数理科学特論Ａ～画像数学～　第４回

シリーズ２：直交変換による画像圧縮

(1) 主成分分析と Karhunen-Loéve 変換

「画像数学」の第２シリーズは，「直交変換による画像圧縮」について説明します．これは，画像のデータを「見た目に影響の大きい」成分と「あまり影響がない」成分に分け，あまり影響がない成分を省くことによって，画像の見た目の印象をあまり変えることなく画像のデータ量を減らすという方法です．このシリーズは，(1) データの「影響の大きい成分」と「あまり影響がない成分」という考え方を導くための主成分分析と Karuhunen-Loéve(KL) 変換，(2) KL 変換，フーリエ変換などを一般的にとらえる行列のユニタリー変換，(3) ユニタリー変換の１つコサイン変換を使った画像圧縮と JPEG 規格の３回に分けて講義します.

「重要な成分」と「あまり重要でない成分」

この節では，説明のために画素が２つしかない画像を考えます．そして，この２画素の画像をたくさん取り扱うという状況を考えてみましょう．当然，いろいろな画素値の組み合わせがあります．そこで，２つの画素値を x_1, x_2 として，これらの画像の分布を x_1, x_2 を軸とする座標平面に描いてみます．これが図１のようになったとしましょう．+ 印１つが１つの画像に対応します．図１の場合では，x_1, x_2 のどちらの方向の分散も大きく，各画像の違いを表現するのに x_1, x_2 とも意味のある役割を果たしていることがわかります．ですから，x_1, x_2 の両軸とも省略することはできません.

しかし，もしも画像の分布が図２のようであればどうでしょう．この場合，各画像の画素 x_1 の値はさまざまに異なっているのに対して，画素 x_2 はどの画像もあまり変わらない値になっています．したがって，図２の各画像の違いを表現するのには画素 x_1 だけが重要で，画素 x_2 はさほど重要ではなく，何か１つの値（例えば各画像の画素 x_2 の平均）で置き換えてしまってもさほど困らないということになります.

このような画像の分布を，図１のように画像が分布している場合でも作ることはできないでしょうか？それは，図３のように軸を回転することによって可能になります．x_1, x_2 軸を回転した z_1, z_2 軸では，z_1 方向の値の分散が最大となっています．このように画素 x_1, x_2 を新たな画素 z_1, z_2 に変換した場合，画素 z_1 は「重要な成分」，画素 z_2 は「あまり重要でない成分」ということになります.

主成分分析

このような変換後の座標軸，すなわち画素値 z_1, z_2 を求めてみましょう．元の画素値 x_1, x_2 に対し

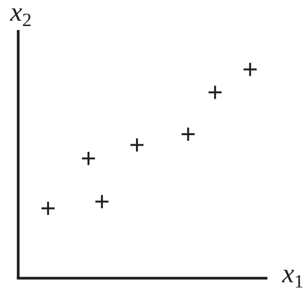

図１．２画素の画像の分布の例 (1)

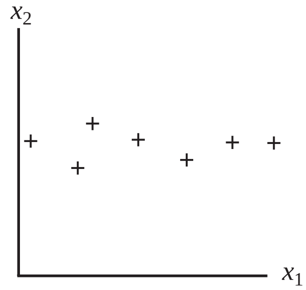

図２．２画素の画像の分布の例 (2)

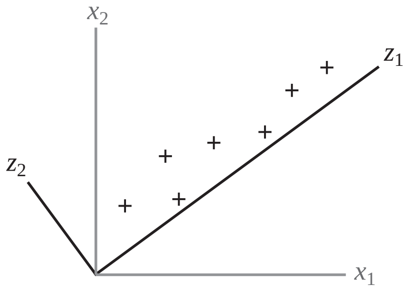

図３．軸の回転

て，新しい画素値 z_1 が

$$z_1 = a_1 x_1 + a_2 x_2 \tag{1}$$

で表されるものとします*)．i 番目の画像の画素 x_1, x_2 の値を x_{1i}, x_{2i} とすると，画素 x_1, x_2 の値の平均 $\overline{x_1}, \overline{x_2}$，分散 s_{11}, s_{22}，共分散 $s_{12} = s_{21}$ は，画像の数を n として

