離散数理工学 (2025 年度後学期)

第 11 回

高次元 (5)：点配置と次元削減

岡本 吉央　(電気通信大学)
okamotoy@uec.ac.jp

2026 年 1 月 20 日

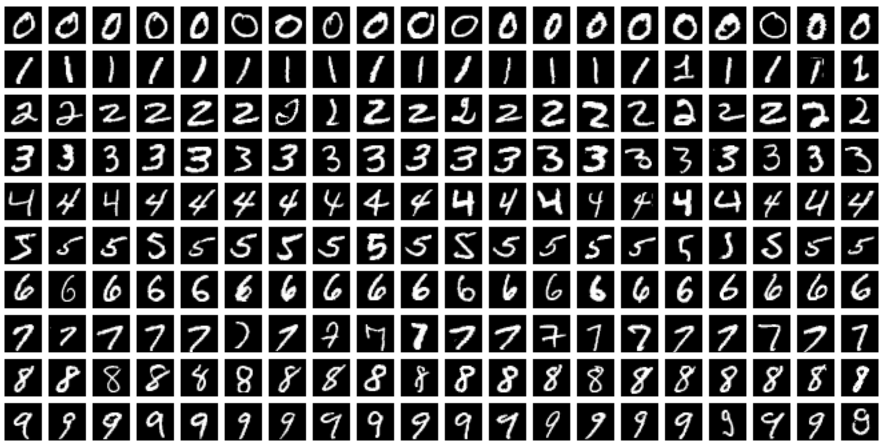
最終更新：2026 年 1 月 12 日　12:01

[復習] なぜ高次元？

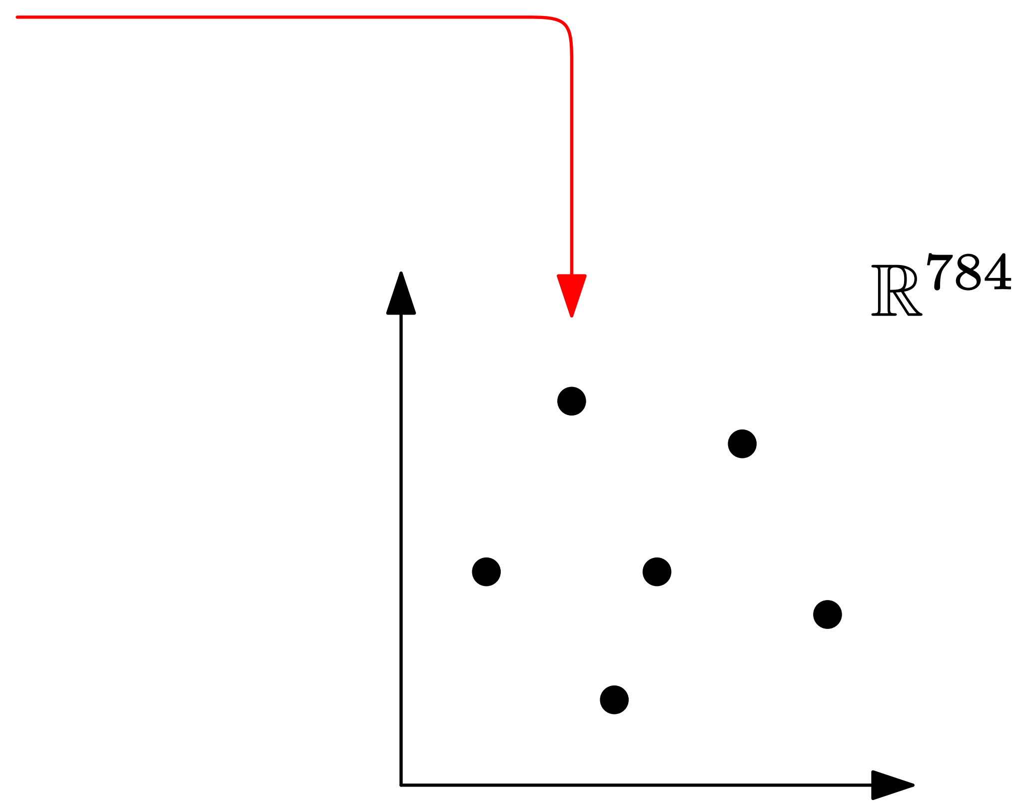
データサイエンスを離散幾何・計算幾何でとらえると

次元 $\approx$ 属性の総数

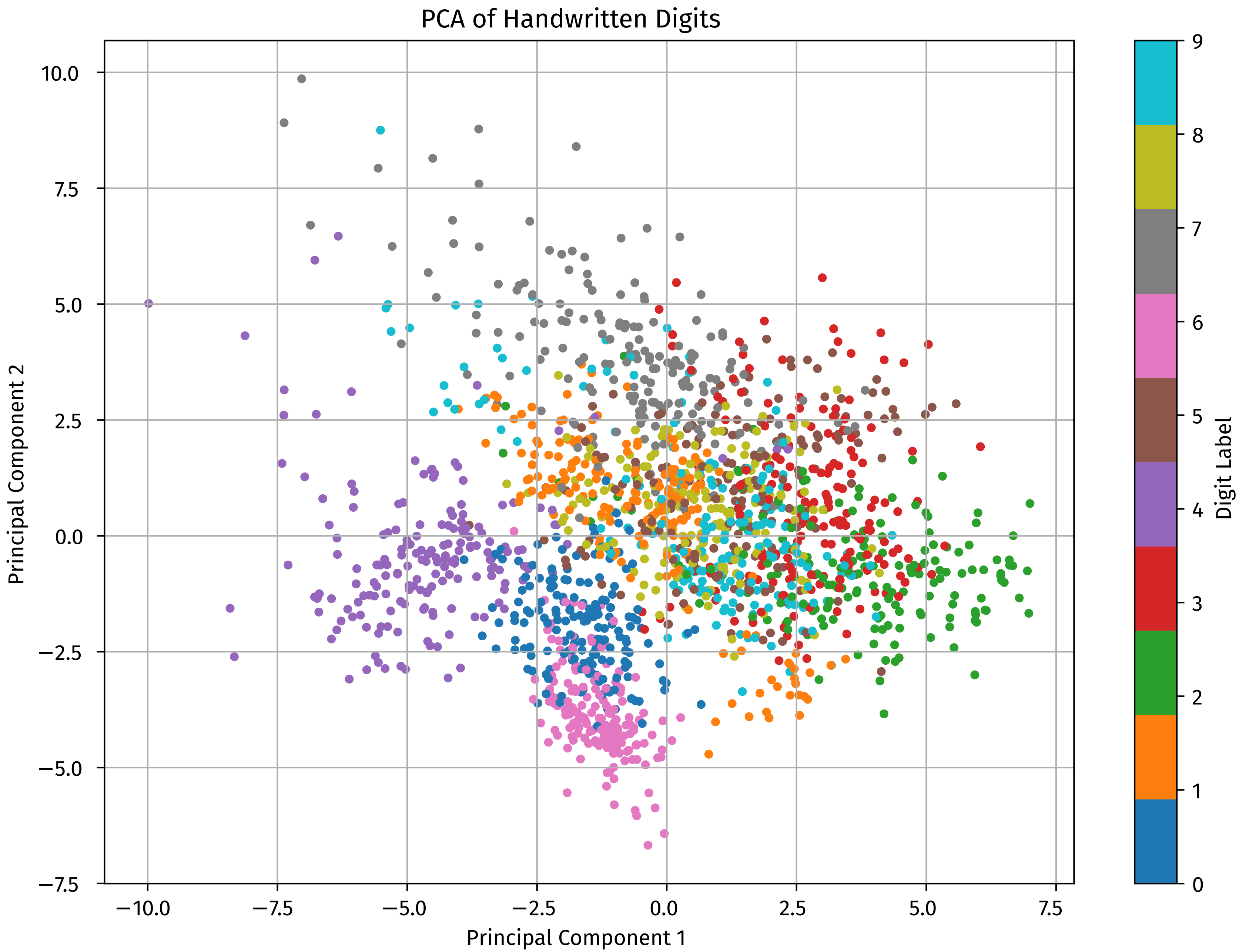
MNIST データベース (手書き文字認識)



28×28 ピクセル



次元削減：例


PCA of Handwritten Digits
Principal Component 1
Principal Component 2
Digit Label
−10.0
−7.5
−5.0
−2.5
0.0
2.5
5.0
7.5
10.0
0
1
2
3
4
5
6
7
8
9

