

第1章 システム生物学と多変量解析

本章の目的は、PLS 回帰を理解し、Janes らの仕事を追試することである。

1.1 概要

PLS 回帰は以下の基礎知識に支えられている。

PLS 回帰	
主成分回帰 (PCR)	
重回帰 (MLR)	主成分分析 (PCA)

よって、この概念図の下から順に取り扱う。また、特異値分解 (singular value decomposition) を知っているともこれらの手法を統一的に理解することができる。

1.2 準備 1: 特異値分解

任意の $m \times n$ 行列 \mathbf{A} は以下のような3つの行列 (それぞれ「回転・引き延ばし・回転」に対応) の積の形に書くことができる。

$$\mathbf{A}_{m \times n} = \mathbf{U}_{m \times m} \mathbf{\Sigma}_{m \times n} \mathbf{V}^T_{n \times n}$$

\mathbf{U} を左特異行列 (left singular matrix)、 \mathbf{V} を右特異行列、 $\mathbf{\Sigma}$ を特異値行列と呼ぶ。

$$\mathbf{U}_{m \times m} = \left(\begin{array}{c|c} \mathbf{A} \text{ の列空間の基底} & \mathbf{A} \text{ の左零空間} \\ \hline m \times r & m \times (m-r) \end{array} \right)$$
$$\mathbf{\Sigma}_{m \times n} = \left(\begin{array}{cc|c} \sigma_1 & 0 & 0 \\ & \ddots & 0 \\ 0 & & \sigma_r \\ \hline 0 & & 0 \end{array} \right)$$

$$\mathbf{V}_{n \times n}^T = \left(\begin{array}{c} \mathbf{A} \text{ の行空間の基底} \\ r \times n \\ \hline \mathbf{A} \text{ の零空間} \\ (n-r) \times n \end{array} \right)$$

\mathbf{U} 、 \mathbf{V} はともに正規直交行列 ($\mathbf{U}^T \mathbf{U} = \mathbf{I}$ 、 $\mathbf{V}^T \mathbf{V} = \mathbf{I}$) である。
 $m \times m$ $n \times n$

演習: 紙と鉛筆でできる特異値分解 以下の手順に従って、行列

$$\mathbf{A} = \begin{pmatrix} 1 & 2 \\ -1 & -2 \end{pmatrix}$$

を $\mathbf{A} = \mathbf{U}\mathbf{\Sigma}\mathbf{V}^T$ の形に特異値分解しなさい。

- (1) $(\mathbf{A}^T \mathbf{A})\mathbf{V} = \mathbf{V}(\mathbf{\Sigma}^T \mathbf{\Sigma})$ を示しなさい。(ヒント: $(\mathbf{A}\mathbf{B})^T = \mathbf{B}^T \mathbf{A}^T$)
- (2) (1) の結果より、 \mathbf{V} は $\mathbf{A}^T \mathbf{A}$ の固有ベクトルから構成されていることがわかる。この性質を利用して、 \mathbf{V} のうち「 \mathbf{A} の行空間の基底」に相当する部分を求めなさい。ただし、 $\mathbf{V}^T \mathbf{V} = \mathbf{I}$ である。
- (3) \mathbf{V} のうち「 \mathbf{A} の零空間」に相当する部分を求めなさい。 $\mathbf{V}^T \mathbf{V} = \mathbf{I}$ を満たすように正規化すること。(ヒント: \mathbf{A} の零空間は $\mathbf{A}\mathbf{K} = \mathbf{0}$ を満たすような行列 \mathbf{K} として求めることができる。)
- (4) (1) の結果より、 $\mathbf{\Sigma}$ の特異値成分 $\sigma_1, \dots, \sigma_r$ は $\mathbf{A}^T \mathbf{A}$ の固有値の平方根である。この性質を利用して $\mathbf{\Sigma}$ を求めなさい。
- (5) (1) と同様にして、 $(\mathbf{A}\mathbf{A}^T)\mathbf{U} = \mathbf{U}(\mathbf{\Sigma}\mathbf{\Sigma}^T)$ を示しなさい。
- (6) (2)(3) を参考にして \mathbf{U} を求めなさい。 $\mathbf{U}^T \mathbf{U} = \mathbf{I}$ に留意すること。
- (7) $\mathbf{A} = \mathbf{U}\mathbf{\Sigma}\mathbf{V}^T$ が成り立っていることを確認しなさい。

解答

(1)

$$\begin{aligned} (\mathbf{A}^T \mathbf{A})\mathbf{V} &= (\mathbf{V}\mathbf{\Sigma}^T \mathbf{U}^T)(\mathbf{U}\mathbf{\Sigma}\mathbf{V}^T)\mathbf{V} \\ &= \mathbf{V}\mathbf{\Sigma}^T \mathbf{\Sigma} \quad (\because \mathbf{U}^T \mathbf{U} = \mathbf{I}, \mathbf{V}^T \mathbf{V} = \mathbf{I}) \end{aligned}$$

(2) $\mathbf{A}^T \mathbf{A} = \begin{pmatrix} 1 & 2 \\ -1 & -2 \end{pmatrix}^T \begin{pmatrix} 1 & 2 \\ -1 & -2 \end{pmatrix}$ の固有値・固有ベクトルを求める。

$$\begin{aligned} |\mathbf{A}^T \mathbf{A} - \mathbf{I}\lambda| &= 0 \\ \begin{vmatrix} 2-\lambda & 4 \\ 4 & 8-\lambda \end{vmatrix} &= 0 \\ (2-\lambda)(8-\lambda) - 16 &= 0 \\ \lambda(\lambda-10) &= 0 \\ \therefore \lambda &= 0, 10 \end{aligned}$$

$$\begin{pmatrix} 2 & 4 \\ 4 & 8 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = 10 \begin{pmatrix} x \\ y \end{pmatrix}$$

この連立方程式を解くと $y = 2x$ となるので、

$$\begin{pmatrix} x \\ y \end{pmatrix} = k \begin{pmatrix} 1 \\ 2 \end{pmatrix}$$

さらに正規化して

$$\begin{pmatrix} x \\ y \end{pmatrix} = \frac{1}{\sqrt{5}} \begin{pmatrix} 1 \\ 2 \end{pmatrix}$$

(3) $\mathbf{AK} = \mathbf{0}$ となる行列 \mathbf{K} を求める。

1.3 準備 2: 行列・ベクトルの微分

行列・ベクトルで書かれた関数にも微分公式がある。覚えておくと重宝する。

$$\begin{aligned} d(\mathbf{Ax}) &= \mathbf{A}d\mathbf{x} \\ d(\mathbf{Ax} + \mathbf{b})^T \mathbf{C}(\mathbf{Dx} + \mathbf{e}) &= \{(\mathbf{Ax} + \mathbf{b})^T \mathbf{C} \mathbf{D} + (\mathbf{Dx} + \mathbf{e})^T \mathbf{C}^T \mathbf{A}\} d\mathbf{x} \end{aligned}$$

1.4 共分散最大化と特異値分解

$\text{Cov}(\mathbf{X}\mathbf{d}, \mathbf{Y}\mathbf{e}) = \mathbf{d}^T \mathbf{X}^T \mathbf{Y} \mathbf{e}$ を最大化するようなベクトル \mathbf{d} 、 \mathbf{e} はそれぞれ $\mathbf{X}^T \mathbf{Y}$ を特異値分解した際の第1左特異ベクトル、右特異ベクトルに等しい。ただし、 $\mathbf{d}^T \mathbf{d} = 1$ 、 $\mathbf{e}^T \mathbf{e} = 1$ とする。

$$\begin{aligned}\mathbf{X}^T \mathbf{Y} &= \mathbf{U} \mathbf{\Sigma} \mathbf{V}^T \\ &= \mathbf{u}_1 \sigma_1 \mathbf{v}_1^T + \cdots + \mathbf{u}_r \sigma_r \mathbf{v}_r^T\end{aligned}$$

$$\mathbf{d} = a_1 \mathbf{u}_1 + \cdots + a_r \mathbf{u}_r$$

$$\mathbf{e} = b_1 \mathbf{v}_1 + \cdots + b_r \mathbf{v}_r$$

$$\begin{aligned}\mathbf{d}^T \mathbf{X}^T \mathbf{Y} \mathbf{e} &= \mathbf{d}^T (\mathbf{u}_1 \sigma_1 \mathbf{v}_1^T + \cdots + \mathbf{u}_r \sigma_r \mathbf{v}_r^T) \mathbf{e} \\ &= (a_1 \mathbf{u}_1 + \cdots + a_r \mathbf{u}_r)^T \cdot (\mathbf{u}_1 \sigma_1 \mathbf{v}_1^T + \cdots + \mathbf{u}_r \sigma_r \mathbf{v}_r^T) \cdot (b_1 \mathbf{v}_1 + \cdots + b_r \mathbf{v}_r) \\ &= a_1 \sigma_1 b_1 + \cdots + a_r \sigma_r b_r \\ &= \mathbf{a}^T \mathbf{\Sigma} \mathbf{b}\end{aligned}$$

