5. 多変量時間領域解析その１：EOF 解析

Empirical orthogonal functions (EOFs)は，データを要約するための代表的な手段である.日本語に訳すと,経験的直交関数展開である.多変量解析の分野では,主成分分析(Principal component analysis)と呼ばれる.この手法は，多数のデータから，モードと呼ばれる小数個の時間・空間関数を抽出し，次元を減らす事によってもとのデータの持つ意味を理解しやすくしようという手法である．大気海洋の研究ではデータ解析でもモデルでも広く用いられている．

EOF 解析は主として，一つの物理量（例えば，SLP や SST）に用いられる．二つ以上の物理量の場合でも EOF を，結合 **EOF(combined EOF)**という方法で用いることができる．しかし二つの物理量については，一般的に**正準相関分析(Canonical Correlation Analsysis)**あるいは**特異値分析(Singular Value Decomposition Analysis)**を用いる方が良い．ただし，これらの手法は３つ以上の物理量については適用し得ないので，その場合はやはり結合 EOF を用いる．

5.1. 独立なモードへの分解

もとのデータを$(N \times P)$の行列 $\mathbf{Z}$ で表そう．ここで N, P はそれぞれ,時間と空間方向のデータ数であるとする．ただし，$\mathbf{Z}$ の各々の空間点での時間方向の平均はゼロであるとする(平均ゼロの偏差データを扱うと仮定)．このデータを記述する場合に，いくつかの主要な特徴を客観的に抽出できれば非常に効果的な記述がし易い．EOF はこのような分解手法の代表格である．

EEOF では，観測されたデータを，次のように表現する

$$\mathbf{Z} = \sum_{m=1}^{M} \vec{t}_m \vec{x}^T{}_m = \mathbf{T}\mathbf{X}^T \tag{5.1}$$

ここで m はモードの番号，x_m は第 m モードの空間構造（EOF とも呼ばれる）で，N 要素の列ベクトルである．t_m は P 要素の列ベクトルで，**時間関数（主成分(principal component (PC))，負荷量(load)**，**スコア(score)**とも呼ばれる）と呼ばれる．モードの総数である M は N, P のうち小さい方の数である．(5.1)式と同じ内容を行列表現で書けば，

$$\mathbf{Z} = \mathbf{T}\mathbf{X}^T \tag{5.2}$$

と表される．ここで $\mathbf{T}$ は$(N \times M)$の行列，$\mathbf{X}$ は$(P \times M)$の行列である．

いずれの表現を用いるのでも，EOF は時間・空間構造の線形結合に元々のデータを分解する方法であり，通常説明する分散の大きい最初のいくつかのモードだけに着目する．

(5.1)(5.2)のように表現するというだけでは，各モードの時間・空間構造は一意には求まらず，いくつかの条件を加えなくてはならない．通常 EOF 解析では異なるモードの時間関数は直交する，

$$\vec{t}_{m_1}{}^T \vec{t}_{m_2} = \begin{cases} \lambda_{m_1} & (m_1 = m_2) \\ 0, & (m_1 \neq m_2) \end{cases} \tag{5.3}$$

ことを要請する．時間関数が直交するとは，無相関であるということである．一方線形なシステムはある条件を満足する場合には，例えば弦の振動のように，互いに独立な固有モードの重ね合わせで表現されることは広く知られている．物理的に独立であれば統計的には無相関として表される．そこで，EOF 解析は物理システムに内在する固有モードを検出する手法であることが期待され，その暗黙の期待の元で研究が行なわれる．しかしこの期待は残念ながら，後に示すように必ずしも正しくはない．なぜなら，各モードが独立という仮定だけでは，実際にモードを求めるには不充分で，さらにモードの性質を仮定しなくてはならない．この後の仮定が妥当かどうかは，物理的なモードの性質を知らなくては判断することはできない（従って自然現象についてはデータ解析のみからでは不可知である）．

5.2. 共分散行列のモード分解表現

EOF の理論は$(p \times p)$の**共分散行列(covariance matrix), V,**

$$\mathbf{V} = \mathbf{Z}^T\mathbf{Z} \qquad \Leftrightarrow \quad v(p_1, p_2) = \sum_{n=1}^{N} z(n, p_1) z(n, p_2) \tag{5.4}$$