今回の目標

今日の目標

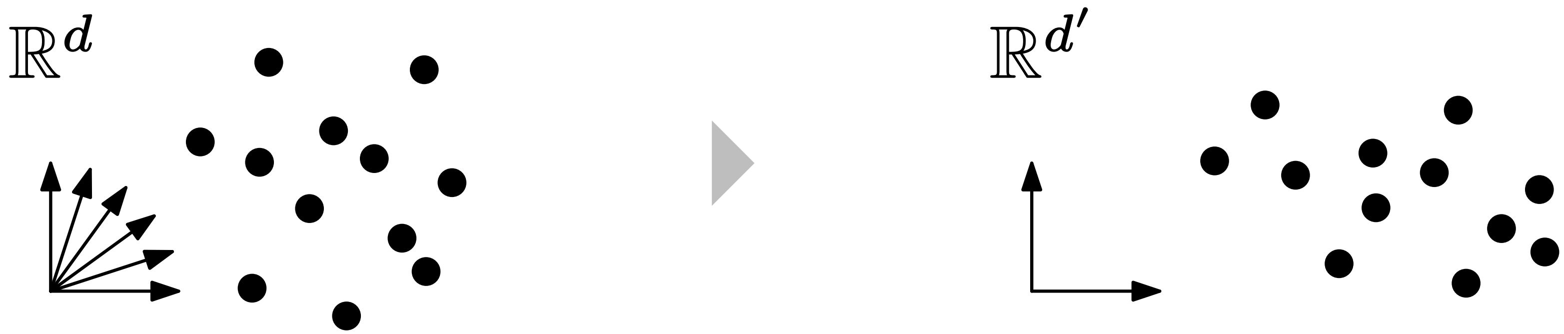
- 次元削減の手法として，主成分分析を使える
- 主成分分析と固有値・固有ベクトルを関連付けられる

目次

次元削減

問題：次元削減

入力： 点 $\boldsymbol{x}_1, \boldsymbol{x}_2, \ldots, \boldsymbol{x}_n \in \mathbb{R}^d$，目標次元 d' $(< d)$
出力： 点 $\boldsymbol{x}'_1, \boldsymbol{x}'_2, \ldots, \boldsymbol{x}'_n \in \mathbb{R}^{d'}$
目標： 出力 $\boldsymbol{x}'_1, \boldsymbol{x}'_2, \ldots, \boldsymbol{x}'_n$ が入力 $\boldsymbol{x}_1, \boldsymbol{x}_2, \ldots, \boldsymbol{x}_n$ を
うまく 説明する

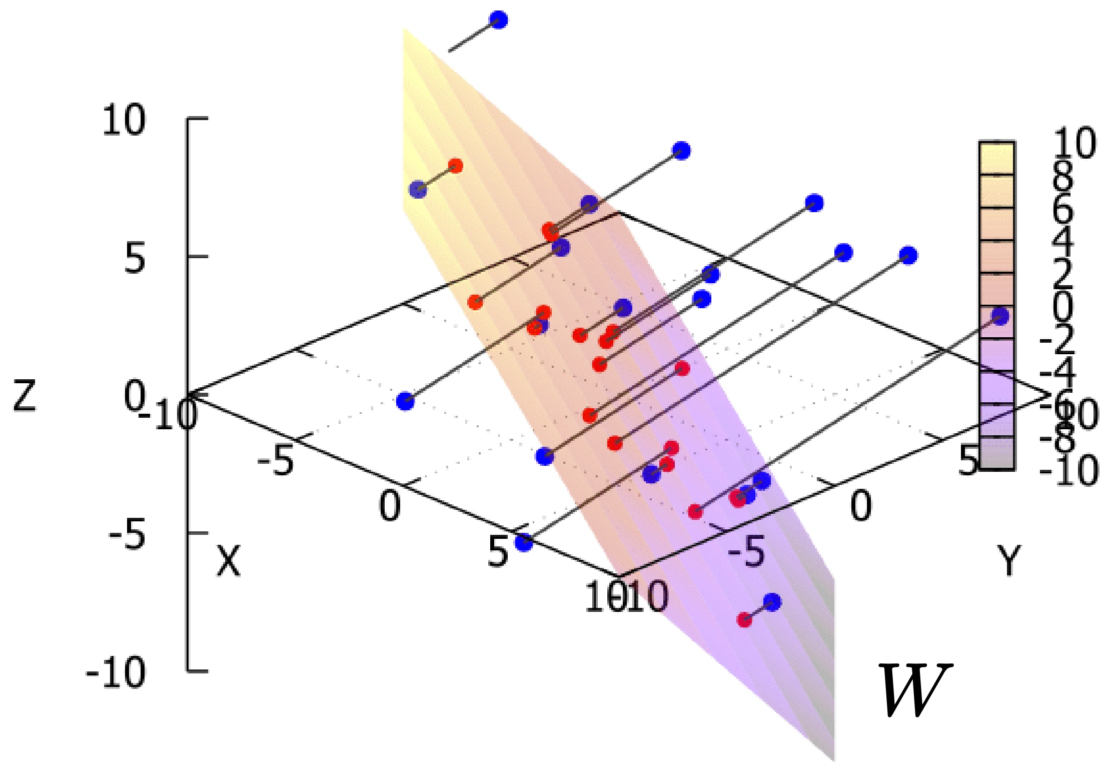


教訓

アルゴリズムを考えるときは，「入力」と「出力」と
その仕様を明示する

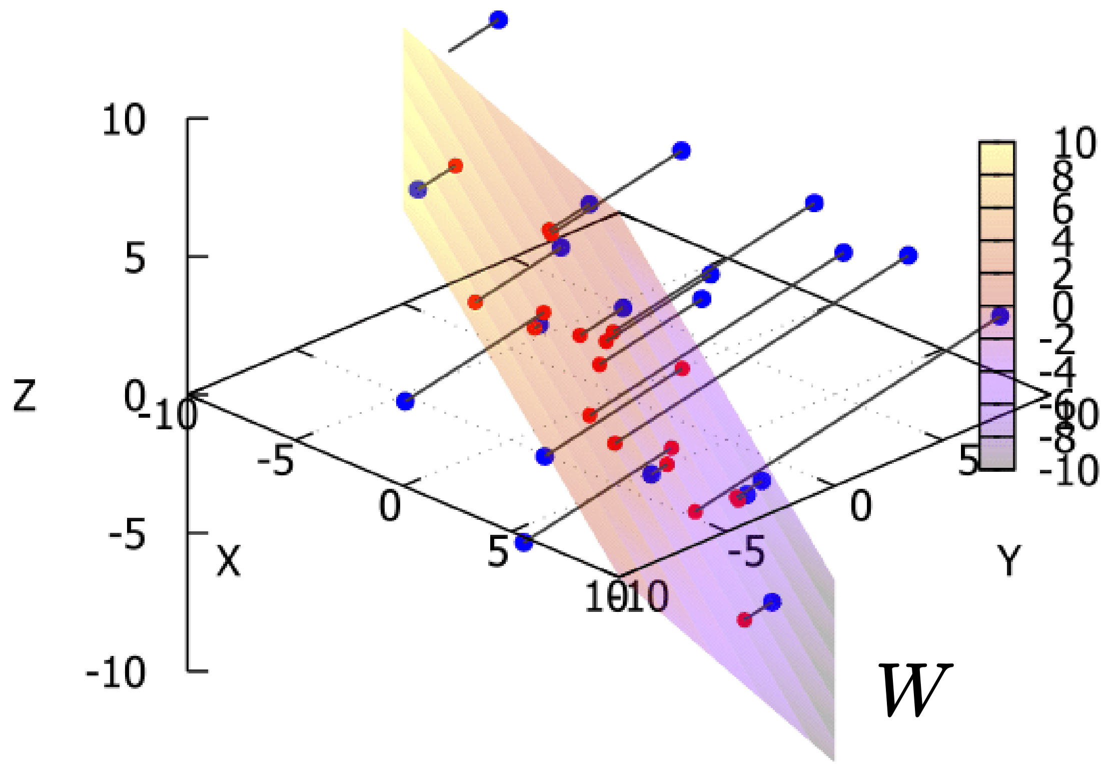
アイディア (1)

$\mathbb{R}^d$ の $\mathbb{R}^{d'}$ 次元アフィン部分空間 W へ直交射影を行う



アイディア (2)

W への距離の二乗の和が最小となるようにする



行列を用いた設定

入力：$\boldsymbol{x}_1, \boldsymbol{x}_2, \ldots, \boldsymbol{x}_n \in \mathbb{R}^d$

