18. 統計的データ処理

- 観測データ
  - m 個の項目$x_1, \dots, x_m$についての n 組のデータが, 以下のように得られたとする

$$\{x_1(1), \dots, x_m(1)\}, \{x_1(2), \dots, x_m(2)\}, \dots, \{x_1(n), \dots, x_m(n)\}.$$

これらのデータを次のように表すことにしよう

$$X = \begin{bmatrix} x_1(1) & \dots & x_m(1) \\ \vdots & & \vdots \\ x_1(n) & \dots & x_m(n) \end{bmatrix} = \begin{bmatrix} \boldsymbol{x}_1 & \dots & \boldsymbol{x}_m \end{bmatrix} = \begin{bmatrix} \boldsymbol{d}^T(1) \\ \vdots \\ \boldsymbol{d}^T(n) \end{bmatrix},$$

ただし

$$\boldsymbol{x}_i = \begin{bmatrix} x_i(1) \\ \vdots \\ x_i(n) \end{bmatrix}$$

は項目 x_i に関するデータのベクトルであり,

$$\boldsymbol{d}(k) = \begin{bmatrix} x_1(k) \\ \vdots \\ x_m(k) \end{bmatrix}$$

は全項目についての k 番目のサンプルである

18. 統計的データ処理

- 基本統計量
  - 平均ベクトル

$$\bar{\boldsymbol{d}} = \frac{1}{n}\sum_{k=1}^{n} \boldsymbol{d}(k) = \begin{bmatrix} \frac{1}{n}\sum_{k=1}^{n} x_1(k) \\ \vdots \\ \frac{1}{n}\sum_{k=1}^{n} x_m(k) \end{bmatrix} = \begin{bmatrix} \bar{x}_1 \\ \vdots \\ \bar{x}_m \end{bmatrix}$$

  - 平均偏差行列

$$\tilde{X} = X - \begin{bmatrix} \bar{\boldsymbol{d}}^T \\ \vdots \\ \bar{\boldsymbol{d}}^T \end{bmatrix} = \begin{bmatrix} \boldsymbol{d}^T(1) - \bar{\boldsymbol{d}}^T \\ \vdots \\ \boldsymbol{d}^T(n) - \bar{\boldsymbol{d}}^T \end{bmatrix} = \begin{bmatrix} x_1(1) - \bar{x}_1 & \dots & x_m(1) - \bar{x}_m \\ \vdots & & \vdots \\ x_1(n) - \bar{x}_1 & \dots & x_m(n) - \bar{x}_m \end{bmatrix}$$

18. 統計的データ処理

- 標本共分散行列

$$S = \frac{1}{n}\tilde{X}^T\tilde{X}$$

この行列の (i,j) 成分は

$$s_{ij} = \frac{1}{n}\sum_{k=1}^{n}(x_i(k) - \bar{x}_i)(x_j(k) - \bar{x}_j)$$

であり, これは項目 x_i と項目 x_j の標本共分散を表す

- 相関係数

$$r_{ij} = \frac{s_{ij}}{\sqrt{s_{ii}}\sqrt{s_{jj}}}, \ -1 \leq r_{ij} \leq 1.$$

18. 統計的データ処理

- 重回帰分析
  - 従属変数, 目的変数 y を独立変数, 説明変数 $x_1, \ldots, x_m$ を用いて表現する

$$y = a_0 + a_1 x_1 + \cdots + a_m x_m.$$

$x_1, \ldots, x_m$ と y についての n 組のデータがあるとしよう

$$\{x_1(1), \ldots, x_m(1); y(1)\}, \ldots, \{x_1(n), \ldots, x_m(n); y(n)\}.$$

データ間の関係が次式で表されると仮定する

$$y(k) = a_0 + a_1 x_1(k) + \cdots + a_m x_m(k) + \varepsilon(k),\ k = 1, \ldots, n,$$

ここで $\varepsilon(k)$ は測定雑音（測定誤差）である

18. 統計的データ処理

関係式

$$y(k) = a_0 + a_1 x_1(k) + \cdots + a_m x_m(k) + \varepsilon(k),\ k = 1, \ldots, n,$$

