で書くと、

$$(AX)_{ik} = \sum_{\ell=1}^n A_{i\ell} X_{\ell k}$$

なので、

$$J(A) = \frac{1}{2} \sum_{i=1}^n \sum_{k=1}^N \left( X'_{ik} - \sum_{\ell=1}^n A_{i\ell} X_{\ell k} \right)^2$$

である。

これを  $A_{ij}$  で微分する。 $A_{ij}$  は

$$\sum_{\ell=1}^n A_{i\ell} X_{\ell k}$$

のうち  $\ell = j$  の項にだけ現れるので、

$$\frac{\partial J}{\partial A_{ij}} = - \sum_{k=1}^N \left( X'_{ik} - \sum_{\ell=1}^n A_{i\ell} X_{\ell k} \right) X_{jk}$$

となる。

最小値では、すべての  $i, j$  について

$$\frac{\partial J}{\partial A_{ij}} = 0$$

が成り立つ。したがって、

$$\sum_{k=1}^N \left( \sum_{\ell=1}^n A_{i\ell} X_{\ell k} - X'_{ik} \right) X_{jk} = 0$$

である。

この式を行列でまとめると、

$$(AX - X')X^\top = 0$$

となる。したがって

$$AXX^\top = X'X^\top \tag{1}$$

を得る。この式を正規方程式 (normal equations) と呼ぶ。

もし  $XX^\top$  が逆行列を持つならば (等価なこととして、 $X$  が行フルランクならば<sup>2)</sup>)、

$$A = X'X^\top (XX^\top)^{-1}$$

---

2) ノイズやランダムな摂動が入る実データでは、観測時点の数が変数の数に比べて十分多い時には  $X$  は行フルランクになることが多い。

である。行フルランクな  $X$  については

$$X^+ = X^\top (XX^\top)^{-1}$$

が成り立つので、Moore-Penrose inverse を用いて、

$$A = X'X^+$$

とも書ける。

また、 $X$  が行フルランクでない場合には、最小二乗問題

$$\min_A \|X' - AX\|_F^2$$

の最適解は一般に一意ではない。実際、ある最適解を  $A_0$  とし、 $CX = 0$  を満たす行列  $C$  を加えると、

$$(A_0 + C)X = A_0X$$

であるため、 $A_0 + C$  も同じ予測誤差を与える。このように最適解が複数存在する場合でも、Moore-Penrose 疑似逆行列を用いた式 (2.5) は最適解の一つを与え、特にすべての最適解の中で

$$\|A\|_F$$

が最小となる解である。つまり、データから決まらない方向に余計な成分を入れない、最小ノルムの代表解を選んでいる（次節）。

## 2.6 SVD から見た最小ノルム解

ここでは、SVD を用いて

$$A_* = X'X^+$$

が最小二乗解であり、さらにその中で最小ノルム解であることを確認する。

$X$  のコンパクト SVD を

$$X = U_r \Sigma_r V_r^\top$$

と書く。このとき Moore-Penrose 疑似逆行列は

$$X^+ = V_r \Sigma_r^{-1} U_r^\top$$

である。したがって

$$A_* = X'X^+ = X'V_r \Sigma_r^{-1} U_r^\top$$

である。

まず、 $A_\star$  が正規方程式 (1) を満たすことを確認する。

$$XX^\top = U_r \Sigma_r^2 U_r^\top$$

なので、正規方程式 (1) の左辺は

$$\mathit{LHS} = A_\star XX^\top = X' V_r \Sigma_r^{-1} U_r^\top U_r \Sigma_r^2 U_r^\top = X' V_r \Sigma_r U_r^\top.$$

