多変量解析に必要な基礎知識

概要

- 線形代数
  - ベクトル、行列の演算
  - 固有値・固有ベクトル
- 統計・確率
  - 平均、分散、共分散、相関
  - t分布
- 微分
  - 偏微分

式がわからなくなったときは、小さいベクトル(1行2列や2行1列)や行列(2行2列)で考えるとわかりやすい

線形代数の基礎知識(1)

- ベクトルの和と差

$$\begin{bmatrix} a & b \end{bmatrix} + \begin{bmatrix} p & q \end{bmatrix} = \begin{bmatrix} a+p & b+q \end{bmatrix}$$

Rでの一例：c(1,2)+c(3,4)

$$\begin{bmatrix} a \\ b \end{bmatrix} + \begin{bmatrix} p \\ q \end{bmatrix} = \begin{bmatrix} a+p \\ b+q \end{bmatrix}$$

R：t(c(1,2))+t(c(3,4))

$$\begin{bmatrix} a & b \end{bmatrix} - \begin{bmatrix} p & q \end{bmatrix} = \begin{bmatrix} a-p & b-q \end{bmatrix}$$

R：c(1,2)-c(3,4)

$$\begin{bmatrix} a \\ b \end{bmatrix} - \begin{bmatrix} p \\ q \end{bmatrix} = \begin{bmatrix} a-p \\ b-q \end{bmatrix}$$

R：t(c(1,2))-t(c(3,4))

- 行列の和と差

$$\begin{bmatrix} a & b \\ c & d \end{bmatrix} + \begin{bmatrix} p & q \\ r & s \end{bmatrix} = \begin{bmatrix} a+p & b+q \\ c+r & d+s \end{bmatrix}$$

R：matrix(c(1,2,3,4),nrow=2, byrow=T)+matrix(c(5,6,7,8),nrow=2, byrow=T)

$$\begin{bmatrix} a & b \\ c & d \end{bmatrix} - \begin{bmatrix} p & q \\ r & s \end{bmatrix} = \begin{bmatrix} a-p & b-q \\ c-r & d-s \end{bmatrix}$$

R：matrix(c(1,2,3,4),nrow=2, byrow=T)-matrix(c(5,6,7,8),nrow=2, byrow=T)

線形代数の基礎知識(2)

- ベクトルの掛け算

$$\begin{bmatrix} a & b \end{bmatrix}\begin{bmatrix} c \\ d \end{bmatrix} = ac + bd$$

$\underline{1}\times 2$　$2\times\underline{1}$　　$\underline{1}\times\underline{1}$ (1)

R：c(1,2)%*%t(c(3,4))

$$\begin{bmatrix} a \\ b \end{bmatrix}\begin{bmatrix} c & d \end{bmatrix} = \begin{bmatrix} ac & ad \\ bc & bd \end{bmatrix}$$

$\underline{2}\times 1$　$1\times\underline{2}$　　$\underline{2}\times\underline{2}$

R : t(c(1,2))%*%c(3,4)

- 行列の掛け算

$$\begin{bmatrix} \boxed{a \quad b} \\ c \quad d \end{bmatrix}\begin{bmatrix} p & q \\ r & s \end{bmatrix} = \begin{bmatrix} \boxed{ap+br} & aq+bs \\ cp+dr & cq+ds \end{bmatrix}$$

$\underline{2}\times 2$　　$2\times\underline{2}$　　$\underline{2}\times\underline{2}$

R : matrix(c(1,2,3,4),nrow=2, byrow=T)%*%matrix(c(5,6,7,8),nrow=2, byrow=T)

線形代数の基礎知識(3)

- 行列の掛け算の注意点

$$\mathbf{A} = \begin{bmatrix} a & b \\ c & d \end{bmatrix} \quad \mathbf{P} = \begin{bmatrix} p & q \\ r & s \end{bmatrix}$$ のとき、

$$\mathbf{AP} = \begin{bmatrix} a & b \\ c & d \end{bmatrix} \begin{bmatrix} p & q \\ r & s \end{bmatrix} = \begin{bmatrix} ap+br & aq+bs \\ cp+dr & cq+ds \end{bmatrix}$$

$$\mathbf{PA} = \begin{bmatrix} p & q \\ r & s \end{bmatrix} \begin{bmatrix} a & b \\ c & d \end{bmatrix} = \begin{bmatrix} pa+qc & pb+qd \\ ra+sc & rb+sd \end{bmatrix}$$

行列の掛け算は、掛ける順番を変えると一致しないことに注意

線形代数の基礎知識(4)

- 転置行列

$$A=\begin{bmatrix} a & b \\ c & d \end{bmatrix}$$ のとき、Aの転置行列は $$A'=\begin{bmatrix} a & c \\ b & d \end{bmatrix}$$

R : t(matrix(c(1,2,3,4),nrow=2, byrow=T))

- 対称行列

行列とその転置行列が一致する　$B=B'$

$$B=\begin{bmatrix} a & b \\ b & d \end{bmatrix}$$ のとき、Bの転置行列は $$B'=\begin{bmatrix} a & b \\ b & d \end{bmatrix}$$

線形代数の基礎知識(5)

- 逆行列

$A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$ のとき、Aの逆行列は $A^{-1} = \dfrac{1}{ad - bc}\begin{bmatrix} d & -b \\ -c & a \end{bmatrix}$

R : solve(matrix(c(1,2,3,4),nrow=2, byrow=T))

- 単位行列

$$I = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

R : diag(2)

- 単位行列の性質(1)

行列Aと単位行列Iを掛けると行列Aになる

$$AI = \begin{bmatrix} a & b \\ c & d \end{bmatrix}\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} a & b \\ c & d \end{bmatrix} = A$$

単位行列の積では、AI=IAが成り立つ

線形代数の基礎知識(6)

- 単位行列の性質(2)

行列Aとその逆行列を掛けると単位行列Iになる

$$AA^{-1} = \frac{1}{ad-bc}\begin{bmatrix} a & b \\ c & d \end{bmatrix}\begin{bmatrix} d & -b \\ -c & a \end{bmatrix}$$

$$= \frac{1}{ad-bc}\begin{bmatrix} ad-bc & -ab+ab \\ cd-dc & -bc+ad \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = I$$

- 固有値・固有ベクトル

$\begin{bmatrix} a & b \\ c & d \end{bmatrix}\begin{bmatrix} x \\ y \end{bmatrix} = \lambda \begin{bmatrix} x \\ y \end{bmatrix}$ を満たすλを固有値、$\begin{bmatrix} x \\ y \end{bmatrix}$ を固有ベクトルという

R : eigen(matrix(c(1,2,3,4),nrow=2, byrow=T))

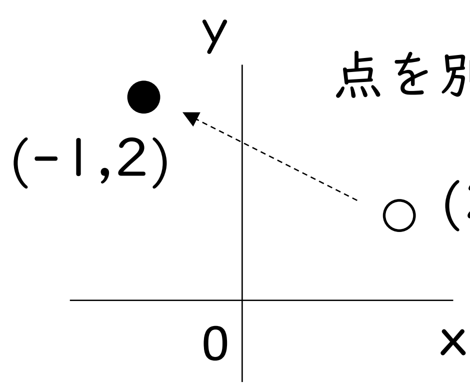
線形代数の参考知識(1)