は行列形式で表すことができる

$$\underbrace{\begin{bmatrix} y(1) \\ \vdots \\ y(n) \end{bmatrix}}_{\boldsymbol{y}} = \underbrace{\begin{bmatrix} 1 & x_1(1) & \ldots & x_m(1) \\ \vdots & \vdots & & \vdots \\ 1 & x_1(n) & \ldots & x_m(n) \end{bmatrix}}_{X} \underbrace{\begin{bmatrix} a_0 \\ a_1 \\ \vdots \\ a_m \end{bmatrix}}_{\boldsymbol{a}} + \underbrace{\begin{bmatrix} \varepsilon(1) \\ \vdots \\ \varepsilon(n) \end{bmatrix}}_{\boldsymbol{\varepsilon}},$$

$$\boldsymbol{y} = X\boldsymbol{a} + \boldsymbol{\varepsilon}.$$

ここで $\varepsilon(k),\ k = 1, \ldots, n$ は互いに統計的に独立で，同じ確率分布から得られるとする

上の式の中のベクトル $\boldsymbol{a} = \begin{bmatrix} a_0 \\ a_1 \\ \vdots \\ a_m \end{bmatrix}$ の値を推定したい

18. 統計的データ処理

– 最小2乗法

- $\underline{\mathrm{E}}[\boldsymbol{\varepsilon}] = \mathbf{0},\ \mathrm{E}[\boldsymbol{\varepsilon}\boldsymbol{\varepsilon}^T] = \sigma^2 I$ と仮定する

  期待値

- 次の $e(\boldsymbol{a})$ を最小にする $\boldsymbol{a}$ を求める $e(\boldsymbol{a}) = (\boldsymbol{y} - X\boldsymbol{a})^T(\boldsymbol{y} - X\boldsymbol{a}).$

$$
\begin{aligned}
e(\boldsymbol{a}) &= (\boldsymbol{y}^T - \boldsymbol{a}^T X^T)(\boldsymbol{y} - X\boldsymbol{a}) \\
&= \boldsymbol{y}^T\boldsymbol{y} - \boldsymbol{y}^T X\boldsymbol{a} - \boldsymbol{a}^T X^T \boldsymbol{y} + \boldsymbol{a}^T X^T X\boldsymbol{a}
\end{aligned}
$$

$$
\begin{aligned}
\frac{\partial}{\partial \boldsymbol{a}} e(\boldsymbol{a}) &= \mathbf{0} - (\boldsymbol{y}^T X)^T - X^T\boldsymbol{y} + X^T X\boldsymbol{a} + (\boldsymbol{a}^T X^T X)^T \\
&= -X^T\boldsymbol{y} - X^T\boldsymbol{y} + X^T X\boldsymbol{a} + X^T X\boldsymbol{a} \\
&= 2(X^T X\boldsymbol{a} - X^T\boldsymbol{y})
\end{aligned}
$$

したがって，次式を満たす $\hat{a}$ を求める必要がある

$$X^T X\hat{\boldsymbol{a}} = X^T\boldsymbol{y}$$

この式は正規方程式と呼ばれる

18. 統計的データ処理

行列 $X^T X$ が正則のとき

$$\hat{\boldsymbol{a}} = \underbrace{(X^T X)^{-1} X^T}_{X\text{の疑似逆行列}} \boldsymbol{y}.$$

次の式

$$\begin{aligned}\hat{\boldsymbol{a}} &= (X^T X)^{-1} X^T (X\boldsymbol{a} + \boldsymbol{\varepsilon}) \\ &= (X^T X)^{-1} X^T X \boldsymbol{a} + (X^T X)^{-1} X^T \boldsymbol{\varepsilon} \\ &= \boldsymbol{a} + (X^T X)^{-1} X^T \boldsymbol{\varepsilon},\end{aligned}$$

がなりたつので，以下の性質が得られる

$$\begin{aligned}\mathrm{E}[\hat{\boldsymbol{a}}] &= \boldsymbol{a} + \mathrm{E}[(X^T X)^{-1} X^T \boldsymbol{\varepsilon}] \\ &= \boldsymbol{a} + (X^T X)^{-1} X^T \underbrace{\mathrm{E}[\boldsymbol{\varepsilon}]}_{\mathbf{0}} \\ &= \boldsymbol{a}.\end{aligned}$$

