射影と最小二乗法

数理工学 第 2 回

小林 健/ Ken Kobayashi

東京科学大 工学院 経営工学系

2025 年 10 月 06 日

目次

目次

部分空間: 和とスカラー倍に閉じた空間

定義 (部分空間)

線形空間 V の (空でない) 部分集合 $W \subseteq V$ が次の条件を満たすとき,
W を V の**部分空間 (subspace)** という:

(1) 任意の $\boldsymbol{u}, \boldsymbol{v} \in W$ に対して, $\boldsymbol{u} + \boldsymbol{v} \in W$, (和に関して閉じている)

(2) 任意の $\boldsymbol{v} \in W$, $c \in \mathbb{R}$ に対して, $c\boldsymbol{v} \in W$. (スカラー倍に関して閉じている)

- n 個の m 次実ベクトル $\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_n \in \mathbb{R}^m$ の線形結合で表せるベクトル全体の集合

$$W = \{c_1\boldsymbol{v}_1 + c_2\boldsymbol{v}_2 + \cdots + c_n\boldsymbol{v}_n \mid c_1, c_2, \ldots, c_n \in \mathbb{R}\}$$

  は $\mathbb{R}^m$ の部分空間である.

- W を $\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_n$ によって**張られる**部分空間といい,

$$W = \langle \boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_n \rangle$$

  と表す.

部分空間の基底

定義 (線形独立と線形従属)

線形空間 V 上の元 $\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_n \in V$ に対して,

$$c_1\boldsymbol{v}_1 + c_2\boldsymbol{v}_2 + \cdots + c_n\boldsymbol{v}_n = \boldsymbol{0} \quad \Longrightarrow \quad c_1 = c_2 = \cdots = c_n = 0$$

が成り立つとき, $\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_n$ は**線形独立 (linear independent)** であるという.

線形独立でないベクトルの組を**線形従属 (linear dependent)** という.

定義 (部分空間の基底)

線形空間 V の部分空間 W に対して, $\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_n \in W$ が次の条件を満たすとき, $\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_n$ を W の**基底 (base)** という:

(1) $W = \langle \boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_n \rangle$,

(2) $\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_n$ は線形独立である.

部分空間への射影

定義 (直交射影)

内積空間 V 上の部分空間 $W \subseteq V$ と, ベクトル $\boldsymbol{v} \in V$ に対して,

$$\forall \boldsymbol{y} \in W,\ ((\boldsymbol{v} - \boldsymbol{p}) \cdot \boldsymbol{y}) = 0$$

となる $\boldsymbol{p} \in W$ を $\boldsymbol{v}$ の W への**射影 (projection)** という.

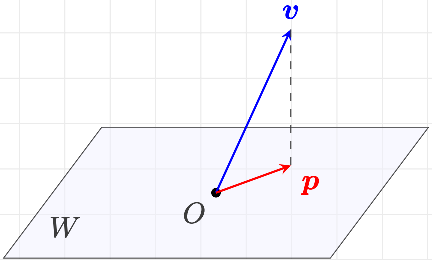

Figure: ベクトル $\boldsymbol{v}$ の部分空間 W への射影 $\boldsymbol{p}$

射影の意味づけ

定理 (射影は部分空間上で最も「近い」)

内積空間 V 上の部分空間を $W \subseteq V$, ベクトル $\boldsymbol{v} \in V$ の W への射影を $\boldsymbol{p} \in W$ とする.
このとき以下が成り立つ:

$$\forall \boldsymbol{y} \in W, \ \|\boldsymbol{v} - \boldsymbol{y}\| \geq \|\boldsymbol{v} - \boldsymbol{p}\|.$$

証明

- 任意の $\boldsymbol{y} \in W$ に対して,

$$\|\boldsymbol{v} - \boldsymbol{y}\|^2 = \|(\boldsymbol{v} - \boldsymbol{p}) + (\boldsymbol{p} - \boldsymbol{y})\|^2 = \|\boldsymbol{v} - \boldsymbol{p}\|^2 + 2((\boldsymbol{v} - \boldsymbol{p}) \cdot (\boldsymbol{p} - \boldsymbol{y})) + \|\boldsymbol{p} - \boldsymbol{y}\|^2.$$

