

6. 多変量時間領域解析その2 : SVD 解析 1

6.1. SVD 解析の方法 1

6.2. SVD の計算例 2

6.3. Matlab での SVD プログラム 4

6.4. parallel formalism 4

6.5. その他 5

6.6. References 5

6. 多変量時間領域解析その2 : SVD 解析

SVD 解析 (特異値分解解析) は、異なる二つの物理量 (例えば SLP と SST) の間の関係を見出す手法である。標準的な SVD 解析では、二つの物理量間の相互共分散行列を、行列の特異値分解を行うことによって、各々の物理量の特異ベクトルを得る。

6.1. SVD 解析の方法

二つの物理量 A, B があり、物理量 A のデータは $Z_A (N \times P_A)$ 、他方は $Z_B (N \times P_B)$ という配列にデータが格納されているとしよう。ここで N は時間サンプル数、 $P_A(P_B)$ は A (B) の空間グリッド数である。時間サンプルの数 N は二つの物理量で同じでなくてはならない。

相互共分散行列(cross-covariance matrix) は、

$$C = Z_A^T Z_B \tag{6.1}$$

と定義される、 $(P_A \times P_B)$ の配列である。この行列を特異値分解、

$$C = X_A L X_B^T \tag{6.2}$$

することによって、物理量 A (B) について**特異ベクトル(singular vector)** $X_A (X_B)$ と、対角行列 L とが求められる。L の成分は**特異値(singular value)**と呼ばれる。この X_A, X_B はその m 列目が、SVD 解析の第 m モードの空間構造を表現する。それぞれの物理量の時間係数を T_A, T_B を、

$$T_A = Z_A X_A, T_B = Z_B X_B \tag{6.3}$$

と定義する。また X_A, X_B が正規化された直交行列 ($X_A^T X_A = X_A X_A^T = X_B^T X_B = X_B X_B^T = E$, E は単位行列)であることから、

$$Z_A = T_A X_A^T, Z_B = T_B X_B^T \tag{6.4}$$

と表すことができる。

異なる物理量の時間係数については、

$$T_A^T T_B = (Z_A X_A)^T (Z_B X_B) = X_A^T Z_A^T Z_B X_B = X_A^T C X_B = X_A^T X_A L X_B^T X_B$$

$$\therefore \boxed{T_A^T T_B = L} \tag{6.5}$$

およびこの転置を取った

$$\mathbf{T}_B^T \mathbf{T}_A = \mathbf{L} \tag{6.6}$$

が成立する。L は対角成分だけがゼロではない行列なので、異なるモードについての物理量 A と物理量 B の時間係数は直交する。一方同じ物理量の時間係数は、モードが異なっても直交しないことに注意が必要である。SVD で得られる第一モードは、異なる物理量間の時間係数が最大の共分散を持つように選ばれる。

SVD 解析で時間関数が得られるならば、時間関数 T_A, T_B と元のデータ Z_A, Z_B との相関あるいは回帰マップを求めることで空間構造が得られる。この際 T_A と Z_A あるいは、 T_B と Z_B というように同じ物理量について相関マップを求めたものを同質相関マップ(homogeneous correlation map)と呼び、 T_A と Z_B あるいは、 T_B と Z_A のように異なる物理量について相関マップを求めたものを異質相関マップ(heterogeneous correlation map)と呼ぶ。同様に回帰マップについても、同質と異質の二つが得られる。通常異質か同質のどちらかのマップのみを示すことが多く、一般に異質の方が相関係数・回帰係数の値が小さいので、保守的に異質マップを示す方を推奨する。

EOF では、全分散つまり自己共分散行列の対角成分の和について、説明される分散の割合を評価した。SVD 解析では、相互共分散行列の各要素の 2 乗の総和、すなわちフロベニウスのノルムの二乗が、どれだけ各々のモードで説明されるかを評価する。付録で説明するように、行列のフロベニウスのノルムの二乗は、特異値の二乗和で与えられる。すなわち、

$$\|\mathbf{C}\|_F^2 = \sum_{p_1=1}^{P_1} \sum_{p_2=1}^{P_2} c^2(p_1, p_2) = \sum_{m=1}^R l_m^2 \tag{6.7}$$