すなわち，得られる $\hat{a}$ の期待は真値 $\boldsymbol{a}$ に等しい
このようなとき，$\hat{a}$ は不偏推定量であるといわれる

18. 統計的データ処理

- 例
  - データ

$$X = \begin{bmatrix} 1 & 1 \\ 1 & 2 \\ 1 & 3 \\ 1 & 4 \end{bmatrix}, \quad \boldsymbol{y} = \begin{bmatrix} 5 \\ 7 \\ 9 \\ 11 \end{bmatrix}$$

  - $\boldsymbol{a}$ の推定

$$X^T X = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & 2 & 3 & 4 \end{bmatrix} \begin{bmatrix} 1 & 1 \\ 1 & 2 \\ 1 & 3 \\ 1 & 4 \end{bmatrix} = \begin{bmatrix} 4 & 10 \\ 10 & 30 \end{bmatrix}, \, X^T \boldsymbol{y} = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & 2 & 3 & 4 \end{bmatrix} \begin{bmatrix} 5 \\ 7 \\ 9 \\ 11 \end{bmatrix} = \begin{bmatrix} 32 \\ 90 \end{bmatrix}$$

$$(X^T X)^{-1} = \frac{1}{4 \cdot 30 - 10 \cdot 10} \begin{bmatrix} 30 & -10 \\ -10 & 4 \end{bmatrix} = \begin{bmatrix} \frac{3}{2} & -\frac{1}{2} \\ -\frac{1}{2} & \frac{1}{5} \end{bmatrix}$$

$$\hat{\boldsymbol{a}} = (X^T X)^{-1} X^T \boldsymbol{y} = \begin{bmatrix} \frac{3}{2} & -\frac{1}{2} \\ -\frac{1}{2} & \frac{1}{5} \end{bmatrix} \begin{bmatrix} 32 \\ 90 \end{bmatrix} = \begin{bmatrix} 3 \\ 2 \end{bmatrix}$$

18. 統計的データ処理

- 最尤推定

データベクトル $\boldsymbol{y}$ の確率分布を定めるパラメータベクトル $\boldsymbol{a}$ を，条件付き確率密度関数 $f(\boldsymbol{y}|\boldsymbol{a})$ を最大化することによって求める

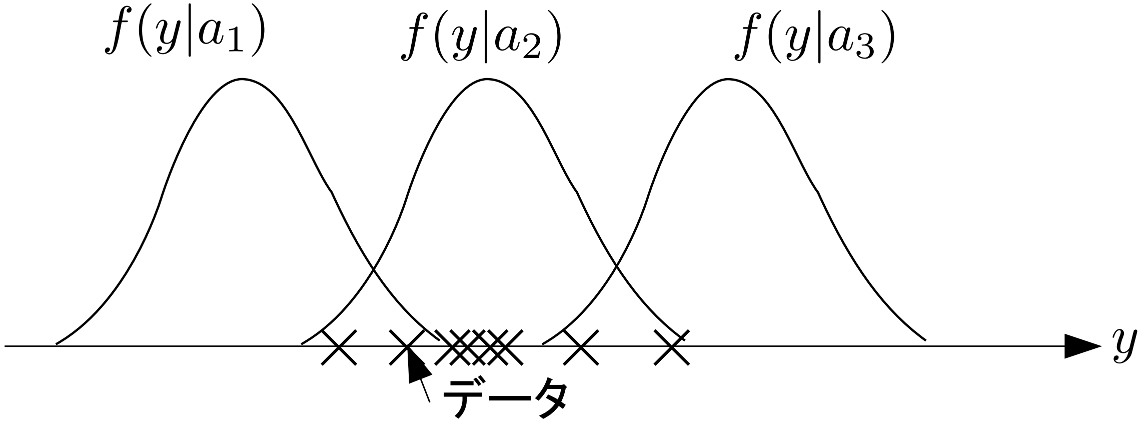
あるデータベクトル $\boldsymbol{y}$ が得られて，$f(\boldsymbol{y}|\boldsymbol{a})$ に代入されると，$f(\boldsymbol{y}|\boldsymbol{a})$ は $\boldsymbol{a}$ の関数となる
この関数を