- $\boldsymbol{p}$ は $\boldsymbol{v}$ の W への射影なので, $\boldsymbol{v} - \boldsymbol{p}$ は W 上の任意のベクトル $\boldsymbol{y}$ と直交する.
- よって $\boldsymbol{p} - \boldsymbol{y} \in W$ より $((\boldsymbol{v} - \boldsymbol{p}) \cdot (\boldsymbol{p} - \boldsymbol{y})) = 0$ であり,

$$\|\boldsymbol{v} - \boldsymbol{y}\|^2 = \|\boldsymbol{v} - \boldsymbol{p}\|^2 + \|\boldsymbol{p} - \boldsymbol{y}\|^2 \geq \|\boldsymbol{v} - \boldsymbol{p}\|^2.$$

目次

射影の求め方

ベクトル $\boldsymbol{v}$ の部分空間 W への射影 $\boldsymbol{p}$ を求める

方針

- $\boldsymbol{p}$ は W 上のベクトルであるため, W の基底 $\boldsymbol{y}_1, \ldots, \boldsymbol{y}_k$ を用いて

$$\boldsymbol{p} = \lambda_1 \boldsymbol{y}_1 + \lambda_2 \boldsymbol{y}_2 + \cdots + \lambda_k \boldsymbol{y}_k$$

  と表現できる．ここで $\lambda_1, \lambda_2, \ldots, \lambda_k \in \mathbb{R}$ は係数.

- 射影の定義から

$$\forall \boldsymbol{w} \in W, \ ((\boldsymbol{v} - \boldsymbol{p}) \cdot \boldsymbol{w}) = 0$$

  となる $\lambda_1, \lambda_2, \ldots, \lambda_k$ を求めればよい.

🤔 **「W 上のすべてのベクトルと直交する」という条件をどう扱うか?**

部分空間上のベクトルと直交する条件

補題

部分空間 W の基底を $\boldsymbol{y}_1, \boldsymbol{y}_2, \ldots, \boldsymbol{y}_k$ とする．このとき，

$$\forall \boldsymbol{w} \in W,\ (\boldsymbol{z} \cdot \boldsymbol{w}) = 0 \quad \Longleftrightarrow \quad (\boldsymbol{z} \cdot \boldsymbol{y}_1) = (\boldsymbol{z} \cdot \boldsymbol{y}_2) = \cdots = (\boldsymbol{z} \cdot \boldsymbol{y}_k) = 0.$$

証明 $(\Rightarrow)$ は明らか．$(\Leftarrow)$ を示す．

- $\boldsymbol{y}_1, \boldsymbol{y}_2, \ldots, \boldsymbol{y}_k$ は W の基底なので，任意の $\boldsymbol{w} \in W$ は次のように書ける：

$$\boldsymbol{w} = \lambda_1 \boldsymbol{y}_1 + \lambda_2 \boldsymbol{y}_2 + \cdots + \lambda_k \boldsymbol{y}_k \quad (\lambda_1, \ldots, \lambda_k \in \mathbb{R}).$$

- 内積 $(\boldsymbol{z} \cdot \boldsymbol{w})$ を計算すると

$$\begin{aligned} (\boldsymbol{z} \cdot \boldsymbol{w}) &= (\boldsymbol{z} \cdot (\lambda_1 \boldsymbol{y}_1 + \lambda_2 \boldsymbol{y}_2 + \cdots + \lambda_k \boldsymbol{y}_k)) \\ &= \lambda_1 (\boldsymbol{z} \cdot \boldsymbol{y}_1) + \lambda_2 (\boldsymbol{z} \cdot \boldsymbol{y}_2) + \cdots + \lambda_k (\boldsymbol{z} \cdot \boldsymbol{y}_k) \qquad (\because \text{内積の線形性}) \\ &= 0 \end{aligned}$$