$$\begin{cases} \overline{x_1} = \dfrac{1}{n}\sum_{i=1}^{n} x_{1i}, \quad \overline{x_2} = \dfrac{1}{n}\sum_{i=1}^{n} x_{2i} \\ s_{11} = \dfrac{1}{n}\sum_{i=1}^{n} \left(x_{1i} - \overline{x_1}\right)^2, \quad s_{22} = \dfrac{1}{n}\sum_{i=1}^{n} \left(x_{2i} - \overline{x_2}\right)^2 \\ s_{12} = s_{21} = \dfrac{1}{n}\sum_{i=1}^{n} \left(x_{1i} - \overline{x_1}\right)\left(x_{2i} - \overline{x_2}\right) \end{cases} \tag{2}$$

で表されます．共分散 $s_{12} = s_{21}$ を（画素 x_1 の値の標準偏差）×（画素 x_2 の値の標準偏差）すなわち $\sqrt{s_{11}}\sqrt{s_{22}}$ で割ったものが相関係数で，２つの軸の相関がないとき共分散は０となります．これは，次のように直感的に説明できます**)．図４のように $\overline{x_1}, \overline{x_2}$ を仮の座標軸としたとき，第１・３象限では (2) 式の $(x_{1i} - \overline{x_1})(x_{2i} - \overline{x_2})$ は正となり，第２・４象限では負となります．図４(b) のように，相関がないときは，２つの差の積の正負が打ち消しあって０になります．

さて，新しい画素値の分散 $V(z_1)$ は，(1)(2) 式から次のように表されます．

$$\begin{aligned} V(z_1) &= \frac{1}{n}\sum_{i=1}^{n} \left(z_{1i} - \overline{z_1}\right)^2 \\ &= \frac{1}{n}\sum_{i=1}^{n} \left\{(a_1 x_{1i} + a_2 x_{2i}) - (a_1 \overline{x_1} + a_2 \overline{x_2})\right\}^2 \\ &= \frac{1}{n}\sum_{i=1}^{n} \left\{a_1(x_{1i} - \overline{x_1}) + a_2(x_{2i} - \overline{x_2})\right\}^2 \\ &= \frac{1}{n}\sum_{i=1}^{n} \left\{a_1^2(x_{1i} - \overline{x_1})^2 + 2a_1a_2(x_{1i} - \overline{x_1})(x_{2i} - \overline{x_2}) + a_2^2(x_{2i} - \overline{x_2})^2\right\} \\ &= a_1^2\left\{\frac{1}{n}\sum_{i=1}^{n} (x_{1i} - \overline{x_1})^2\right\} + 2a_1a_2\frac{1}{n}\sum_{i=1}^{n} (x_{1i} - \overline{x_1})(x_{2i} - \overline{x_2}) + a_2^2\left\{\frac{1}{n}\sum_{i=1}^{n} (x_{2i} - \overline{x_2})^2\right\} \\ &= a_1^2 s_{11} + 2a_1a_2 s_{12} + a_2^2 s_{22} \end{aligned} \tag{3}$$

したがって，この $V(z_1)$ を最大にする a_1, a_2 を求めればよいわけです．ここで，θ_1, θ_2 を z_1 軸が x_1, x_2 軸となす角として

$$a_1 = \cos\theta_1, \ a_2 = \cos\theta_2 \tag{4}$$

とおくと，(a_1, a_2) は新しい座標軸 z_1 の方向余弦ということになり，a_1, a_2 は

*) $z_1 = a_1(x_1 - \overline{x_1}) + a_2(x_2 - \overline{x_2})$ とする場合もしばしばあります．分散共分散行列は変わりません．