$$L(\boldsymbol{a}) = f(\boldsymbol{y}|\boldsymbol{a}),$$

と表す．これは尤度関数と呼ばれる

条件付き確率密度関数 $f(\boldsymbol{y}|\boldsymbol{a})$ は，パラメータ $\boldsymbol{a}$ の値が与えられた時の $\boldsymbol{y}$ の確率分布を表す
ある $\boldsymbol{y}$ の値が得られたとき，この $\boldsymbol{y}$ の値は高い確率密度を持つので（雑に言うと，高い確率を持つので）得られたのだと考えることは，自然である

18. 統計的データ処理



データは，それが高い確率密度を持つから得られたと考える
したがって，上の図では，$f(y|a_2)$ が最も適切な確率密度関数であり，したがって，a_2 が最も適切な値である

18. 統計的データ処理

式

$$\boldsymbol{y} = X\boldsymbol{a} + \boldsymbol{\varepsilon},$$

において，$\boldsymbol{\varepsilon}$ は，平均ベクトルが $\mathbf{0}$ で共分散行列が $\sigma^2 I$ の n 次元ガウス分布（正規分布）から得られたものとしよう

すると，$\boldsymbol{y}$ は，平均ベクトル $X\boldsymbol{a}$，共分散行列 $\sigma^2 I$ の n 次元ガウス分布に従う

ある与えられた X に対して，データベクトル $\boldsymbol{y}$ を得たとしよう．このとき，尤度関数は次式となる

$$L(\boldsymbol{a}) = f(\boldsymbol{y}|X, \boldsymbol{a}) = \frac{1}{(2\pi\sigma^2)^{n/2}} \exp\left\{-\frac{1}{2\sigma^2}(\boldsymbol{y} - X\boldsymbol{a})^T(\boldsymbol{y} - X\boldsymbol{a})\right\}.$$

18. 統計的データ処理

$L(\boldsymbol{a})$ の最大化は $\ell(\boldsymbol{a}) = \log L(\boldsymbol{a})$ の最大化と等価である

ここで

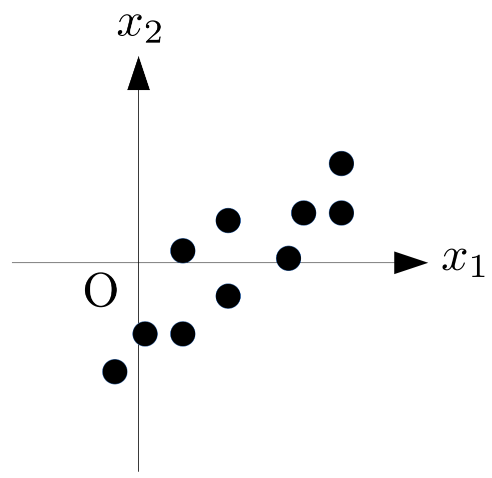
$$\ell(\boldsymbol{a}) = \underbrace{\log \frac{1}{(2\pi\sigma^2)^{n/2}}}_{\text{定数}} - \frac{1}{2\sigma^2}(\boldsymbol{y} - X\boldsymbol{a})^T(\boldsymbol{y} - X\boldsymbol{a}),$$

であるので， $(\boldsymbol{y} - X\boldsymbol{a})^T(\boldsymbol{y} - X\boldsymbol{a})$ が最小のときに， $\ell(\boldsymbol{a})$ は最大となる

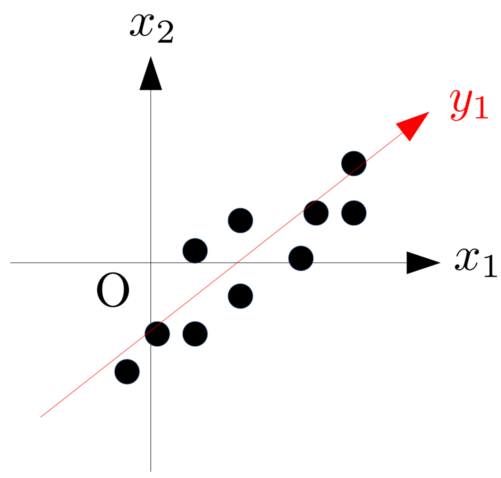
したがって，この，測定雑音がガウス分布である場合，最尤推定量は最小2乗推定量と一致する