$\mathbf{d}^T \mathbf{d} = 1$ 、 $\mathbf{e}^T \mathbf{e} = 1$ より、

$$\begin{aligned}\mathbf{d}^T \mathbf{d} &= (a_1 \mathbf{u}_1 + \cdots + a_r \mathbf{u}_r)^T (a_1 \mathbf{u}_1 + \cdots + a_r \mathbf{u}_r) \\ &= a_1^2 + \cdots + a_r^2 \\ &= \mathbf{a}^T \mathbf{a}\end{aligned}$$

$$\therefore \mathbf{d}^T \mathbf{d} = \mathbf{a}^T \mathbf{a} = 1$$

同様にして $\mathbf{e}^T \mathbf{e} = \mathbf{b}^T \mathbf{b} = 1$ と求まる。

よって、 $\mathbf{d}^T \mathbf{d} = 1$ 、 $\mathbf{e}^T \mathbf{e} = 1$ を条件とする $\text{Cov}(\mathbf{X}\mathbf{d}, \mathbf{Y}\mathbf{e}) = \mathbf{d}^T \mathbf{X}^T \mathbf{Y} \mathbf{e}$ の最大化問題は、 $\mathbf{a}^T \mathbf{a} = 1$ 、 $\mathbf{b}^T \mathbf{b} = 1$ を条件とする $\text{Cov}(\mathbf{X}\mathbf{d}, \mathbf{Y}\mathbf{e}) = \mathbf{a}^T \mathbf{\Sigma} \mathbf{b}$ の最大化問題に帰着できる。

$$Q = \mathbf{a}^T \mathbf{\Sigma} \mathbf{b} - \lambda_1 (\mathbf{a}^T \mathbf{a} - 1) - \lambda_2 (\mathbf{b}^T \mathbf{b} - 1)$$

ベクトル・行列の微分公式、

$$\frac{d}{d\mathbf{x}}(\mathbf{x}^T \mathbf{x}) = 2\mathbf{x}$$

$$\frac{d}{d\mathbf{x}}(\mathbf{a}^T \mathbf{x}) = \mathbf{a}$$

を用いると、

$$\begin{cases} \frac{\partial Q}{\partial \mathbf{a}} = \Sigma \mathbf{b} - 2\lambda_1 \mathbf{a} = 0 \\ \frac{\partial Q}{\partial \mathbf{b}} = \Sigma^T \mathbf{a} - 2\lambda_2 \mathbf{b} = 0 \end{cases}$$

これを整理すると

$$\begin{cases} \Sigma \mathbf{b} = 2\lambda_1 \mathbf{a} \\ \mathbf{a}^T \Sigma = 2\lambda_2 \mathbf{b}^T \end{cases}$$

上の式の両辺に左から \mathbf{a}^T 、下の式の両辺に右から \mathbf{b} を掛ける。

$$\begin{cases} \mathbf{a}^T \Sigma \mathbf{b} = 2\lambda_1 \\ \mathbf{a}^T \Sigma \mathbf{b} = 2\lambda_2 \end{cases}$$

よって $\lambda_1 = \lambda_2$ とわかる。

$$\lambda = 2\lambda_1 = 2\lambda_2$$

とおいて式を書き換えると¹、

$$\begin{cases} \Sigma \mathbf{b} = \lambda \mathbf{a} \\ \Sigma^T \mathbf{a} = \lambda \mathbf{b} \end{cases}$$

上式を下式に代入して

$$\begin{aligned} \frac{1}{\lambda} \Sigma^T \Sigma \mathbf{b} &= \lambda \mathbf{b} \\ \therefore \Sigma^T \Sigma \mathbf{b} &= \lambda^2 \mathbf{b} \end{aligned}$$

この式が意味することは、ベクトル \mathbf{b} の要素 b_1, \dots, b_r は

¹証明を一度完成させてみると、ここで係数2を掛けておくと以下の話がすっきりすることがわかる。

$$\begin{pmatrix} \sigma_1^2 b_1 \\ \vdots \\ \sigma_r^2 b_r \end{pmatrix} = \begin{pmatrix} \lambda^2 b_1 \\ \vdots \\ \lambda^2 b_r \end{pmatrix}$$

が常に成立するように定まるということである。この条件を満たす \mathbf{b} は、

$$\begin{pmatrix} 1 \\ 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ 1 \end{pmatrix},$$

のように、1を1つだけ含み、残りは0というベクトルである。ベクトル \mathbf{a} についても同様である。したがって、

$$\mathbf{d}^T \mathbf{X}^T \mathbf{Y} \mathbf{e} = a_1 \sigma_1 b_1 + \cdots + a_r \sigma_r b_r$$

より、 $\mathbf{d}^T \mathbf{X}^T \mathbf{Y} \mathbf{e}$ が取りうる値は、 $\sigma_1, \dots, \sigma_r$ である。このうち最大の値を持つものは、特異値の定義より σ_1 であるから、 $\mathbf{d}^T \mathbf{X}^T \mathbf{Y} \mathbf{e}$ を最大化する \mathbf{a}, \mathbf{b} は、

$$\mathbf{a} = \begin{pmatrix} 1 \\ 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \mathbf{b} = \begin{pmatrix} 1 \\ 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix}$$

である。

$$\begin{aligned} \mathbf{d} &= a_1 \mathbf{u}_1 + \cdots + a_r \mathbf{u}_r \\ \mathbf{e} &= b_1 \mathbf{v}_1 + \cdots + b_r \mathbf{v}_r \end{aligned}$$

より、共分散を最大化する \mathbf{d}, \mathbf{e} の条件は、

$$\begin{aligned} \mathbf{d} &= \mathbf{u}_1 \\ \mathbf{e} &= \mathbf{v}_1 \end{aligned}$$

すなわち、 \mathbf{d}, \mathbf{e} がそれぞれ、第1左特異ベクトル、第1右特異ベクトルである場合である。よって、左右の第1特異ベクトルを用いると、潜在変数ベクトル同士の共分散が最大になる。

1.4.1 Moore-Penrose の一般逆行列

Moore-Penrose の一般逆行列 (Moore-Penrose generalized inverse) は重回帰 (MLR) と主成分回帰 (PCR) を統一的視点から理解する上で欠かせない。

$$\mathbf{A}^{\#} = \mathbf{V} \mathbf{\Sigma}'^{-1} \mathbf{U}^T$$

$n \times m$ $n \times n$ $n \times m$ $m \times m$

ただし、

$$\mathbf{\Sigma}'^{-1} = \left(\begin{array}{cc|c} \frac{1}{\sigma_1} & 0 & 0 \\ & \ddots & \\ 0 & & \frac{1}{\sigma_r} \\ \hline 0 & & 0 \end{array} \right)$$