**) 共分散については，私の「統計データ解析Ｂ」(2000 年度前期）第９回の講義録を見てください．

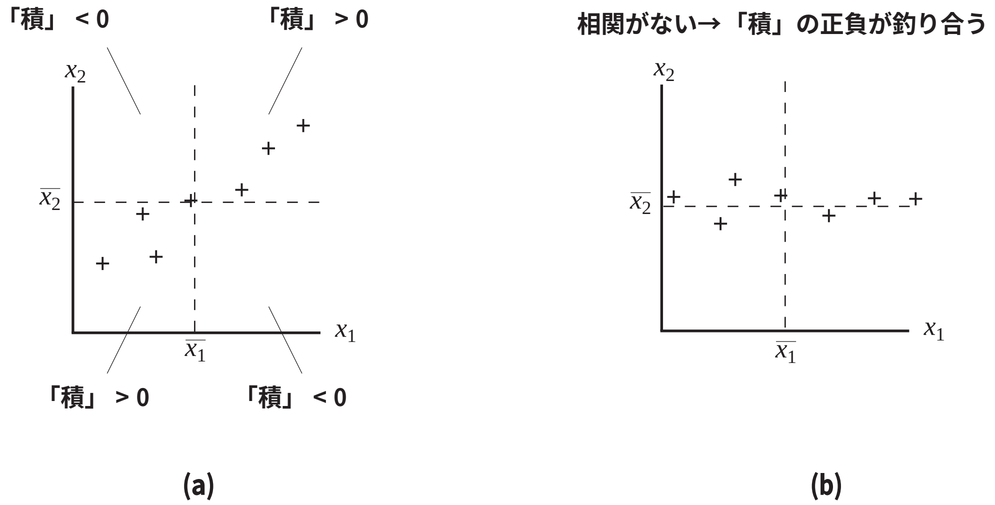

図４．共分散の意味．(a) 各象限での「積」$(x_{1i}-\overline{x_1})(x_{2i}-\overline{x_2})$の正負．(b) 相関がない時．

$$a_1^2+a_2^2=1 \tag{5}$$

を満たします．したがって，問題は (5) 式の条件のもとでの (3) 式の $V(z_1)$ の最大化ということになります．

このような制約条件付き最大化問題は，Lagrange の未定乗数法によって解くことができます．これによれば，この問題は未定乗数を λ とおいて

$$F(a_1, a_2, \lambda)=a_1^2 s_{11}+2a_1a_2s_{12}+a_2^2s_{22}-\lambda(a_1^2+a_2^2-1) \tag{6}$$

を最大化する制約条件なしの最大化問題に帰着されます．これを解くため，F を a_1, a_2, λ でそれぞれ偏微分して，それらが０に等しいとおくと

$$\begin{cases} \dfrac{\partial F}{\partial a_1}=2a_1s_{11}+2a_2s_{12}-2a_1\lambda=0 \\ \dfrac{\partial F}{\partial a_2}=2a_2s_{22}+2a_1s_{12}-2a_2\lambda=0 \\ \dfrac{\partial F}{\partial \lambda}=-\lambda(a_1^2+a_2^2-1)=0 \end{cases} \tag{7}$$

となります．ここで第３式は (5) 式と同じですからすでに満たされています．残りの式からは，

$$\begin{cases} a_1s_{11}+a_2s_{12}=a_1\lambda \\ a_2s_{22}+a_1s_{12}=a_2\lambda \end{cases} \tag{8}$$

という関係が得られます．これを行列を使って書くと

$$\begin{pmatrix} s_{11} & s_{12} \\ s_{12} & s_{22} \end{pmatrix}\begin{pmatrix} a_1 \\ a_2 \end{pmatrix}=\lambda\begin{pmatrix} a_1 \\ a_2 \end{pmatrix} \tag{9}$$

となります．ここで左辺の行列は，**分散共分散行列 (covariance matrix)** とよばれています．

(9) 式はすなわち「分散共分散行列の固有値を求める」問題で，これを満たす λ は固有値

(eigenvalue)，$\begin{pmatrix} a_1 \\ a_2 \end{pmatrix}$は固有ベクトル (eigenvector) とよばれています．このようにして画素値 x_1, x_2 から得られる新しい画素値 z_1 の分散 $V(z_1)$ は，(9) 式から

$$\begin{aligned} a_1 s_{11} + a_2 s_{12} &= \lambda a_1 \\ a_1 s_{12} + a_2 s_{22} &= \lambda a_2 \end{aligned} \tag{10}$$

で，

$$\begin{array}{ll} & \begin{cases} a_1 s_{11} + a_2 s_{12} = \lambda a_1 \\ a_1 s_{12} + a_2 s_{22} = \lambda a_2 \end{cases} \\ \begin{array}{l} \text{（両辺}\times a_1\text{）} \\ \text{（両辺}\times a_2\text{）} \end{array} & \begin{cases} a_1^2 s_{11} + a_1 a_2 s_{12} = \lambda a_1^2 \\ a_1 a_2 s_{12} + a_2^2 s_{22} = \lambda a_2^2 \end{cases} \\ \text{（辺々を加える）} & a_1^2 s_{11} + 2a_1 a_2 s_{12} + a_2^2 s_{22} = \lambda(a_1^2 + a_2^2) \end{array} \tag{11}$$