18. 統計的データ処理

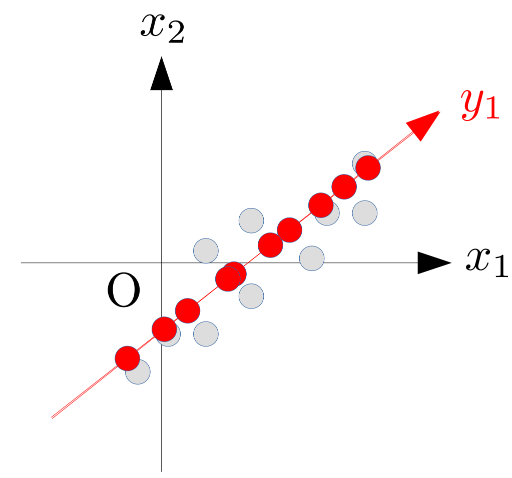
- 主成分分析
  - 高次元データが持つ情報をできるだけ失わないようにして，次元を下げる



元の**2次元**データ



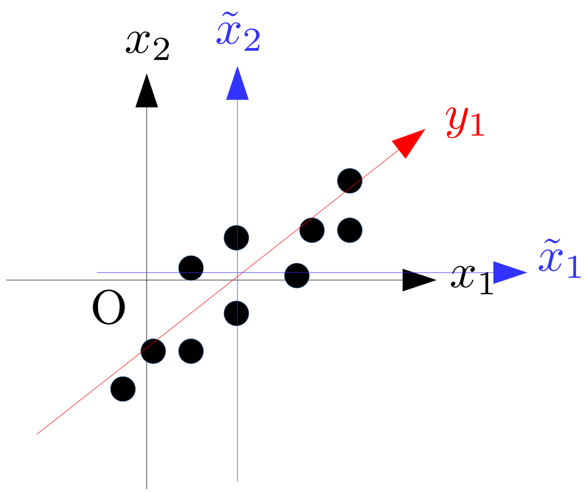
1次元のy_1軸を考える



y_1軸方向成分だけでもデータの分布をある程度把握できる

理由：y_1軸方向のデータのばらつき（分散）が大きい

18. 統計的データ処理



- データ分布の中心が原点になるように座標軸を移動させる

$$\tilde{x}_i(k) = x_i(k) - \bar{x}_i,\ \bar{x}_i = \frac{1}{n}\sum_{k=1}^{n} x_i(k)$$

- 新しい座標軸を作る

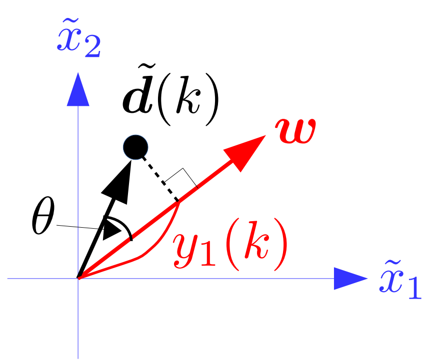
$$y_1(k) = \sum_{i=1}^{m} w_i \tilde{x}_i(k) = \boldsymbol{w}^T \tilde{\boldsymbol{d}}(k),$$

$$\tilde{\boldsymbol{d}}(k) = \begin{bmatrix} \tilde{x}_1(k) \\ \vdots \\ \tilde{x}_m(k) \end{bmatrix} = \begin{bmatrix} x_1(k) - \bar{x}_1 \\ \vdots \\ x_m(k) - \bar{x}_m \end{bmatrix} = \boldsymbol{d}(k) - \bar{\boldsymbol{d}},$$

$$i = 1, \ldots, m,\ k = 1, \ldots, n$$

新しい座標軸の方向を決めればよいので，ベクトル$\boldsymbol{w}$の大きさは任意．
そこで，$\|\boldsymbol{w}\|_2 = \sqrt{\boldsymbol{w}^T \boldsymbol{w}} = 1$とする

$$\boldsymbol{w}^T \tilde{\boldsymbol{d}}(k) = \|\boldsymbol{w}\|_2 \|\tilde{\boldsymbol{d}}(k)\|_2 \cos\theta = \|\tilde{\boldsymbol{d}}(k)\|_2 \cos\theta$$