一方、

$$X^\top = V_r \Sigma_r U_r^\top$$

より、正規方程式 (1) の右辺は

$$\mathit{RHS} = X' X^\top = X' V_r \Sigma_r U_r^\top.$$

したがって、 $A_\star$  が正規方程式を満たすことは確認できた。

次に、この解が最小ノルム解であることを見る。 $U = (U_r \ U_0)$  を直交行列となるように取る。ここで  $U_0$  は  $U_r$  に直交する方向の正規直交基底である。任意の最小二乗解  $A$  は正規方程式

$$AXX^\top = X'X^\top$$

を満たす。 $X = U_r \Sigma_r V_r^\top$  を代入すると

$$A U_r \Sigma_r^2 U_r^\top = X' V_r \Sigma_r U_r^\top$$

である。右から  $U_r$  をかけて整理すると

$$A U_r = X' V_r \Sigma_r^{-1}$$

を得る。したがって、最小二乗解  $A$  の  $U_r$  方向での作用は一意に決まる。

一方、 $U_0^\top X = 0$  なので、 $U_0$  方向での作用は  $AX$  に現れない。よって、最小二乗解は

$$A = \underbrace{X' V_r \Sigma_r^{-1} U_r^\top}_{=A_\star} + C U_0^\top$$

と書ける。ここで  $C$  は任意の行列である。

また、

$$A = A_\star + C U_0^\top$$

なので、Frobenius norm の展開より

$$\|A\|_F^2 = \|A_\star + C U_0^\top\|_F^2$$

$$= \|A_\star\|_F^2 + \|CU_0^\top\|_F^2 + 2 \operatorname{tr}(A_\star(CU_0^\top)^\top). \quad (2)$$

ここで交差項は

$$\operatorname{tr}(A_\star(CU_0^\top)^\top) = \operatorname{tr}(A_\star U_0 C^\top)$$

である。一方、

$$A_\star = X'V_r\Sigma_r^{-1}U_r^\top$$

であり、 $U_r^\top U_0 = 0$  なので

$$A_\star U_0 = X'V_r\Sigma_r^{-1}U_r^\top U_0 = 0.$$

したがって交差項は消え、

$$\|A\|_F^2 = \|A_\star\|_F^2 + \|CU_0^\top\|_F^2$$

となる。したがって、 $\|A\|_F$  が最小になるのは  $C = 0$  のときである。

ゆえに、

$$A_\star = X'X^+$$

は最小二乗解の中で Frobenius norm が最も小さい解である。

## 2.7 DMD モードと固有値の意味

推定した行列  $A$  の固有値・固有ベクトルを見ると、システムの時間発展を分解できる。

$$A\phi_j = \lambda_j\phi_j.$$

ここで、 $\phi_j$  は DMD モード、 $\lambda_j$  は DMD 固有値である。

ざっくり言うと、DMD モード  $\phi_j$  は「空間的なパターン」、固有値  $\lambda_j$  は「そのパターンが時間とともにどう変化するか」を表す。

条件	解釈
$ \lambda_j  > 1$	そのモードは成長する。
$ \lambda_j  < 1$	そのモードは減衰する。
$ \lambda_j  \approx 1$	そのモードは持続しやすい。
$\lambda_j$ が複素数	振動や周期性を表す。

**見方.** DMD は、複雑な時系列を「成長・減衰・振動するパターンの足し合わせ」として理解する方法だと考えると分かりやすい。

### 3 実務的には

---

実務で DMD を使うときは、単に

$$A = X'X^+$$

を計算するだけでは不十分なことが多い。現実のデータには、ノイズ、外れ値、サンプル不足、モデルのミスマッチがあるからである。

そこで重要になるのが、Bootstrapping と Regularization である。

#### 3.1 Bootstrapping

Bootstrapping は、データを再サンプリングして同じ分析を何度も行う方法である。DMD では、たとえば次のように使う。

1. 元の時系列データから複数のサンプルを作る。
2. それぞれに対して DMD を実行する。
3. 得られた固有値、モード、予測結果のばらつきを見る。

これにより、次のようなことが分かる。

- ▶ある DMD モードが安定して出てくるか。
- ▶固有値の推定がどれくらい不確かか。
- ▶予測結果にどれくらい幅があるか。
- ▶ノイズやサンプルの取り方に敏感すぎないか。