ベクトルに行列を掛ける操作 $\begin{bmatrix} a & b \\ c & d \end{bmatrix}\begin{bmatrix} x \\ y \end{bmatrix}$

は、点を別の場所に移すことを意味している

<u>例</u>

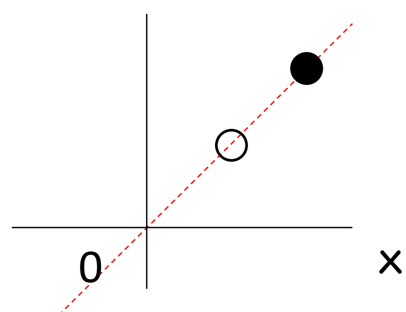
$$\begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}\underbrace{\begin{bmatrix} 2 \\ 1 \end{bmatrix}}_{\text{移動前}}=\underbrace{\begin{bmatrix} -1 \\ 2 \end{bmatrix}}_{\text{移動後}}$$



点を別の場所に移す

固有ベクトルは、

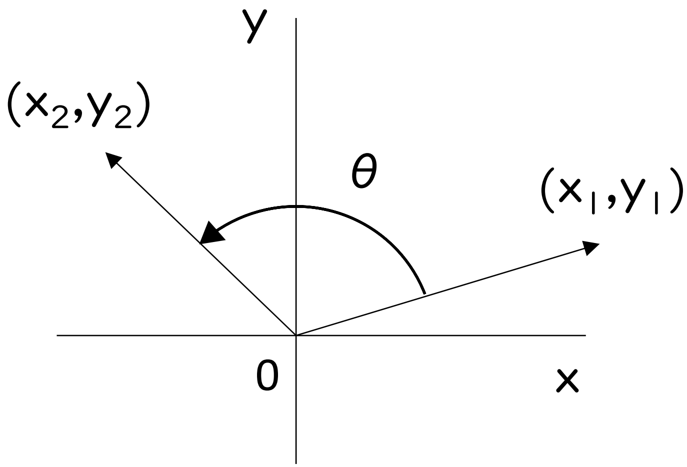
$$\begin{bmatrix} a & b \\ c & d \end{bmatrix}\begin{bmatrix} x \\ y \end{bmatrix}=\lambda\begin{bmatrix} x \\ y \end{bmatrix}$$



同一線上に点を移す
特殊なベクトル

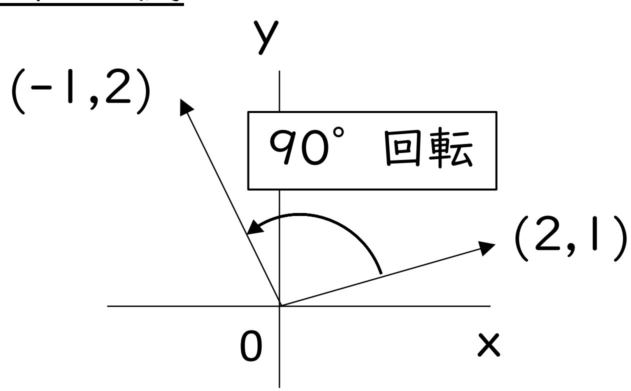
線形代数の参考知識(2)

- 回転行列 $\begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}$ (2変数の場合)



$$\begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix} \begin{bmatrix} x_1 \\ y_1 \end{bmatrix} = \begin{bmatrix} x_2 \\ y_2 \end{bmatrix}$$

<u>回転の例</u>



$$\begin{bmatrix} \cos 90^\circ & -\sin 90^\circ \\ \sin 90^\circ & \cos 90^\circ \end{bmatrix} \underbrace{\begin{bmatrix} 2 \\ 1 \end{bmatrix}}_{\text{回転前}} = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} 2 \\ 1 \end{bmatrix} = \underbrace{\begin{bmatrix} -1 \\ 2 \end{bmatrix}}_{\text{回転後}}$$

統計の基礎知識(1)

- 平均

$$\mu = \frac{1}{n}\sum_{i=1}^{n} x_i = \frac{x_1 + x_2 + \cdots + x_n}{n}$$

R：mean(c(1,2,3,4,5))

- 分散

(不偏)分散 $$\sigma^2 = \frac{1}{n-1}\sum_{i=1}^{n}\left(x - x_i\right)^2$$

R：var(c(1,2,3,4,5))

- 標準偏差

(不偏)標準偏差 $$\sigma = \sqrt{\frac{1}{n-1}\sum_{i=1}^{n}\left(x - x_i\right)^2}$$

R：sd(c(1,2,3,4,5))

統計の基礎知識(2)

- 共分散

$$\mathrm{cov}(x,y)=\frac{1}{n-1}\sum_{i=1}^{n}\left(x-x_i\right)\left(y-y_i\right)$$

R： x <- c(1,2,3,4,5); y <- c(3,2,1,4,5)
cov(x,y)

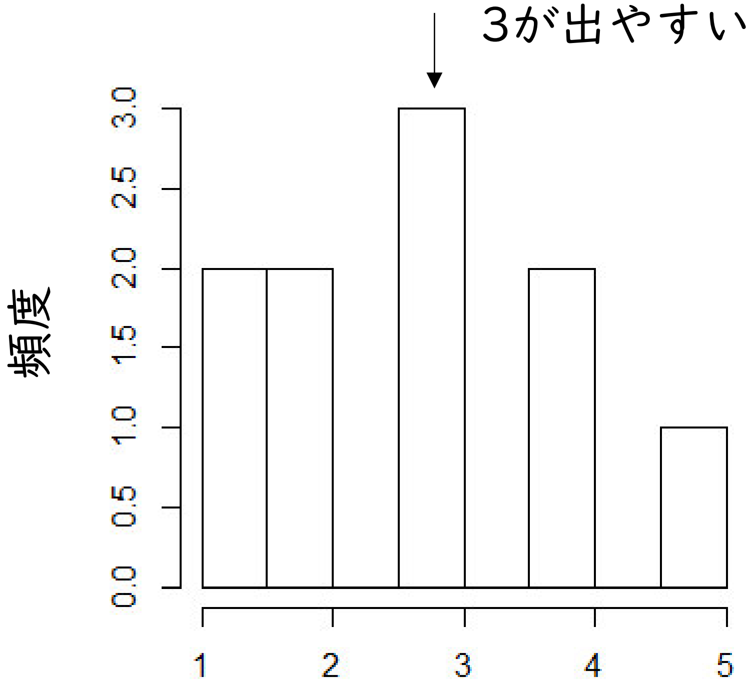
- (ピアソン)相関係数

$$\mathrm{co}\,rr(x,y)=\frac{\mathrm{cov}(x,y)}{\sqrt{\mathrm{var}(x)}\sqrt{\mathrm{var}(y)}}=\frac{\frac{1}{n-1}\sum_{i=1}^{n}\left(x-x_i\right)\left(y-y_i\right)}{\sqrt{\frac{1}{n-1}\sum_{i=1}^{n}\left(x-x_i\right)^2}\sqrt{\frac{1}{n-1}\sum_{i=1}^{n}\left(y-y_i\right)^2}}$$