連立1次方程式 $\mathbf{Ax} = \mathbf{b}$ において \mathbf{A} の行数が列数よりも大きいとき (縦長)、 $\hat{\mathbf{x}} = \mathbf{A}^{\#}\mathbf{b}$ は二乗誤差 $\|\mathbf{Ax} - \mathbf{b}\|^2$ を最小にするような解である。 \mathbf{A} の行数が列数よりも少ないとき (横長) には、 $\hat{\mathbf{x}} = \mathbf{A}^{\#}\mathbf{b}$ は解となり得る \mathbf{x} のうち、ノルム $\|\mathbf{x}\|$ が最小のものを返す。

1.4.2 化学量論行列の特異値分解

$$\mathbf{N} = \mathbf{U}\mathbf{\Sigma}\mathbf{V}^T$$

\mathbf{U} は列空間の基底ベクトルと、左零空間の基底ベクトル。
 \mathbf{V} は行空間の基底ベクトルと、左零空間の基底ベクトル。

列空間の基底

$$\frac{d}{dt} \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_m \end{pmatrix} = \begin{pmatrix} c_{11} \\ c_{21} \\ \vdots \\ c_{m1} \end{pmatrix} v_1 + \begin{pmatrix} c_{12} \\ c_{22} \\ \vdots \\ c_{m2} \end{pmatrix} v_2 + \cdots + \begin{pmatrix} c_{1n} \\ c_{2n} \\ \vdots \\ c_{mn} \end{pmatrix} v_n$$

において、このパスウェイで現れるすべての組み合わせの $(c_{1i} \ c_{2i} \ \cdots \ c_{mi})^T$ を生成する基底のこと。

行空間の基底

$$\frac{dx_i}{dt} = \begin{pmatrix} r_1 & r_2 & \cdots & r_n \end{pmatrix} \begin{pmatrix} v_1 \\ v_2 \\ \vdots \\ v_m \end{pmatrix}$$

において、このパスウェイで現れるすべての組み合わせの $(r_1 \ r_2 \ \cdots \ r_n)$ を生成する基底のこと。

(右) 零空間 化学量論行列 \mathbf{N} について、

$$\mathbf{NK} = \mathbf{0}$$

となるような行列 \mathbf{K} のこと。 \mathbf{K} の要素である列ベクトルの線形結合によって、あらゆる流束分布を記述できる。

左零空間 化学量論行列 \mathbf{N} について、

$$\mathbf{GN} = \mathbf{0}$$

を満たす行列 \mathbf{G} のこと。物質の保存関係を表す。

1.5 重回帰 (MLR)

1.5.1 定義

被説明変数 y 、説明変数 x_1, \dots, x_n の組について、 m 回の測定を行ったとする。 y と x の関係を線形で表すとき、

$$\mathbf{y} = \mathbf{Xa} + \boldsymbol{\varepsilon}$$

$$\begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_m \end{pmatrix} = \begin{pmatrix} 1 & x_{11} & \cdots & x_{1n} \\ 1 & x_{21} & \cdots & x_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ 1 & x_{m1} & \cdots & x_{mn} \end{pmatrix} \begin{pmatrix} a_0 \\ a_1 \\ \vdots \\ a_n \end{pmatrix} + \begin{pmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \vdots \\ \varepsilon_m \end{pmatrix}$$

残差平方和 Q を最小にするような係数 \mathbf{a} を求めることが重回帰分析の核心である。

1.5.2 残差平方和を最小にする係数の導出

$$\begin{aligned} Q &= \sum_{i=1}^m \varepsilon_i^2 \\ &= \sum_{i=1}^m \{y_i - (a_0 + a_1 x_{i1} + \cdots + a_n x_{in})\}^2 \\ &= (\mathbf{y} - \mathbf{Xa})^T (\mathbf{y} - \mathbf{Xa}) \end{aligned}$$

Q を最小にする \mathbf{a} を求めればよい。すなわち、 $\frac{dQ}{d\mathbf{a}} = 0$ を満たす \mathbf{a} を求める²。
ベクトル・行列の微分公式、

$$d(\mathbf{Ax} + \mathbf{b})^T \mathbf{C} (\mathbf{Dx} + \mathbf{e}) = \{(\mathbf{Ax} + \mathbf{b})^T \mathbf{C} \mathbf{D} + (\mathbf{Dx} + \mathbf{e})^T \mathbf{C}^T \mathbf{A}\} dx$$

より、 $\mathbf{C} = \mathbf{I}, \mathbf{D} = \mathbf{A}, \mathbf{e} = \mathbf{b}$ とおくと、以下の公式を得る。

$$\begin{aligned} d(\mathbf{Ax} + \mathbf{b})^T (\mathbf{Ax} + \mathbf{b}) &= \{(\mathbf{Ax} + \mathbf{b})^T \mathbf{A} + (\mathbf{Ax} + \mathbf{b})^T \mathbf{A}\} dx \\ &= 2(\mathbf{Ax} + \mathbf{b})^T \mathbf{A} dx \end{aligned}$$

これを $Q = (\mathbf{y} - \mathbf{Xa})^T (\mathbf{y} - \mathbf{Xa})$ にあてはめて、

$$\begin{aligned} dQ &= d\{(\mathbf{y} - \mathbf{Xa})^T (\mathbf{y} - \mathbf{Xa})\} \\ &= -2(\mathbf{y} - \mathbf{Xa})^T \mathbf{X} dx \quad (\text{前記の微分公式を使って導出した}) \end{aligned}$$

$\frac{dQ}{d\mathbf{a}} = 0$ より、残差平方和を最小にする \mathbf{a} は、

$$-2(\mathbf{y} - \mathbf{Xa})^T \mathbf{X} = 0$$

を満たす。これをさらに変形すると、

²本当は極値と最小値が一致することを示す必要があるが、省略する。

$$\begin{aligned}
-2(\mathbf{y} - \mathbf{X}\mathbf{a})^T \mathbf{X} &= 0 \\
(\mathbf{y} - \mathbf{X}\mathbf{a})^T \mathbf{X} &= 0 \\
\mathbf{X}^T(\mathbf{y} - \mathbf{X}\mathbf{a}) &= 0 && (\because (\mathbf{AB})^T = \mathbf{B}^T \mathbf{A}^T) \\
\mathbf{X}^T \mathbf{y} &= \mathbf{X}^T \mathbf{X} \mathbf{a}
\end{aligned}$$

$$\therefore \mathbf{a} = (\mathbf{X}^T \mathbf{X})^{-1} \mathbf{X}^T \mathbf{y}$$

1.5.3 Moore-Penrose の一般逆行列による表記

これにて y を予測する式 $y = a_0 + a_1 x_1 + \dots + a_n x_n$ を得ることができた。Moore-Penrose の一般逆行列を用いて書くと、

$$\hat{\mathbf{a}} = \mathbf{X}^\# \mathbf{y}$$

となる。

$\mathbf{X}^T \mathbf{X}$ に逆行列が存在する場合 (= \mathbf{X} の列が互いに線形独立である場合)、

$$\mathbf{X}^\# = (\mathbf{X}^T \mathbf{X})^{-1} \mathbf{X}^T$$

である。

$\mathbf{X}\mathbf{X}^T$ に逆行列が存在する場合 (= \mathbf{X} の行が互いに線形独立である場合)、

$$\mathbf{X}^\# = \mathbf{X}^T (\mathbf{X}\mathbf{X}^T)^{-1}$$

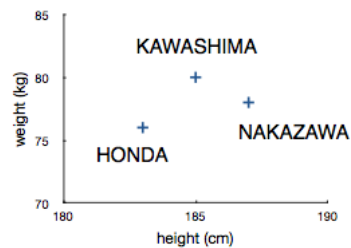
である。

1.6 主成分分析 (PCA)