**実務メモ.** Bootstrapping の目的は、単に 1 つの答えを出すことではない。その答えがどれくらい安定しているか、どれくらい信用できるかを見ることである。

特に時系列データでは、データ点を完全にランダムにシャッフルすると時間構造が壊れることがある。そのため、連続した時間区間をまとめて再サンプリングする block bootstrap のような考え方が使われることもある。

#### 3.2 Regularization

Regularization、つまり正則化は、過学習や数値的不安定性を抑えるための工夫である。

DMD では、疑似逆行列を計算するときに小さい特異値が問題になる。小さい特異値を反転すると、ノイズが大きく増幅される。そこで、代表的には次のような方法を使う。

### 3.2.1 低ランク近似

SVD を使って、重要な成分だけを残す。

$$X \approx U_r \Sigma_r V_r^\top.$$

ここで  $r$  は残すランクである。小さい特異値に対応する成分を捨てることで、ノイズの影響を減らす。

**実務メモ.** DMD における最も基本的な正則化の一つは、SVD による低ランク近似である。ランクを大きくしすぎるとノイズを拾い、小さくしすぎると重要な構造を捨てる。

### 3.2.2 Ridge 正則化

最小二乗問題にペナルティを加える。

$$\min_A \|X' - AX\|_F^2 + \alpha \|A\|_F^2.$$

ここで  $\alpha > 0$  は正則化の強さである。 $\alpha$  が大きいほど、行列  $A$  が過度に大きくなることを防ぐ。

### 3.2.3 Sparse 正則化

場合によっては、行列  $A$  の多くの要素がゼロに近いと仮定することもある。これは「少数の重要な相互作用だけがある」と考える場合に有効である。

ただし、スパース性を入れると解釈しやすくなる一方で、モデルの仮定も強くなる。

## 4 発展的な内容

---

ここでは、DMD の周辺にある重要な考え方を簡単に紹介する。

### 4.1 Kalman Filter

Kalman Filter は、状態空間モデルに基づく推定手法である。DMD では基本的に

$$x_{t+1} = Ax_t$$

のような決定論的なモデルを考える。一方、Kalman Filter では

$$x_{t+1} = Ax_t + w_t, \quad y_t = Cx_t + v_t$$

のように書く。

ここで、 $x_t$  は真の状態、 $y_t$  は観測値、 $w_t$  はプロセスノイズ、 $v_t$  は観測ノイズ、 $C$  は状態から観測への写像である。

**発展メモ.** Kalman Filter は、**確率過程と観測誤差を考慮できる**。DMD で推定した  $A$  を状態遷移行列として使い、Kalman Filter でノイズを考慮した状態推定を行う、という考え方もできる。

現実には、私たちが観測しているデータは真の状態そのものではないことが多い。センサー誤差、欠測、外乱などがある。Kalman Filter は「モデルからの予測」と「実際の観測による修正」を繰り返す方法である。

## 4.2 Koopman Operator

DMD を少し深く理解するには、Koopman Operator の考え方が重要になる。

通常、非線形の力学系は

$$x_{t+1} = F(x_t)$$

のように書かれる。ここで  $F$  は非線形な関数である。

Koopman Operator の考え方では、状態  $x$  そのものではなく、状態を観測する関数  $g(x)$  に注目する。たとえば、 $g(x) = x$  だけでなく、 $g(x) = x^2$ 、 $g(x) = \sin x$ 、 $g(x) = x_1 x_2$  のような観測量を考える。

Koopman Operator は、観測量の時間発展を表す作用素である。

$$(Kg)(x) = g(F(x)).$$

**発展メモ.** 元の力学系  $F$  が非線形でも、Koopman Operator  $K$  は線形作用素になる。ただし、その代わりに、通常は無限次元の関数空間を考える必要がある。

DMD は、この Koopman Operator を有限次元で近似していると見ることができる。つまり、DMD は単なる線形回帰としても理解できるが、より深くは「非線形力学系を、観測量の空間で線形に見る方法」とも理解できる。