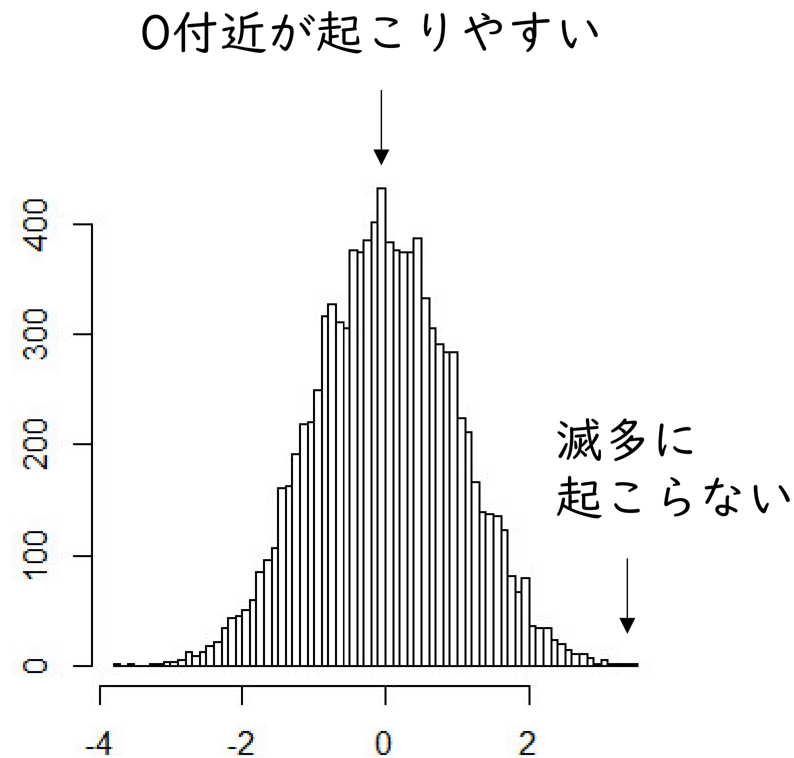
R： x <- c(1,2,3,4,5); y <- c(3,2,1,4,5)
cor.test(x,y)

確率の基礎知識(1)

ヒストグラムは、それぞれの値がどのくらいの頻度で起こるかを表したグラフ



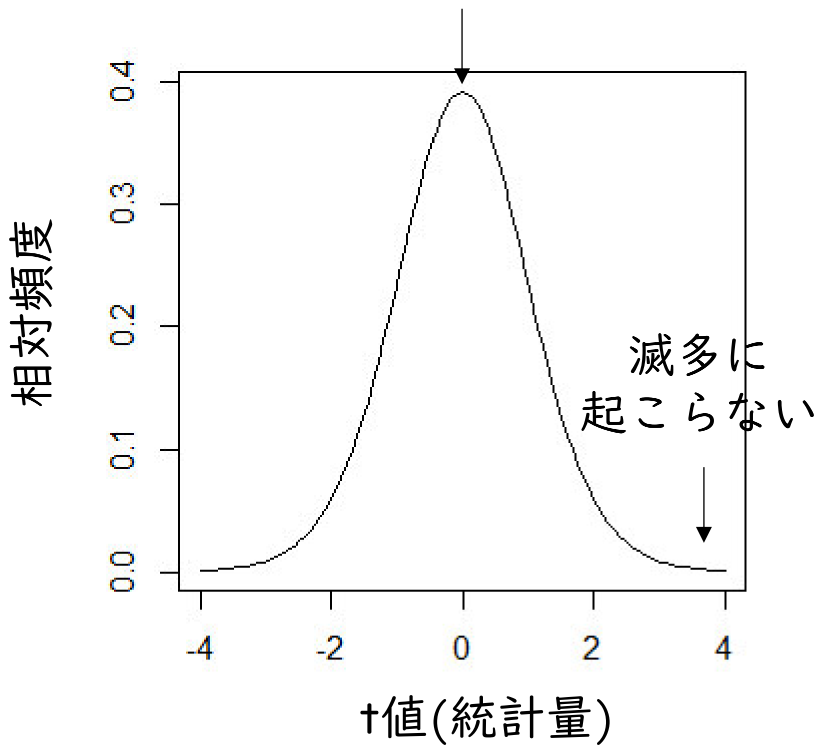
R : x <- c(1,2,2,3,3,4,5,4,3,1)
hist(x, breaks=6)



R : hist(rnorm(10000), breaks=100)

確率の基礎知識(2)

- t分布



Student t検定(2群の平均値の差)

$$t = \frac{\left|\bar{X} - \bar{Y}\right|}{\sqrt{\frac{(m-1)U_x + (n-1)U_y}{m+n-2}\left(\frac{1}{m} + \frac{1}{n}\right)}}$$

相関係数

$$t = \frac{|r|\sqrt{n-2}}{\sqrt{1-r^2}}$$

t分布に従うことを利用して、
差や相関の起こりやすさを調べる
ことが出来る

p-value：データから計算した差や相関が、
どの程度の確率で起こるかの指標

微分の基礎知識(1)

- 微分

f(x)=ax+bのxでの微分　df=a

- 偏微分

f(x,y)=ax+byのxまたはyでの偏微分

$$\frac{\partial f(x,y)}{\partial x} = a \qquad \frac{\partial f(x,y)}{\partial y} = b$$

微分の基礎知識(2)

- ベクトルの偏微分

t=Xw

$$\begin{bmatrix} t_1 \\ t_2 \end{bmatrix} = \begin{bmatrix} x_{11} & x_{12} \\ x_{21} & x_{22} \end{bmatrix} \begin{bmatrix} w_1 \\ w_2 \end{bmatrix} \longrightarrow \begin{aligned} t_1 &= x_{11}w_1 + x_{12}w_2 \\ t_2 &= x_{21}w_1 + x_{22}w_2 \end{aligned}$$

wで偏微分

$$\frac{\partial t_1}{\partial w_1} = x_{11} \quad \frac{\partial t_1}{\partial w_2} = x_{12}$$

$$\frac{\partial t_2}{\partial w_1} = x_{21} \quad \frac{\partial t_2}{\partial w_2} = x_{22}$$