である．よって $\boldsymbol{z}$ は W 上の全てのベクトルと直交する．

射影の求め方

- 先の補題から
$$\forall \boldsymbol{w} \in W,\ ((\boldsymbol{v}-\boldsymbol{p}) \cdot \boldsymbol{w}) = 0 \quad \Longleftrightarrow \quad ((\boldsymbol{v}-\boldsymbol{p}) \cdot \boldsymbol{y}_i) = 0\ (i = 1, \ldots, k)$$
である.
- $\boldsymbol{p} = \displaystyle\sum_{j=1}^{k} \lambda_j \boldsymbol{y}_j$ であることから
$$((\boldsymbol{v}-\boldsymbol{p}) \cdot \boldsymbol{y}_i) = \left(\left(\boldsymbol{v} - \sum_{j=1}^{k} \lambda_j \boldsymbol{y}_j\right) \cdot \boldsymbol{y}_i\right) = 0 \quad (i = 1, 2, \ldots, k) \tag{1}$$
を満たす $\lambda_1, \ldots, \lambda_k$ を求めれば射影 $\boldsymbol{p}$ が求まる.

- 内積の線形性から

$$\left(\left(\boldsymbol{v} - \sum_{j=1}^{k} \lambda_j \boldsymbol{y}_j\right) \cdot \boldsymbol{y}_i\right) = (\boldsymbol{y}_i \cdot \boldsymbol{v}) - \sum_{j=1}^{k} \lambda_j (\boldsymbol{y}_i \cdot \boldsymbol{y}_j) \quad (i = 1, 2, \ldots, k)$$

- これを式 (1) に代入すると,

$$\text{式 (1)} \quad \Longleftrightarrow \quad \begin{cases} \lambda_1(\boldsymbol{y}_1 \cdot \boldsymbol{y}_1) + \lambda_2(\boldsymbol{y}_1 \cdot \boldsymbol{y}_2) + \cdots + \lambda_k(\boldsymbol{y}_1 \cdot \boldsymbol{y}_k) = (\boldsymbol{y}_1 \cdot \boldsymbol{v}) \\ \lambda_1(\boldsymbol{y}_2 \cdot \boldsymbol{y}_1) + \lambda_2(\boldsymbol{y}_2 \cdot \boldsymbol{y}_2) + \cdots + \lambda_k(\boldsymbol{y}_2 \cdot \boldsymbol{y}_k) = (\boldsymbol{y}_2 \cdot \boldsymbol{v}) \\ \qquad \vdots \\ \lambda_1(\boldsymbol{y}_k \cdot \boldsymbol{y}_1) + \lambda_2(\boldsymbol{y}_k \cdot \boldsymbol{y}_2) + \cdots + \lambda_k(\boldsymbol{y}_k \cdot \boldsymbol{y}_k) = (\boldsymbol{y}_k \cdot \boldsymbol{v}) \end{cases}$$

$$\Longleftrightarrow \quad \begin{pmatrix} (\boldsymbol{y}_1 \cdot \boldsymbol{y}_1) & (\boldsymbol{y}_1 \cdot \boldsymbol{y}_2) & \cdots & (\boldsymbol{y}_1 \cdot \boldsymbol{y}_k) \\ (\boldsymbol{y}_2 \cdot \boldsymbol{y}_1) & (\boldsymbol{y}_2 \cdot \boldsymbol{y}_2) & \cdots & (\boldsymbol{y}_2 \cdot \boldsymbol{y}_k) \\ \vdots & \vdots & \ddots & \vdots \\ (\boldsymbol{y}_k \cdot \boldsymbol{y}_1) & (\boldsymbol{y}_k \cdot \boldsymbol{y}_2) & \cdots & (\boldsymbol{y}_k \cdot \boldsymbol{y}_k) \end{pmatrix} \begin{pmatrix} \lambda_1 \\ \lambda_2 \\ \vdots \\ \lambda_k \end{pmatrix} = \begin{pmatrix} (\boldsymbol{y}_1 \cdot \boldsymbol{v}) \\ (\boldsymbol{y}_2 \cdot \boldsymbol{v}) \\ \vdots \\ (\boldsymbol{y}_k \cdot \boldsymbol{v}) \end{pmatrix} \qquad (2)$$