1.6.1 定義

相関の強い複数の変数を 1 変数にまとめる方法である。例えば以下のような身長・体重のデータがあったとする。

	身長 (cm)	体重 (kg)
川島さん	185	80
中澤さん	187	78
本田さん	183	76

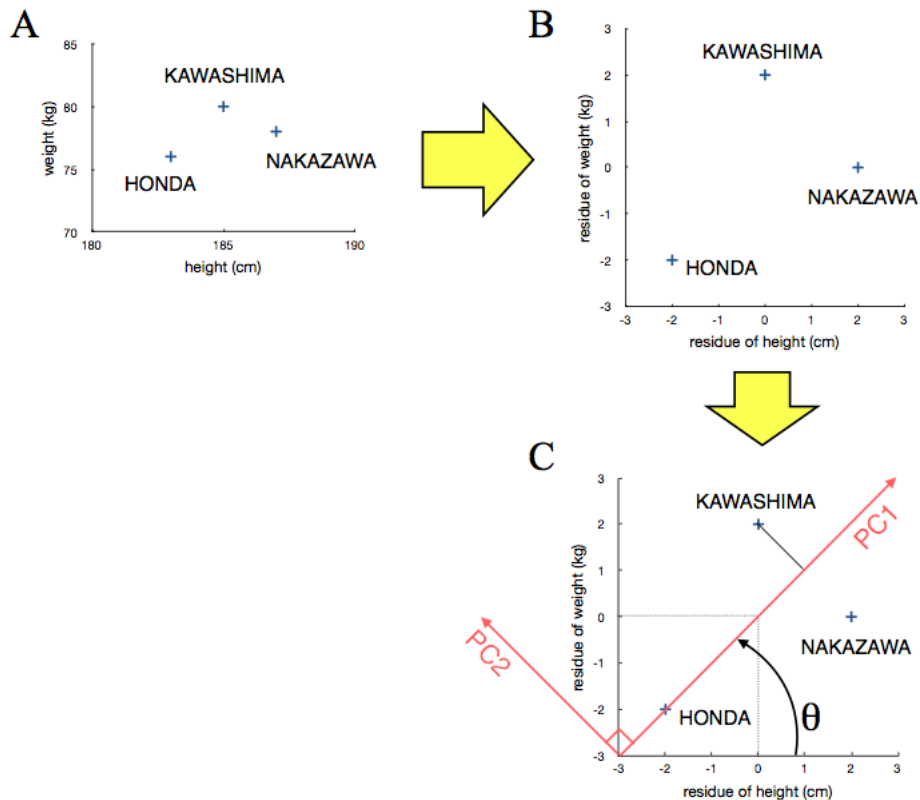


身長と体重をそれぞれ横軸・縦軸に取った散布図上の位置で個々人の体格がわかる。しかし、身長と体重の間には一定の相関が見られるので、これらを体格を表す1つの指標 t_{PC1} にまとめてしまうことを考える。身長、体重をそれぞれ x_1, x_2 で表すと、

$$t_{PC1} = w_1x_1 + w_2x_2$$

で表すことを考える。

どのような指標にまとめるのが合理的だろうか？主成分分析では、座標変換を行う。この例で言うと、身長と体重の直交座標系を角度 θ だけ回転する。 w_1, w_2 の値は t_{PC1} の分散を最大化するように定める。



分散最大化が最小二乗法(回帰分析)と似ていると思った人も多いだろう。実際、データからあらかじめ平均値が差し引かれていれば(中心化)、分散最大化は最小二乗法(回帰分析)と一致する。すなわち、主成分軸と回帰直線が一致する。

(利点) 多次元のデータであっても2次元に縮約し、平面にプロットすることが可能になる。

\mathbf{X} から各列ごとに列の平均値を差し引いたものを $\hat{\mathbf{X}}$ とすると、 \mathbf{X} の共分散行列は $\hat{\mathbf{X}}^T \hat{\mathbf{X}}$ と表せる。 $\hat{\mathbf{X}}^T \hat{\mathbf{X}}$ の固有ベクトルを **loading** と呼ぶ。

残差を残す

$$x_{ij} = x_{ij}^* - \bar{x}_j$$

データ行列

$$\mathbf{X} = \begin{pmatrix} x_{11} & x_{12} & \cdots & x_{1n} \\ x_{21} & x_{22} & \cdots & x_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ x_{m1} & x_{m2} & \cdots & x_{mn} \end{pmatrix}$$

さきほどの例で言うと、

	身長 (cm)	体重 (kg)	
川島さん	185	80	→ $\mathbf{X} = \begin{pmatrix} 185 & 80 \\ 187 & 78 \\ 183 & 76 \end{pmatrix}$
中澤さん	187	78	
本田さん	183	76	
主成分			

$$t_{11} = w_{11}x_{11} + w_{21}x_{12} + \cdots + w_{n1}x_{1n}$$

ただし $\mathbf{w}_1^T \mathbf{w}_1 = 1$

$$t_{m1} = \mathbf{x}_m \mathbf{w}_1$$

この分散を最大化する \mathbf{w}_1 を求める。

1.6.2 結合係数は共分散行列の固有ベクトルであることの証明

結合係数 \mathbf{w}_1 は、共分散行列の固有ベクトルとして求められる。以下はその証明である。

$$\mathbf{x}_m = \begin{pmatrix} x_{m1} & x_{m2} & \cdots & x_{mn} \end{pmatrix}, \mathbf{w}_1 = \begin{pmatrix} w_{11} \\ \vdots \\ w_{n1} \end{pmatrix}$$

$$\mathbf{t}_1 = \begin{pmatrix} t_{11} \\ t_{21} \\ \vdots \\ t_{n1} \end{pmatrix}, \mathbf{t}_1 = \mathbf{X}\mathbf{w}_1$$

$$\begin{aligned} \bar{t}_1 &= \frac{1}{m} \sum_{i=1}^m t_{i1} \\ &= \frac{1}{m} \sum_{i=1}^m \mathbf{x}_i \mathbf{w}_1 \\ &= \frac{1}{m} \sum_{i=1}^m \sum_{j=1}^n x_{ij} w_{j1} \\ &= \frac{1}{m} \sum_{j=1}^n w_{j1} \left(\sum_{i=1}^m x_{ij} \right) \\ &= \frac{1}{m} \sum_{j=1}^n w_{j1}(0) \\ &= 0 \end{aligned}$$

$$\begin{aligned} \sigma_{\mathbf{t}_1}^2 &= \frac{1}{m-1} \mathbf{t}_1^T \mathbf{t}_1 \\ &= \frac{1}{m-1} (\mathbf{X}\mathbf{w}_1)^T (\mathbf{X}\mathbf{w}_1) \\ &= \mathbf{w}_1^T \left(\frac{1}{m-1} \mathbf{X}^T \mathbf{X} \right) \mathbf{w}_1 \\ &= \mathbf{w}_1^T \mathbf{V} \mathbf{w}_1 \\ &\geq 0 \end{aligned}$$

Lagrange の未定乗数法 $\mathbf{w}_1^T \mathbf{w}_1 = 1$ のもとで $\sigma_{\mathbf{t}_1}^2$ の極値を求める。

$$\begin{aligned} Q(\mathbf{w}_1) &= \sigma_{\mathbf{t}_1}^2 - \lambda(\mathbf{w}_1^T \mathbf{w}_1 - 1) \\ &= \mathbf{w}_1^T \mathbf{V} \mathbf{w}_1 - \lambda(\mathbf{w}_1^T \mathbf{w}_1 - 1) \end{aligned}$$

微分公式 $d(\mathbf{Ax} + \mathbf{b})^T \mathbf{C}(\mathbf{Dx} + \mathbf{e}) = \{(\mathbf{Ax} + \mathbf{b})^T \mathbf{C} \mathbf{D} + (\mathbf{Dx} + \mathbf{e})^T \mathbf{C}^T \mathbf{A}\} d\mathbf{x}$ を用いると、

$$\begin{aligned}d(\mathbf{w}_1^T \mathbf{V} \mathbf{w}_1) &= 2\mathbf{w}_1^T \mathbf{V} d\mathbf{w}_1 \\d(\mathbf{w}_1^T \mathbf{w}_1) &= 2\mathbf{w}_1^T d\mathbf{w}_1\end{aligned}$$

よって、極値条件は

$$\begin{aligned}dQ(\mathbf{w}_1) &= 2\mathbf{w}_1^T \mathbf{V} d\mathbf{w}_1 - \lambda 2\mathbf{w}_1^T d\mathbf{w}_1 \\ &= 0\end{aligned}$$