となります．この式の左辺は (3) 式のとおり $V(z_1)$ で，右辺に (5) 式を代入すると

$$V(z_1) = \lambda \tag{12}$$

が得られます．

(9) 式の２変数のベクトルの固有値問題では，固有値と固有ベクトルの組み合わせが２組得られます．この計算は新しい画素値 z_1 の分散 $V(z_1)$ を最大にするもので，(12) 式で示したようにその値は固有値λですから，２組のうち固有値の大きいほうの組み合わせから，最大の分散をもつ画素値 z_1 が得られます．これを**第１主成分**といい，「もっとも重要な成分」ということができます．
ところで，(9) 式のとおり分散共分散行列は対称行列で，対称行列の固有ベクトルは直交することが知られています（線形代数学の教科書を見てください）．したがって，もう１つの固有値に対応する固有ベクトルから (1) 式の計算でもうひとつの新しい画素値 z を求め，これを z_2 とすれば，新しい画素値 z_1, z_2 で新たな直交座標が得られ，図３のような座標軸の回転が行えることがわかります．この手法を**主成分分析 (principal component analysis, PCA)** といいます．

主成分分析と行列の対角化

ここまでは画素が２つの画像を考えましたが，では画素が p 個ある画像の場合を考えてみましょう．画像が画素 $x_1, x_2, \ldots, x_p$ からなるとするとき，変換後の画素を

$$z = a_1 x_1 + a_2 x_2 + \cdots + a_p x_p \tag{13}$$

で表して，z の分散を最大にすることを考えます．この問題は，前節と同様に

$$\begin{pmatrix} s_{11} & s_{12} & \cdots & s_{1p} \\ s_{21} & s_{22} & \cdots & s_{2p} \\ \vdots & & \ddots & \\ s_{p1} & s_{p2} & \cdots & s_{pp} \end{pmatrix} \begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_p \end{pmatrix} = \lambda \begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_p \end{pmatrix} \tag{14}$$

という分散共分散行列の固有値問題になります *)．ここで s_{ij} は x_i と x_j の共分散を意味します．また，前節と同様に (14) 式で求められる p 個の固有値は各々変換後の画素 z の分散になります **)．そ

*), **) くわしくは，私の「情報統計学 II」第７回の講義録を見てください．

こで，固有値を大きいほうから $\lambda_1, \lambda_2, ..., \lambda_p$ とし，対応する変換後の画素を $z_1, z_2, ..., z_p$ とすると，z_1 が「分散が最大の成分」すなわち「もっとも重要な成分」で，z_2 はそれに直交する成分の中で分散が最大の成分，以下番号が進むにつれ重要さがだんだん落ちてゆくことになります *)．このとき z_k を**第 *k* 主成分**とよびます．

　さて，固有値 l_i に対応する固有ベクトルを $(a_{1(i)}, a_{2(i)}, ..., a_{p(i)})'$ とすると **)，変換後の「k 番目に重要な」画素値すなわち第 k 主成分 z_k は (13) 式から

$$z_k = a_{1(k)}x_1 + a_{2(k)}x_2 + \cdots + a_{p(k)}x_p = \begin{pmatrix} a_{1(k)} & a_{2(k)} & \cdots & a_{p(k)} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_p \end{pmatrix} \tag{15}$$

とあらわされます．さらに，その固有ベクトルを対応する固有値の大きさの順に並べた行列を P とします．すなわち，

$$P = \begin{pmatrix} a_{1(1)} & a_{1(2)} & \cdots & a_{1(p)} \\ a_{2(1)} & a_{2(2)} & & a_{2(p)} \\ \vdots & & \ddots & \\ a_{p(1)} & a_{p(2)} & \cdots & a_{p(p)} \end{pmatrix} \tag{16}$$

です．このとき，分散共分散行列を S とすると (14) 式から

$$S\begin{pmatrix} a_{1(k)} \\ a_{2(k)} \\ \vdots \\ a_{p(k)} \end{pmatrix} = \lambda_k \begin{pmatrix} a_{1(k)} \\ a_{2(k)} \\ \vdots \\ a_{p(k)} \end{pmatrix} \quad (k = 1, 2, ..., p) \tag{17}$$