$\rightsquigarrow$ $\lambda_1, \ldots, \lambda_k$ **に関する線形方程式系**

- 行列とベクトルの表記を導入して

$$K = \begin{pmatrix} (\boldsymbol{y}_1 \cdot \boldsymbol{y}_1) & (\boldsymbol{y}_1 \cdot \boldsymbol{y}_2) & \cdots & (\boldsymbol{y}_1 \cdot \boldsymbol{y}_k) \\ (\boldsymbol{y}_2 \cdot \boldsymbol{y}_1) & (\boldsymbol{y}_2 \cdot \boldsymbol{y}_2) & \cdots & (\boldsymbol{y}_2 \cdot \boldsymbol{y}_k) \\ \vdots & \vdots & \ddots & \vdots \\ (\boldsymbol{y}_k \cdot \boldsymbol{y}_1) & (\boldsymbol{y}_k \cdot \boldsymbol{y}_2) & \cdots & (\boldsymbol{y}_k \cdot \boldsymbol{y}_k) \end{pmatrix}, \quad \boldsymbol{\lambda} = \begin{pmatrix} \lambda_1 \\ \lambda_2 \\ \vdots \\ \lambda_k \end{pmatrix}, \quad \boldsymbol{b} = \begin{pmatrix} (\boldsymbol{y}_1 \cdot \boldsymbol{v}) \\ (\boldsymbol{y}_2 \cdot \boldsymbol{v}) \\ \vdots \\ (\boldsymbol{y}_k \cdot \boldsymbol{v}) \end{pmatrix}$$

  とする (行列 K を**グラム行列 (Gram matrix)** という).
- $K, \boldsymbol{\lambda}, \boldsymbol{b}$ を用いると先の線形方程式系 (2) は

$$K\boldsymbol{\lambda} = \boldsymbol{b} \tag{3}$$

  と書ける.

💡 **線形方程式系 (3) の解 $\boldsymbol{\lambda}$ が求まれば射影が求まる**

グラム行列の性質

定理 (グラム行列の正則性)

ベクトル $\boldsymbol{y}_1, \boldsymbol{y}_2, \ldots, \boldsymbol{y}_k$ が線形独立のとき，グラム行列 K は正則である．

証明　$\boldsymbol{c} = (c_1, c_2, \ldots, c_k)^T \in \mathbb{R}^k$ として，

$$K\boldsymbol{c} = \boldsymbol{0} \implies \boldsymbol{c} = \boldsymbol{0}$$

を示せばよい (∵ 正方行列が正則行列 $\Leftrightarrow$ その列ベクトルは線形独立)．

- ベクトル $K\boldsymbol{c}$ の各成分に注目すると，

$$K\boldsymbol{c} = \boldsymbol{0} \iff (\boldsymbol{y}_i \cdot \boldsymbol{y}_1)c_1 + (\boldsymbol{y}_i \cdot \boldsymbol{y}_2)c_2 + \cdots + (\boldsymbol{y}_i \cdot \boldsymbol{y}_k)c_k = 0 \quad (i = 1, 2, \ldots, k)$$
$$\iff \left(\boldsymbol{y}_i \cdot \sum_{j=1}^{k} c_j \boldsymbol{y}_j\right) = 0 \quad (i = 1, 2, \ldots, k)$$

  である．

- よって $K\boldsymbol{c} = \boldsymbol{0}$ のとき $\displaystyle\sum_{j=1}^{k} c_j \boldsymbol{y}_j$ は W の基底のベクトルすべてと直交．

- よって $\displaystyle\sum_{j=1}^{k} c_j \boldsymbol{y}_j$ は W 上のすべてのベクトルと直交. 自身も W 上のベクトルなので

$$\left(\sum_{j=1}^{k} c_j \boldsymbol{y}_j \cdot \sum_{j=1}^{k} c_j \boldsymbol{y}_j\right) = 0,$$

すなわち

$$\sum_{j=1}^{k} c_j \boldsymbol{y}_j = 0$$

が成り立つ ($\because$ 内積の定義).