18. 統計的データ処理

$\boldsymbol{w}^T\boldsymbol{w}=1$ の条件の下で，y_1の，データから求めた分散（標本分散）が最大になる$\boldsymbol{w}$を求める

$$
\begin{aligned}
s_{y1} &= \frac{1}{n}\sum_{k=1}^{n}\left(y_1(k)-\bar{y}_1\right)^2 = \frac{1}{n}\sum_{k=1}^{n}\left(\boldsymbol{w}^T\tilde{\boldsymbol{d}}(k)-0\right)^2 \\
&= \frac{1}{n}\sum_{k=1}^{n}\left(\boldsymbol{w}^T\tilde{\boldsymbol{d}}(k)\right)\left(\boldsymbol{w}^T\tilde{\boldsymbol{d}}(k)\right)^T \\
&= \frac{1}{n}\sum_{k=1}^{n}\boldsymbol{w}^T\left(\boldsymbol{d}(k)-\bar{\boldsymbol{d}}\right)\left(\boldsymbol{d}(k)-\bar{\boldsymbol{d}}\right)^T\boldsymbol{w} \\
&= \boldsymbol{w}^T\left[\frac{1}{n}\sum_{k=1}^{n}\left(\boldsymbol{d}(k)-\bar{\boldsymbol{d}}\right)\left(\boldsymbol{d}(k)-\bar{\boldsymbol{d}}\right)^T\right]\boldsymbol{w} \\
&= \boldsymbol{w}^T S\boldsymbol{w}
\end{aligned}
$$

18. 統計的データ処理

- Sは対称行列であり，次のように対角化できる（139ページ参照）

$$S = T\Lambda T^{-1},\ \Lambda = T^{-1}ST$$

$T = \begin{bmatrix} \boldsymbol{t}_1 & \dots & \boldsymbol{t}_m \end{bmatrix}$：$S$の固有ベクトルを並べた行列

$$T^{-1} = T^T$$

$$\therefore T^T T = \begin{bmatrix} {\boldsymbol{t}_1}^T \\ \vdots \\ {\boldsymbol{t}_m}^T \end{bmatrix} \begin{bmatrix} \boldsymbol{t}_1 & \dots & \boldsymbol{t}_m \end{bmatrix} = \begin{bmatrix} 1 & 0 & \dots & 0 \\ 0 & 1 & \dots & 0 \\ & & \ddots & \\ 0 & 0 & \ddots & 0 \\ 0 & 0 & \dots & 1 \end{bmatrix},$$

$${\boldsymbol{t}_i}^T \boldsymbol{t}_j = \begin{cases} 1, & i = j, \\ 0, & i \neq j \end{cases}$$

$$\Lambda = \begin{bmatrix} \lambda_1 & 0 & \dots & 0 \\ 0 & \lambda_2 & \dots & 0 \\ & & \ddots & \\ 0 & 0 & \ddots & 0 \\ 0 & 0 & \dots & \lambda_m \end{bmatrix},\ \lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_m$$

18. 統計的データ処理

$\boldsymbol{u}$を $\boldsymbol{w} = T\boldsymbol{u}$ と定義すると

$$\boldsymbol{w}^T S \boldsymbol{w} = (T\boldsymbol{u})^T S (T\boldsymbol{u}) = \boldsymbol{u}^T T^T S T \boldsymbol{u} = \boldsymbol{u}^T T^{-1} S T \boldsymbol{u} = \boldsymbol{u}^T \Lambda \boldsymbol{u}$$