前処理：平均を $\mathbf{0}$ にする

$$\sum_{i=1}^{n} \boldsymbol{x}_i = \mathbf{0}$$

行列を用いた設定

入力：$\boldsymbol{x}_1, \boldsymbol{x}_2, \ldots, \boldsymbol{x}_n \in \mathbb{R}^d$

前処理：平均を $\mathbf{0}$ にする

$$\sum_{i=1}^{n} \boldsymbol{x}_i = \mathbf{0}$$

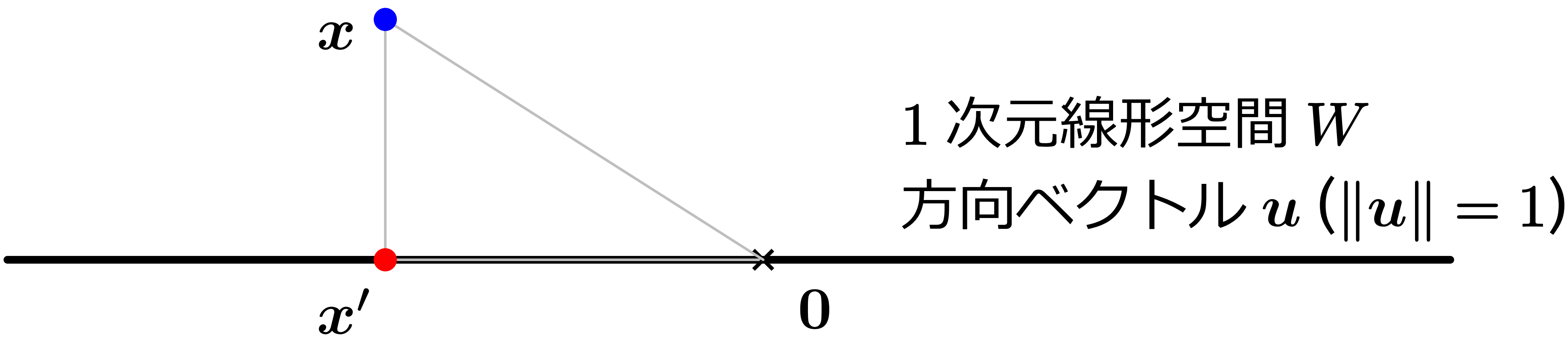
行列に変換

$$X = \begin{array}{cc} \begin{bmatrix} \text{——} & \boldsymbol{x}_1^{\mathrm{T}} & \text{——} \\ \text{——} & \boldsymbol{x}_2^{\mathrm{T}} & \text{——} \\ & & \\ \text{——} & \boldsymbol{x}_n^{\mathrm{T}} & \text{——} \end{bmatrix} & {\color{blue} n} \\ {\color{blue} d} & \end{array}$$

$$\mathbf{1}^{\mathrm{T}} X = \mathbf{0}^{\mathrm{T}}$$

直線への射影

<u>手始めに</u>：$d' = 1$ の場合 (直線への射影) を考える



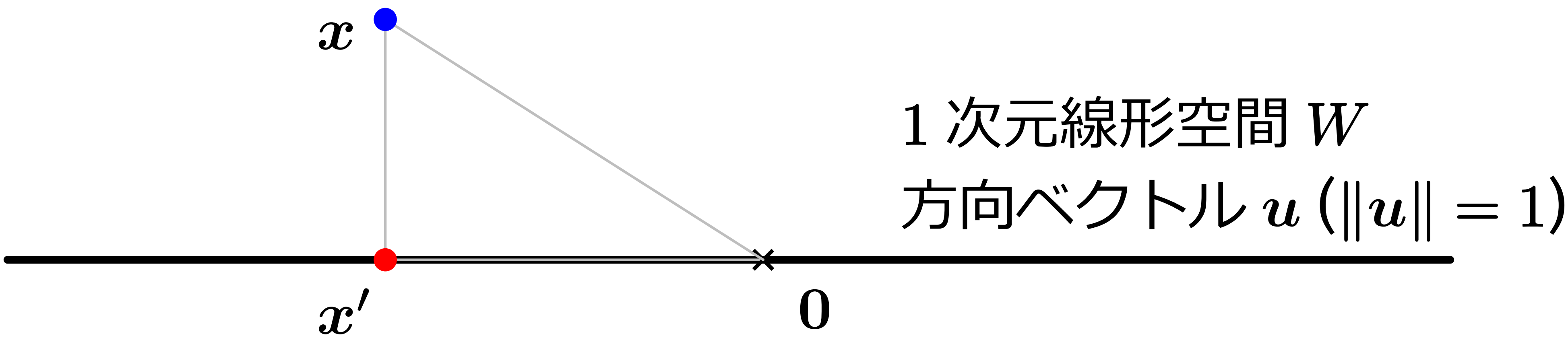
$\boldsymbol{x}' = \alpha\boldsymbol{u}$ とすると ($\alpha \in \mathbb{R}$)

$$\|\boldsymbol{x} - \boldsymbol{x}'\|^2 = \|\boldsymbol{x} - \alpha\boldsymbol{u}\|^2 = (\boldsymbol{x} - \alpha\boldsymbol{u})^{\mathrm{T}}(\boldsymbol{x} - \alpha\boldsymbol{u})$$
$$= \|\boldsymbol{x}\|^2 - 2\alpha\boldsymbol{x}^{\mathrm{T}}\boldsymbol{u} + \alpha^2\|\boldsymbol{u}\|^2$$

$\therefore \|\boldsymbol{x} - \boldsymbol{x}'\|^2$ は $\alpha = \boldsymbol{x}^{\mathrm{T}}\boldsymbol{u}$ のとき最小値 $\|\boldsymbol{x}\|^2 - (\boldsymbol{x}^{\mathrm{T}}\boldsymbol{u})^2$ をとる

直線への射影

手始めに：$d' = 1$ の場合 (直線への射影) を考える



$\boldsymbol{x}' = \alpha\boldsymbol{u}$ とすると ($\alpha \in \mathbb{R}$)

$$\|\boldsymbol{x} - \boldsymbol{x}'\|^2 = \|\boldsymbol{x} - \alpha\boldsymbol{u}\|^2 = (\boldsymbol{x} - \alpha\boldsymbol{u})^{\mathrm{T}}(\boldsymbol{x} - \alpha\boldsymbol{u})$$
$$= \|\boldsymbol{x}\|^2 - 2\alpha\boldsymbol{x}^{\mathrm{T}}\boldsymbol{u} + \alpha^2\|\boldsymbol{u}\|^2$$

$\therefore$ $\|\boldsymbol{x} - \boldsymbol{x}'\|^2$ は $\alpha = \boldsymbol{x}^{\mathrm{T}}\boldsymbol{u}$ のとき最小値 $\|\boldsymbol{x}\|^2 - (\boldsymbol{x}^{\mathrm{T}}\boldsymbol{u})^2$ をとる

$\boldsymbol{x}$ の W への射影 $\boldsymbol{x}'$ は $\boldsymbol{x}' = (\boldsymbol{x}^{\mathrm{T}}\boldsymbol{u})\boldsymbol{u}$ で与えられる

最適化問題として考える

考える最適化問題

変数： $\boldsymbol{u} \in \mathbb{R}^d$

目的： $\sum_{i=1}^{n}(\|\boldsymbol{x}_i\|^2 - (\boldsymbol{x}_i^{\mathrm{T}}\boldsymbol{u})^2)$ の最小化

制約： $\|\boldsymbol{u}\| = 1$

最適化問題として考える

考える最適化問題

変数： $\boldsymbol{u} \in \mathbb{R}^d$

目的： $\displaystyle\sum_{i=1}^{n}(\|\boldsymbol{x}_i\|^2 - (\boldsymbol{x}_i^{\mathrm{T}}\boldsymbol{u})^2)$ の最小化

制約： $\|\boldsymbol{u}\| = 1$

$\downarrow$ $\|\boldsymbol{x}_i\|^2$ は入力だけから決まる

考える最適化問題 (上の問題と等価)

変数： $\boldsymbol{u} \in \mathbb{R}^d$

目的： $\displaystyle\sum_{i=1}^{n}(\boldsymbol{x}_i^{\mathrm{T}}\boldsymbol{u})^2$ の最大化

制約： $\|\boldsymbol{u}\| = 1$

最適化問題として考える

考える最適化問題

変数： $\boldsymbol{u} \in \mathbb{R}^d$

目的： $\displaystyle\sum_{i=1}^{n}(\|\boldsymbol{x}_i\|^2 - (\boldsymbol{x}_i^{\mathrm{T}}\boldsymbol{u})^2)$ の最小化