- $\boldsymbol{y}_1, \boldsymbol{y}_2, \ldots, \boldsymbol{y}_k$ は線形独立であるから,

$$\sum_{j=1}^{k} c_j \boldsymbol{y}_j = 0 \Longrightarrow \boldsymbol{c} = \boldsymbol{0}$$

であり, 最終的に $K\boldsymbol{c} = \boldsymbol{0} \Rightarrow \boldsymbol{c} = \boldsymbol{0}$ を得て K が正則行列であることが示される.

射影の求め方

解きたい線形方程式系

$$K\boldsymbol{\lambda} = \boldsymbol{b}$$

- 先に示した命題から K は正則行列.
- よって上の線形方程式系は唯一の解をもち, その解は次のように書ける:

$$\boldsymbol{\lambda} = K^{-1}\boldsymbol{b}.$$

- 線形方程式系の解 $\boldsymbol{\lambda} = (\lambda_1, \lambda_2, \ldots, \lambda_k)^T$ を用いて,

$$\boldsymbol{p} = \lambda_1\boldsymbol{y}_1 + \lambda_2\boldsymbol{y}_2 + \cdots + \lambda_k\boldsymbol{y}_k$$

とすると, $\boldsymbol{p}$ は $\boldsymbol{v}$ の部分空間 W への射影となる.

$v \in V$ の部分空間 W への射影 p を求める手順

(1) $K = \begin{pmatrix} (\boldsymbol{y}_1 \cdot \boldsymbol{y}_1) & (\boldsymbol{y}_1 \cdot \boldsymbol{y}_2) & \cdots & (\boldsymbol{y}_1 \cdot \boldsymbol{y}_k) \\ (\boldsymbol{y}_2 \cdot \boldsymbol{y}_1) & (\boldsymbol{y}_2 \cdot \boldsymbol{y}_2) & \cdots & (\boldsymbol{y}_2 \cdot \boldsymbol{y}_k) \\ \vdots & \vdots & \ddots & \vdots \\ (\boldsymbol{y}_k \cdot \boldsymbol{y}_1) & (\boldsymbol{y}_k \cdot \boldsymbol{y}_2) & \cdots & (\boldsymbol{y}_k \cdot \boldsymbol{y}_k) \end{pmatrix}$ を計算する.

(2) $\boldsymbol{b} = \begin{pmatrix} (\boldsymbol{y}_1 \cdot \boldsymbol{v}) \\ (\boldsymbol{y}_2 \cdot \boldsymbol{v}) \\ \vdots \\ (\boldsymbol{y}_k \cdot \boldsymbol{v}) \end{pmatrix}$ を計算する.

(3) $\boldsymbol{\lambda} = K^{-1}\boldsymbol{b}$ とする.

(4) $\boldsymbol{p} = \sum_{j=1}^{k} \lambda_j \boldsymbol{y}_j$ とする.

正規直交基底を用いた射影の求め方

基底が正規直交基底の場合，射影は簡単に求まる

定義 (正規直交基底)

部分空間 W の基底 $\boldsymbol{y}_1, \boldsymbol{y}_2, \ldots, \boldsymbol{y}_k$ の各ベクトルのノルムが 1 で互いに直交するとき，**正規直交基底 (orthonormal basis)** という．

正規直交基底の特徴　正規直交基底のグラム行列は単位行列である:

$$K = \begin{pmatrix} (\boldsymbol{y}_1 \cdot \boldsymbol{y}_1) & (\boldsymbol{y}_1 \cdot \boldsymbol{y}_2) & \cdots & (\boldsymbol{y}_1 \cdot \boldsymbol{y}_k) \\ (\boldsymbol{y}_2 \cdot \boldsymbol{y}_1) & (\boldsymbol{y}_2 \cdot \boldsymbol{y}_2) & \cdots & (\boldsymbol{y}_2 \cdot \boldsymbol{y}_k) \\ \vdots & \vdots & \ddots & \vdots \\ (\boldsymbol{y}_k \cdot \boldsymbol{y}_1) & (\boldsymbol{y}_k \cdot \boldsymbol{y}_2) & \cdots & (\boldsymbol{y}_k \cdot \boldsymbol{y}_k) \end{pmatrix} = \begin{pmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 \end{pmatrix} = I$$