を中心に展開される．共分散行列の各要素は空間２点(p1, p2)の間の共分散である．共分散行列は対称行列で，その対角成分は（自己）共分散である．ここで通常の共分散で行う N-1 で割る処理は式の簡便さのために，取り入れていないことに注意．N-1 で割っても割らなくても本質的な議論は変わらない．

共分散行列を(5.1)(5.2)のような時間・空間関数展開を用いて表現すれば，

$$\mathbf{V} = \mathbf{Z}^T\mathbf{Z} = \left(\sum_{i=m_1}^{M} \vec{t}_{m_1} \vec{x}_{m_1}{}^T \right)^T \left(\sum_{j=m_2}^{M} \vec{t}_{m_2} \vec{x}_{m_2}{}^T \right) = \left(\sum_{m_1=1}^{M} \vec{x}_{m_1} \vec{t}_{m_1}{}^T \right) \left(\sum_{m_2=1}^{M} \vec{t}_{m_2} \vec{x}_{m_2}{}^T \right) \tag{5.5}$$

ここで，時間関数が(5.3)で示されるように直行するのであれば，

$$\mathbf{V} = \sum_{m=1}^{M} \lambda_m \vec{x}_m \vec{x}_m{}^T, \ \lambda_m = \vec{t}_m{}^T \vec{t}_m = \left|\vec{t}_m\right|^2 \tag{5.6}$$

と表される．

自己共分散の空間方向の総和 $\sum_{p=1}^{P}\sum_{n=1}^{N} z(n,p)^2$ を全分散(total variance)と呼び，考えている領域全体での変動のエネルギーを表す量である．自己共分散は共分散行列の対角成分であるから，全分散は共分散行列のトレース(trace 対角成分の和)である．このことは次式でも確認できる．

$$trace(\mathbf{V}) = trace(\mathbf{Z}^T\mathbf{Z}) = trace\left(\sum_{n=1}^{N} z(n,p_1) z(n,p_2)\right) = \sum_{p=1}^{P}\sum_{n=1}^{N} z(n,p)^2 \tag{5.7}$$

時間関数が直交し，したがって(5.6)のように共分散行列が書けるのであれば，そのトレースは，

$$\begin{aligned} trace(\mathbf{V}) &= \sum_{i=1}^{P} v(i,i) = \sum_{i=1}^{P}\left(\sum_{m=1}^{M}\lambda_m x_m(i) x_m(i)\right) = \sum_{m=1}^{M}\lambda_m \sum_{i=1}^{P} x_m^2(i) \\ &= \sum_{m=1}^{M}\lambda_m \left|\vec{x}_m\right|^2 \end{aligned} \tag{5.8}$$

と，後で出てくる空間関数の直交の有無によらず，書くことができる．さらに空間関数が規格化されているなら

$$trace(\mathbf{V}) = \sum_{m=1}^{M}\lambda_m \tag{5.9}$$

つまり，m 番目のモードが持つ分散は λ_m または $\lambda_m\left|\vec{x}_m\right|^2$ であって，総分散はその和である．特定のモードが説明する分散の全分散に対する比を，**分散の割合(explained variance)**と呼び，

$$\frac{\lambda_m}{\sum_{m=1}^{M}\lambda_m} \tag{5.10}$$

で表す．

5.3. EOF 解析

通常の EOF 解析では，共分散行列の固有ベクトルとして空間構造を求める．この場合明らかに空間構造は互いに直交し，また次節で示すように第一モードは時間関数が直交する線形結合の中で最も大きな分散を説明するという特徴を持つ．

EOF の空間構造を求める，固有値・固有ベクトル問題は

$$\mathbf{V}\vec{x}_m = \lambda_m \vec{x}_m \tag{5.11}$$

である．固有値を対角成分に持つ対角行列を $\mathbf{\Lambda}$ と書けば，上式と同じ内容を

$$\mathbf{VX} = \mathbf{X\Lambda} \tag{5.12}$$

と表現できる．固有ベクトルは互いに直交するので，固有ベクトルが規格化されているのであれば，(5.2)(5.12)から時間関数は

$$\mathbf{T} = \mathbf{T}\mathbf{X}^T\mathbf{X} = \mathbf{Z}\mathbf{X} \tag{5.13}$$