より、

$$\mathbf{w}_1^T \mathbf{V} = \lambda \mathbf{w}_1^T$$

両辺を転置すると、

$$\mathbf{V} \mathbf{w}_1 = \lambda \mathbf{w}_1$$

となる (\mathbf{V} は対称行列なので $\mathbf{V} = \mathbf{V}^T$)。よって \mathbf{w}_1 は 共分散行列 \mathbf{V} の固有ベクトルとして求められる。また、

$$\begin{aligned}\sigma_{t_1}^2 &= \mathbf{w}_1^T \mathbf{V} \mathbf{w}_1 \\ &= \lambda\end{aligned}$$

より、第1主成分の分散 $\sigma_{t_1}^2$ は共分散行列 \mathbf{V} の固有値、ラグランジュ乗数 λ に等しい。

1.6.3 寄与率

ある主成分の説明能力を評価する尺度として、「寄与率」がある。第 i 主成分の寄与率は、次のように定義される。

$$\begin{aligned} \text{寄与率} &= \frac{\sigma_{\mathbf{t}_i}^2}{\sigma_{\mathbf{t}_1}^2 + \sigma_{\mathbf{t}_2}^2 + \cdots + \sigma_{\mathbf{t}_n}^2} \\ &= \frac{\lambda_i}{\lambda_1 + \lambda_2 + \cdots + \lambda_n} \end{aligned}$$

$\sigma_{\mathbf{t}_i}^2$ は第 i 主成分軸上でのデータの分散である。これを分散の総和で割ったものが第 i 主成分の寄与率である。別の言い方をすれば、データの全分散のうち、第 i 主成分で説明できる割合を表す量が寄与率である。

また、先に証明したように、 $\sigma_{\mathbf{t}_i}^2$ は共分散行列 \mathbf{V} の第 i 固有値に等しい。これを利用すると、第 i 主成分の寄与率は、第 i 固有値を全固有値の和で割ったものとして書ける。

1.6.4 行列で表記する

$$\mathbf{t}_i = \mathbf{X}\mathbf{w}_i$$

$$\mathbf{T} = \begin{pmatrix} \mathbf{t}_1 & \mathbf{t}_2 & \cdots & \mathbf{t}_k \end{pmatrix}$$

$$\mathbf{P} = \begin{pmatrix} \mathbf{p}_1 & \mathbf{p}_2 & \cdots & \mathbf{p}_k \end{pmatrix}$$

$$\mathbf{T} = \mathbf{X}\mathbf{P}$$

$$\mathbf{X} = \mathbf{T}\mathbf{P}^T$$

1.6.5 特異値分解との関係

主成分スコア、loading とともに特異値分解と比較すると、以下のようになる。

$$\begin{aligned} \mathbf{X} &= \mathbf{U} \mathbf{\Sigma} \mathbf{V}^T \\ &= \mathbf{T} \mathbf{P}^T \quad (\mathbf{U}\mathbf{\Sigma} = \mathbf{T}, \quad \mathbf{V} = \mathbf{P}) \end{aligned}$$

\mathbf{T} が主成分スコア行列、 \mathbf{P} が loading 行列 (共分散行列の固有ベクトル) である。

1.6.6 演習

以下に示す3名分の身体測定データから、「身体の大きさ」を表す主成分を求めたい。

	身長 (cm)	体重 (kg)
川島さん	185	80
中澤さん	187	78
本田さん	183	76

次の設問に順に答え、「身体の大きさ」を表す主成分スコアを求めなさい。

- (1) 上の身体測定データをデータ行列にまとめなさい。
- (2) 各人の身長、体重から平均値をそれぞれ差し引きなさい。この行列を \mathbf{X} とする。
- (3) $\mathbf{X}^T \mathbf{X}$ を計算しなさい。
- (4) $\mathbf{X}^T \mathbf{X}$ の固有値、固有ベクトルを計算しなさい。
- (5) 第1主成分について、各人の主成分スコアを求めなさい。
- (6) 第1主成分の寄与率を求めなさい。

1.7 主成分回帰 (PCR)

1.7.1 着想

\mathbf{X} の行数が列数よりも小さいとき (横長)、 $\mathbf{X}^\#$ は解となり得る \mathbf{a} のうち、ノルム $\|\mathbf{a}\|$ が最小のものを返す。これには生物学的意味は無いので、使えない。そこで、 \mathbf{X} の行数が列数よりも小さいとき、主成分分析を用いて変数を減らし、 \mathbf{X} の列数を減らす。

1.7.2 Moore-Penrose の一般逆行列で表す

$$\begin{aligned} \mathbf{Y} &= \mathbf{X} \mathbf{B} \\ &= \mathbf{T} \mathbf{P}^T \mathbf{B} \quad (\mathbf{X} = \mathbf{T} \mathbf{P}^T) \end{aligned}$$

のとき \mathbf{B} の

$$\hat{\mathbf{B}}_{PCR} = \mathbf{P}\mathbf{T}^{\#}\mathbf{Y}$$

$$\mathbf{Y} = \mathbf{X}\hat{\mathbf{B}}_{PCR}$$

1.8 PLS 回帰

基本となるアイデア PCRを改良したものがPLSである。PLSの基本アイデアは以下の通り。

- (1) 説明変数群に関するデータ行列 \mathbf{X} と被説明変数に関するデータ行列 \mathbf{Y} のそれぞれの主成分スコア (に近いもの) \mathbf{T}, \mathbf{U} を考え、 \mathbf{T} と \mathbf{U} の間の共分散を最大化する。
- (2) \mathbf{X}, \mathbf{Y} それぞれの主成分スコア (に近いもの) \mathbf{T}, \mathbf{U} を導く際、 \mathbf{X} の共分散行列 (\mathbf{X}^T と \mathbf{X} との積) ではなく、 \mathbf{X} と \mathbf{Y} の共分散行列 (\mathbf{X}^T と \mathbf{Y} との積) を使う。これにより、 \mathbf{X} と \mathbf{Y} の相関を加味した主成分スコア (に近いもの) になる。

PLS 回帰係数の導き方 それでは PLS 回帰の方法に従って回帰係数を導く手順を見てみよう。ここでの \mathbf{X}, \mathbf{Y} は中心化されているものとする (Höskuldsson 1988)。

- (1) $\mathbf{X}^T\mathbf{Y}$ を特異値分解し、第1左特異ベクトル \mathbf{w} 、第1右特異ベクトル \mathbf{c} を求める。特異値分解の性質より、 \mathbf{w} と \mathbf{c} は $\mathbf{X}\mathbf{w}$ と $\mathbf{Y}\mathbf{c}$ の共分散を最大化する (Höskuldsson 1988)。
- (2) $\mathbf{w}^T\mathbf{w} = 1$ 、 $\mathbf{c}^T\mathbf{c} = 1$ となるよう \mathbf{w} 、 \mathbf{c} を正規化する。
- (3)
$$\left. \begin{array}{l} \mathbf{t} = \mathbf{X}\mathbf{w} \\ \mathbf{u} = \mathbf{Y}\mathbf{c} \end{array} \right\} \text{より、スコア } \mathbf{t}, \mathbf{u} \text{ を求める。}$$
- (4) $\mathbf{t}^T\mathbf{t} = 1$ 、 $\mathbf{u}^T\mathbf{u} = 1$ となるよう \mathbf{t} 、 \mathbf{u} を正規化する。
- (5) $\mathbf{u} = b\mathbf{t}$ の線形関係を仮定して、回帰係数 b を式 $b = \mathbf{u}^T\mathbf{t}$ より推定する。

$$\begin{aligned} \mathbf{u}^T &= b\mathbf{t}^T \\ \mathbf{u}^T\mathbf{t} &= b\mathbf{t}^T\mathbf{t} \\ \therefore \mathbf{u}^T\mathbf{t} &= b \end{aligned}$$