この考え方を拡張したものに Extended DMD、略して EDMD がある。EDMD では、状態  $x$  だけでなく、さまざまな特徴量や基底関数を使って、より豊かな空間で線形モデルを作る。

### 4.3 Sparse Identification of Nonlinear Dynamics, SINDy

SINDy は、非線形力学系の支配方程式をデータから発見する方法である。

DMD が基本的に

$$x_{t+1} \approx Ax_t$$

という線形モデルを仮定するのに対して、SINDy では

$$\frac{dx}{dt} = f(x)$$

の形の非線形方程式を探す。

ただし、完全に自由な非線形関数  $f$  を探すのは難しい。そこで SINDy では、候補となる関数をあらかじめたくさん用意する。たとえば、 $1$ 、 $x$ 、 $x^2$ 、 $x^3$ 、 $\sin x$ 、 $\cos x$  などである。

そして、

$$\frac{dx}{dt} \approx \Theta(x)\xi$$

と書く。ここで、 $\Theta(x)$  は候補関数を並べたライブラリ、 $\xi$  は各候補関数の係数である。

**発展メモ.** SINDy の重要な仮定は、**本当に必要な項は少数である**ということである。多くの候補関数の中から、重要なものだけをスパース回帰で選ぶ。

DMD と SINDy の違いは次のように整理できる。

手法	基本仮定	得られるもの
DMD	状態は線形写像で進む	線形時間発展行列、モード、固有値
SINDy	支配方程式は少数の項で書ける	非線形の微分方程式

SINDy は解釈しやすい方程式を得られる可能性がある一方で、微分の推定やノイズの影響に注意が必要である。

## 5 まとめ

---

DMD は、時系列データから

$$x_{t+1} \approx Ax_t$$

という線形の力学系を推定する方法である。基本的には

$$X' \approx AX$$

という最小二乗問題であり、その解は

$$A = X'X^+$$

と疑似逆行列を使って書ける。

### 重要ポイント

DMD は、時系列データから線形の時間発展を最小二乗で推定する方法である。

ただし、実データではノイズやサンプル不足により、疑似逆行列の計算が不安定になりやすい。そのため、低ランク近似、Ridge 正則化、Bootstrapping などが重要になる。

また、DMD は因果推論の手法ではない。同時相関だけを見るよりは時間構造を使っているが、DMD で得られた行列をそのまま因果関係と解釈するのは危険である。

発展的には、Kalman Filter、Koopman Operator、SINDy が DMD と関連している。

- ▶ **Kalman Filter**: 確率過程と観測誤差を考慮した状態推定。
- ▶ **Koopman Operator**: 非線形力学系を観測量の空間で線形に見る理論。
- ▶ **SINDy**: スパース回帰によって非線形支配方程式を発見する方法。

