数理統計 補助資料
～多変量解析のための固有値分解の復習～

2023年度2学期:　月曜1限, 水曜3限
担当教員:　石垣　司

多変量解析

複数の変数から発生したデータを利用して, 変数間の関係性を分析する手法の総称
- 可視化: データが含む情報を効率的に提示する手段
- 予測・判別: 未知データの予測や判別
- 仮説検証: データと計量モデルを用いて仮説の正しさを検証
- 発見: データ中に存在する(かもしれない)パターンを発見




本授業で扱う手法
- 重回帰分析, 主成分分析, 判別分析, 因子分析など

固有値と固有ベクトル

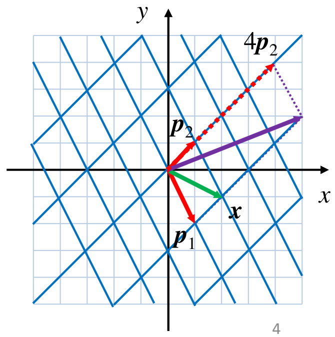
固有値問題
- $N \times N$の正方行列 A に対して, 方向を変化させないベクトルの組 $\{\boldsymbol{p}_1, \dots, \boldsymbol{p}_N\}$ とスカラーの組 $\{\lambda_1, \dots, \lambda_N\}$
$$A\boldsymbol{p}_i = \lambda_i \boldsymbol{p}_i$$
を求める問題
- λ_i を A の固有値, $\boldsymbol{p}_i$ を λ_i に対する固有ベクトルと呼ぶ

固有方程式
$$|A - \lambda I| = \boldsymbol{0}$$
の解 $\lambda(\lambda = \lambda_1, \dots, \lambda_N)$ は固有値となる
- 固有値が求まれば、それに対応する固有ベクトルも求まる
- 問題:次の行列の固有値と固有ベクトルを求めなさい
$$A = \begin{bmatrix} 3 & 1 \\ 2 & 2 \end{bmatrix}$$

固有値分解による対角化の解釈

例: 標準基底での線形変換 $A = \begin{bmatrix} 3 & 1 \\ 2 & 2 \end{bmatrix}$ は, ベクトル $\boldsymbol{x} = [2 \quad -1]$ をどのような変換するか？
- 標準基底での表現 $A\boldsymbol{x} = \begin{bmatrix} 3 & 1 \\ 2 & 2 \end{bmatrix}\begin{bmatrix} 2 \\ -1 \end{bmatrix} = \begin{bmatrix} 5 \\ 2 \end{bmatrix}$
- 固有ベクトルの組の行列 P を基底とした表現
$$P^{-1}AP\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}_{\boldsymbol{p}_1\boldsymbol{p}_2} = \begin{bmatrix} 1 & 0 \\ 0 & 4 \end{bmatrix}$$

標準基底での線形変換 A はベクトル $\boldsymbol{x}$ を
固有ベクトル$\boldsymbol{p}_1 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}_{\boldsymbol{p}_1\boldsymbol{p}_2}$ の方向へ1倍
固有ベクトル$\boldsymbol{p}_2 = \begin{bmatrix} 0 \\ 1 \end{bmatrix}_{\boldsymbol{p}_1\boldsymbol{p}_2}$ の方向へ4倍
する変換を施す



2次形式の行列による表現

2次形式（変数の2次の項“のみ”からなる多項式）

- $x_1, \dots, x_M$：　M 個の変数
- $a_{11}, \dots, a_{MM}$：　$M \times M$ 個の係数

$$f_2(x_1, \dots, x_M) = \sum_{i=1}^{M} \sum_{j=1}^{M} a_{ij} x_i x_j$$
$$= a_{11}x_1^2 +, \dots, +a_{MM}x_M^2 +, \dots, +2a_{12}x_1x_2 +, \dots, +2a_{(M-1)M}x_{M-1}x_M$$