である。そこで、相互共分散行列の 2 乗の総和が、あるモードで説明される寄与率を、二乗共分散寄与率(squared covariance fraction, SCF)と呼び、

$$SCF(m) = \frac{l_m^2}{\|\mathbf{C}\|_F^2} = \frac{l_m^2}{\sum_{r=1}^R l_r^2} \tag{6.8}$$

と定義する。また、最初のいくつかの SCF の和を累積二乗分散寄与率(cumulative squared covariance fraction, CSCF)と呼ぶ。

6.2. SVD の計算例

SVD の結果を、1 つのモードについて図示する方法で標準的なのは、下図のように、それぞれの空間パターンを二つのパネルに示し、時間関数を 1 パネルに示すことであろう。空間パターンは、異質回帰・異質相関・同質回帰・同質相関のいずれかのマップを用いる。筆者自身は異質マップを用いることが多い。これらの図の中に、SCF と時系列相互の相関係数、及び各々の物理量が他の物理量の時間関数でどれだけ説明されるか(下図では VARE, VARIance Fraction がそれ)を示すとよいだろう。この表現方法は、元々この方法を提案したグループの論文(Wallace et al. 1992)に準拠している。

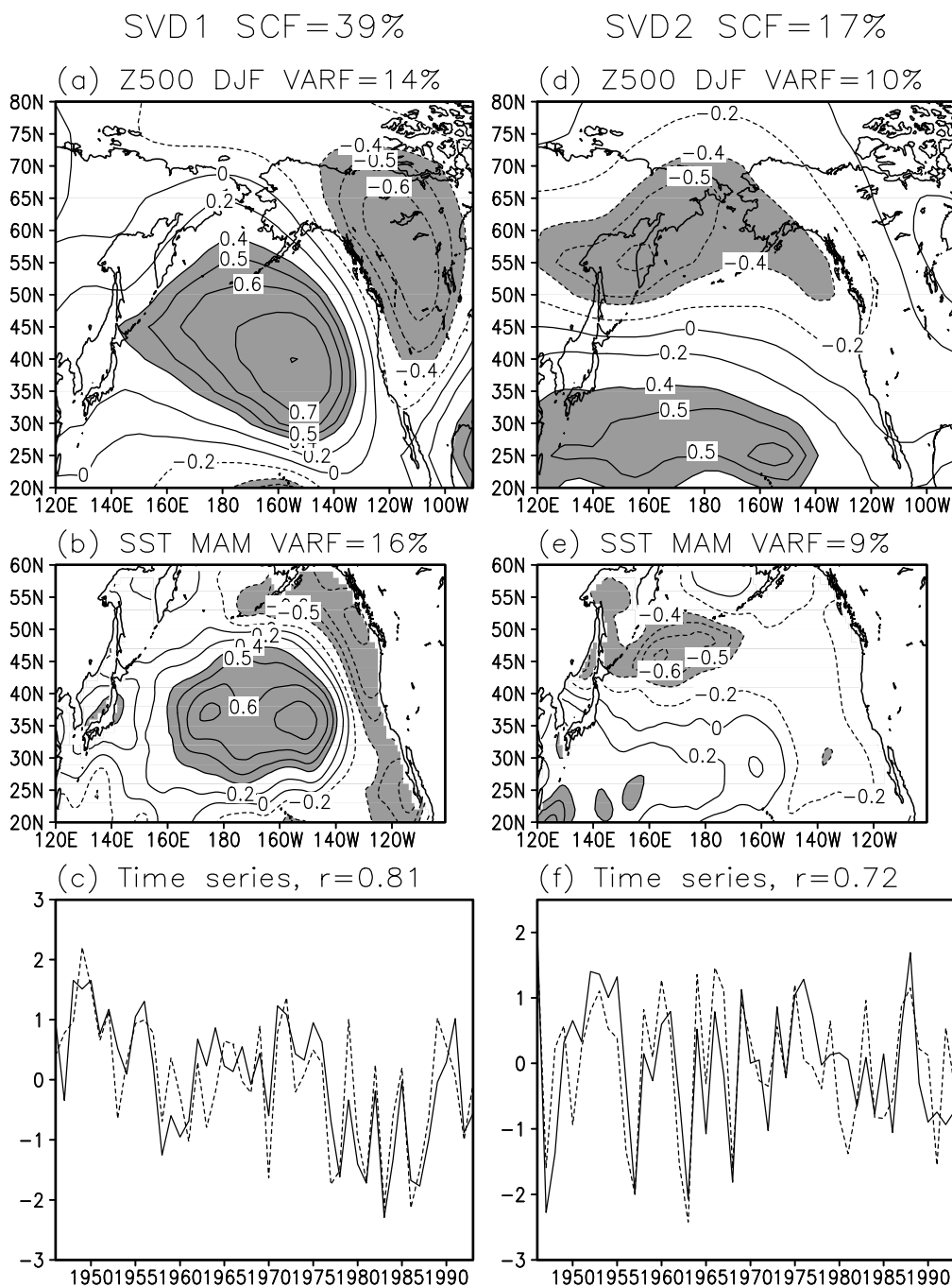


図 6.1. The first (left panels) and second (right panels) SVD mode of winter (Dec.–Feb.) 500 hPa height and spring (May–Mar.) SST over North Pacific for 1946–1993. (a,d) Heterogeneous correlation map for 500 hPa height. (b,e) Heterogeneous correlation map for SST. Contour intervals of correlation maps are 0.2 where the absolute correlation coefficient is smaller than 0.4 and 0.1 otherwise. Shading indicates grid points where the absolute heterogeneous correlation is greater than 0.4. (c,f) Time series of the normalized expansion coefficients for 500 hPa height (dashed curve) and SST (solid curve). (after Minobe and

Mantua 1999)

6.3. Matlab での SVD プログラム

Matlab では SVD の計算は EOF 同様非常に容易である .

```
% データの設定
tIng=100; spsz_a=3; spsz_b=5;
common_tsrs=randn(tIng,1); %物理量 A, B で共通の時系列