## A Singular Value Decomposition と Moore–Penrose 疑似逆行列

---

この付録では、DMD でよく用いられる特異値分解

$$X = U\Sigma V^{\top}$$

について説明する。特に、コンパクト SVD

$$X = U_r \Sigma_r V_r^{\top}$$

が何を意味するのかを整理する。

### A.1 SVD とは何か

任意の行列

$$X \in \mathbb{R}^{n \times N}$$

に対して、直交行列

$$U \in \mathbb{R}^{n \times n}, \quad V \in \mathbb{R}^{N \times N}$$

と、対角成分が非負の行列

$$\Sigma \in \mathbb{R}^{n \times N}$$

を用いて

$$X = U\Sigma V^{\top}$$

と分解できる。この分解を **Singular Value Decomposition**、または **SVD** という。

ここで、 $\Sigma$  の対角成分

$$\sigma_1, \sigma_2, \dots$$

を **特異値** という。通常は

$$\sigma_1 \geq \sigma_2 \geq \dots \geq 0$$

となるように並べる。

行列  $U$  の列ベクトルを

$$u_1, u_2, \dots, u_n$$

とし、行列  $V$  の列ベクトルを

$$v_1, v_2, \dots, v_N$$

と書く。すると、SVD は直感的には次のような意味を持つ。

$$\text{入力側の方向 } v_i \quad \mapsto \quad \text{出力側の方向 } u_i$$

より具体的には、 $\sigma_i > 0$  のとき

$$Xv_i = \sigma_i u_i$$

が成り立つ。つまり、 $X$  は入力側の方向  $v_i$  を、出力側の方向  $u_i$  に写し、その大きさを  $\sigma_i$  倍する。

## A.2 なぜ SVD が存在するのか

SVD が存在する理由は、行列

$$X^T X$$

に注目すると分かりやすい。まず、 $X^T X$  は対称行列である。実際、

$$(X^T X)^T = X^T X$$

である。また、任意のベクトル  $z$  に対して

$$z^T X^T X z = \|Xz\|_2^2 \geq 0$$

なので、 $X^T X$  は半正定値である。対称行列は直交対角化できるため、ある直交行列  $V$  と非負の固有値を並べた対角行列

$$\Lambda = \text{diag}(\lambda_1, \dots, \lambda_N)$$

によって

$$X^T X = V \Lambda V^T$$

と書ける。ここで  $V$  の列ベクトル  $v_i$  は、 $X^T X$  の固有ベクトルであり、互いに正規直交するよう選ぶ。固有値が縮退している場合、固有ベクトルの選び方は一意ではないが、縮退した固有空間の中で正規直交基底を選ぶことができる。したがって

$$v_i^T v_j = \delta_{ij}$$

としてよい。固有ベクトル  $v_i$  に対応する固有値を

$$\lambda_i = \sigma_i^2$$

と書く。すると

$$X^T X v_i = \sigma_i^2 v_i$$

である。ここで、 $\sigma_i > 0$  のものだけを考える。このとき

$$u_i = \frac{Xv_i}{\sigma_i}$$

と定義できる。すると

$$Xv_i = \sigma_i u_i$$

が成り立つ。さらに、このようにして得られる  $u_i$  たちは互いに正規直交する。実際、 $\sigma_i, \sigma_j > 0$  のとき

$$\begin{aligned} u_i^\top u_j &= \frac{(Xv_i)^\top (Xv_j)}{\sigma_i \sigma_j} \\ &= \frac{v_i^\top X^\top X v_j}{\sigma_i \sigma_j} \\ &= \frac{v_i^\top (\sigma_j^2 v_j)}{\sigma_i \sigma_j} \\ &= \frac{\sigma_j^2}{\sigma_i \sigma_j} v_i^\top v_j = \delta_{ij}. \end{aligned} \tag{3}$$

最後の等号では、 $v_i$  たちが正規直交系をなすことを用いた。したがって、 $\sigma_i > 0$  に対応する  $u_i$  たちも正規直交系をなす。正の特異値の個数を

$$r = \text{rank}(X)$$

とする。正の特異値に対応する右特異ベクトルを並べて

$$V_r = \begin{bmatrix} | & | & & | \\ v_1 & v_2 & \cdots & v_r \\ | & | & & | \end{bmatrix} \in \mathbb{R}^{N \times r},$$

左特異ベクトルを並べて

$$U_r = \begin{bmatrix} | & | & & | \\ u_1 & u_2 & \cdots & u_r \\ | & | & & | \end{bmatrix} \in \mathbb{R}^{n \times r},$$

さらに

$$\Sigma_r = \text{diag}(\sigma_1, \dots, \sigma_r) \in \mathbb{R}^{r \times r}$$