$\rightsquigarrow$ **射影の計算が簡単になる**

$\boldsymbol{v} \in V$ の部分空間 W への射影 $\boldsymbol{p}$ を求める手順 (正規直交基底の場合)

(1) $K = \begin{pmatrix} (\boldsymbol{y}_1 \cdot \boldsymbol{y}_1) & (\boldsymbol{y}_1 \cdot \boldsymbol{y}_2) & \cdots & (\boldsymbol{y}_1 \cdot \boldsymbol{y}_k) \\ (\boldsymbol{y}_2 \cdot \boldsymbol{y}_1) & (\boldsymbol{y}_2 \cdot \boldsymbol{y}_2) & \cdots & (\boldsymbol{y}_2 \cdot \boldsymbol{y}_k) \\ \vdots & \vdots & \ddots & \vdots \\ (\boldsymbol{y}_k \cdot \boldsymbol{y}_1) & (\boldsymbol{y}_k \cdot \boldsymbol{y}_2) & \cdots & (\boldsymbol{y}_k \cdot \boldsymbol{y}_k) \end{pmatrix}$ を計算する．**正規直交基底の場合，$K = I$**

(2) $\boldsymbol{b} = \begin{pmatrix} (\boldsymbol{y}_1 \cdot \boldsymbol{v}) \\ (\boldsymbol{y}_2 \cdot \boldsymbol{v}) \\ \vdots \\ (\boldsymbol{y}_k \cdot \boldsymbol{v}) \end{pmatrix}$ を計算する．

(3) $\boldsymbol{\lambda} = K^{-1}\boldsymbol{b}$ とする．　**正規直交基底の場合，$\boldsymbol{\lambda} = \boldsymbol{b}$**

(4) $\boldsymbol{p} = \sum_{j=1}^{k} \lambda_j \boldsymbol{y}_j$ とする．　**正規直交基底の場合，$\boldsymbol{p} = \sum_{j=1}^{k} (\boldsymbol{y}_j \cdot \boldsymbol{v})\boldsymbol{y}_j$**

正規直交基底を用いると内積·ノルムの計算も簡単

- $\boldsymbol{y}_1, \boldsymbol{y}_2, \dots, \boldsymbol{y}_k \in W$ を部分空間 W の正規直交基底とする.
- 2 つのベクトル $\boldsymbol{u}, \boldsymbol{v} \in W$ が

$$\boldsymbol{u} = \sum_{i=1}^{k} a_i \boldsymbol{y}_i, \quad \boldsymbol{v} = \sum_{i=1}^{k} b_i \boldsymbol{y}_i$$

と表されるとする.

- このとき

$$\boxed{\text{内積}} \quad (\boldsymbol{u} \cdot \boldsymbol{v}) = \left(\sum_{i=1}^{k} a_i \boldsymbol{y}_i \cdot \sum_{i=1}^{k} b_i \boldsymbol{y}_i \right) = \sum_{i=1}^{k} \sum_{j=1}^{k} a_i b_j (\boldsymbol{y}_i \cdot \boldsymbol{y}_j) = \sum_{i=1}^{k} a_i b_i$$

$$\boxed{\text{ノルム}} \quad \|\boldsymbol{u}\| = \sqrt{(\boldsymbol{u} \cdot \boldsymbol{u})} = \sqrt{\sum_{i=1}^{k} {a_i}^2}.$$

$\rightsquigarrow$ **係数だけで内積とノルムが計算できる.**

目次

実ベクトル空間 $\mathbb{R}^m$ での射影

- $\mathbb{R}^m$ の部分空間 W の基底を $\boldsymbol{y}_1, \boldsymbol{y}_2, \ldots, \boldsymbol{y}_k \in \mathbb{R}^m$ とする.
- $\boldsymbol{v} \in \mathbb{R}^m$ の W への射影 $\boldsymbol{p} \in W$ を求めたい. (内積は標準内積 $(\boldsymbol{x} \cdot \boldsymbol{y}) = \boldsymbol{x}^T \boldsymbol{y}$ を用いる)