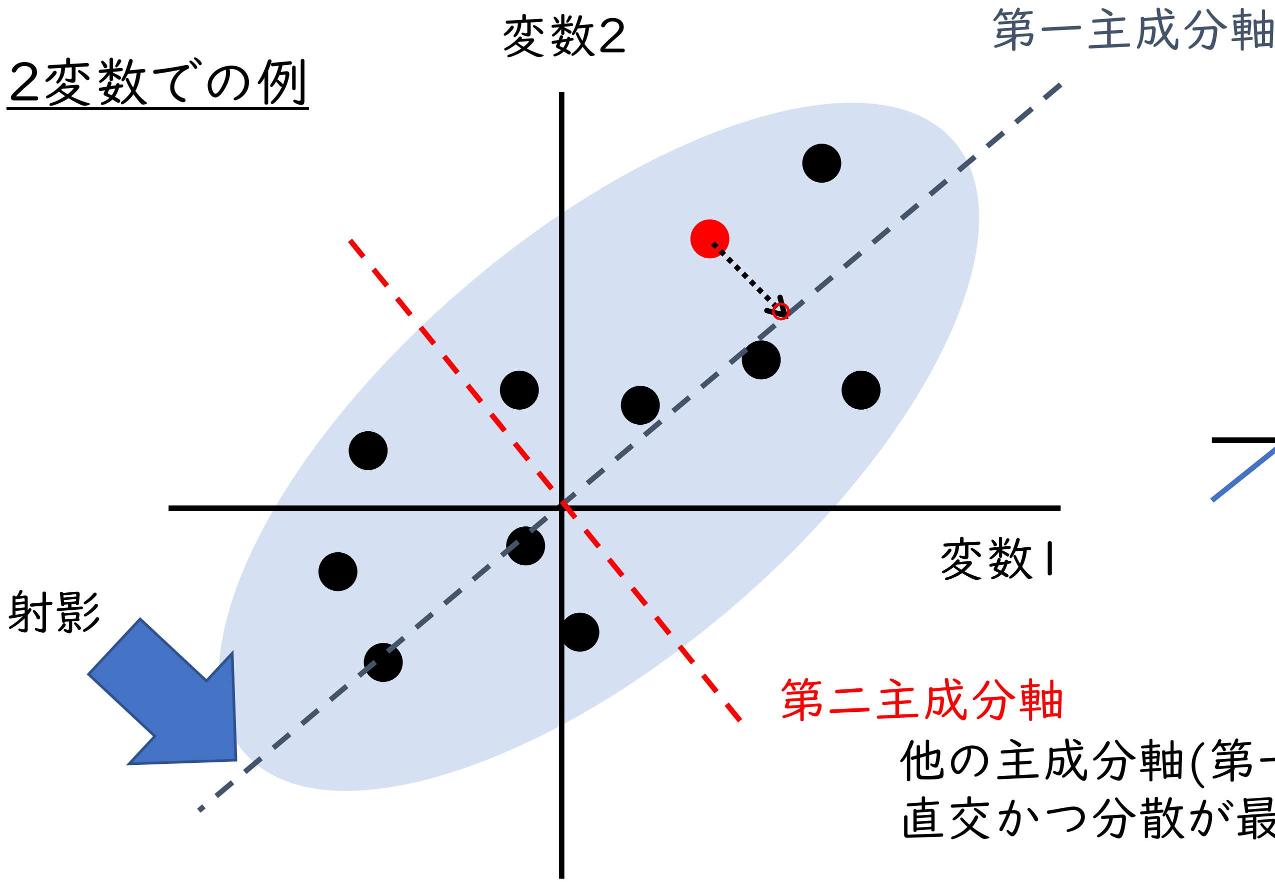
$$\begin{bmatrix} \dfrac{\partial t_1}{\partial w_1} & \dfrac{\partial t_1}{\partial w_2} \\ \dfrac{\partial t_2}{\partial w_1} & \dfrac{\partial t_2}{\partial w_2} \end{bmatrix} = \begin{bmatrix} x_{11} & x_{12} \\ x_{21} & x_{22} \end{bmatrix}$$

ベクトル**t**のベクトル**w**による偏微分は、 $\dfrac{\partial \mathbf{t}}{\partial \mathbf{w}} = \mathbf{X}$

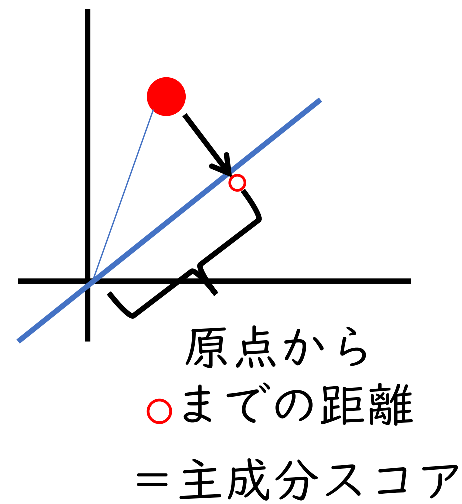
主成分分析の理論(詳細)

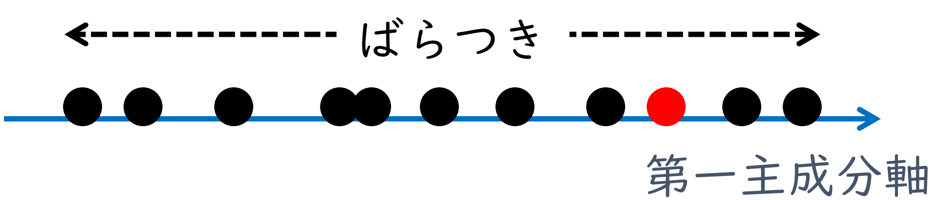
主成分分析は、群情報を利用しない方法

2変数での例



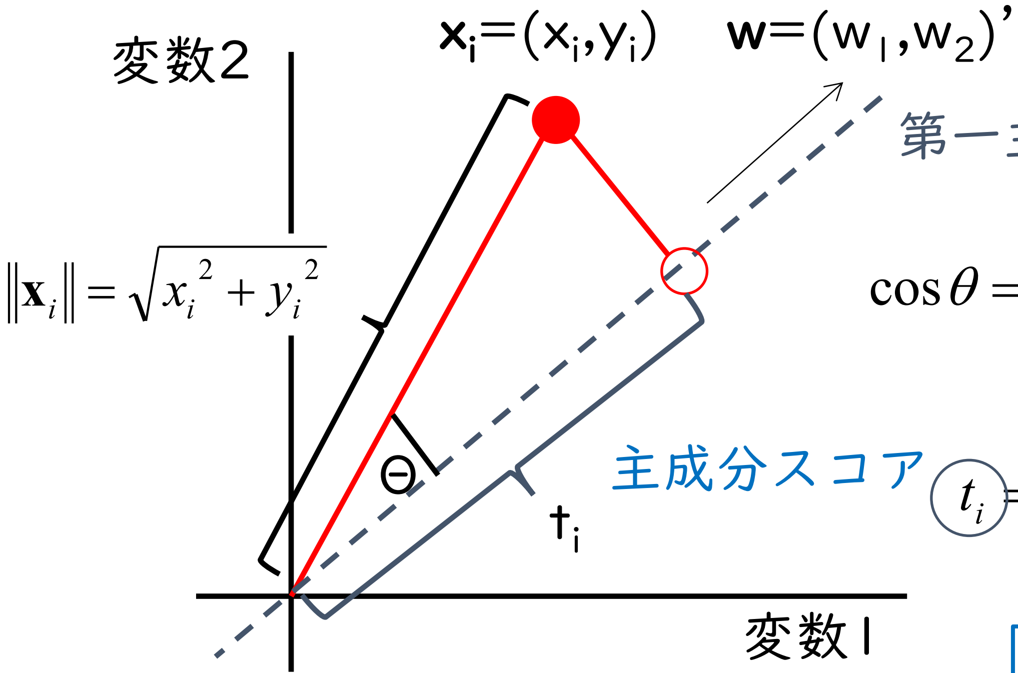
他の主成分軸(第一主成分軸)と
直交かつ分散が最大方向





主成分スコアの分散が最大に
なるように主成分軸を決める

主成分分析は、スコアの分散を最大化する方法



$$\cos\theta = \frac{\mathbf{x_i}\cdot\mathbf{w}}{\|\mathbf{x_i}\|\|\mathbf{w}\|}$$