例：$f_2(x, y) = ax^2 + bxy + cy^2$

2次形式の行列による表現

- $\boldsymbol{x} = [x_1 \ \dots \ x_M]^T, A = \begin{bmatrix} a_{11} & \cdots & a_{1M} \\ \vdots & \ddots & \vdots \\ a_{M1} & \cdots & a_{MM} \end{bmatrix}$

$$f_2(\boldsymbol{x}) = \boldsymbol{x}^T A \boldsymbol{x}$$

対称行列と2次形式

対称行列

- $A = A^T$ を満たす正方行列

例：$A = \begin{bmatrix} x & a & b \\ a & y & c \\ b & c & z \end{bmatrix}$

特に, $A \in \mathbb{R}^{M \times M}$ のとき実対象行列とよぶ

2次形式（$\boldsymbol{x}^T A \boldsymbol{x}$）中の行列 A と対称行列

- 行列 A を実対称行列とすると任意の2次形式 f は $f = \boldsymbol{x}^T A \boldsymbol{x}$ と表現できる。
- また, その実対称行列 A はただ1つに定まる

対称行列と直交行列

対称行列 $A = A^T$

- 例：$A = \begin{bmatrix} 1 & 2 \\ 2 & 4 \end{bmatrix}, B = \begin{bmatrix} 2 & 4 & 6 \\ 4 & 3 & 1 \\ 6 & 1 & 5 \end{bmatrix}$
- 性質：対称行列の異なる固有値に対する固有ベクトルは直交する

直交行列 $UU^T = U^T U = I$ または $U^{-1} = U^T$

- 例：$A = \frac{1}{\sqrt{5}} \begin{bmatrix} 2 & 1 \\ 1 & -2 \end{bmatrix}$, $B = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}$
- 性質：正方行列 U の各列のベクトルが直交していて, かつ, ノルムが1に正規化されている

2次形式と行列の定値性

2次形式　$\boldsymbol{x}^T A \boldsymbol{x}$ (復習)

- A は対称行列。$\boldsymbol{x}$ は任意のベクトル
- 例：$x^2 + 4xy + 5y^2 = [x \quad y] \begin{bmatrix} 1 & 2 \\ 2 & 5 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \boldsymbol{x}^T A \boldsymbol{x}$

行列の定値性

零ベクトル以外の任意の x について次の条件を満たすとき、対称行列 A を次のように呼ぶ

- $\boldsymbol{x}^T A \boldsymbol{x} > 0$ のとき、正定値行列　(⇔すべての固有値が正)
- $\boldsymbol{x}^T A \boldsymbol{x} \geq 0$ のとき、半正定値行列 (⇔すべての固有値が非負)
- $\boldsymbol{x}^T A \boldsymbol{x} \leq 0$ のとき、半負定値行列 (⇔すべての固有値が非正)
- $\boldsymbol{x}^T A \boldsymbol{x} < 0$ のとき、負定値行列　(⇔すべての固有値が負)

- 対称行列の定値性の(固有値を用いない)判別法(シルベスタの定理)

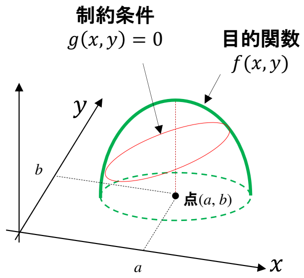
正定値行列 ⇔ 首座小行列式がすべて正
負定値行列 ⇔ 首座小行列式が交互に負、正と繰り返す

制約付き最適化問題の性質

制約無し最適化問題の解との関係

制約付き最大化問題の解 $f(a_1,\cdots,a_N)$
制約無し最大化問題の解 $f(b_1,\cdots,b_N)$

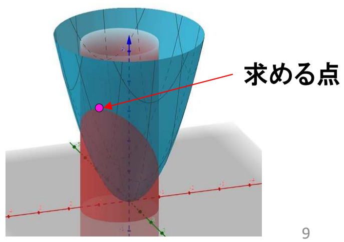
– $f(a_1,\cdots,a_N) \le f(b_1,\cdots,b_N)$