とおく。各  $i = 1, \dots, r$  について

$$Xv_i = \sigma_i u_i$$

が成り立つので、まとめて

$$XV_r = U_r\Sigma_r$$

と書ける。ここで、 $\sigma_i = 0$  に対応する右特異ベクトルは  $\ker X$  に属するため、 $X$  によってゼロに写る。したがって、 $X$  の再構成には正の特異値に対応する部分だけを残せばよい。よって

$$X = U_r\Sigma_rV_r^\top$$

を得る。これがコンパクト SVD である。

### A.3 Moore–Penrose Inverse

コンパクト SVD を使うと、Moore–Penrose Inverse は簡単に書ける。

$$X = U_r\Sigma_rV_r^\top$$

のとき、

$$X^+ = V_r\Sigma_r^{-1}U_r^\top$$

である。ここで

$$\Sigma_r^{-1} = \text{diag}\left(\frac{1}{\sigma_1}, \dots, \frac{1}{\sigma_r}\right)$$

である。この  $X^+$  が Moore–Penrose pseudoinverse の定義を満たしていることは容易に確認できる。ゼロ特異値はコンパクト SVD では最初から除かれているため、 $1/0$  のような割り算は現れない。これが、疑似逆行列を SVD で定義すると見通しがよい理由である。

### A.4 フルランクの場合の明示公式

Moore–Penrose 疑似逆行列は一般には SVD を用いて

$$X = U_r\Sigma_rV_r^\top, \quad X^+ = V_r\Sigma_r^{-1}U_r^\top$$

と書ける。ここでは、 $X$  が列フルランクまたは行フルランクの場合に、より簡単な形で  $X^+$  を書くことを確認する。

#### A.4.1 列フルランクの場合

まず、

$$X \in \mathbb{R}^{n \times N}$$

が列フルランクである場合を考える。これは

$$\text{rank}(X) = N$$

という意味であり、 $X$  の列ベクトルが一次独立であることを表す。このとき  $N \leq n$  である。

列フルランクなら、

$$X^T X \in \mathbb{R}^{N \times N}$$

は正則である。実際、任意の  $z \in \mathbb{R}^N$  に対して

$$z^T X^T X z = \|Xz\|_2^2$$

である。もし  $z \neq 0$  なら、列フルランク性より  $Xz \neq 0$  なので、

$$z^T X^T X z > 0$$

である。したがって  $X^T X$  は正定値であり、逆行列を持つ。

このとき、Moore–Penrose 疑似逆行列は

$$X^+ = (X^T X)^{-1} X^T$$

と書ける。

実際、

$$X^+ X = (X^T X)^{-1} X^T X = I_N$$

である。つまり、列フルランクの場合、 $X^+$  は  $X$  の左逆行列のように働く。

この公式は SVD から確認できる。列フルランクのとき  $r = N$  なので、

$$X = U_N \Sigma_N V_N^T$$

と書ける。このとき

$$X^T X = V_N \Sigma_N^2 V_N^T$$

であるから、

$$(X^T X)^{-1} X^T = V_N \Sigma_N^{-2} V_N^T V_N \Sigma_N U_N^T = V_N \Sigma_N^{-1} U_N^T = X^+.$$

#### A.4.2 行フルランクの場合

次に、 $X$  が行フルランクである場合を考える。これは

$$\text{rank}(X) = n$$

という意味であり、 $X$  の行ベクトルが一次独立であることを表す。このとき  $n \leq N$  である。  
行フルランクなら、

$$XX^T \in \mathbb{R}^{n \times n}$$

は正則である。実際、任意の  $y \in \mathbb{R}^n$  に対して

$$y^T XX^T y = \|X^T y\|_2^2$$

である。もし  $y \neq 0$  なら、行フルランク性より  $X^T y \neq 0$  なので、

$$y^T XX^T y > 0$$

である。したがって  $XX^T$  は正定値であり、逆行列を持つ。

このとき、Moore–Penrose 疑似逆行列は

$$X^+ = X^T (XX^T)^{-1}$$

と書ける。

実際、

$$XX^+ = XX^T (XX^T)^{-1} = I_n$$

である。つまり、行フルランクの場合、 $X^+$  は  $X$  の右逆行列のように働く。

SVD から同じ公式が得られる。行フルランクのとき  $r = n$  なので、

$$X = U_n \Sigma_n V_n^T$$

と書ける。このとき

$$XX^T = U_n \Sigma_n^2 U_n^T$$

であるから、

$$X^T (XX^T)^{-1} = V_n \Sigma_n U_n^T U_n \Sigma_n^{-2} U_n^T = V_n \Sigma_n^{-1} U_n^T = X^+.$$