$$t_i = \|\mathbf{x_i}\|\cos\theta = \frac{\|\mathbf{x_i}\|\mathbf{x_i}\cdot\mathbf{w}}{\|\mathbf{x_i}\|\|\mathbf{w}\|} = \mathbf{x_i}\cdot\frac{\mathbf{w}}{\|\mathbf{w}\|}$$

主成分スコアの分散を最大化

$$Var(\mathbf{t}) = \frac{1}{n-1}\sum_{i=1}^{n}\frac{\mathbf{w}'\mathbf{x_i}'\mathbf{x_i}\mathbf{w}}{\|\mathbf{w}\|^2} = \frac{1}{n-1}\frac{\mathbf{w}'\mathbf{X}'\mathbf{X}\mathbf{w}}{\|\mathbf{w}\|^2}$$ が最大になるwを計算

$$\boxed{\frac{1}{n-1}\mathbf{X}'\mathbf{X}\mathbf{w} = \lambda\mathbf{w}}$$

第i主成分軸: i番目に大きい固有値に対応する固有ベクトルw

主成分分析

測定データ(行列) X

$$
\begin{bmatrix}
x_{11} & x_{12} & x_{13} & x_{14} & \cdots & x_{1p} \\
x_{21} & x_{22} & x_{23} & x_{24} & \cdots & x_{2p} \\
\vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\
x_{n1} & x_{n2} & x_{n3} & x_{n4} & \cdots & x_{np}
\end{bmatrix}
$$

サンプル1のデータ

(横)ベクトル

変数1のデータ

(波長、代謝物など)

(縦)ベクトル

主成分分析

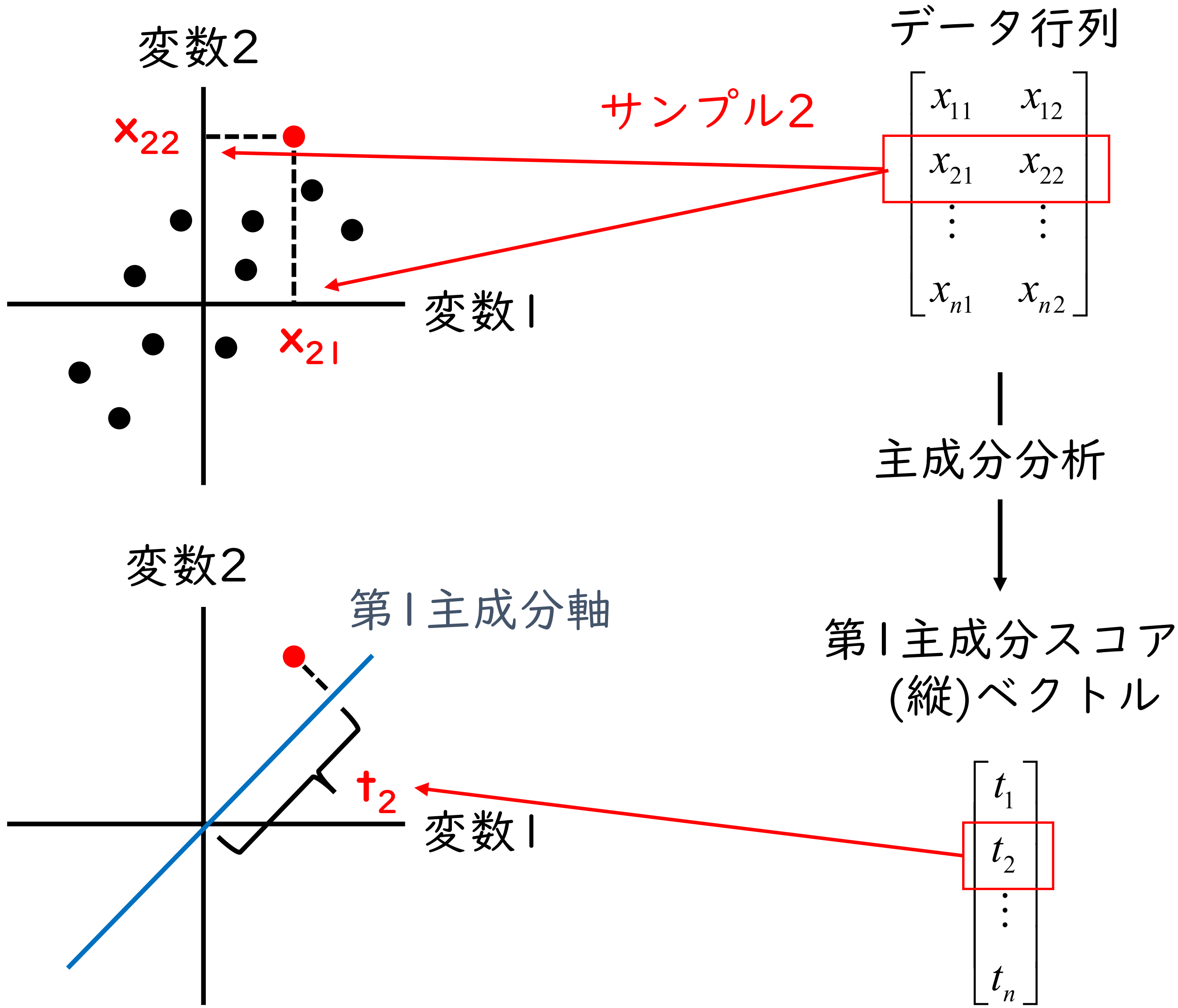
主成分スコア(行列)

$$
\begin{bmatrix}
t_{11} & t_{12} \\
t_{21} & t_{22} \\
\vdots & \vdots \\
t_{n1} & t_{n2}
\end{bmatrix}
$$