がなりたちますから，$k = 1, 2, ..., p$ を合わせると

$$S\begin{pmatrix} a_{1(1)} & a_{1(2)} & \cdots & a_{1(p)} \\ a_{2(1)} & a_{2(2)} & & a_{2(p)} \\ \vdots & & \ddots & \\ a_{p(1)} & a_{p(2)} & \cdots & a_{p(p)} \end{pmatrix} = \begin{pmatrix} a_{1(1)} & a_{1(2)} & \cdots & a_{1(p)} \\ a_{2(1)} & a_{2(2)} & & a_{2(p)} \\ \vdots & & \ddots & \\ a_{p(1)} & a_{p(2)} & \cdots & a_{p(p)} \end{pmatrix} \begin{pmatrix} \lambda_1 & & & 0 \\ & \lambda_2 & & \\ & & \ddots & \\ 0 & & & \lambda_p \end{pmatrix} \tag{18}$$

で，これは

$$\Lambda = \begin{pmatrix} \lambda_1 & & & 0 \\ & \lambda_2 & & \\ & & \ddots & \\ 0 & & & \lambda_p \end{pmatrix} \tag{19}$$

とおくと

$$SP = P\Lambda \quad \text{すなわち} \quad P^{-1}SP = \Lambda \tag{20}$$

と表されます．２画素の場合に (5) 式で行ったように，各固有ベクトルは正規化されており，また S

*) 詳しくは，例えば塩谷「多変量解析概論」を見てください．

**) 記号「′」はベクトル・行列の転置 (transposition) を表します．

は対称行列なので各固有ベクトルは直交しています．すなわち，各固有ベクトルは正規直交基底をなしています．したがって P は正規直交行列です．P が直交行列のとき $P^{-1} = P'$ が成り立つので（これも線形代数学の教科書を見てください），

$$P'SP = \Lambda \quad \text{または} \quad S = P\Lambda P' \tag{21}$$

となります．これを対称行列 S の**対角化 (diagonalization)** とよんでいます．

ところで (15) 式を $k = 1,2, ..., p$ についてまとめると，

$$\begin{pmatrix} z_1 \\ z_2 \\ \vdots \\ z_p \end{pmatrix} = \begin{pmatrix} a_{1(1)} & a_{2(1)} & \cdots & a_{p(1)} \\ a_{1(2)} & a_{2(2)} & & a_{p(2)} \\ \vdots & & \ddots & \\ a_{1(p)} & a_{2(p)} & \cdots & a_{p(p)} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_p \end{pmatrix} = P' \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_p \end{pmatrix} \tag{22}$$

となりますから，行列 P' はもとの画像の画素値 $x_1, x_2,..., x_p$ を新たな画像の画素値 $z_1, z_2,..., z_p$ に変換する行列となります．この形式による画像の変換を，画像の**直交変換 (orthogonal transformation)** とよんでいます．また (21) 式は，もとの画像 $x_1, x_2,..., x_p$ での分散共分散行列が「一度画像 $z_1, z_2,..., z_p$ に変換して（P'）」,「固有値をならべた対角行列 Λ をとり」,「もう一度画像 $x_1, x_2,..., x_p$ にもどる（$P = (P')^{-1}$）」という操作で得られることを示しています．このことは, 画像 $z_1, z_2,..., z_p$ の分散共分散行列が対角行列 L で表され，共分散が全て０，すなわち各画素値 $z_1, z_2,..., z_p$ が互いに無相関であることを意味しています.

Karhunen-Loéve 変換

「第 k 主成分の分散の，分散の合計に対する割合」，すなわち λ_k の $\lambda_1+\lambda_2+...+\lambda_p$ に対する割合を，第 k 主成分の**寄与率**といいます．つまり寄与率は，最初の節で説明した「変換後の画素の重要さ」を表しています．ここで，ある k 以降の主成分の寄与率が０，あるいはほぼ０とみなせる場合を考えてみましょう．例えば最初に述べた２画素の画像の例では，第２主成分の寄与率をほぼ０とみなすことができ，第２主成分に対応する軸の方向の分散がほぼ０，すなわち変換後の画素値 z_2 の分散がほぼ０ということになります．つまり図１のように本来２画素で表現していた各画像については，z_2 は不要で z_2 の平均 $\bar{z}_2$ に置き換えてしまってもよく，z_1 の１変量だけで表現できることになります.