### フルランクの場合の公式

$X \in \mathbb{R}^{n \times N}$  に対して、Moore–Penrose 疑似逆行列は、フルランクの場合には次のように明示的に書ける：

$$\text{rank}(X) = N \implies X^+ = (X^T X)^{-1} X^T \quad \text{列フルランクの場合,}$$

$$\text{rank}(X) = n \implies X^+ = X^T (XX^T)^{-1} \quad \text{行フルランクの場合.}$$

DMD で現れるデータ行列  $X \in \mathbb{R}^{n \times N}$  では、 $N$  が大きく  $X$  が行フルランクである場合、

$$X^+ = X^\top (X X^\top)^{-1}$$

という形が現れる。

## A.5 [補足] コンパクト SVD から通常の SVD へ

DMD の計算にはコンパクト SVD で十分であるが、ここでは通常の SVD の構成方法についても述べておく。

通常の SVD では、 $X \in \mathbb{R}^{n \times N}$  に対して、直交行列

$$U \in \mathbb{R}^{n \times n}, \quad V \in \mathbb{R}^{N \times N},$$

を用いて

$$X = U \Sigma V^\top$$

と書く。ここで

$$\Sigma \in \mathbb{R}^{n \times N}$$

は、 $\Sigma_r$  のほかにゼロ特異値に対応する部分を加えた行列である。

通常の SVD を作るには、 $U_r$  と  $V_r$  をそれぞれ直交行列になるように補完すればよい。具体的には、 $\ker X$  の正規直交基底を並べた行列を

$$V_0 \in \mathbb{R}^{N \times (N-r)}$$

とし、

$$V = (V_r \ V_0)$$

と置く。このとき  $V$  は  $N \times N$  の直交行列である。

同様に、 $\ker X^\top = \text{Range}(X)^\perp$  の正規直交基底を並べた行列を

$$U_0 \in \mathbb{R}^{n \times (n-r)}$$

とし、

$$U = (U_r \ U_0)$$

と置く。このとき  $U$  は  $n \times n$  の直交行列である。

最後に

$$\Sigma = \begin{pmatrix} \Sigma_r & 0 \\ 0 & 0 \end{pmatrix} \in \mathbb{R}^{n \times N}$$

と置けば、

$$U\Sigma V^\top = (U_r \ U_0) \begin{pmatrix} \Sigma_r & 0 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} V_r^\top \\ V_0^\top \end{pmatrix} = U_r \Sigma_r V_r^\top = X$$

となる。

## 参考文献

---

- [1]P. J. Schmid, “Dynamic mode decomposition of numerical and experimental data,” *Journal of Fluid Mechanics*, 656, 5–28, 2010.
- [2]J. H. Tu, C. W. Rowley, D. M. Luchtenburg, S. L. Brunton, and J. N. Kutz, “On dynamic mode decomposition: theory and applications,” *Journal of Computational Dynamics*, 1(2), 391–421, 2014.
- [3]B. O. Koopman, “Hamiltonian systems and transformation in Hilbert space,” *Proceedings of the National Academy of Sciences*, 17(5), 315–318, 1931.
- [4]R. E. Kalman, “A new approach to linear filtering and prediction problems,” *Journal of Basic Engineering*, 82(1), 35–45, 1960.
- [5]S. L. Brunton, J. L. Proctor, and J. N. Kutz, “Discovering governing equations from data by sparse identification of nonlinear dynamical systems,” *Proceedings of the National Academy of Sciences*, 113(15), 3932–3937, 2016.