$$= \begin{bmatrix} u_1 & \dots & u_m \end{bmatrix} \begin{bmatrix} \lambda_1 & 0 & \dots & 0 \\ 0 & \lambda_2 & \dots & 0 \\ & & \ddots & \\ 0 & 0 & & 0 \\ 0 & 0 & \dots & \lambda_m \end{bmatrix} \begin{bmatrix} u_1 \\ \vdots \\ u_m \end{bmatrix} = \sum_{i=1}^{m} \lambda_i {u_i}^2$$

$$\boldsymbol{w}^T \boldsymbol{w} = (T\boldsymbol{u})^T (T\boldsymbol{u}) = \boldsymbol{u}^T T^T T \boldsymbol{u} = \boldsymbol{u} T^{-1} T \boldsymbol{u} = \boldsymbol{u}^T \boldsymbol{u} = 1$$

したがって，やるべきことは

$\boldsymbol{u}^T \boldsymbol{u} = 1$ を満たしつつ，$\boldsymbol{u}^T \Lambda \boldsymbol{u} = \sum_{i=1}^{m} \lambda_i {u_i}^2$ を最大にする

$\boldsymbol{u} = \begin{bmatrix} u_1 \\ \vdots \\ u_m \end{bmatrix}$ を求めること

18. 統計的データ処理

$\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_m$ を考慮すると，解は

$$\boldsymbol{u}_1 = \begin{bmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix} \qquad \boldsymbol{w}_1 = T\boldsymbol{u}_1 = \boldsymbol{t}_1$$

最大固有値λ_1に対応する
固有ベクトル

$\boldsymbol{w}^T S \boldsymbol{w} = \boldsymbol{u}^T \Lambda \boldsymbol{u}$ の最大値はλ_1

18. 統計的データ処理

（別解）

- $\boldsymbol{w}^T\boldsymbol{w}=1$ の条件の下で $\boldsymbol{w}^T S\boldsymbol{w}$ を最大にする$\boldsymbol{w}$は，ラグランジュ乗数を用いて

$$\frac{\partial}{\partial \boldsymbol{w}}\left[\boldsymbol{w}^T S\boldsymbol{w} - v\left(\boldsymbol{w}^T\boldsymbol{w}-1\right)\right]=\mathbf{0}$$

の条件から求められる

- Sは対称行列であることに留意すると

$$\frac{\partial}{\partial \boldsymbol{w}}\left[\boldsymbol{w}^T S\boldsymbol{w} - v\left(\boldsymbol{w}^T\boldsymbol{w}-1\right)\right]=2S\boldsymbol{w}-2v\boldsymbol{w}=\mathbf{0}$$

よって $S\boldsymbol{w}=v\boldsymbol{w}$. すなわち，vはSの固有値，$\boldsymbol{w}$は対応する固有ベクトル

- さらに，$\boldsymbol{w}^T S\boldsymbol{w}=\boldsymbol{w}^T(v\boldsymbol{w})=v\boldsymbol{w}^T\boldsymbol{w}=v$ を最大にするので，vはSの最大固有値 λ_1 ，$\boldsymbol{w}$は λ_1 に対応する固有ベクトル $\boldsymbol{t}_1$
（行列の誘導ノルムの箇所参照）

18. 統計的データ処理

- 以上から得られる $y_1 = \boldsymbol{w}_1{}^T \tilde{\boldsymbol{x}}$ を第1主成分と呼ぶ

- 第2主成分は第1主成分以外の成分の中で，分散が最大のものとする．
  第2主成分が第1主成分以外の成分であることは，第1主成分と第2主成分の方向を定めるベクトルが直交する（ $\boldsymbol{w}_2{}^T \boldsymbol{w}_1 = 0$ ）ことによって保証する

- 第p主成分は，方向を定めるベクトル$\boldsymbol{w}$が $\boldsymbol{w}_1, \ldots, \boldsymbol{w}_{p-1}$ のいずれとも直交するという条件のもとで分散を最大化することによって定める

  $\boldsymbol{w}^T \boldsymbol{w}_i = 0,\ i = 1, \ldots, p-1,\ \|\boldsymbol{w}\|_2 = 1$ の条件の下で