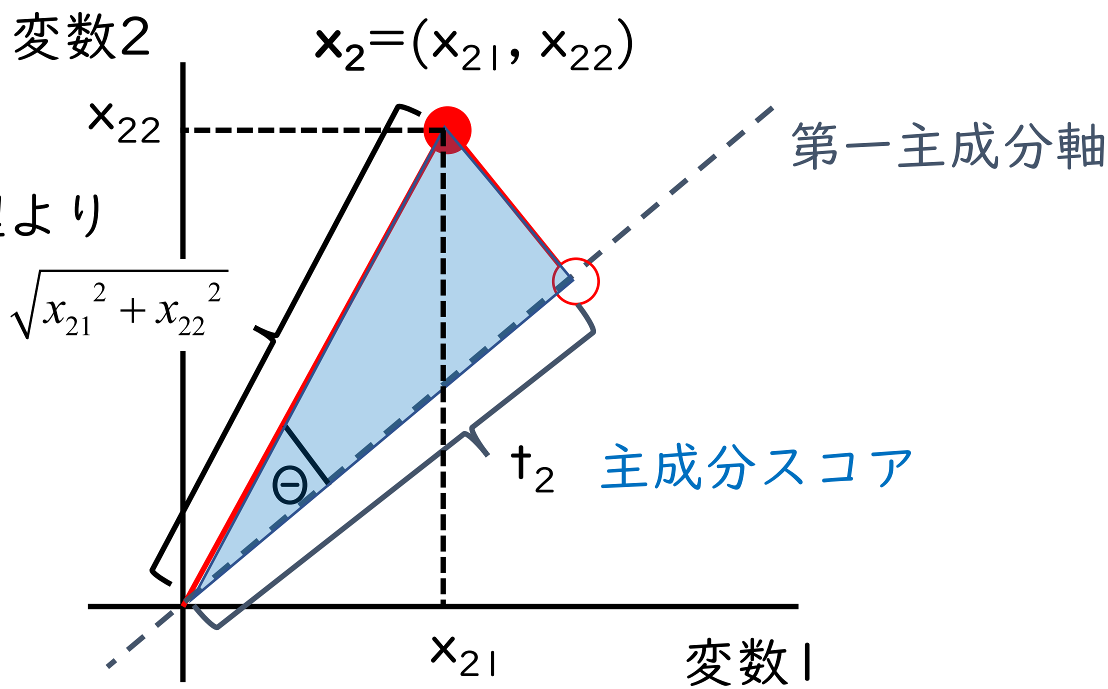
サンプル1のスコア

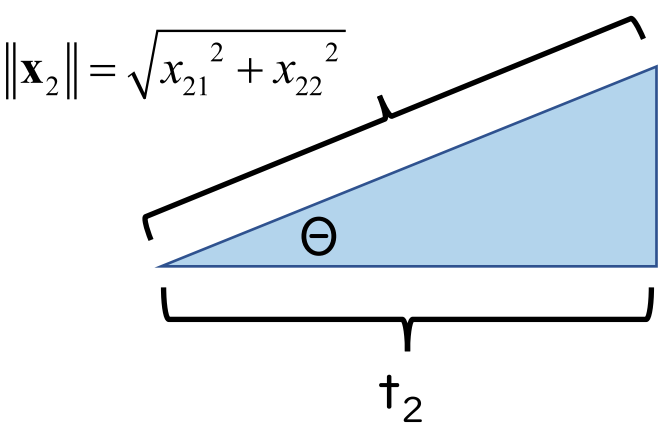
第1主成分 t

2変数での単純な例


変数2
x_{22}
x_{21}
変数1
サンプル2
データ行列
$\begin{bmatrix} x_{11} & x_{12} \\ x_{21} & x_{22} \\ \vdots & \vdots \\ x_{n1} & x_{n2} \end{bmatrix}$
主成分分析
変数2
第1主成分軸
t_2
変数1
第1主成分スコア
(縦)ベクトル
$\begin{bmatrix} t_1 \\ t_2 \\ \vdots \\ t_n \end{bmatrix}$

2変数での単純な例

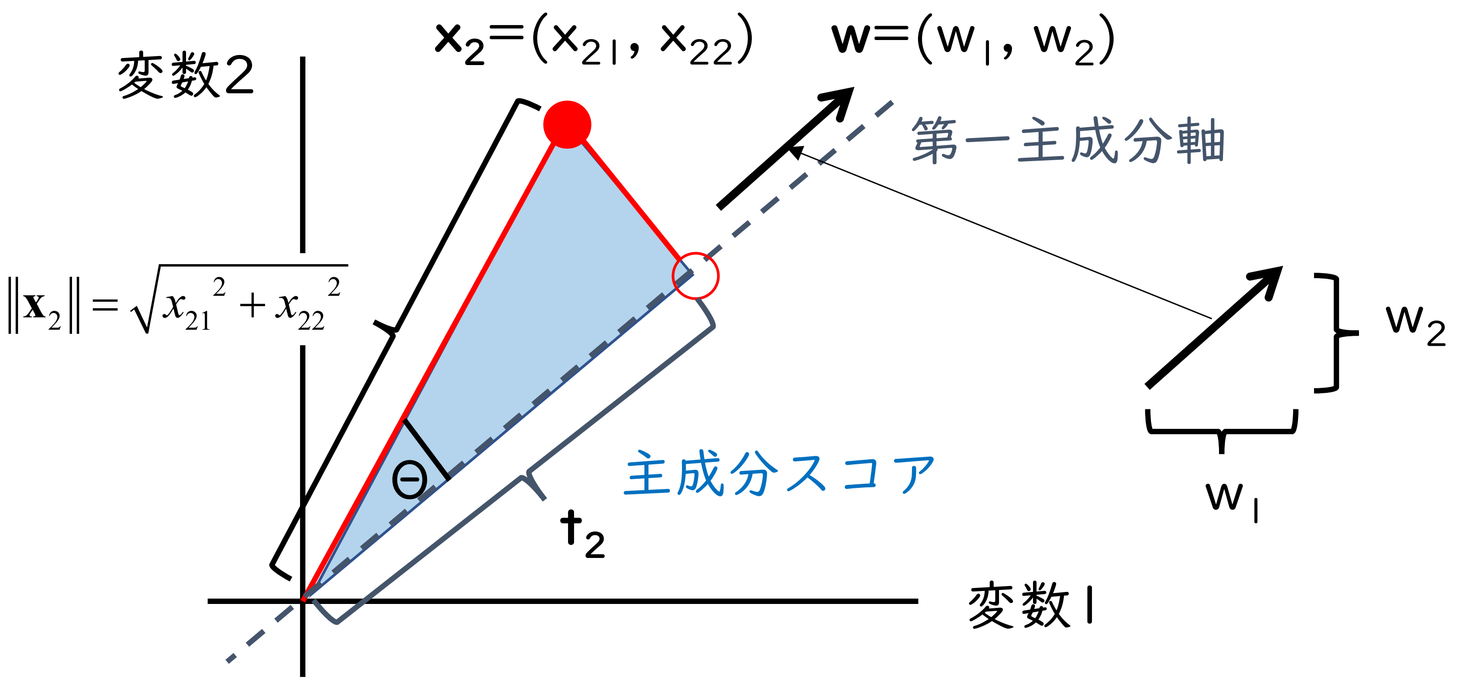


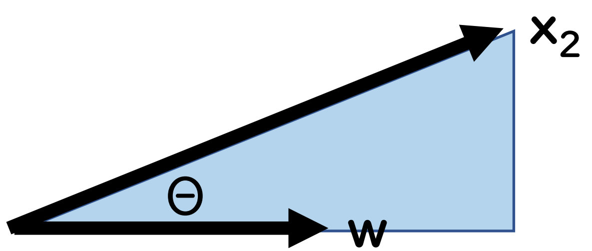


cosの定義から

$$\cos\theta = \frac{t_2}{\|\mathbf{x}_2\|} \longrightarrow t_2 = \underbrace{\|\mathbf{x}_2\|}_{\text{データ}}\underbrace{\cos\theta}_{?}$$