制約： $\|\boldsymbol{u}\| = 1$

↓ $\|\boldsymbol{x}_i\|^2$ は入力だけから決まる

考える最適化問題 (上の問題と等価)

変数： $\boldsymbol{u} \in \mathbb{R}^d$

目的： $\displaystyle\sum_{i=1}^{n}(\boldsymbol{x}_i^{\mathrm{T}}\boldsymbol{u})^2$ の最大化

制約： $\|\boldsymbol{u}\| = 1$

$$\max_{\substack{\boldsymbol{u}\in\mathbb{R}^d \\ \|\boldsymbol{u}\|=1}} \sum_{i=1}^{n}(\boldsymbol{x}_i^{\mathrm{T}}\boldsymbol{u})^2$$

最適化問題として考える (続)

$$\max_{\substack{\boldsymbol{u}\in\mathbb{R}^d \\ \|\boldsymbol{u}\|=1}} \sum_{i=1}^{n} (\boldsymbol{x}_i^{\mathrm{T}}\boldsymbol{u})^2 \quad = \max_{\substack{\boldsymbol{u}\in\mathbb{R}^d \\ \|\boldsymbol{u}\|=1}} (X\boldsymbol{u})^{\mathrm{T}}(X\boldsymbol{u})$$

$$\because X\boldsymbol{u} = \begin{bmatrix} \boldsymbol{x}_1^{\mathrm{T}} \\ \boldsymbol{x}_2^{\mathrm{T}} \\ \vdots \\ \boldsymbol{x}_n^{\mathrm{T}} \end{bmatrix} \boldsymbol{u} = \begin{bmatrix} \boldsymbol{x}_1^{\mathrm{T}}\boldsymbol{u} \\ \boldsymbol{x}_2^{\mathrm{T}}\boldsymbol{u} \\ \vdots \\ \boldsymbol{x}_n^{\mathrm{T}}\boldsymbol{u} \end{bmatrix}$$

$$\max_{\substack{\boldsymbol{u}\in\mathbb{R}^d\\ \|\boldsymbol{u}\|=1}} \sum_{i=1}^{n} (\boldsymbol{x}_i^{\mathrm{T}}\boldsymbol{u})^2 = \max_{\substack{\boldsymbol{u}\in\mathbb{R}^d\\ \|\boldsymbol{u}\|=1}} (X\boldsymbol{u})^{\mathrm{T}}(X\boldsymbol{u})$$

$$\because X\boldsymbol{u} = \begin{bmatrix} \boldsymbol{x}_1^{\mathrm{T}} \\ \boldsymbol{x}_2^{\mathrm{T}} \\ \vdots \\ \boldsymbol{x}_n^{\mathrm{T}} \end{bmatrix} \boldsymbol{u} = \begin{bmatrix} \boldsymbol{x}_1^{\mathrm{T}}\boldsymbol{u} \\ \boldsymbol{x}_2^{\mathrm{T}}\boldsymbol{u} \\ \vdots \\ \boldsymbol{x}_n^{\mathrm{T}}\boldsymbol{u} \end{bmatrix}$$

$$= \max_{\substack{\boldsymbol{u}\in\mathbb{R}^d\\ \|\boldsymbol{u}\|=1}} \boldsymbol{u}^{\mathrm{T}}X^{\mathrm{T}}X\boldsymbol{u}$$

最適化問題として考える (続)

$$\max_{\substack{\boldsymbol{u}\in\mathbb{R}^d \\ \|\boldsymbol{u}\|=1}} \sum_{i=1}^{n} (\boldsymbol{x}_i^{\mathrm{T}}\boldsymbol{u})^2 = \max_{\substack{\boldsymbol{u}\in\mathbb{R}^d \\ \|\boldsymbol{u}\|=1}} (X\boldsymbol{u})^{\mathrm{T}}(X\boldsymbol{u})$$

$$\because X\boldsymbol{u} = \begin{bmatrix} \boldsymbol{x}_1^{\mathrm{T}} \\ \boldsymbol{x}_2^{\mathrm{T}} \\ \vdots \\ \boldsymbol{x}_n^{\mathrm{T}} \end{bmatrix} \boldsymbol{u} = \begin{bmatrix} \boldsymbol{x}_1^{\mathrm{T}}\boldsymbol{u} \\ \boldsymbol{x}_2^{\mathrm{T}}\boldsymbol{u} \\ \vdots \\ \boldsymbol{x}_n^{\mathrm{T}}\boldsymbol{u} \end{bmatrix}$$

$$= \max_{\substack{\boldsymbol{u}\in\mathbb{R}^d \\ \|\boldsymbol{u}\|=1}} \boldsymbol{u}^{\mathrm{T}} X^{\mathrm{T}} X \boldsymbol{u}$$

$$= \max_{\substack{\boldsymbol{u}\in\mathbb{R}^d \\ \|\boldsymbol{u}\|=1}} \boldsymbol{u}^{\mathrm{T}} A \boldsymbol{u}$$

(ただし，$A = X^{\mathrm{T}}X$)

<u>注</u>：A は対称半正定値

[復習] 対称正定値行列

定義：対称正定値行列

正則な対称行列 $M \in \mathbb{R}^{d \times d}$ に対して，次は同値

1. ある正則行列 $C \in \mathbb{R}^{d \times d}$ が存在して，$M = C^{\mathrm{T}} C$
2. M の固有値はすべて正
3. 任意の $\boldsymbol{x} \in \mathbb{R}^{d} - \{\boldsymbol{0}\}$ に対して，$\boldsymbol{x}^{\mathrm{T}} M \boldsymbol{x} > 0$

この性質を持つ M を **対称正定値行列** と呼ぶ

対称半正定値行列

定義：対称半正定値行列

対称行列 $M \in \mathbb{R}^{d \times d}$ に対して，次は同値

1. ある行列 $C \in \mathbb{R}^{r \times d}$ が存在して，$M = C^{\mathrm{T}} C$
2. M の固有値はすべて非負
3. 任意の $\boldsymbol{x} \in \mathbb{R}^{d}$ に対して，$\boldsymbol{x}^{\mathrm{T}} M \boldsymbol{x} \geq 0$

この性質を持つ M を **対称半正定値行列** と呼ぶ

注：対称半正定値行列は正則であるとは限らない

考える最適化問題

考える最適化問題

入力： 実対称半正定値行列 $A \in \mathbb{R}^{d \times d}$

出力： $\displaystyle \max_{\substack{\boldsymbol{u} \in \mathbb{R}^d \\ \|\boldsymbol{u}\|=1}} \boldsymbol{u}^{\mathrm{T}} A \boldsymbol{u}$

考える最適化問題

考える最適化問題

入力： 実対称半正定値行列 $A \in \mathbb{R}^{d \times d}$

出力： $\displaystyle \max_{\substack{\boldsymbol{u} \in \mathbb{R}^d \\ \|\boldsymbol{u}\|=1}} \boldsymbol{u}^{\mathrm{T}} A \boldsymbol{u}$

性質：最適化と固有値・固有ベクトル

実対称半正定値行列 A の最大固有値を $\lambda_1 \geq 0$ として，
λ_1 に対応する単位固有ベクトルを $\boldsymbol{v}_1$ とすると

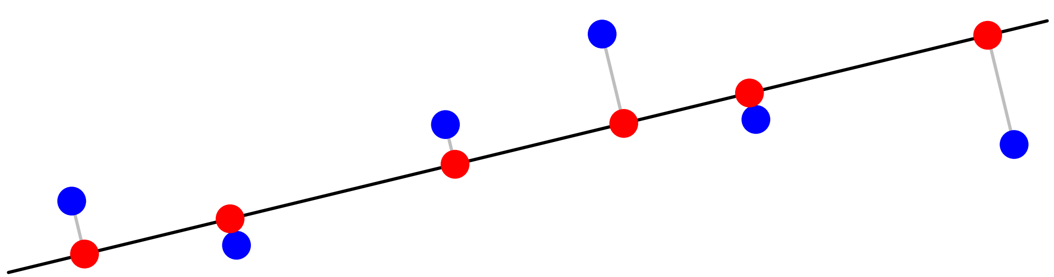
$$\max_{\substack{\boldsymbol{u} \in \mathbb{R}^d \\ \|\boldsymbol{u}\|=1}} \boldsymbol{u}^{\mathrm{T}} A \boldsymbol{u} = \boldsymbol{v}_1^{\mathrm{T}} A \boldsymbol{v}_1 = \lambda_1$$

つまり，$\boldsymbol{v}_1$ が最適解で，λ_1 が最適値

主成分分析：$d' = 1$ の場合

次元削減の手法：主成分分析 (直線への射影)