- (6) $\left. \begin{array}{l} \mathbf{p} = \mathbf{X}^T \mathbf{t} \\ \mathbf{q} = \mathbf{Y}^T \mathbf{u} \end{array} \right\}$ より、 \mathbf{p}, \mathbf{q} を求める。 \mathbf{p}, \mathbf{q} は PCA の loading に相当するが、PCA の loading とは異なり、直交行列ではない。
- (7) \mathbf{X} から $\mathbf{t}\mathbf{p}^T$ を、 \mathbf{Y} から $\mathbf{u}\mathbf{q}^T$ を、それぞれ差し引き、(1)に戻る。PLS 成分の数だけこれを繰り返す。
- (8) 最終的に、 \mathbf{X} と \mathbf{Y} の PLS 回帰係数行列は、 \mathbf{X} 、 \mathbf{Y} それぞれの loadings に相当する \mathbf{P} 、 \mathbf{Q} およびその間をつなぐ回帰係数 \mathbf{B} によって決まる。 \mathbf{P} 、 \mathbf{Q} はベクトル \mathbf{p} 、 \mathbf{q} をそれぞれ並べて作った行列である。また、 \mathbf{B} は、 b を対角要素に並べた行列である。

$$\begin{aligned} \mathbf{Y} &= \mathbf{U}\mathbf{Q}^T \\ &= \mathbf{T}\mathbf{B}\mathbf{Q}^T \\ &= \mathbf{X}(\mathbf{P}^T)\#\mathbf{B}\mathbf{Q}^T \quad (\because \hat{\mathbf{T}} = \mathbf{X}(\mathbf{P}^T)\#) \\ &= \mathbf{X}\mathbf{B}_{\text{PLS}} \quad (\text{ただし } \mathbf{B}_{\text{PLS}} = (\mathbf{P}^T)\#\mathbf{B}\mathbf{Q}^T) \end{aligned}$$

$\hat{\mathbf{T}} = \mathbf{X}(\mathbf{P}^T)\#$ は以下のように示すことができる。

$$\begin{aligned} \mathbf{P} &= \mathbf{X}^T \mathbf{T} \\ \mathbf{P}^T &= \mathbf{T}^T \mathbf{X} \\ \mathbf{T}\mathbf{P}^T &= \mathbf{T}\mathbf{T}^T \mathbf{X} \\ \hat{\mathbf{T}} &= \mathbf{X}(\mathbf{P}^T)\# \end{aligned}$$

1.8.1 演習

EGF、NGF でそれぞれ細胞を刺激した際の ERK の時系列および 30 分経過時点における c-Fos、c-Jun の発現量を測定し、以下のような結果を得た。

	ERK _{t1}	ERK _{t2}	ERK _{t3}		c-Fos _{30min}	c-Jun _{30min}
EGF	0	4	2	EGF	0	1
NGF	2	4	4	NGF	1	2

以下の問いに答えて、ERK の時系列から c-Fos、c-Jun の発現量を予測する PLS 回帰式を求めなさい。PLS 成分についての計算は、第 1 成分まででよい。

- (1) 上の MAPK、IEG の測定結果をそれぞれデータ行列にまとめなさい。これらの行列をそれぞれ \mathbf{X}_0 、 \mathbf{Y}_0 とする。
- (2) \mathbf{X}_0 、 \mathbf{Y}_0 の各列の平均値を列方向に並べた行列をそれぞれ $\bar{\mathbf{X}}$ 、 $\bar{\mathbf{Y}}$ とする。 \mathbf{X}_0 、 \mathbf{Y}_0 から $\bar{\mathbf{X}}$ 、 $\bar{\mathbf{Y}}$ を差し引きなさい。得られた行列を \mathbf{X} 、 \mathbf{Y} とする。
- (3) $\mathbf{X}^T\mathbf{Y}$ を計算しなさい。
- (4) $\mathbf{X}^T\mathbf{Y}$ を特異値分解し、第1左特異ベクトル \mathbf{w} 、第1右特異ベクトル \mathbf{c} をそれぞれ求めなさい。ただし、 $\mathbf{w}^T\mathbf{w} = 1$ 、 $\mathbf{c}^T\mathbf{c} = 1$ となるよう \mathbf{w} 、 \mathbf{c} を正規化すること。
- (5) $\mathbf{t} = \mathbf{X}\mathbf{w}$ 、 $\mathbf{u} = \mathbf{Y}\mathbf{c}$ より、PLS スコアを表すベクトル \mathbf{t} 、 \mathbf{u} をそれぞれ求めなさい。 $\mathbf{t}^T\mathbf{t} = 1$ 、 $\mathbf{u}^T\mathbf{u} = 1$ となるよう \mathbf{t} 、 \mathbf{u} を正規化すること。
- (6) PLS スコア同士の線形関係 $\mathbf{u} = b\mathbf{t}$ の係数 b を求めなさい。 $b = \mathbf{u}^T\mathbf{t}$ より求める。
- (7) $\mathbf{p} = \mathbf{X}^T\mathbf{t}$ 、 $\mathbf{q} = \mathbf{Y}^T\mathbf{u}$ より、PLS loadings を表すベクトル \mathbf{p} 、 \mathbf{q} をそれぞれ求めなさい。
- (8) PLS 成分を第 n 成分まで計算する場合は、 $\mathbf{X} - \mathbf{t}\mathbf{p}^T$ を \mathbf{X} に、 $\mathbf{Y} - \mathbf{u}\mathbf{q}^T$ を \mathbf{Y} にそれぞれ代入して (3) に戻る。 $\mathbf{X} - \mathbf{t}\mathbf{p}^T$ および $\mathbf{Y} - \mathbf{u}\mathbf{q}^T$ を計算しなさい。
- (9) 以下の誘導に従って、PLS 回帰式 $\mathbf{Y} = \mathbf{X}\mathbf{B}_{\text{PLS}}$ の回帰係数行列 \mathbf{B}_{PLS} を求めなさい。
 - (a) \mathbf{P}^T を $\mathbf{P}^T = \mathbf{U}\Sigma\mathbf{V}^T$ の形に特異値分解しなさい。
 - (b) $(\mathbf{P}^T)^\# = \mathbf{V}\Sigma'^{-1}\mathbf{U}^T$ より、 $(\mathbf{P}^T)^\#$ を求めなさい。
 - (c) $\mathbf{B}_{\text{PLS}} = (\mathbf{P}^T)^\#\mathbf{B}\mathbf{Q}^T$ より、回帰係数行列 \mathbf{B}_{PLS} を求めなさい。
- (10) $\mathbf{Y}_0 - \bar{\mathbf{Y}} = (\mathbf{X}_0 - \bar{\mathbf{X}})\mathbf{B}_{\text{PLS}}$ より、もともとの変数である c-Jun や ERK を用いた回帰式を書き出しなさい。
- (11) ERK の時系列データのうち、c-Fos、c-Jun の発現量に強く寄与する時点はどこか。PLS loadings を表すベクトル \mathbf{p} 、 \mathbf{q} の値を吟味して答えなさい。

1.8.2 解答

(1)

$$\mathbf{X}_0 = \begin{pmatrix} 0 & 4 & 2 \\ 2 & 4 & 4 \end{pmatrix}, \quad \mathbf{Y}_0 = \begin{pmatrix} 0 & 1 \\ 1 & 2 \end{pmatrix}$$

(2)

$$\mathbf{X} = \begin{pmatrix} -1 & 0 & -1 \\ 1 & 0 & 1 \end{pmatrix}, \quad \mathbf{Y} = \begin{pmatrix} -\frac{1}{2} & -\frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} \end{pmatrix}$$

(3)

$$\begin{aligned} \mathbf{X}^T \mathbf{Y} &= \begin{pmatrix} -1 & 1 \\ 0 & 0 \\ -1 & 1 \end{pmatrix} \begin{pmatrix} -\frac{1}{2} & -\frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} \end{pmatrix} \\ &= \begin{pmatrix} 1 & 1 \\ 0 & 0 \\ 1 & 1 \end{pmatrix} \end{aligned}$$