2変数での単純な例





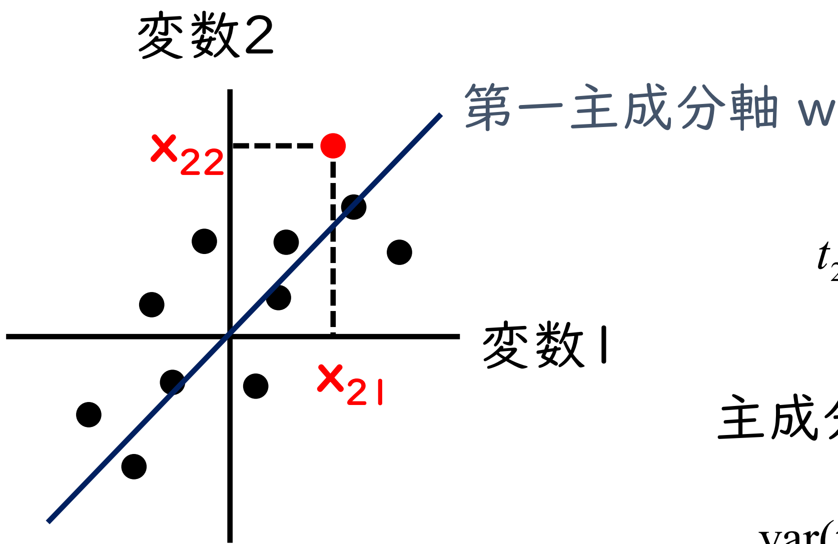
ベクトルの内積の定義より、

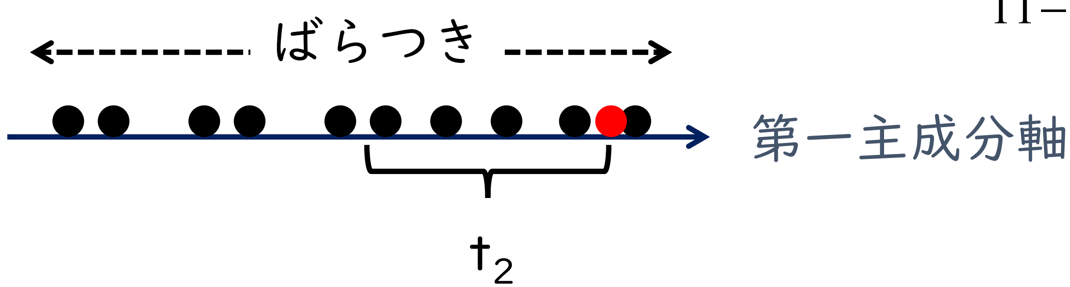
$$\mathbf{x}_2 \cdot \mathbf{w} = \|\mathbf{x}_2\|\|\mathbf{w}\|\cos\theta \longrightarrow \cos\theta = \frac{\mathbf{x}_2 \cdot \mathbf{w}}{\|\mathbf{x}_2\|\|\mathbf{w}\|}$$

前ページの式に代入して

$$t_2 = \|\mathbf{x}_2\|\cos\theta = \|\mathbf{x}_2\|\frac{\mathbf{x}_2 \cdot \mathbf{w}}{\|\mathbf{x}_2\|\|\mathbf{w}\|} = \frac{\mathbf{x}_2 \cdot \mathbf{w}}{\|\mathbf{w}\|}$$

主成分分析は、スコア分散を最大化する方法





$$t_2 = \|\mathbf{x}_2\| \cos\theta = \|\mathbf{x}_2\| \frac{\mathbf{x}_2 \cdot \mathbf{w}}{\|\mathbf{x}_2\| \|\mathbf{w}\|} = \frac{\mathbf{x}_2 \cdot \mathbf{w}}{\|\mathbf{w}\|}$$

主成分スコアの分散

$$\mathrm{var}(\mathbf{t}) = \frac{1}{11-1} t_1^2 + t_2^2 + t_3^2 + \cdots + t_{11}^2 = \frac{1}{n-1} \sum_{i=1}^{n} t_i^{\,2}$$

$$var(\mathbf{t}) = \frac{1}{n-1} \sum_{i=1}^{n} \frac{\mathbf{w}'\mathbf{x_i}'\mathbf{x_i}\mathbf{w}}{\|\mathbf{w}\|^2} = \frac{1}{n-1} \frac{1}{\|\mathbf{w}\|^2} \mathbf{w}' \underline{\left(\sum_{i=1}^{n} \mathbf{x_i}'\mathbf{x_i} \right)} \mathbf{w}$$

主成分分析の導出(1)

$\sum_{i=1}^{n} \mathbf{x_i}'\mathbf{x_i}$ について考える

2サンプル2変数の単純な場合について考える。

$$\sum_{i=1}^{2} \mathbf{x_i}'\mathbf{x_i} = \mathbf{x}_1'\mathbf{x}_1 + \mathbf{x}_2'\mathbf{x}_2 = \begin{bmatrix} x_{11} \\ x_{21} \end{bmatrix} \begin{bmatrix} x_{11} & x_{21} \end{bmatrix} + \begin{bmatrix} x_{12} \\ x_{22} \end{bmatrix} \begin{bmatrix} x_{12} & x_{22} \end{bmatrix}$$

$$= \begin{bmatrix} {x_{11}}^2 & x_{11}x_{21} \\ x_{21}x_{11} & {x_{21}}^2 \end{bmatrix} + \begin{bmatrix} {x_{12}}^2 & x_{12}x_{22} \\ x_{22}x_{12} & {x_{22}}^2 \end{bmatrix} = \begin{bmatrix} {x_{11}}^2 + {x_{12}}^2 & x_{11}x_{21} + x_{12}x_{22} \\ x_{21}x_{11} + x_{22}x_{12} & {x_{21}}^2 + {x_{22}}^2 \end{bmatrix}$$

ここで、

$$X = \begin{array}{cc} \left[\begin{array}{cc} x_{11} & x_{12} \\ x_{21} & x_{22} \end{array}\right] & \begin{array}{l} \text{サンプル1} \\ \end{array} \\ \begin{array}{l} \text{変数1} \end{array} & \end{array}$$