から容易に求めることができる．

一方逆に空間構造が直交するのであれば, 固有値問題に帰着することも次のように証明できる．

$$\mathbf{V}\vec{x}_m = \left(\sum_{i=m}^{M}\lambda_m\vec{x}_m\vec{x}_m{}^T\right)\vec{x}_m = \lambda_m\vec{x}_m\vec{x}_m{}^T\vec{x}_m = \left|\vec{x}_m\right|^2\lambda_m\vec{x}_m \tag{5.14}$$

ここで空間パターンが規格化されていると約束すれば,

$$\mathbf{V}\vec{x}_m = \lambda_m\vec{x}_m \tag{5.15}$$

と固有値問題に帰着する．

共分散行列 **V** は実対称行列である．一般に実対象行列の固有値・固有関数はすべて実数値を取る．また，実対角行列は，その固有ベクトルを用いて実対角行列に変換することができる．すなわち，

$$\mathbf{X}^T\mathbf{V}\mathbf{X} = \mathbf{\Lambda} \tag{5.16}$$

である．

例えば北太平洋上の冬季の SLP の EOF は次のようになる．

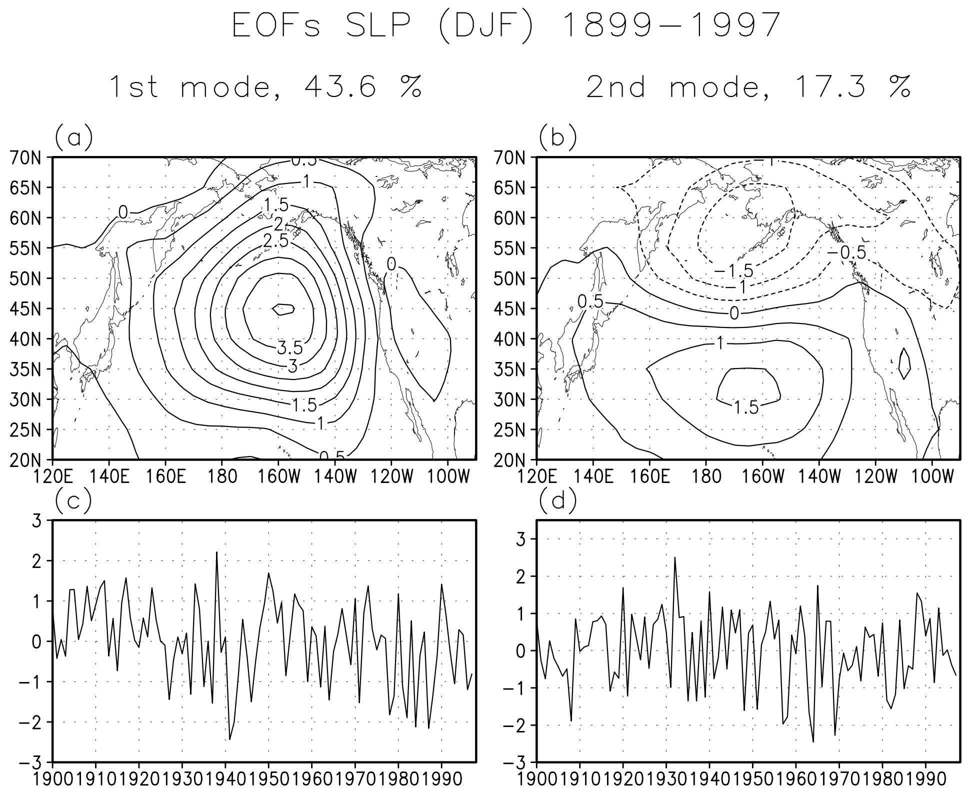

Fig. 5. Spatial patterns of the first (a) and second (b) EOF modes and their temporal coefficients (c and d) for the SLP in winter season (Dec.–Feb.) over the North Pacific (120°E–90°W) (after *Minobe and Mantua* 1999).