主成分分析では，なるべく第１,２，といった番号の若い主成分が分散がなるべく大きくなるように変換しています．いいかえれば，最後のほうの番号の主成分は寄与率がなるべく小さくなるようにしているわけで，最後のほうの番号の主成分を捨てて変換後の画素数を減らしたとき，元の画像との誤差が最小になります．主成分分析には，このようにして，もとのデータのもつ情報をなるべく損なわずにデータ量を減らすことができるという側面があります.

そこで，いま p 画素の画像が多数あるとします．これを伝達するのに，１度に $p/2$ 画素しか使えないとしましょう．このときに情報をなるべく損なわずに伝達するには，どうすればよいでしょうか？それはここまで述べたように，画像を主成分に変換し，第 $(p/2)$ 主成分だけを伝達して，それ以外の成分は各成分の平均値だけを１回だけ伝達しておけばよいことがわかります．もとの画素値を各主成分に変換する (22) 式の直交変換を，この意味で用いるとき **Karhunen-Loéve 変換（KL 変換）**とよんでいます.

この主成分を受け取ったほうでは，主成分への変換の逆変換（z から x への変換）を行って，元の画像にもどします．この逆変換は

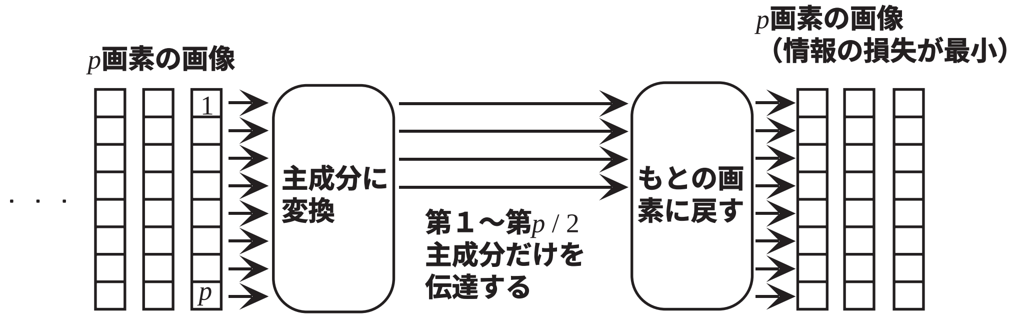

図５．KL 変換による画像の圧縮

$$\begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_p \end{pmatrix} \fallingdotseq \left(P'\right)^{-1} \begin{pmatrix} z_1 \\ z_2 \\ \vdots \\ z_{p/2} \\ \bar{z}_{p/2+1} \\ \vdots \\ \bar{z}_p \end{pmatrix} = P \begin{pmatrix} z_1 \\ z_2 \\ \vdots \\ z_{p/2} \\ \bar{z}_{p/2+1} \\ \vdots \\ \bar{z}_p \end{pmatrix} \tag{23}$$

となります．こうすると，画素の数を 1/2 にしたときに情報の損失が最小になります．

しかし，この方法を用いるには，取り扱う全ての画像を観察して分散共分散行列を求める必要がありますが，一般にはそれは不可能です．つまり，どんなデータが入ってくるかわからなければ，*KL* 変換はできません．仮に，ここまで述べてきたような「１つの画素・画素対の，各画像間での分散・共分散」を「１つの画像にある各画素・画素対の，１つの画像内での分散・共分散」で代用できるという性質（エルゴード性 (ergodicity) といいます）[*)] がいま取り扱っている画像にあれば，分散共分散行列が計算できますが，これは通常画像についてはあまり成り立たない性質です．

そこで，主成分のかわりに適当な基底ベクトルをあらかじめ決めておき，情報の損失が最小ではなくともなるべく少なくする方法が広く用いられています．これについて次回，次々回の講義で説明します．

今日の参考文献

塩谷，多変量解析概論，朝倉書店 ISBN4-254-12544-5
田中・脇本，多変量統計解析法，現代数学社 ISBN4-7687-0154-X

---

*) くわしくは，例えば M. Petrou and P. Bosdogianni, *Image Processing The Fundamentals* を見てください．