z_a=randn(tIng,spsz_a)+common_tsrs*[0.5 1 0.5]; %
z_b=randn(tIng,spsz_b)+common_tsrs*[1 0.5 0 -0.5 -1];
% z_a=randn(tIng,spsz_a)+common_tsrs*[1 0 0];
% z_b=randn(tIng,spsz_b)+common_tsrs*[1 0 0 0 0];

% SVD の計算
crscov=z_a'*z_b; % 相互共分散行列の計算
[x_a,s,x_b]=svds(crscov,3); % 特異値分解の実行
t_a=z_a*x_a; % 物理量 A についての時間係数の計算
t_b=z_b*x_b; % 物理量 B についての時間係数の計算
corr_PCs =crr(t_a,t_b,'mat')

% 結果の表示とチェック
tdm=1:1:tIng;
plot(tdm,t_a(:,1),tdm,t_b(:,1),tdm,common_tsrs)
corr_PCs =crr(t_a,common_tsrs,'mat')
corr_PCs =crr(t_b,common_tsrs,'mat')
```

6.4. dual formalism

SVD についても dual formalism によって、空間のグリッド数よりも時間のサンプル数が少ない場合には、高速に計算することが可能である。この計算方法は、Kuroda (1998)が提案した。以下ではこの手法を概観しよう。

まず、(6.4)から \mathbf{X}_A が直交行列であるから、

$$\mathbf{Z}_A \mathbf{Z}_A^T = \mathbf{T}_A \mathbf{X}_A^T (\mathbf{T}_A \mathbf{X}_A^T)^T = \mathbf{T}_A \mathbf{X}_A^T \mathbf{X}_A \mathbf{T}_A^T = \mathbf{T}_A \mathbf{T}_A^T \tag{6.9}$$

この式に右から \mathbf{T}_B をかければ、

$$(\mathbf{Z}_A \mathbf{Z}_A^T) \mathbf{T}_B = \mathbf{T}_A (\mathbf{T}_A^T \mathbf{T}_B) = \mathbf{T}_A \mathbf{L} \tag{6.10}$$

同様に、

$$(\mathbf{Z}_B \mathbf{Z}_B^T) \mathbf{T}_A = \mathbf{T}_B \mathbf{L} \tag{6.11}$$

したがって，(6.10)に右から \mathbf{L}^{-1} をかけて得られる $\mathbf{T}_A = (\mathbf{Z}_A \mathbf{Z}_A^T) \mathbf{T}_B \mathbf{L}^{-1}$ を(6.11)に代入して，

$$\boxed{(\mathbf{Z}_B \mathbf{Z}_B^T) (\mathbf{Z}_A \mathbf{Z}_A^T) \mathbf{T}_B = \mathbf{T}_B \mathbf{L}^2} \quad (6.12)$$