(4) $\mathbf{X}^T \mathbf{Y} = \mathbf{U} \mathbf{\Sigma} \mathbf{V}^T$ と特異値分解し、 $\mathbf{U} = (\mathbf{w}_1 \mathbf{w}_2 \dots)$, $\mathbf{V} = (\mathbf{c}_1 \mathbf{c}_2 \dots)$ とおくと、 \mathbf{V} について $\{(\mathbf{X}^T \mathbf{Y})^T \mathbf{X}^T \mathbf{Y}\} \mathbf{V} = \mathbf{V} (\mathbf{\Sigma}^T \mathbf{\Sigma})$ となることから、

$$\begin{aligned} \begin{pmatrix} 1 & 0 & 1 \\ 1 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 1 \\ 0 & 0 \\ 1 & 1 \end{pmatrix} &= \begin{pmatrix} 2 & 2 \\ 2 & 2 \end{pmatrix} \\ \begin{vmatrix} 2 - \lambda & 2 \\ 2 & 2 - \lambda \end{vmatrix} &= 0 \end{aligned}$$

より、

$$\begin{aligned} (2 - \lambda)^2 - 4 &= 0 \\ \lambda^2 - 4\lambda &= 0 \\ \lambda &= 4, 0 \end{aligned}$$

$$\begin{aligned} \begin{pmatrix} 2 & 2 \\ 2 & 2 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} &= 4 \begin{pmatrix} x \\ y \end{pmatrix} \\ x &= y \\ \therefore \begin{pmatrix} x \\ y \end{pmatrix} &= k \begin{pmatrix} 1 \\ 1 \end{pmatrix} \end{aligned}$$

これを正規化する。

$$\begin{aligned} \sqrt{2k^2} &= 1 \\ k &= \frac{1}{\sqrt{2}} \\ \therefore \mathbf{c}_1 &= \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 1 \end{pmatrix} \end{aligned}$$

同様にして、 \mathbf{U} について $\{\mathbf{X}^T \mathbf{Y} (\mathbf{X}^T \mathbf{Y})^T\} \mathbf{U} = \mathbf{U} (\boldsymbol{\Sigma} \boldsymbol{\Sigma}^T)$ となることから、

$$\begin{pmatrix} 1 & 1 \\ 0 & 0 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 1 \\ 1 & 0 & 1 \end{pmatrix} = \begin{pmatrix} 2 & 0 & 2 \\ 0 & 0 & 0 \\ 2 & 0 & 2 \end{pmatrix}$$
$$\begin{vmatrix} 2-\lambda & 0 & 2 \\ 0 & -\lambda & 0 \\ 2 & 0 & 2-\lambda \end{vmatrix} = 0$$

より、

$$\begin{aligned} -\lambda(2-\lambda)^2 + 4\lambda &= 0 \\ -\lambda(\lambda^2 - 4\lambda + 4) + 4\lambda &= 0 \\ -\lambda(\lambda^2 - 4\lambda) &= 0 \\ \lambda &= 4, 0, 0 \end{aligned}$$

$$\begin{pmatrix} 2 & 0 & 2 \\ 0 & 0 & 0 \\ 2 & 0 & 2 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = 4 \begin{pmatrix} x \\ y \\ z \end{pmatrix}$$
$$x = z, y = 0$$

$$\therefore \begin{pmatrix} x \\ y \\ z \end{pmatrix} = k \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}$$

これを正規化する。

$$\begin{aligned} \sqrt{2k^2} &= 1 \\ k &= \frac{1}{\sqrt{2}} \\ \therefore \mathbf{w}_1 &= \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix} \end{aligned}$$

(5) $\mathbf{t} = \mathbf{X}\mathbf{w}$ より、

$$\begin{aligned}\mathbf{t} &= \begin{pmatrix} -1 & 0 & -1 \\ 1 & 0 & 1 \end{pmatrix} \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix} \\ &= \frac{1}{\sqrt{2}} \begin{pmatrix} -2 \\ 2 \end{pmatrix}\end{aligned}$$

これを正規化して

$$\mathbf{t} = \frac{1}{\sqrt{2}} \begin{pmatrix} -1 \\ 1 \end{pmatrix}$$

同様にして $\mathbf{u} = \mathbf{Y}\mathbf{c}$ より、

$$\begin{aligned}\mathbf{u} &= \begin{pmatrix} -\frac{1}{2} & -\frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} \end{pmatrix} \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 1 \end{pmatrix} \\ &= \frac{1}{\sqrt{2}} \begin{pmatrix} -1 \\ 1 \end{pmatrix}\end{aligned}$$

(6) $\mathbf{b} = \mathbf{u}^T \mathbf{t}$ より、

$$\begin{aligned}\mathbf{b} &= \frac{1}{\sqrt{2}} \begin{pmatrix} -1 & 1 \end{pmatrix} \frac{1}{\sqrt{2}} \begin{pmatrix} -1 \\ 1 \end{pmatrix} \\ &= 1\end{aligned}$$

(このような計算をしなくても気付くが、念のため計算した。)

(7) $\mathbf{p} = \mathbf{X}^T \mathbf{t}$ より、

$$\begin{aligned}\mathbf{p} &= \begin{pmatrix} -1 & 1 \\ 0 & 0 \\ -1 & 1 \end{pmatrix} \frac{1}{\sqrt{2}} \begin{pmatrix} -1 \\ 1 \end{pmatrix} \\ &= \begin{pmatrix} \sqrt{2} \\ 0 \\ \sqrt{2} \end{pmatrix}\end{aligned}$$

同様にして $\mathbf{q} = \mathbf{Y}^T \mathbf{u}$ より、

$$\begin{aligned}\mathbf{q} &= \begin{pmatrix} -\frac{1}{2} & \frac{1}{2} \\ -\frac{1}{2} & \frac{1}{2} \end{pmatrix} \frac{1}{\sqrt{2}} \begin{pmatrix} -1 \\ 1 \end{pmatrix} \\ &= \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 1 \end{pmatrix}\end{aligned}$$

(8) $\mathbf{X} - \mathbf{t}\mathbf{p}^T = 0$

(9) 回帰係数行列 \mathbf{B}_{PLS} を求める。

(a) \mathbf{p}^T を特異値分解する。 $\mathbf{p}^T = \begin{pmatrix} \sqrt{2} & 0 & \sqrt{2} \end{pmatrix}$ より、

$$\begin{aligned}(\mathbf{p}^T \mathbf{p})\mathbf{U} &= \mathbf{U}(\boldsymbol{\Sigma}\boldsymbol{\Sigma}^T) \\ 4\mathbf{U} &= \mathbf{U}(\boldsymbol{\Sigma}\boldsymbol{\Sigma}^T)\end{aligned}$$

$$\begin{aligned}|4 - \lambda| &= 0 \\ \therefore \lambda &= 4, \quad \mathbf{u}_1 = 1\end{aligned}$$

また、

$$\mathbf{p}\mathbf{p}^T = \begin{pmatrix} 2 & 0 & 2 \\ 0 & 0 & 0 \\ 2 & 0 & 2 \end{pmatrix}$$

$$\begin{vmatrix} 2 - \lambda & 0 & 2 \\ 0 & -\lambda & 0 \\ 2 & 0 & 2 - \lambda \end{vmatrix} = 0$$

$$\begin{aligned}-\lambda(2 - \lambda)^2 + 4\lambda &= 0 \\ -\lambda(\lambda^2 - 4\lambda + 4) + 4\lambda &= 0 \\ -\lambda(\lambda^2 - 4\lambda) &= 0 \\ \therefore \lambda &= 4, 0, 0\end{aligned}$$

$$\begin{pmatrix} 2 & 0 & 2 \\ 0 & 0 & 0 \\ 2 & 0 & 2 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = 4 \begin{pmatrix} x \\ y \\ z \end{pmatrix}$$