とすると、

$$X'X = \begin{bmatrix} x_{11} & x_{21} \\ x_{12} & x_{22} \end{bmatrix} \begin{bmatrix} x_{11} & x_{12} \\ x_{21} & x_{22} \end{bmatrix} = \begin{bmatrix} {x_{11}}^2 + {x_{21}}^2 & x_{11}x_{12} + x_{21}x_{22} \\ x_{12}x_{11} + x_{22}x_{21} & {x_{12}}^2 + {x_{22}}^2 \end{bmatrix}$$

主成分分析の導出(2)

$$var(\mathbf{t}) = \frac{1}{n-1}\frac{1}{\|\mathbf{w}\|^2}\mathbf{w}'\underline{\left(\sum_{i=1}^{n}\mathbf{x_i}'\mathbf{x_i}\right)}\mathbf{w} = \frac{1}{n-1}\frac{\mathbf{w}'\underline{\underline{\mathbf{X}'\mathbf{X}}}\mathbf{w}}{\|\mathbf{w}\|^2}$$

この値を最大化するようなwを求めたい

$$\frac{1}{n-1}\frac{\boxed{\mathbf{w}'\mathbf{X}'\mathbf{X}\mathbf{w}}}{\|\mathbf{w}\|^2}$$

$\longleftarrow$ 最大

$\longleftarrow$ $\|\mathbf{w}\|^2 = 1$ で固定

とすると、ラグランジュ乗数法より、

“https://ja.wikipedia.org/wiki/ラグランジュの未定乗数法”

$$J = \frac{1}{n-1}\mathbf{w}'\mathbf{X}'\mathbf{X}\mathbf{w} + \lambda(1-\|\mathbf{w}\|^2)$$

Jが最大になるwを求めれば良い

主成分分析の導出(3)

Jをwで偏微分して0とすると

$$\frac{\partial J}{\partial w} = \frac{1}{n-1} 2\mathbf{X'Xw} - 2\lambda \mathbf{w} = 0$$

整理すると、主成分分析は次の式で書ける

$$\frac{1}{n-1}\mathbf{X'Xw} = \lambda \mathbf{w}$$

よって、行列 $\underline{\frac{1}{n-1}\mathbf{X'X}}$ の固有値・固有ベクトルを計算すればよい

分散・共分散行列
データをautoscalingした時は相関係数行列

変数1と2の相関係数

$$\begin{bmatrix} r_{11} & r_{12} \\ r_{21} & r_{22} \end{bmatrix}$$

Rでの解析

Rでの解析(ベクトルの代入と演算)

- ベクトルの代入
  - X <- c(1,2,3,4,5)
  - Y <- c(6,7,8,9,10)

- ベクトルの四則演算
  - ベクトルの足し算　X+Y
  - ベクトルの引き算　X–Y
  - ベクトルの掛け算　t(X)%*%Y

$$\begin{bmatrix}1 & 2 & 3 & 4 & 5\end{bmatrix}\begin{bmatrix}6\\7\\8\\9\\10\end{bmatrix}=130$$

- ベクトルの要素同士の計算
  - 掛け算　X*Y
  - 割り算　X/Y

$$\begin{bmatrix}6 & 14 & 24 & 36 & 50\end{bmatrix}$$

Rの基本(行列と行列の掛け算)

- 行列

matrix(c(1,2,3,4),nrow=2, byrow=T)

$$\begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}$$

- 行列の掛け算

matrix(c(1,2,3,4),nrow=2,byrow=T)
%*%matrix(c(5,6,7,8),nrow=2, byrow=T)

$$\begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}\begin{bmatrix} 5 & 6 \\ 7 & 8 \end{bmatrix}$$

- 行列の要素ごとの掛け算

matrix(c(1,2,3,4),nrow=2, byrow=T)
*matrix(c(5,6,7,8),nrow=2, byrow=T)

Rstudioでのベクトルと行列の演算の結果

```
> X <- c(1,2,3,4,5)
> Y <- c(6,7,8,9,10)
> X+Y
[1]  7  9 11 13 15
> X-Y
[1] -5 -5 -5 -5 -5
> t(X)%*%Y
     [,1]
[1,]  130
> X*Y
[1]  6 14 24 36 50
> X/Y
[1] 0.1666667 0.2857143 0.3750000 0.4444444 0.5000000
> matrix(c(1,2,3,4),nrow=2, byrow=T)
     [,1] [,2]
[1,]    1    2
[2,]    3    4
> matrix(c(1,2,3,4),nrow=2,byrow=T)%*%matrix(c(5,6,7,8),nrow=2,byrow=T)
     [,1] [,2]
[1,]   19   22
[2,]   43   50
> matrix(c(1,2,3,4),nrow=2,byrow=T)*matrix(c(5,6,7,8),nrow=2,byrow=T)
     [,1] [,2]
[1,]    5   12
[2,]   21   32
> 
```

ベクトルと行列の掛け算

- ベクトルの掛け算

$$\begin{bmatrix} a & b \end{bmatrix} \begin{bmatrix} c \\ d \end{bmatrix} = ac + bd$$

$\underline{1} \times 2$ $\quad 2 \times \underline{1}$ $\quad \underline{1} \times \underline{1}$ (1)

R：t(c(1,2))%*%c(3,4)

$$\begin{bmatrix} a \\ b \end{bmatrix} \begin{bmatrix} c & d \end{bmatrix} = \begin{bmatrix} ac & ad \\ bc & bd \end{bmatrix}$$

$\underline{2} \times 1$ $\quad 1 \times \underline{2}$ $\quad \underline{2} \times \underline{2}$

R : c(1,2)%*%t(c(3,4))

- 行列の掛け算

$$\begin{bmatrix} \boxed{a \quad b} \\ c \quad d \end{bmatrix} \begin{bmatrix} \boxed{\begin{matrix} p \\ r \end{matrix}} & \begin{matrix} q \\ s \end{matrix} \end{bmatrix} = \begin{bmatrix} \boxed{ap + br} & aq + bs \\ cp + dr & cq + ds \end{bmatrix}$$

$\underline{2} \times 2$ $\quad 2 \times \underline{2}$ $\quad \underline{2} \times \underline{2}$

R : matrix(c(1,2,3,4),nrow=2, byrow=T)%*%matrix(c(5,6,7,8),nrow=2, byrow=T)

Rでの主成分分析(prcompを使わず計算)

- データの準備

```
library(chemometrics)
data(NIR)
X0 <- NIR$xNIR # NIRS
X <- scale(X0)
```

- 固有ベクトル

```
w <- eigen(t(X)%*%X/(nrow(X)-1))
v1 <- w$vectors[,1]
```

- 特異値分解

```
w <- svd(X/sqrt(nrow(X)-1))
v1 <- w$v[,1]
```

Rでの主成分分析(prcompを使わず計算)

- 主成分スコア

```
PC1_score <- X%*%v1 # PC1
v2 <- w$vectors[,2]
PC2_score <- X%*%v2 # PC2
```

- 寄与率

```
PC_score <- X%*%w$vectors
cr_var <- apply(PC_score,2,var)/
            sum(apply(PC_score,2,var)) # 寄与率
cumcr <- cumsum(cr_var) # 累積寄与率
```