制約付き最適化問題では
点(a,b)が最適点ではない

具体例

– $\max_{x,y}\ x^2+y^2$
– $\mathrm{s.t.}\ x^2+y^2-x-y-1=0$



制約無しの場合 x^2+y^2 の最大化は ∞
制約付きの場合は解あり

ラグランジュ未定乗数法

目的関数 $f(x_1,\cdots,x_N)$
制約条件 $g(x_1,\cdots,x_N)=0$
ラグランジュ関数 $L=f(x_1,\cdots,x_N)-\lambda g(x_1,\cdots,x_N)$
– ラグランジュ乗数: $\lambda \in \mathbb{R}$
– f,g は一階微分可能
– $\frac{\partial L(a_1,\cdots,a_N)}{\partial x_i}$ は L を x_i で偏微分し，その後 $(a_1,\cdots,a_N)$ を代入を表記

制約条件を満たす $f(x_1,\cdots,x_N)$ が点 $(a_1,\cdots,a_N)$ で極値をとり，かつ，$\left[\frac{\partial g(a_1,\cdots,a_N)}{\partial x_1}\ ,\cdots,\ \frac{\partial g(a_1,\cdots,a_N)}{\partial x_N}\right]^T \neq 0$ ならば

$$\Rightarrow \frac{\partial L(a_1,\cdots,a_N)}{\partial x_1}=,\cdots,\frac{\partial L(a_1,\cdots,a_N)}{\partial x_N}=\frac{\partial L(a_1,\cdots,a_N)}{\partial \lambda}=0$$

$f(x_1,\cdots,x_N)$ が制約条件の下で極値となるための必要条件

2次形式の最適化問題と固有値問題

目的関数が2次形式の制約付き最適化問題

– 目的関数 $\boldsymbol{x}^T A \boldsymbol{x}$
– 制約条件 $\boldsymbol{x}^T \boldsymbol{x}=1$

– 固有値問題に帰着する check!

– A が半正定値対称行列のとき，最も大きな(小さな)値をもつ固有値に対する固有ベクトルが $\boldsymbol{x}^T A\boldsymbol{x}$ を最大(最小)とし，その固有値の値は $\boldsymbol{x}^T A\boldsymbol{x}$ に一致する

半正定値対称行列: $\boldsymbol{x}^T A\boldsymbol{x} \ge 0$ を満たす対称行列 A
– $f(\boldsymbol{x})=\boldsymbol{x}^T A\boldsymbol{x}$ とすると関数 $f(\boldsymbol{x})$ は凸関数
$\because$ 定理: $f(\boldsymbol{x})$ は凸関数 $\Leftrightarrow$ $f(\boldsymbol{x})$ のヘッセ行列は半正定値
$f(\boldsymbol{x})$ のヘッセ行列: $H=2A$

分散共分散行列

#メモ 図のデータは分散共分散行列 V の多変量正規分布から1000個の乱数を発生

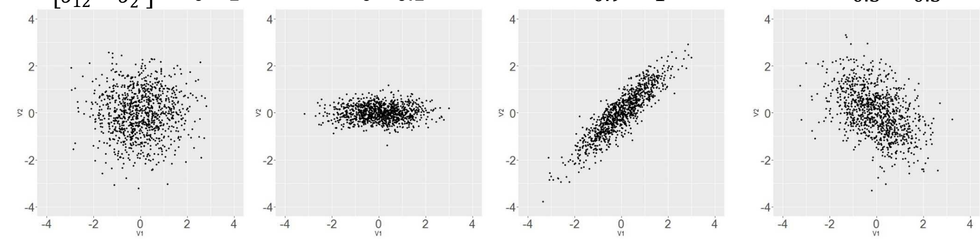
分散共分散行列 V

– $x_1,\dots,x_M$ を確率変数，期待値を E で表すとき、

$$V=\begin{bmatrix} E[(x_1-E[x_1])(x_1-E[x_1])] & \cdots & E[(x_M-E[x_M])(x_1-E[x_1])] \\ \vdots & \ddots & \vdots \\ E[(x_M-E[x_M])(x_1-E[x_1])] & \cdots & E[(x_M-E[x_M])(x_M-E[x_M])] \end{bmatrix}$$