1. $\boldsymbol{x}_1, \ldots, \boldsymbol{x}_n$ の和が $\mathbf{0}$ となるように平行移動する
2. $\boldsymbol{x}_1^{\mathrm{T}}, \ldots, \boldsymbol{x}_n^{\mathrm{T}}$ を行とする行列 $X \in \mathbb{R}^{n \times d}$ を作る
3. $A = X^{\mathrm{T}} X$ を作る
4. A の最大固有値に対する単位固有ベクトル $\boldsymbol{v}_1$ を求める
5. 各 i に対して，$\boldsymbol{x}_i$ を $\boldsymbol{x}_i^{\mathrm{T}} \boldsymbol{v}_1 \in \mathbb{R}$ に写す



証明 (1)：最適値 = 最大固有値

証明 (最適値 $\geq$ 最大固有値)：

単位固有ベクトル $\boldsymbol{v}_1 \in \mathbb{R}^d$ は $\|\boldsymbol{v}_1\| = 1$ を満たすので

$$\max_{\substack{\boldsymbol{u} \in \mathbb{R}^d \\ \|\boldsymbol{u}\|^2 = 1}} \boldsymbol{u}^{\mathrm{T}} A \boldsymbol{u} \geq \boldsymbol{v}_1^{\mathrm{T}} A \boldsymbol{v}_1 = \boldsymbol{v}_1^{\mathrm{T}} (\lambda_1 \boldsymbol{v}_1) = \lambda_1 (\boldsymbol{v}_1^{\mathrm{T}} \boldsymbol{v}_1) = \lambda_1$$

証明 (1)：最適値 = 最大固有値

証明 (最適値 $\geq$ 最大固有値)：

単位固有ベクトル $\boldsymbol{v}_1 \in \mathbb{R}^d$ は $\|\boldsymbol{v}_1\| = 1$ を満たすので

$$\max_{\substack{\boldsymbol{u} \in \mathbb{R}^d \\ \|\boldsymbol{u}\|^2 = 1}} \boldsymbol{u}^{\mathrm{T}} A \boldsymbol{u} \geq \boldsymbol{v}_1^{\mathrm{T}} A \boldsymbol{v}_1 = \boldsymbol{v}_1^{\mathrm{T}} (\lambda_1 \boldsymbol{v}_1) = \lambda_1 (\boldsymbol{v}_1^{\mathrm{T}} \boldsymbol{v}_1) = \lambda_1$$

証明 (最適値 $\leq$ 最大固有値)：

- $\|\boldsymbol{u}\| = 1$ を満たす任意のベクトル $\boldsymbol{u} \in \mathbb{R}^d$ を考える
- A の固有値を $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_d$ とし，
  $\{\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_d\}$ が $\mathbb{R}^d$ の正規直交基底となるように
  λ_i に対応する固有ベクトル $\boldsymbol{v}_i$ をとれる
  (注：A が対称であるとき，これが可能)

証明 (最適値 $\leq$ 最大固有値)：

- $\boldsymbol{u} = \sum_{i=1}^{d} \alpha_i \boldsymbol{v}_i$ とする
- このとき，$1 = \|\boldsymbol{u}\|^2 = \left\| \sum_i \alpha_i \boldsymbol{v}_i \right\|^2 = \sum_i \alpha_i^2$

$$\left\| \sum_i \alpha_i \boldsymbol{v}_i \right\|^2 = \left(\sum_i \alpha_i \boldsymbol{v}_i \right)^{\mathrm{T}} \left(\sum_j \alpha_j \boldsymbol{v}_j \right)$$

$$= \sum_i \sum_j \alpha_i \alpha_j \underbrace{\boldsymbol{v}_i^{\mathrm{T}} \boldsymbol{v}_j}_{= \begin{cases} 1 & (i = j) \\ 0 & (i \neq j) \end{cases}} = \sum_i \alpha_i^2$$

証明 (2)：最適値 = 最大固有値

証明 (最適値 $\leq$ 最大固有値)：

- $\boldsymbol{u} = \sum_{i=1}^{d} \alpha_i \boldsymbol{v}_i$ とする
- このとき，$1 = \|\boldsymbol{u}\|^2 = \left\| \sum_i \alpha_i \boldsymbol{v}_i \right\|^2 = \sum_i \alpha_i^2$
- したがって

$$
\begin{aligned}
\boldsymbol{u}^{\mathrm{T}} A \boldsymbol{u} &= \left(\sum_{i=1}^{d} \alpha_i \boldsymbol{v}_i \right)^{\mathrm{T}} A \left(\sum_{j=1}^{d} \alpha_j \boldsymbol{v}_j \right) \\
&= \sum_{i=1}^{d} \sum_{j=1}^{d} \alpha_i \alpha_j \boldsymbol{v}_i^{\mathrm{T}} A \boldsymbol{v}_j = \sum_{i=1}^{d} \sum_{j=1}^{d} \alpha_i \alpha_j \lambda_j \boldsymbol{v}_i^{\mathrm{T}} \boldsymbol{v}_j \\
&= \sum_{i=1}^{d} \alpha_i^2 \lambda_i \leq \sum_{i=1}^{d} \alpha_i^2 \lambda_1 = \lambda_1
\end{aligned}
$$

□

関連する用語だけ導入する

定義：レイリー商

実対称行列 $A \in \mathbb{R}^{d \times d}$ とベクトル $\boldsymbol{x} \in \mathbb{R}^d - \{\mathbf{0}\}$ に対する **レイリー商** とは，次の比のこと

$$\frac{\boldsymbol{x}^{\mathrm{T}} A \boldsymbol{x}}{\boldsymbol{x}^{\mathrm{T}} \boldsymbol{x}}$$

注：レイリー商は $\boldsymbol{x}$ のスカラー倍に関して不変

つまり，$\boldsymbol{x}$ を $c\boldsymbol{x}$ に変えると $(c \in \mathbb{R} - \{0\})$

$$\frac{(c\boldsymbol{x})^{\mathrm{T}} A (c\boldsymbol{x})}{(c\boldsymbol{x})^{\mathrm{T}} (c\boldsymbol{x})} = \frac{c^2 (\boldsymbol{x}^{\mathrm{T}} A \boldsymbol{x})}{c^2 (\boldsymbol{x}^{\mathrm{T}} \boldsymbol{x})} = \frac{\boldsymbol{x}^{\mathrm{T}} A \boldsymbol{x}}{\boldsymbol{x}^{\mathrm{T}} \boldsymbol{x}}$$

レイリー商

関連する用語だけ導入する

定義：レイリー商

実対称行列 $A \in \mathbb{R}^{d\times d}$ とベクトル $\boldsymbol{x} \in \mathbb{R}^d - \{\boldsymbol{0}\}$ に対する **レイリー商** とは，次の比のこと

$$\frac{\boldsymbol{x}^{\mathrm{T}} A \boldsymbol{x}}{\boldsymbol{x}^{\mathrm{T}} \boldsymbol{x}}$$

注：レイリー商は $\boldsymbol{x}$ のスカラー倍に関して不変

つまり，$\boldsymbol{x}$ を $c\boldsymbol{x}$ に変えると ($c \in \mathbb{R} - \{0\}$)

$$\frac{(c\boldsymbol{x})^{\mathrm{T}} A (c\boldsymbol{x})}{(c\boldsymbol{x})^{\mathrm{T}} (c\boldsymbol{x})} = \frac{c^2(\boldsymbol{x}^{\mathrm{T}} A \boldsymbol{x})}{c^2(\boldsymbol{x}^{\mathrm{T}} \boldsymbol{x})} = \frac{\boldsymbol{x}^{\mathrm{T}} A \boldsymbol{x}}{\boldsymbol{x}^{\mathrm{T}} \boldsymbol{x}}$$

したがって，$\displaystyle \max_{\boldsymbol{x} \in \mathbb{R}^d - \{\boldsymbol{0}\}} \frac{\boldsymbol{x}^{\mathrm{T}} A \boldsymbol{x}}{\boldsymbol{x}^{\mathrm{T}} \boldsymbol{x}} = A$ の最大固有値