5.4.　EOF が説明する分散は最大

任意の空間構造の中で EOF の第一モードは最大の分散を説明する．EOF の第二モードは，第一モードに伴う変動を取り除いた後に，やはり最大の分散を説明する．この著しい特徴は EOF を強力なツールとしている．例えば，北極振動が北半球の SLP 変動の EOF 第一モードであるということは，それ以上に大きなエネルギーをもつ変動を見出すことはできないということである．ここではこの性質の証明をしよう．

まず，$\vec{e}$ を空間グリッド数と同じ要素をもつ単位ベクトル（つまり空間パターン，p×1 行列）としよう．この空間パターンへの観測値の投影は，$\mathbf{Z}\vec{e}$ であり，これで説明される分散は，$(\mathbf{Z}\vec{e})^T(\mathbf{Z}\vec{e}) = \vec{e}^T\mathbf{Z}^T\mathbf{Z}\vec{e}$ である．ここで $\mathbf{Z}^T\mathbf{Z} \equiv \mathbf{V}$ は上でも出てきた分散行列である．

同様に $\vec{e}\,' \equiv \vec{e} + \delta\vec{e}$ という $\vec{e}$ に微小なベクトル $\delta\vec{e}$ が加えられたベクトル $\vec{e}\,'$ によって説明される分散は

$$\begin{aligned}\vec{e}\,'^T\,\mathbf{V}\vec{e}\,' &= \left(\vec{e}^T + \delta\vec{e}^T\right)\mathbf{V}\left(\vec{e} + \delta\vec{e}\right) \\ &= \vec{e}^T\mathbf{V}\vec{e} + 2(\delta e)^T\,\mathbf{V}\vec{e} + \delta\vec{e}^T\mathbf{V}\delta\vec{e}\end{aligned} \tag{5.17}$$

ここで，$\mathbf{V}$ が対称行列なので，$\vec{a}^T\mathbf{V}\vec{b} = \vec{b}^T\mathbf{V}\vec{a}$ であることを利用している．今はじめのベクトル $\vec{e}$ で説明される分散が最大であるとすると，$\vec{e}$ に任意の微小なベクトルを加えても説明される分散は変化しない（１階微分までを考えれば，２階微分まで考えればわずかに減少する）ので，$\vec{e}^T\mathbf{V}\vec{e} = \vec{e}\,'^T\,\mathbf{V}\vec{e}\,'$ となるであろう．さらに，(5.17)の第３項は $\delta\vec{e}$ が微小であるので，第２項よりも小さく，結局

$$\left(\delta\vec{e}\right)^T\mathbf{V}\vec{e} = 0 \tag{5.18}$$

である．

さらに，$\vec{e}\,'$ が単位ベクトルであれば，$\vec{e}$ と $\delta\vec{e}$ は直交しなくてはならないので，

$$\left(\delta\vec{e}\right)^T\vec{e} = 0 \tag{5.19}$$

が成り立つ．(5.19)に未定定数 λ をかけて(5.18)との和を取れば，

$$\left(\delta\vec{e}\right)^T\left\{\mathbf{V}\vec{e} - \lambda\vec{e}\right\} = 0$$

{} 内が成り立たなくてはならないので，結局

$$\mathbf{V}\vec{e} = \lambda\vec{e} \tag{5.20}$$

が十分条件である．（必要条件であることも証明できる Proisendorfer p. 28）

5.5.　回転 EOF (REOF)

EOF 第一モードは，時間関数が直行する線形結合モードの中で((5.2)(5.3))，最大のエネルギーを説明するモードである，ということは少ないモードで観測データが説明できるという点では利点であるけれど，常に望ましいとは限らない．EOF 解析は多くの場合単にデータを縮約するという