を得る．また同様に(6.11)の \mathbf{T}_B を(6.10)に代入すれば，

$$\boxed{(\mathbf{Z}_A \mathbf{Z}_A^T) (\mathbf{Z}_B \mathbf{Z}_B^T) \mathbf{T}_A = \mathbf{T}_A \mathbf{L}^2} \quad (6.13)$$

となる．(6.12)(6.13)で SVD 解析の時間係数を求めることができる．行列 \mathbf{Z}_A の大きさ ($N \times P_A$) は，行列 \mathbf{Z}_B の大きさは ($N \times P_B$) であるので，(6.12)(6.13)はどちらも大きさ N の行列の固有値問題である．すなわち，通常の計算方法では， $P_A \times P_B$ の大きさの特異値分解を計算する必要があるのに対し，dual formalism では $N \times N$ の大きさの固有値問題を2回計算すればよい．

ただし多くの SVD の応用においては，空間構造自体（特異ベクトル）ではなく，同質あるいは異質の相関・回帰マップを示すので，特異ベクトル自体を求める必要は必ずしもない．もし空間パターン $\mathbf{X}_A, \mathbf{X}_B$ を求めるのであれば，dual formalism では時間関数から空間構造を得る必要がある．上で注意したように同質な時間関数は直行しないので，この操作は EOF の互いに直行する時間関数から空間パターンを得るよりは，少々やっかいである．空間構造の求め方は次のように導出できる．まず，(6.4)の $\mathbf{Z}_A = \mathbf{T}_A \mathbf{X}_A^T$ から，

$$\mathbf{X}_A^T = \mathbf{T}_A^{-1} \mathbf{Z}_A \quad (6.14)$$

(6.6)から

$$\mathbf{T}_A^{-1} = \mathbf{L}^{-1} \mathbf{T}_B^T \quad (6.15)$$

なので

$$\begin{aligned} \mathbf{X}_A^T &= \mathbf{L}^{-1} \mathbf{T}_B^T \mathbf{Z}_A \\ \therefore \mathbf{X}_A^T &= \mathbf{Z}_A^T \mathbf{T}_B \mathbf{L}^{-1} \end{aligned} \quad (6.16)$$

が得られる．また，同様に，

$$\mathbf{X}_B^T = \mathbf{Z}_B^T \mathbf{T}_A \mathbf{L}^{-1} \quad (6.17)$$

となる．(6.16) (6.17)から空間構造（特異ベクトル）を求めることができる．

6.5. その他

SVD も EOF と同様，空間関数が直行するという特徴をもっている．この特徴が強く結果に影響していないかどうかを確認するために，領域を変更して SVD を計算することは有効であろう．また，Cheng and Dunkerton (1995)は回転 EOF と同様な，回転 SVD を提案している．

SVD 解析の問題点については，Newman and Sardeshmukh (1995)が論じている．

6.6. References

- Bretherton, C. S., C. Smith, and J. M. Wallace, 1992: An intercomparison of methods for finding coupled patterns in climate data. *J. Climate*, 5, 541-560, 1992.
- Björnsson H. and S. A. Venegas, 1997: A Manual for EOF and SVD Analyses of Climate Data, Centre for Climate and GlobalChange Research, Report No. 97-1, pp. 52, 1997.

- Cheng, X. and T. J. Dunkerton, 1995: Orthogonal rotation of spatial patterns derived from singular value decomposition analysis. *J. Climate*, 5, 541-560.
- Kuroda, Y. 1998: An effective SVD calculation method for climate analysis. *J. Meteor. Soc. Japan*, 76, 649-655.
- Minobe, S., and N. Mantua, 1999: Interdecadal modulation of interannual atmospheric and oceanic variability over the North Pacific, *Progress in Oceanography*, 43, 163-192.
- Newman, M., and P. D. Sardeshmukh, 1995: A caveat concerning singular value decomposition. *J. Climate*, 8, 352-360, 1995.
- Wallace, J. M., C. Smith, and C. S. Bretherton, 1992: Singular value decomposition of wintertime sea surface temperature and 500-mb height anomalies, *J. Climate*, 5, 562-576.