より、 $x = z, y = 0$ であるから、

$$\mathbf{v}_1 = k \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}$$

正規化して、

$$\mathbf{v}_1 = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}$$

よって、 \mathbf{p}^T の特異値分解は次のようになる。

$$\therefore \mathbf{p}^T = 1 \cdot \begin{pmatrix} 2 & 0 & 0 \end{pmatrix} \begin{pmatrix} \frac{1}{\sqrt{2}} & 0 & \frac{1}{\sqrt{2}} \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

(b) \mathbf{p}^T の一般逆行列 $(\mathbf{p}^T)^\#$ を求める。 $\mathbf{A}^\# = \mathbf{V}\Sigma'^{-1}\mathbf{U}^T$ より、

$$\begin{aligned} (\mathbf{p}^T)^\# &= \begin{pmatrix} \frac{1}{\sqrt{2}} & 0 & 0 \\ 0 & 0 & 0 \\ \frac{1}{\sqrt{2}} & 0 & 0 \end{pmatrix} \begin{pmatrix} \frac{1}{2} \\ 0 \\ 0 \end{pmatrix} \cdot 1 \\ &= \frac{1}{2\sqrt{2}} \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix} \end{aligned}$$

(c) 回帰係数行列 \mathbf{B}_{PLS} を求める。 $\mathbf{B}_{\text{PLS}} = (\mathbf{P}^T)^\# \mathbf{BQ}^T$ より

$$\begin{aligned} \mathbf{B}_{\text{PLS}} &= \frac{1}{2\sqrt{2}} \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix} \cdot 1 \cdot \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 \end{pmatrix} \\ &= \frac{1}{4} \begin{pmatrix} 1 & 1 \\ 0 & 0 \\ 1 & 1 \end{pmatrix} \end{aligned}$$

(10)

$$\begin{aligned}
\mathbf{Y}_0 - \bar{\mathbf{Y}} &= (\mathbf{X}_0 - \bar{\mathbf{X}})\mathbf{B}_{\text{PLS}} \\
\mathbf{Y}_0 &= \mathbf{X}_0\mathbf{B}_{\text{PLS}} - \bar{\mathbf{X}}\mathbf{B}_{\text{PLS}} + \bar{\mathbf{Y}} \\
\begin{pmatrix} \text{Fos}_{\text{EGF}} & \text{Jun}_{\text{EGF}} \\ \text{Fos}_{\text{NGF}} & \text{Jun}_{\text{NGF}} \end{pmatrix} &= \begin{pmatrix} \text{ERK}_{\text{EGF}}^{t1} & \text{ERK}_{\text{EGF}}^{t2} & \text{ERK}_{\text{EGF}}^{t3} \\ \text{ERK}_{\text{NGF}}^{t1} & \text{ERK}_{\text{NGF}}^{t2} & \text{ERK}_{\text{NGF}}^{t3} \end{pmatrix} \frac{1}{4} \begin{pmatrix} 1 & 1 \\ 0 & 0 \\ 1 & 1 \end{pmatrix} \\
&\quad - \begin{pmatrix} 1 & 4 & 3 \\ 1 & 4 & 3 \end{pmatrix} \frac{1}{4} \begin{pmatrix} 1 & 1 \\ 0 & 0 \\ 1 & 1 \end{pmatrix} + \begin{pmatrix} \frac{1}{2} & \frac{3}{2} \\ \frac{1}{2} & \frac{3}{2} \end{pmatrix}
\end{aligned}$$

$$\therefore \begin{cases} \text{Fos}_{\text{EGF}} &= \frac{1}{4}(\text{ERK}_{\text{EGF}}^{t1} + \text{ERK}_{\text{EGF}}^{t3}) - \frac{1}{2} \\ \text{Jun}_{\text{EGF}} &= \frac{1}{4}(\text{ERK}_{\text{EGF}}^{t1} + \text{ERK}_{\text{EGF}}^{t3}) + \frac{1}{2} \\ \text{Fos}_{\text{NGF}} &= \frac{1}{4}(\text{ERK}_{\text{NGF}}^{t1} + \text{ERK}_{\text{NGF}}^{t3}) - \frac{1}{2} \\ \text{Jun}_{\text{NGF}} &= \frac{1}{4}(\text{ERK}_{\text{NGF}}^{t1} + \text{ERK}_{\text{NGF}}^{t3}) + \frac{1}{2} \end{cases}$$

(11) $\mathbf{p} = \begin{pmatrix} \sqrt{2} \\ 0 \\ \sqrt{2} \end{pmatrix}$ より、 \mathbf{X} の第 1 列、第 3 列の値が \mathbf{t} , \mathbf{u} に寄与する。

$\mathbf{q} = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 1 \end{pmatrix}$ より、 \mathbf{Y} の第 1 列、第 2 列に対して \mathbf{u} は均等に寄与する。

よって、ERK 時系列データ (\mathbf{X}) のうち、c-Fos、c-Jun の発現量 (\mathbf{Y}) に最も寄与している時点は $t1$ と $t3$ (第 1 列と第 3 列) である。

1.8.3 まとめ

1.9 シグナル伝達系の多変量解析

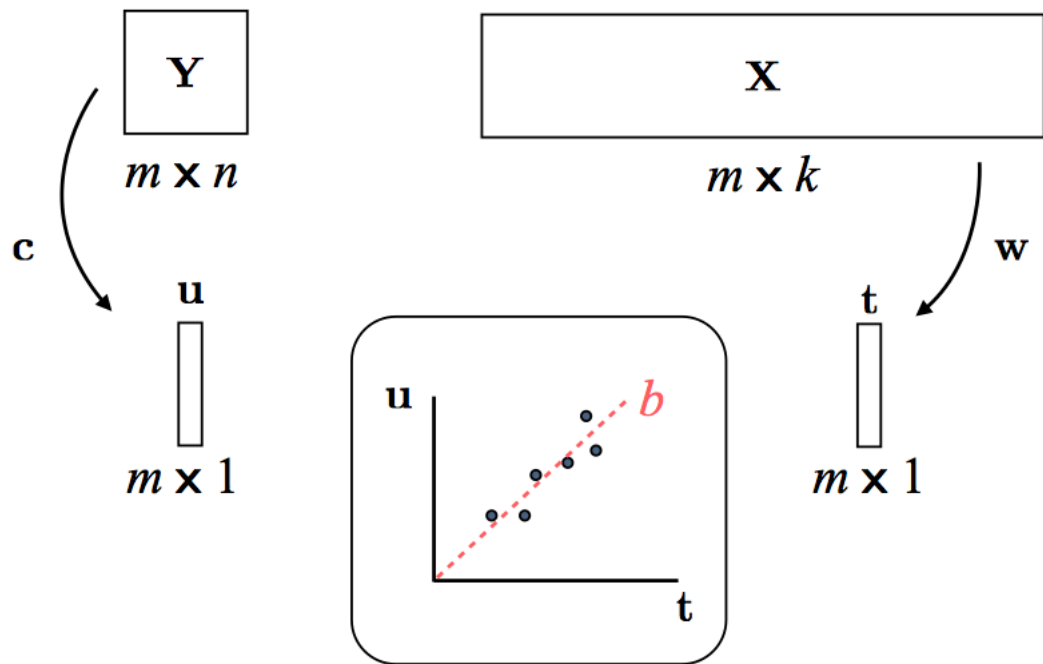
1.9.1 Janes and Yaffe のデータ

Metrics

ダウンロード stimuli は Cell、responses は Science。

1.9.2 R を使う

Mevik and Wehrens "The pls Package: Principal Component and Parital Least Squares Regression in R", J. Stat. Soft. 18(2): , 2007.



1.9.3 Janes and Yaffe のデータに PLS 回帰を適用

1.10 Further reading

Abdi, H. (2007) Partial least square regression (PLS regression). In N.J. Salkind (ed.): Encyclopedia of Measurement and Statistics. Thousand Oaks, CA, pp. 740-744. (本書における PLS 回帰の解説はこの論文に基づいている。当該論文 PDF ファイルは著者である Abdi 教授のホームページから入手できる <http://utdallas.edu/herve/>)

Jane and Yaffe 京都大学大学院工学研究科化学工学専攻プロセスシステム工学研究室・加納学氏のテキスト