いままでの議論を模倣すると，次が分かる

性質：レイリー商と最大/最小固有値

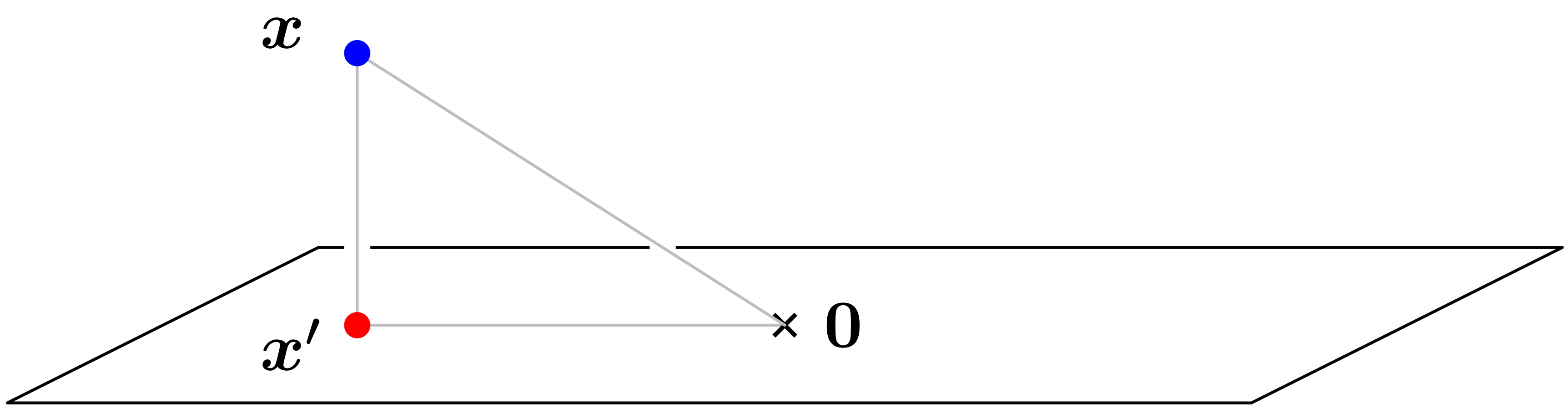
実対称行列 $A \in \mathbb{R}^{d\times d}$ の最大固有値 λ_1，最小固有値 λ_d は次を満たす

$$\lambda_1 = \max_{\boldsymbol{x}\in\mathbb{R}^d-\{\boldsymbol{0}\}} \frac{\boldsymbol{x}^{\mathrm{T}}A\boldsymbol{x}}{\boldsymbol{x}^{\mathrm{T}}\boldsymbol{x}}$$

$$\lambda_d = \min_{\boldsymbol{x}\in\mathbb{R}^d-\{\boldsymbol{0}\}} \frac{\boldsymbol{x}^{\mathrm{T}}A\boldsymbol{x}}{\boldsymbol{x}^{\mathrm{T}}\boldsymbol{x}}$$

目次

アフィン部分空間への射影



d' 次元線形空間 W

正規直交基底 $\{\boldsymbol{u}_1, \boldsymbol{u}_2, \ldots, \boldsymbol{u}_{d'}\}$

$\boldsymbol{x}' = \sum_{j=1}^{d'} \alpha_j \boldsymbol{u}_j$ とすると

$$\begin{aligned}
\|\boldsymbol{x} - \boldsymbol{x}'\|^2 &= \|\boldsymbol{x} - \textstyle\sum_j \alpha_j \boldsymbol{u}_j\|^2 \\
&= \|\boldsymbol{x}\|^2 - 2\textstyle\sum_j \alpha_j \boldsymbol{x}^{\mathrm{T}} \boldsymbol{u}_j + \textstyle\sum_j \alpha_j^2 \|\boldsymbol{u}_j\|^2 \\
&= \|\boldsymbol{x}\|^2 - \textstyle\sum_j (\boldsymbol{x}^{\mathrm{T}} \boldsymbol{u}_j)^2 \quad (\alpha_j = \boldsymbol{x}^{\mathrm{T}} \boldsymbol{u}_j \text{ のとき})
\end{aligned}$$

最適化問題として考える

考える最適化問題

変数： $\boldsymbol{u}_1, \boldsymbol{u}_2, \ldots, \boldsymbol{u}_{d'} \in \mathbb{R}^d$

目的： $\displaystyle\sum_{i=1}^{n}\sum_{j=1}^{d'}(\|\boldsymbol{x}_i\|^2 - (\boldsymbol{x}_i^{\mathrm{T}}\boldsymbol{u}_j)^2)$ の最小化

制約： $\|\boldsymbol{u}_j\| = 1,\ \boldsymbol{u}_j^{\mathrm{T}}\boldsymbol{u}_k = 0\ (j \neq k)$

最適化問題として考える

考える最適化問題

変数： $\boldsymbol{u}_1, \boldsymbol{u}_2, \ldots, \boldsymbol{u}_{d'} \in \mathbb{R}^d$

目的： $\sum_{i=1}^{n} \sum_{j=1}^{d'} (\|\boldsymbol{x}_i\|^2 - (\boldsymbol{x}_i^{\mathrm{T}} \boldsymbol{u}_j)^2)$ の最小化

制約： $\|\boldsymbol{u}_j\| = 1, \ \ \boldsymbol{u}_j^{\mathrm{T}} \boldsymbol{u}_k = 0 \, (j \neq k)$

↓ $\|\boldsymbol{x}_i\|^2$ は入力だけから決まる

考える最適化問題 (上の問題と等価)

変数： $\boldsymbol{u}_1, \boldsymbol{u}_2, \ldots, \boldsymbol{u}_{d'} \in \mathbb{R}^d$

目的： $\sum_{i=1}^{n} \sum_{j=1}^{d'} (\boldsymbol{x}_i^{\mathrm{T}} \boldsymbol{u}_j)^2$ の最大化

制約： $\|\boldsymbol{u}_j\| = 1, \ \ \boldsymbol{u}_j^{\mathrm{T}} \boldsymbol{u}_k = 0 \, (j \neq k)$

最適化問題として考える

考える最適化問題

変数： $\boldsymbol{u}_1, \boldsymbol{u}_2, \ldots, \boldsymbol{u}_{d'} \in \mathbb{R}^d$

目的： $\displaystyle\sum_{i=1}^{n}\sum_{j=1}^{d'}(\|\boldsymbol{x}_i\|^2 - (\boldsymbol{x}_i^{\mathrm{T}}\boldsymbol{u}_j)^2)$ の最小化

制約： $\|\boldsymbol{u}_j\| = 1,\ \ \boldsymbol{u}_j^{\mathrm{T}}\boldsymbol{u}_k = 0\,(j \neq k)$

$\downarrow$ $\|\boldsymbol{x}_i\|^2$ は入力だけから決まる

考える最適化問題 (上の問題と等価)

変数： $\boldsymbol{u}_1, \boldsymbol{u}_2, \ldots, \boldsymbol{u}_{d'} \in \mathbb{R}^d$

目的： $\displaystyle\sum_{i=1}^{n}\sum_{j=1}^{d'}(\boldsymbol{x}_i^{\mathrm{T}}\boldsymbol{u}_j)^2$ の最大化

制約： $\|\boldsymbol{u}_j\| = 1,\ \ \boldsymbol{u}_j^{\mathrm{T}}\boldsymbol{u}_k = 0\,(j \neq k)$

$$\max_{\substack{\boldsymbol{u}_1,\ldots,\boldsymbol{u}_{d'}\in\mathbb{R}^d \\ \|\boldsymbol{u}_j\|=1 \\ \boldsymbol{u}_j^{\mathrm{T}}\boldsymbol{u}_k=0\ (j\neq k)}} \sum_{i=1}^{n}\sum_{j=1}^{d'}(\boldsymbol{x}_i^{\mathrm{T}}\boldsymbol{u}_j)^2$$

最適化問題として考える (続)

以下，$d' = 2$ として，アイディアを説明する

$$
\begin{aligned}
&\max_{\substack{\boldsymbol{u}_1, \boldsymbol{u}_2 \in \mathbb{R}^d \\ \|\boldsymbol{u}_j\| = 1 \\ \boldsymbol{u}_1^{\mathrm{T}} \boldsymbol{u}_2 = 0}} \sum_{i=1}^{n} ((\boldsymbol{x}_i^{\mathrm{T}} \boldsymbol{u}_1)^2 + (\boldsymbol{x}_i^{\mathrm{T}} \boldsymbol{u}_2)^2) \\
&= \max_{\substack{\boldsymbol{u}_1, \boldsymbol{u}_2 \in \mathbb{R}^d \\ \|\boldsymbol{u}_j\| = 1 \\ \boldsymbol{u}_1^{\mathrm{T}} \boldsymbol{u}_2 = 0}} (\boldsymbol{u}_1^{\mathrm{T}} X^{\mathrm{T}} X \boldsymbol{u}_1) + (\boldsymbol{u}_2^{\mathrm{T}} X^{\mathrm{T}} X \boldsymbol{u}_2) \\
&= \max_{\substack{\boldsymbol{u}_1, \boldsymbol{u}_2 \in \mathbb{R}^d \\ \|\boldsymbol{u}_j\| = 1 \\ \boldsymbol{u}_1^{\mathrm{T}} \boldsymbol{u}_2 = 0}} (\boldsymbol{u}_1^{\mathrm{T}} A \boldsymbol{u}_1) + (\boldsymbol{u}_2^{\mathrm{T}} A \boldsymbol{u}_2) \qquad \text{(ただし，} A = X^{\mathrm{T}} X\text{)}
\end{aligned}
$$