よりも，明示的に書かれなくても独立な変動を取り出すという目的で使われることが多い．物理的なモードがエネルギーを最も説明する可能なモードよりも，実際には少ないエネルギーしか説明しないのであれば，EOF 解析は物理的なモードの重ね合わせでより大きなエネルギーを説明しようとする．つまり EOF 解析で得られるモードは，必ずしも物理的に妥当なものではない．

EOF 解析ではまた，説明するエネルギーを最大化するために，広い領域に分布する空間構造を同定する傾向にある．さらに，EOF 解析では異なる空間構造は，直交するがこの条件を物理的に正当化することはできない．

これらの問題を念頭において，しばしば**回転 EOF(Rotated EOF, REOF)**と呼ばれる手法が用いられる．回転 EOF とは空間パターン（固有ベクトル）を，長さ一定のまま向きを回転させて，新しい固有ベクトルの組を作る方法である．回転 EOF で得られる，空間パターンは直交しない．また，回転 EOF は局在化した変動を捉えやすいという特徴を持つ．

5.6. EOF の Matlab プログラム

```
% サンプルデータの作成
tlng=1000; spcsz=3;
spcgvn=[ 1 1 1; -1 0 1;  0 1 0]'; % サンプルデータその１
% spcgvn=[ 1 0 0; 0 0 1;  1 1 1]'; % サンプルデータその２
for p=1:spcsz
  spcgvn(:,p)=spcgvn(:,p)./sqrt(sum(spcgvn(:,p).^2));
end
timgvn=randn(tlng,3);
z1=timgvn(:,1)*spcgvn(:,1)';
z2=timgvn(:,2)*spcgvn(:,2)';
z3=timgvn(:,3)*spcgvn(:,3)';
z1=z1./sqrt(sum(sum(z1.*z1)))*sqrt(0.40);
z2=z2./sqrt(sum(sum(z2.*z2)))*sqrt(0.35);
z3=z3./sqrt(sum(sum(z3.*z3)))*sqrt(0.25);
z=z1+z2+z3;
explvargvn(1)=sum(sum(z1.*z1))/sum(sum(z.*z));
explvargvn(2)=sum(sum(z2.*z2))/sum(sum(z.*z));
explvargvn(3)=sum(sum(z3.*z3))/sum(sum(z.*z));

% EOF の計算
covmat=z'*z;                       % 共分散行列の計算
[spc,eig_valmat]=eig(covmat);      % 固有値・固有ベクトルの計算
eig_val=diag(eig_valmat);          % 固有値を対角行列からベクトルに変更
[eig_val,ind]=sort(eig_val);       % 固有値を昇順に並び替え
eig_val=flipud(eig_val);           % 固有値を降順に並び替え
ind=flipud(ind);                   % 降順の固有値の番号を記録
```

```
spc=spc(:,ind);                     % 空間パターンを固有値降順に揃える
tim=z*spc;                          % 時間関数を計算

% 結果の確認
explvargvn
explvar=eig_val'./sum(eig_val)
spcgvn
spc
corr_PCs =crr(tim,timgvn,'mat')  % これは Matlab の関数では無い
```

5.6.1. 結果その１

```
explvargvn =0.4051   0.3545   0.2532
explvar =   0.5126   0.3540   0.1334

spcgvn =    0.5774  -0.7071        0
            0.5774        0   1.0000
            0.5774   0.7071        0


spc =       0.4640  -0.6834  -0.5637
            0.7981   0.0464   0.6008
            0.3844   0.7286  -0.5669

corr_PCs=   0.8264  -0.0682   0.5320
            0.0361   0.9977   0.0279
           -0.5618  -0.0012   0.8464
```

5.6.2. 結果その２

```
explvargvn = 0.3981   0.3483   0.2488
explvar =    0.5830   0.3626   0.0544

spcgvn =     1.0000        0   0.5774
                  0        0   0.5774
                  0   1.0000   0.5774

spc =        0.7752  -0.6121   0.1564
             0.2302   0.0431  -0.9722
             0.5883   0.7896   0.1743
```

```
corr_PCs=    0.6599    0.4641    0.6097
            -0.6254    0.7617    0.0906
             0.4177    0.4529   -0.7864
```