- 基底のベクトルを並べた行列を $A = (\boldsymbol{y}_1, \boldsymbol{y}_2, \ldots, \boldsymbol{y}_k) \in \mathbb{R}^{m \times k}$ とすると,

$$K = \begin{pmatrix} \boldsymbol{y}_1^T \boldsymbol{y}_1 & \boldsymbol{y}_1^T \boldsymbol{y}_2 & \cdots & \boldsymbol{y}_1^T \boldsymbol{y}_k \\ \boldsymbol{y}_2^T \boldsymbol{y}_1 & \boldsymbol{y}_2^T \boldsymbol{y}_2 & \cdots & \boldsymbol{y}_2^T \boldsymbol{y}_k \\ \vdots & \vdots & \ddots & \vdots \\ \boldsymbol{y}_k^T \boldsymbol{y}_1 & \boldsymbol{y}_k^T \boldsymbol{y}_2 & \cdots & \boldsymbol{y}_k^T \boldsymbol{y}_k \end{pmatrix} = A^T A \in \mathbb{R}^{k \times k}, \ \boldsymbol{b} = \begin{pmatrix} \boldsymbol{y}_1^T \boldsymbol{v} \\ \boldsymbol{y}_2^T \boldsymbol{v} \\ \vdots \\ \boldsymbol{y}_k^T \boldsymbol{v} \end{pmatrix} = A^T \boldsymbol{v} \in \mathbb{R}^k.$$

- よって射影を求めるために解く線形方程式系 $K\boldsymbol{\lambda} = \boldsymbol{b}$ は次のように書ける:

$$A^T A \boldsymbol{\lambda} = A^T \boldsymbol{v}. \tag{4}$$

- 式 (4) は**正規方程式 (normal equation)** とよばれる.

実ベクトル空間 $\mathbb{R}^m$ での射影 (続き)

- $\boldsymbol{y}_1, \boldsymbol{y}_2, \ldots, \boldsymbol{y}_k$ は線形独立なので, A^TA は正則行列である.
- 正規方程式 (4) の解は唯一であり, その解は次のように書ける:

$$\boldsymbol{\lambda} = (A^TA)^{-1}A^T\boldsymbol{v}.$$

- よって, 射影 $\boldsymbol{p}$ は

$$\boldsymbol{p} = \sum_{j=1}^{k} \lambda_j \boldsymbol{y}_j = A\boldsymbol{\lambda} = \underbrace{A(A^TA)^{-1}A^T}_{\text{射影行列という}}\boldsymbol{v}.$$

  と求まる.
- 特に $\boldsymbol{y}_1, \boldsymbol{y}_2, \ldots, \boldsymbol{y}_k$ が正規直交基底の場合, $A^TA = I$ なので

$$\boldsymbol{p} = AA^T\boldsymbol{v}.$$

  となる.

最小 2 乗法

$\boxed{\text{記法}}$ 行列 A の列空間 (列ベクトルが張る部分空間) を $R(A)$ と書く.

- $A \in \mathbb{R}^{m\times n}$, $\boldsymbol{b} \in \mathbb{R}^m$ を所与の定数とした次の線形方程式系を解くことを考える:

$$A\boldsymbol{x} = \boldsymbol{b}.$$

- $A\boldsymbol{x} \in R(A)$ より, $\boldsymbol{b} \notin R(A)$ のときこの解は存在しない.

$\rightsquigarrow$ **$\boldsymbol{b} \notin R(A)$ の場合, $\|A\boldsymbol{x} - \boldsymbol{b}\|^2$ が小さくなるような $\boldsymbol{x}$ を求めることを考える**

- これは $\boldsymbol{b}$ の $R(A)$ への射影を求めることにほかならならず,

$$\hat{\boldsymbol{x}} = (A^T A)^{-1} A^T \boldsymbol{b}$$

とすると, $A\hat{\boldsymbol{x}}$ は $\boldsymbol{b}$ の $R(A)$ への射影である.