<u>注</u>：A は対称半正定値

考える最適化問題

入力： 実対称半正定値行列 $A \in \mathbb{R}^{d \times d}$

出力：

$$\max_{\substack{\boldsymbol{u}_1, \boldsymbol{u}_2 \in \mathbb{R}^d \\ \|\boldsymbol{u}_j\| = 1 \\ \boldsymbol{u}_1^{\mathrm{T}} \boldsymbol{u}_2 = 0}} (\boldsymbol{u}_1^{\mathrm{T}} A \boldsymbol{u}_1) + (\boldsymbol{u}_2^{\mathrm{T}} A \boldsymbol{u}_2)$$

考える最適化問題

考える最適化問題

入力： 実対称半正定値行列 $A \in \mathbb{R}^{d\times d}$

出力： $\displaystyle \max_{\substack{u_1, u_2 \in \mathbb{R}^d \\ \|u_j\|=1 \\ u_1^\mathrm{T} u_2 = 0}} (u_1^\mathrm{T} A u_1) + (u_2^\mathrm{T} A u_2)$

性質：最適化と固有値・固有ベクトル

A の最大・第 2 最大固有値を $\lambda_1, \lambda_2 \geq 0$ として，
λ_1, λ_2 に対応する単位固有ベクトルを v_1, v_2 とすると
(ただし，$v_1^\mathrm{T} v_2 = 0$)

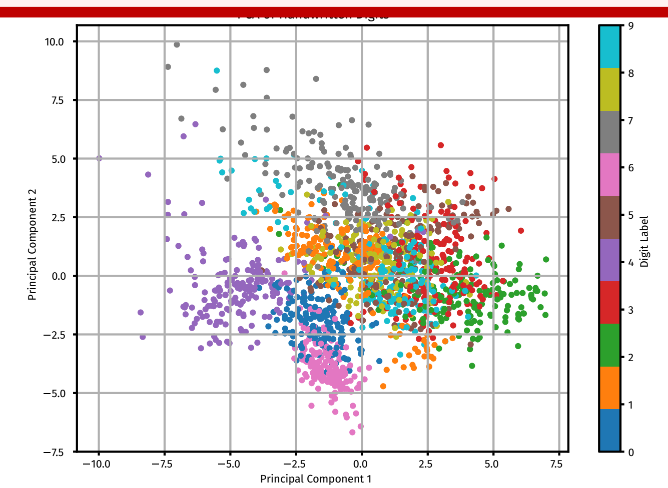
$$
\begin{aligned}
\max_{\substack{u_1, u_2 \in \mathbb{R}^d \\ \|u_j\|=1 \\ u_1^\mathrm{T} u_2 = 0}} (u_1^\mathrm{T} A u_1) + (u_2^\mathrm{T} A u_2) &= v_1^\mathrm{T} A v_1 + v_2^\mathrm{T} A v_2 \\
&= \lambda_1 + \lambda_2
\end{aligned}
$$

次元削減の手法：主成分分析 (平面への射影)

1. $\boldsymbol{x}_1, \ldots, \boldsymbol{x}_n$ の和が $\boldsymbol{0}$ となるように平行移動する
2. $\boldsymbol{x}_1^{\mathrm{T}}, \ldots, \boldsymbol{x}_n^{\mathrm{T}}$ を行とする行列 $X \in \mathbb{R}^{n \times d}$ を作る
3. $A = X^{\mathrm{T}} X$ を作る
4. A の最大固有値，第 2 最大固有値に対する単位固有ベクトル $\boldsymbol{v}_1, \boldsymbol{v}_2$ を求める (ただし，$\boldsymbol{v}_1^{\mathrm{T}} \boldsymbol{v}_2 = 0$ とする)
5. 各 i に対して，$\boldsymbol{x}_i$ を $(\boldsymbol{x}_i^{\mathrm{T}} \boldsymbol{v}_1, \boldsymbol{x}_i^{\mathrm{T}} \boldsymbol{v}_2) \in \mathbb{R}^2$ に写す

$d' > 2$ の場合も同様

(第 d' 最大固有値まで考える)



証明 (1)：最適値 $= \lambda_1 + \lambda_2$

以下，証明は発展的内容

証明 (最適値 $\geq \lambda_1 + \lambda_2$)：

単位固有ベクトル $\boldsymbol{v}_1, \boldsymbol{v}_2 \in \mathbb{R}^d$ を

$\|\boldsymbol{v}_1\|, \|\boldsymbol{v}_2\| = 1$ と $\boldsymbol{v}_1^{\mathrm{T}} \boldsymbol{v}_2 = 0$ を満たすようにとると，

$$\max_{\substack{\boldsymbol{u}_1, \boldsymbol{u}_2 \in \mathbb{R}^d \\ \|\boldsymbol{u}_j\| = 1 \\ \boldsymbol{u}_1^{\mathrm{T}} \boldsymbol{u}_2 = 0}} (\boldsymbol{u}_1^{\mathrm{T}} A \boldsymbol{u}_1) + (\boldsymbol{u}_2^{\mathrm{T}} A \boldsymbol{u}_2) \geq \boldsymbol{v}_1^{\mathrm{T}} A \boldsymbol{v}_1 + \boldsymbol{v}_2^{\mathrm{T}} A \boldsymbol{v}_2 = \lambda_1 + \lambda_2$$

復習

実対称行列 $X \in \mathbb{R}^{d \times d}$ の異なる固有値に対応する固有ベクトルは直交する

証明 (2)：最適値 $= \lambda_1 + \lambda_2$

証明 (最適値 $\leq \lambda_1 + \lambda_2$)：

- A は次のように対角化できる

$$A = \underbrace{\begin{bmatrix} \\ v_1 & \cdots & v_d \\ \\ \end{bmatrix}}_{V} \underbrace{\begin{bmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_d \end{bmatrix}}_{D} \underbrace{\begin{bmatrix} v_1^{\mathrm{T}} \\ \vdots \\ v_d^{\mathrm{T}} \end{bmatrix}}_{V^{\mathrm{T}}}$$

ただし，$v_1, \ldots, v_d$ は A の固有ベクトルで
$VV^{\mathrm{T}} = V^{\mathrm{T}}V = E_d$

証明 (2)：最適値 $= \lambda_1 + \lambda_2$

証明 (最適値 $\leq \lambda_1 + \lambda_2$)：

- A は次のように対角化できる

$$A = \underbrace{\begin{bmatrix} \boldsymbol{v}_1 & \cdots & \boldsymbol{v}_d \end{bmatrix}}_{V} \underbrace{\begin{bmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_d \end{bmatrix}}_{D} \underbrace{\begin{bmatrix} \boldsymbol{v}_1^{\mathrm{T}} \\ \vdots \\ \boldsymbol{v}_d^{\mathrm{T}} \end{bmatrix}}_{V^{\mathrm{T}}}$$

ただし，$\boldsymbol{v}_1, \ldots, \boldsymbol{v}_d$ は A の固有ベクトルで
$VV^{\mathrm{T}} = V^{\mathrm{T}}V = E_d$

- 行列 $U \in \mathbb{R}^{d \times 2}$ を次のように定義する

$$U = \begin{bmatrix} \boldsymbol{u}_1 & \boldsymbol{u}_2 \end{bmatrix}$$

$\|\boldsymbol{u}_j\| = 1$，$\boldsymbol{u}_1^{\mathrm{T}}\boldsymbol{u}_2 = 0$ のとき，$U^{\mathrm{T}}U = E_2$