5.7. EOF と回転 EOF との比較

EOF の一族の中で，普通の EOF はもっとも広く使われている手法であり，その問題点も知られている．まず，EOF と REOF のどちらを使うべきかという問題がある．上で述べたように，EOF は最大の分散を第一モードでするので，広く分布するモードを検出しようとする．一方 REOF はより局在化したモードを検出しようとする．その結果同じデータに適用しても，古くから知られているように異なる結果が得られる. 最近話題の現象について, Dommenget and Latif (2002)は EOF と REOF をそれぞれ計算し，次のような相違があることを示している．

表 5.1

	EOF 1	EOF2	REOF1	REOF2
北半球 SLP	北極振動	太平洋が強い北極振動もどき	北大西洋振動	アリューシャン低気圧
熱帯大西洋 SST	ギニア湾に局大を持つ全領域同符号	南北ダイポール	ギニア湾に局在化した mono pole	EOF2 のダイポールの北側の局の位置に存在する mono pole
熱帯インド洋 SST	領域一様	東西ダイポール	南西に mono-pole	西に mono-pole

EOF と REOF とどちらかが他に対して優れているとは言えない[1]．もし物理的なモードが EOF で捕らえやすい広い分布を持つのであれば EOF を使う方が適切であるし，物理的なモードが REOF で検出しようとする局在化した分布を持つのであれば REOF を使う方が適切である．

5.8. 実用上のいくつかのポイント

5.8.1. 共分散行列か相関行列か

EOF は通常共分散行列について求められるが，相関行列に関して求める場合もある．相関行列は，各グリッドでの時系列を分散１に規格化しておくことと等価である．すなわち，相関行列を用いるとは元々存在する振幅の空間的な大小を無視することである．これは，空間的な振幅の大小によらず, どの程度広い領域に変動が分布しているかどうかを問題にしようという立場である．この立場は着目したい現象が比較的振幅が小さいものの広く分布している場合に有効である．一

[1] EOF あるいは REOF をよく使う人は，多分それぞれ自分が使っている手法が何らかの点で他に対して優れていると思って使っているのであろう． REOF のプログラムを持っていないために，EOF を使うしかないということもありそうだが．いずれにしても，広く一般的にどちらが優れているというコンセンサスは無い．

方共分散行列を用いる EOF は，空間的な振幅の分布も重要な情報であると考える立場である．

5.8.2. 規格化

EOF の理論では上で見たように，通常空間パターンが大きさが１に規格化され，時間関数の大きさがそのモードの分散に相当することになっている．しかし空間パターンが物理的に意味を持つ量である方が，解釈上有利である．そこで，実用上は，

$$\vec{x}_i\,' = x_i\sqrt{\lambda_i}, \quad \vec{t}_i\,' = t_i/\sqrt{\lambda_i} \tag{5.21}$$

と再規格化を行うことが一般的である．この規格化によって（簡単のためにプライムを落として），

$$\left|\vec{x}_i\right|^2 = \lambda_i, \quad \left|\vec{t}_i\right|^2 = 1 \tag{5.22}$$

となる．

空間構造の単位は元々のデータの単位となって，その２乗がそのグリッド点で i 番目のモードによって説明される分散となる(時間関数の分散は今１である)．また，相関行列による EOF の場合はこのように規格化することで，空間構造の値はその場所での観測値と i 番目の PC との相関係数となる．

5.8.3. Dual formalism

通常 EOF を求める対象のデータは時間方向のグリッド数よりも，空間方向のグリッド数の方がはるかに多い．例えば全球５度×５度で月毎の表面温度データが５０年間あるとしよう．データの空間グリッド数は 72×36=2592 で，時間グリッド数は 600 である．このような $P>>N$ である場合には，もとのデータ行列 $\mathbf{Z}(N\times P)$についての共分散行列 $\mathrm{V}(P\times P)$の固有値問題ではなく，$\mathbf{Z}^T$ の共分散行列 $\mathbf{L}(N\times N)=\mathbf{ZZ}^T$ の固有値門題から，EOF を得る方法が知られている．これを dual formalism あるいは parallel algorithm と呼ぶ．共分散行列の固有値問題は

$$\mathbf{VX}=\mathbf{X\Lambda}$$

である.この式に対して,両辺左から $\mathbf{Z}$ をかけると.