対角成分は各 x_i の分散(σ_i^2)
対角成分以外の要素は各 x_i と x_j の共分散(σ_{ij})(変数間の線形関係の強さを表す)

$V=\begin{bmatrix} \sigma_1^2 & \sigma_{12} \\ \sigma_{12} & \sigma_2^2 \end{bmatrix}=\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$ $V=\begin{bmatrix} 1 & 0 \\ 0 & 0.1 \end{bmatrix}$ $V=\begin{bmatrix} 1 & 0.9 \\ 0.9 & 1 \end{bmatrix}$ $V=\begin{bmatrix} 1 & -0.5 \\ -0.5 & 0.5 \end{bmatrix}$



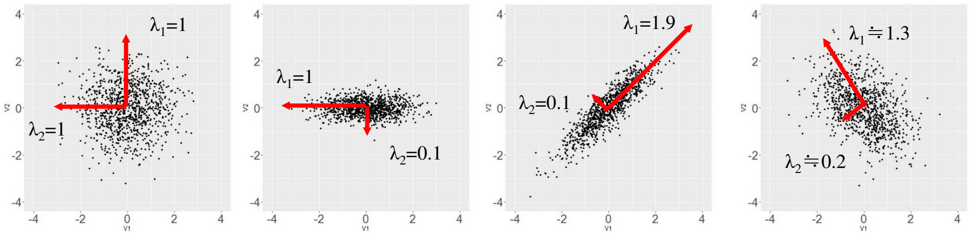
分散共分散行列と固有値問題

分散共分散行列 V は半正定値対称行列

#メモ 多変量解析の手法である
主成分分析や判別分析の数学的本質

– 各固有ベクトルは直交する

$V = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$　$V = \begin{bmatrix} 1 & 0 \\ 0 & 0.1 \end{bmatrix}$　$V = \begin{bmatrix} 1 & 0.9 \\ 0.9 & 1 \end{bmatrix}$　$V = \begin{bmatrix} 1 & -0.5 \\ -0.5 & 0.5 \end{bmatrix}$



$\boldsymbol{p}_1 = k\begin{bmatrix} 0 \\ 1 \end{bmatrix}, \boldsymbol{p}_2 = k\begin{bmatrix} -1 \\ 0 \end{bmatrix}$　$\boldsymbol{p}_1 = k\begin{bmatrix} -1 \\ 0 \end{bmatrix}, \boldsymbol{p}_2 = k\begin{bmatrix} 0 \\ 1 \end{bmatrix}$　$\boldsymbol{p}_1 = k\begin{bmatrix} 0.7 \\ 0.7 \end{bmatrix}, \boldsymbol{p}_2 = k\begin{bmatrix} -0.7 \\ 0.7 \end{bmatrix}$　$\boldsymbol{p}_1 = k\begin{bmatrix} -0.85 \\ 0.53 \end{bmatrix}, \boldsymbol{p}_2 = k\begin{bmatrix} -0.53 \\ -0.85 \end{bmatrix}$

– V の最大固有値に対応する固有ベクトルは, データと直線のズレの2乗の期待値を最小にする直線と同じ方向を向く

演習問題

1. 対称行列 $A = \begin{bmatrix} 3 & 1 \\ 1 & 3 \end{bmatrix}$ に関する次の2次形式の制約付き最適化問題を解き, $x^T A x$ を最大とするベクトル x_1 と最小とするベクトル x_2 をそれぞれ求めなさい。

– 目的関数　$\boldsymbol{x}^T A \boldsymbol{x}$

– 制約条件　$\boldsymbol{x}^T \boldsymbol{x} = 1$