- このとき，

$$U^{\mathrm{T}}AU = \begin{bmatrix} \boldsymbol{u}_1^{\mathrm{T}} \\ \boldsymbol{u}_2^{\mathrm{T}} \end{bmatrix} A \begin{bmatrix} \boldsymbol{u}_1 & \boldsymbol{u}_2 \end{bmatrix} = \begin{bmatrix} \boldsymbol{u}_1^{\mathrm{T}} \\ \boldsymbol{u}_2^{\mathrm{T}} \end{bmatrix} \begin{bmatrix} A\boldsymbol{u}_1 & A\boldsymbol{u}_2 \end{bmatrix}$$
$$= \begin{bmatrix} \boldsymbol{u}_1^{\mathrm{T}}A\boldsymbol{u}_1 & \boldsymbol{u}_1^{\mathrm{T}}A\boldsymbol{u}_2 \\ \boldsymbol{u}_2^{\mathrm{T}}A\boldsymbol{u}_1 & \boldsymbol{u}_2^{\mathrm{T}}A\boldsymbol{u}_2 \end{bmatrix}$$

- したがって，$\boldsymbol{u}_1^{\mathrm{T}}A\boldsymbol{u}_1 + \boldsymbol{u}_2^{\mathrm{T}}A\boldsymbol{u}_2 = \mathrm{tr}(U^{\mathrm{T}}AU)$

定義：トレイス，跡

実対称行列 $X \in \mathbb{R}^{d \times d}$ の **トレイス** とは，

$$\mathrm{tr}(X) = \sum_{i=1}^{d} x_{ii}$$

証明 (4)：最適値 $= \lambda_1 + \lambda_2$

- すでに導入した記法と性質を使うと

$$
\begin{aligned}
\mathrm{tr}(U^\mathrm{T} A U) &= \mathrm{tr}(U^\mathrm{T} V D V^\mathrm{T} U) \\
&= \mathrm{tr}(Z^\mathrm{T} D Z) \quad \text{（ただし，} Z = V^\mathrm{T} U \in \mathbb{R}^{d \times 2}\text{）} \\
&= \sum_{i=1}^{d} \lambda_i (z_{i1}^2 + z_{i2}^2)
\end{aligned}
$$

証明 (4)：最適値 $= \lambda_1 + \lambda_2$

- すでに導入した記法と性質を使うと

$$
\begin{aligned}
\mathrm{tr}(U^{\mathrm{T}} A U) &= \mathrm{tr}(U^{\mathrm{T}} V D V^{\mathrm{T}} U) \\
&= \mathrm{tr}(Z^{\mathrm{T}} D Z) \quad \text{(ただし，} Z = V^{\mathrm{T}} U \in \mathbb{R}^{d \times 2}\text{)} \\
&= \sum_{i=1}^{d} \lambda_i (z_{i1}^2 + z_{i2}^2)
\end{aligned}
$$

- ここで，$Z^{\mathrm{T}} Z = U^{\mathrm{T}} V V^{\mathrm{T}} U = U^{\mathrm{T}} U = E_2$

  特に，$\mathrm{tr}(Z^{\mathrm{T}} Z) = \mathrm{tr}(I_2) = 2$ なので，$\displaystyle\sum_{i=1}^{d} (z_{i1}^2 + z_{i2}^2) = 2$

補題

いまの設定で，$0 \leq z_{i1}^2 + z_{i2}^2 \leq 1$

補題

いまの設定で，$0 \leq z_{i1}^2 + z_{i2}^2 \leq 1$

<u>補題の証明</u>：$z_{i1}^2 + z_{i2}^2$ は ZZ^{T} の第 i, i 成分

- $ZZ^{\mathrm{T}} = V^{\mathrm{T}} U U^{\mathrm{T}} V$
- $P = UU^{\mathrm{T}}$ は次を満たす
  - $P^{\mathrm{T}} = P$
  - $P^2 = P$
    $(\because P^2 = (UU^{\mathrm{T}})(UU^{\mathrm{T}}) = U(U^{\mathrm{T}}U)U^{\mathrm{T}} = UU^{\mathrm{T}} = P)$
- したがって，P の固有値は 0 か 1
  $(\because P$ の固有値 α に対応する固有ベクトルを z とすると，
  $\alpha z = Pz = P^2 z = P(\alpha z) = \alpha^2 z$ で，$\alpha = \alpha^2)$

証明 (6)：最適値 $= \lambda_1 + \lambda_2$

- $ZZ^{\mathrm{T}} = V^{\mathrm{T}}UU^{\mathrm{T}}V = V^{\mathrm{T}}PV$
- $\therefore ZZ^{\mathrm{T}}$ の固有値も 0 か 1
  $$(\because \det(ZZ^{\mathrm{T}} - \lambda E) = \det(V^{\mathrm{T}}PV - \lambda V^{\mathrm{T}}V) = \det(V^{\mathrm{T}})\det(P - \lambda E)\det(V) = \det(P - \lambda E))$$
- $\therefore$ 任意のベクトル $\boldsymbol{x} \in \mathbb{R}^d$ (ただし，$\|\boldsymbol{x}\| = 1$) に対して
  $$0 \leq \boldsymbol{x}^{\mathrm{T}}(ZZ^{\mathrm{T}})\boldsymbol{x} \leq 1$$
- $\boldsymbol{x} = \boldsymbol{e}_i$ (第 i 基本ベクトル) とすると，
  $$0 \leq ZZ^{\mathrm{T}} \text{ の第 } i, i \text{ 成分} \leq 1$$
- したがって，
  $$0 \leq z_{i1}^2 + z_{i2}^2 \leq 1 \qquad \square$$

証明 (7) : 最適値 $= \lambda_1 + \lambda_2$

- いままでの議論から次が成り立つ

$$\max_{\substack{\boldsymbol{u}_1, \boldsymbol{u}_2 \in \mathbb{R}^d \\ \|\boldsymbol{u}_j\| = 1 \\ \boldsymbol{u}_1^{\mathrm{T}} \boldsymbol{u}_2 = 0}} (\boldsymbol{u}_1^{\mathrm{T}} A \boldsymbol{u}_1) + (\boldsymbol{u}_2^{\mathrm{T}} A \boldsymbol{u}_2)$$

$$\leq \max_{Z \in \mathbb{R}^{d \times 2}} \left\{ \sum_i \lambda_i (z_{i1}^2 + z_{i2}^2) \;\middle|\; \begin{array}{l} 0 \leq z_{i1}^2 + z_{i2}^2 \leq 1 \quad \forall\, i, \\ \sum_i (z_{i1}^2 + z_{i2}^2) = 2 \end{array} \right\}$$

証明 (7)：最適値 $= \lambda_1 + \lambda_2$

- いままでの議論から次が成り立つ

$$\max_{\substack{\boldsymbol{u}_1, \boldsymbol{u}_2 \in \mathbb{R}^d \\ \|\boldsymbol{u}_j\| = 1 \\ \boldsymbol{u}_1^{\mathrm{T}} \boldsymbol{u}_2 = 0}} (\boldsymbol{u}_1^{\mathrm{T}} A \boldsymbol{u}_1) + (\boldsymbol{u}_2^{\mathrm{T}} A \boldsymbol{u}_2)$$

$$\leq \max_{Z \in \mathbb{R}^{d \times 2}} \left\{ \sum_i \lambda_i (z_{i1}^2 + z_{i2}^2) \,\middle|\, \begin{array}{l} 0 \leq z_{i1}^2 + z_{i2}^2 \leq 1 \quad \forall\, i, \\ \sum_i (z_{i1}^2 + z_{i2}^2) = 2 \end{array} \right\}$$

$$z_{11}^2 + z_{12}^2 = 1, z_{21}^2 + z_{22}^2 = 1$$

$$z_{i1}^2 + z_{i2}^2 = 0 \quad (i \neq 1, 2)$$

$$= \lambda_1 + \lambda_2 \qquad \square$$

レイリー商と第 2 最大固有値

性質：レイリー商と第 2 最大固有値

実対称行列 $A \in \mathbb{R}^{d \times d}$ の第 2 最大固有値 λ_2 は次を満たす

$$\lambda_2 = \max_{\substack{\boldsymbol{x} \in \mathbb{R}^d - \{\mathbf{0}\} \\ \boldsymbol{v}_1^{\mathrm{T}} \boldsymbol{x} = 0}} \frac{\boldsymbol{x}^{\mathrm{T}} A \boldsymbol{x}}{\boldsymbol{x}^{\mathrm{T}} \boldsymbol{x}}$$

ただし，$\boldsymbol{v}_1$ は最大固有値に対応する固有ベクトル

証明：演習問題

まとめ

今日の目標

- 次元削減の手法として，主成分分析を使える
- 主成分分析と固有値・固有ベクトルを関連付けられる

主成分分析のまとめ

1. データを行列 $X \in \mathbb{R}^{n\times d}$ で表す (平均は $\mathbf{0}$ にする)
2. $X^{\mathrm{T}}X$ の固有値・固有ベクトルを求める
3. 大きい方から順に固有値・固有ベクトルを持ってくる