  $\boldsymbol{w}^T S \boldsymbol{w}$ を最大にする

18. 統計的データ処理

第2主成分を求める問題を対角化後の行列を使って表現すると,

$$\boldsymbol{u}^T\boldsymbol{u}_1 = \begin{bmatrix} u_1 & \dots & u_m \end{bmatrix} \begin{bmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix} = u_1 = 0,\ \boldsymbol{u}^T\boldsymbol{u} = 1$$ の条件の下で

$$\boldsymbol{u}^T \mathit{\Lambda} \boldsymbol{u} = \sum_{i=1}^{m} \lambda_i {u_i}^2 = \sum_{i=2}^{m} \lambda_i {u_i}^2$$ を最大にする

解: $\boldsymbol{u}_2 = \begin{bmatrix} 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix}$, よって $\boldsymbol{w}_2 = T\boldsymbol{u}_2 = \boldsymbol{t}_2$

2番目に値が大きい最大固有値 λ_2 に対応する
固有ベクトル

18. 統計的データ処理

同様の議論を続けると，一般に

$$
\text{解：}\ \boldsymbol{u}_p = \begin{bmatrix} 0 \\ \vdots \\ 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix}
$$

p番目の成分

$$
\boldsymbol{w}_p = T\boldsymbol{u}_p = \boldsymbol{t}_p
$$

p番目に値が大きい最大固有値 λ_p に対応する
固有ベクトル

$$
y_p = {\boldsymbol{w}_p}^T \tilde{\boldsymbol{d}}
$$

y_p 軸方向の分散の値 ${\boldsymbol{w}_p}^T S \boldsymbol{w}_p = \lambda_p$

18. 統計的データ処理

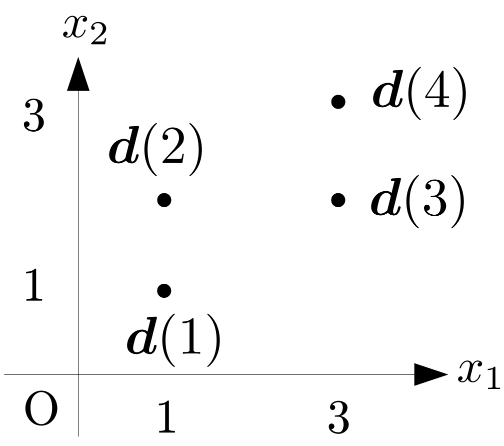
- 各主成分の貢献度は，分散 ${\boldsymbol{w}_p}^T S \boldsymbol{w}_p = \lambda_p$ によって知ることができる
- 貢献度が大きい主成分だけを使って，元のデータの情報を低次元のデータで近似することができる

18. 統計的データ処理

- 例
  - データ

$$\boldsymbol{d}(1) = \begin{bmatrix}1\\1\end{bmatrix}, \boldsymbol{d}(2) = \begin{bmatrix}1\\2\end{bmatrix},$$

$$\boldsymbol{d}(3) = \begin{bmatrix}3\\2\end{bmatrix}, \boldsymbol{d}(4) = \begin{bmatrix}3\\3\end{bmatrix}$$



  - 平均値

$$\bar{\boldsymbol{d}} = \begin{bmatrix}2\\2\end{bmatrix}$$

  - 標本共分散行列

$$S = \frac{1}{4}\begin{bmatrix}1-2 & 1-2 & 3-2 & 3-2\\1-2 & 2-2 & 2-2 & 3-2\end{bmatrix}\begin{bmatrix}-1 & -1\\-1 & 0\\1 & 0\\1 & 1\end{bmatrix} = \begin{bmatrix}1 & \frac{1}{2}\\\frac{1}{2} & \frac{1}{2}\end{bmatrix}$$

18. 統計的データ処理

- Sの固有値，固有ベクトル

$$\lambda_1 = \frac{3+\sqrt{5}}{4} \simeq 1.309, \boldsymbol{t}_1 = \begin{bmatrix} 0.8507 \\ 0.5257 \end{bmatrix} = \boldsymbol{w}_1$$

$$\lambda_2 = \frac{3-\sqrt{5}}{4} \simeq 1.309, \boldsymbol{t}_2 = \begin{bmatrix} 0.5257 \\ -0.8507 \end{bmatrix} = \boldsymbol{w}_2$$