$$\mathbf{ZVX}=\mathbf{ZZ}^T\mathbf{ZX}=\mathbf{LZX}=\mathbf{ZX\Lambda}$$

となるので，ここで $\mathbf{B}=\mathbf{ZX}$ を定義すれば

$$\mathbf{LB}=\mathbf{B\Lambda}$$

という $\mathbf{L}$ の固有値問題となる．これを解いて $\mathbf{B}$ が求まったなら，次に $\mathbf{D}=\mathbf{Z}^T\mathbf{B}$ を求める．$\mathbf{D}$ は

$$\mathbf{D}=\mathbf{Z}^T\mathbf{B}=\mathbf{Z}^T\mathbf{ZX}=\mathbf{VX}=\mathbf{X\Lambda}$$

であるので，最終的に求めるべき $\mathbf{X}$ の成分 x_{ij} は,

$$x_{ij} = \frac{d_{ij}}{\sqrt{\lambda_j}}$$

によって求めることができる. **X** が求まったなら, 通常の方法によって **T** を求めることができる.

5.8.4. 欠損値の処理

欠損値があるデータは用いない，あるいは欠損値をなんらかの方法で補間しておき欠損値の無いデータで EOF を取る，というのもよく用いられる方法である．他の方法は欠損値の存在を許容するもので，欠損値については共分散行列の計算にカウントしない．ただし，共分散は本来

$$\frac{1}{N}\sum_{n=1}^{N} z(n,p_1)z\left(n,p_2\right)$$

と定義されるので，欠損値があるなら最初に割っている n も，二つの空間点の両方で欠損の無いデータの数（n0 としよう）でなくてはならない．これまで紹介してきた書き方は，欠損値がない前提の下に，この n での除算を首尾一貫して用いないものであった．この流儀では，n/n0 を共分散行列の(p_1, p_2)要素にかける必要がある．

5.8.5. 手順

１）EOF を取る対象のデータを読み込む（例えば１０点の冬の気温データ）
２）共分散行列あるいは相関行列を求める
３）共分散行列あるいは相関行列の固有値・固有ベクトル（EOF）を計算する
４）時間関数（PC）を計算する
５）通常固有ベクトルは１に規格化されているので（ライブラリーによってはそうでないかもしれないので，ライブラリ毎に要確認)，時間関数の分散が１になるように，EOF と PC を再規格化する.
６）結果をファイルに出力する.（EOF の場合は結果が時間･空間があるので，コンソール出力のリダイレクトではなく，きちんと open 文でファイルを open して，write(装置番号,書式番号）で個々のファイルに出力した方が良いでしょう.

5.9. EOF **の親戚達の概観**(CEOF,EEOF,REOF,RCEOF,SSA,MSSA)

まず様々な EOF の類縁である手法を概観しよう.

EOF: Empirical Orthogonal Function 経験的直交関数．広義には EOF という手法を示し，狭義には EOF 解析で得られる空間構造を示す．また他の EOF 手法(CEOF, REOF 等)と対比して，実数の非回転 EOF という意味で用いられることもある．科学技術用語にはめずらしく，多様な意味を持つので，どの意味で用いているかが当該の文章を読んでいる際に常に明確にしておく必要がある.

CEOF: complex EOF．EOF の複素数版．伝播成分を表現することができる.

EEOF: Extended EOF，ラグを考慮した共分散行列について EOF を行い，伝播現象を表現できる．通常ラグは３~５など SSA/MSSA に比べて少ない数を取る.