- $\hat{\boldsymbol{x}}$ を**最小 2 乗解**といい, 最小 2 乗解を求めることを**最小 2 乗法**という.

最小 2 乗法の例

- アイスの売上に関して, 次のデータがある:

気温 (°C)	アイスの売上 (個)
10	20
15	30
20	50
25	80
30	100

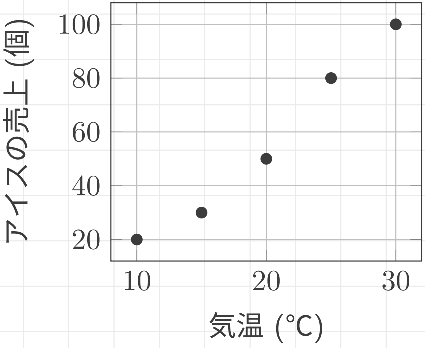

Figure: 売上と気温の散布図

- 気温からアイスの売上を予測するため,
  アイスの売上 b (個) と気温 a (°C) には

$$b = x_1 + x_2 a$$

という関係にあると仮定する.

データから x_1, x_2 を求めるにはどうすればよいか?

最小 2 乗法の例 (続き)

- x_1, x_2 を求めるために,
$$A = \begin{pmatrix} 1 & 10 \\ 1 & 15 \\ 1 & 20 \\ 1 & 25 \\ 1 & 30 \end{pmatrix}, \ \boldsymbol{x} = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}, \ \boldsymbol{b} = \begin{pmatrix} 20 \\ 30 \\ 50 \\ 80 \\ 100 \end{pmatrix}$$
とした線形方程式系 $A\boldsymbol{x} = \boldsymbol{b}$ を考える. $\rightsquigarrow$ **これを満たす解 $\boldsymbol{x}$ は存在しない.**

- そこで最小 2 乗解 $\hat{\boldsymbol{x}}$ を求めることを考えると,
$$\hat{\boldsymbol{x}} = (A^T A)^{-1} A^T \boldsymbol{b} = \begin{pmatrix} -28 \\ 4.2 \end{pmatrix}$$
と求まる (計算過程は省略).
- 気温を用いたアイスの売上の予測式は $\boxed{\hat{b} = -28 + 4.2a}$ と求まる.

最小 2 乗法の例 (続き)

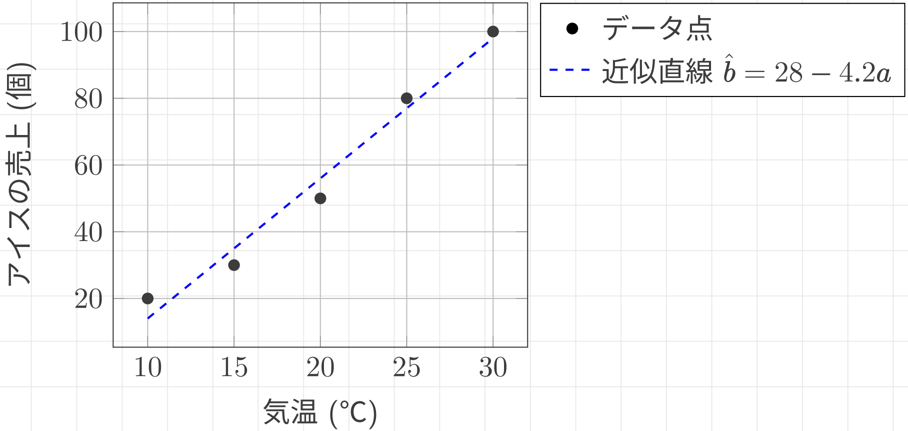

Figure: 売上と気温の散布図と最小 2 乗法による直線近似

最小 2 乗法は機械学習の第一歩

まとめ

講義の振り返り

- 射影の定義と意味づけ
- 射影の計算方法
- 射影の計算としての最小 2 乗法

自宅での復習

- 線形独立･正則行列に関する定義や性質を確認する
- 射影の計算手順と正規直交基底を用いた計算の簡略化を確認する
- 最小 2 乗法における正規方程式の作り方を確認する