REOF: 回転 EOF．一旦 EOF で時間・空間構造を求め，次にその空間構造を基底ベクトルとみて，その基底ベクトルをベクトル空間内で回転させることによって，**互いに直交しない空間構造**とそれに対応する互いに直交する時間関数とを求める．回転手法はいくつかあるが，空間構造が持つ分散の２乗が最大となる(バリマックス回転)Varimax rotation が最も良く用いられる．この手法は,

空間的に直交する通常の EOF では不適切と考えられる空間構造が得られている場合に，空間的な直交という制約を除くので，有効である場合が多い．すでに述べたように通常の EOF は第一モードで説明される分散があらゆる時間関数が直交する展開方法の中で最大なので，REOF の第一モードで説明される分散は EOF の第一モードで説明される分散よりもかならず小さい．

RCEOF: Rotated Complex Empirical Orthogonal Function．REOF の複素数版．

SSA: Singular Spectrum Analysis．単一時系列のラグ共分散行列の固有値展開によって，構造を求める手法．比較的新しく(Ghil and Vautard, 1991)，現在盛んに利用されているが，まだその性質が十分に明らかになっているとは言い難い．

MSSA: Multi-Channel Singular Spectrum Analysis．EEOF で用いるラグをもっと多くしたもの（20~40 など）．SSA の多変量への発展版である．伝播現象の表現が可能．時間スケールが異なる現象を分離できる（他の EOF ではできない）．現在様々な活用がなされている．

これらの手法を利用する上で，いくつかのポイントをまとめておこう．

・上の全ての手法で，時間関数は必ず直交する

・EOF, CEOF では空間構造は必ず直交する．REOF, RCEOF では空間構造は直交しない．

・EOF は最大の分散を第一モードでするので，広く分布するモードを検出しようとする．一方 REOF はより局在化したモードを検出しようとする．

・EOF, REOF では定在振動が表現され, 伝播現象は 90°位相がずれた２つのモードに分離されてしまう．

・CEOF, EEOF, MSSA では伝播現象が表現できる．ただし EEOF,MSSA では利用しているラグの長さよりも現象の周期あるいは代表的な時間スケールが短い場合のみ, 伝播現象が表現し得る．

5.10. EOF 解析の利用上の注意

１．　モードの時系列が独立という EOF の一族で共通に用いられる仮定は，線形システムでは妥当であるけれど，非線形システムでは妥当性は保証されない．この問題を乗り越えるために，様々な非線形解析手法，特に非線形主成分分析(nonlinear Principal Component Analysis, nonlinear PCA)が提案されているが，大気・海洋業界で広く利用されるには至っていない．

２．　EOF では第一モードが説明する分散が最大で，残りの分散の内第二モードが説明する分散はまた最大であるという仮定を置く．これは全変動エネルギーをなるべく少ないモードで表現するには好ましいが, 物理的なモードはより局所的な分布を持つのかもしれない. また, 異なるモードは時間関数だけではなく，空間構造も直交するという物理的に正当化できない制約を受ける．これらの特徴から，第一モードが全領域で同符号，第二モードがダイポールという構造が，人工的に生成されやすいので，この構造が得られた場合は相関・回帰分析などを行なって，構造の妥当性を吟味する方が安全である．

３．　REOF では第一モードが説明する分散の二乗が最大という仮定を置く．局所的なモードを見出しやすい仮定であるが，物理的なモードはより広い分布を持つのかもしれない．

４．　時間的なサンプルの個数が少なければ，EOF の構造は正確に求まらないであろう．North の Rule of Thumb として知られる基準(North et al. 1982)で，統計的な妥当性をチェックすることが望ましい．このサンプリング問題と，注意点２・３で述べた問題とは独立であることに

注意しよう(Dommenget and Latif 2002 参照).

5.11. References

Dommenget, D. and M. Latif, 2002: A Cautionary Note on the Interpretation of EOFs. J. Climate, 15(2), 216-225, 2002.

Preisendorfer, R. W. 1988: Principal component analysis in Meteorology and oceanography. Elsevier, pp. 425. (EOF に関してすごーく詳しく書いてある本．残念ながら絶版で新規入手は困難．私もワシントン大学の知り合いにコピーしてもらった).

North, G. R., T. L. Bell, R. F. Cahalan, and F. J. Moeing, 1982: Samping errors in the estimation of empirical orthogonal functions. Mon. Wea. Rev., 110, 699-706.

Richman, M. B., 1986: Review article: Rotation of principal components. J. Climatol., 6, 293-335.