

Rを用いた主成分分析

宮田 庸一*

平成 22 年 11 月 11 日

*レポート, 論文作成, 授業使用以外での無断引用を禁じる. またレポート, 論文には引用の記載をお願いします.

目次

1	方程式を解く	4
1.1	因数分解を用いる方法	4
1.2	解の計算公式を用いる方法	4
1.3	因数定理を用いる方法	4
2	行列 (Matrix)	4
2.1	行列の和, 差	5
2.2	行列の実数倍	5
2.3	行列の積	5
2.4	単位行列	6
2.5	転置行列	7
3	行列式 (Determination)	7
3.1	高次の行列式	8
3.2	行列を用いた行列式の表し方	8
4	ベクトル	9
4.1	ベクトルの実数倍	10
4.2	ベクトルの成分表示	10
4.3	方向ベクトルを持つ直線	11
5	固有値 (Eigenvalue)	12
5.1	3 次の正方行列に対する固有値, 固有ベクトル	13
6	偏導関数	14
6.1	2 変数関数	14
6.2	偏微分	15
6.3	偏導関数	16
6.4	極大値, 極小値	16
7	ラグランジュの未定乗数法	17
8	標本平均 (Sample mean, Average)	18
9	標本分散 (Sample Variance)	19
9.1	標本分散の意味	19
9.2	標本分散の性質 ... 単位が変われば分散の値も変わる!	21
10	標準化	21
10.1	なぜ標準偏差 s で割る必要があるのですか?	22
10.2	標準化のいいところは何ですか?	23
11	2 つの変数の関係を調べる.	23
11.1	散布図	23
11.2	(標本) 共分散	24
11.3	共分散はなぜ正か負で 2 つの変数の関係を表せるの?	25
11.4	共分散の弱点	26
11.5	相関係数	27

12 主成分分析	29
12.1 分散共分散行列	29
12.2 主成分分析	30
12.3 主成分得点 (主成分スコア)	33
12.4 主成分の解釈	34
12.5 寄与率	34
13 3変数以上の主成分分析	35
13.1 Biplot(バイプロット)	37
13.2 Biplot の解釈	38
14 R を用いた統計処理	41
14.1 データ (csv ファイル) を準備する	41
14.2 データのインポート	42
14.3 主成分分析を行う	42
14.4 R での biplot の目盛りについて	44
14.5 R での注意点	44
14.6 解析結果のエクスポート	44
14.7 princomp がうまくいかない時	44
15 解答	45

1 方程式を解く

1.1 因数分解を用いる方法

【例題 1.1】 $2x^2 - 3x + 1 = 0$ の実数解を求めよ.

解: $(2x - 1)(x - 1) = 0$ より, $x = \frac{1}{2}$ または 1 .

1.2 解の計算公式を用いる方法

【例題 1.2】 2 次方程式 $ax^2 + bx + c = 0$ に対して, その解は $x = \frac{-b \pm \sqrt{b^2 - 4ac}}{2a}$ で与えられる.

問 1.1 $2x^2 - \sqrt{27}x + 3 = 0$ の解を求めよ.

1.3 因数定理を用いる方法

因数定理

多項式 $f(x)$ が $x - a$ で割り切れる $\iff f(a) = 0$

【例題 1.3】 $-6 + 11x - 6x^2 + x^3 = 0$ の実数解を求めよ.

解 $f(x) = x^3 - 6x^2 + 11x - 6$ とおき, ために $x = 1$ を代入すると $f(1) = 1 - 6 + 11 - 6 = 0$ となることより, 因数定理より $f(x)$ は $x - 1$ で割り切れることがわかる. よって $f(x) = (x - 1)(x^2 - 5x + 6)$. 右側の部分は普通に因数分解できるので $f(x) = (x - 1)(x - 2)(x - 3)$ となる. よって求める実数解は $x = 1, 2, 3$.

$$\begin{array}{r} x^2 - 5x + 6 \\ x - 1 \overline{) x^3 - 6x^2 + 11x - 6} \\ \underline{x^3 - x^2} \\ -5x^2 + 11x - 6 \\ \underline{-5x^2 + 5x} \\ 6x - 6 \\ \underline{6x - 6} \\ 0 \end{array}$$

ちなみに多項式の割り算は, 右の図のように数の割り算と同じやり方で行います.

問 1.2 ① $x^3 - 7x + 6 = 0$ ② $2x^3 + 5x^2 + x - 2 = 0$

2 行列 (Matrix)

以下のように数や文字を並べた配列を考えてみます.

$$\begin{pmatrix} -1 \\ 2 \end{pmatrix}, \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}, \begin{pmatrix} 5 & 6 & 7 \\ 0 & 2 & 1 \end{pmatrix}, \begin{pmatrix} a & b \\ c & d \end{pmatrix}.$$

このように数や文字を並べて丸括弧で囲んだものを行列と言います. 括弧の中のそれぞれの数のことを成分 もしくは要素と言います. 行列における横の並びを行, 縦の並びを列と言います.

$$\begin{array}{ccc} & \text{1 列目} & \text{2 列目} & \text{3 列目} \\ \text{1 行目} & \left(\begin{array}{ccc} 5 & 6 & 7 \end{array} \right. \\ \text{2 行目} & \left. \begin{array}{ccc} 0 & 2 & 1 \end{array} \right) \end{array}$$

行の数が m 個あり, 列の数が n 個ある行列を $m \times n$ 行列と言います. 特に行の数と列の数が同じ行列を正方行列と言います.

【例 2.1】 $\begin{pmatrix} -1 \\ 2 \end{pmatrix}$ は 2×1 行列. $\begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}$ は 2×2 の正方行列.

行列は A, B のように大文字, もしくは a のような太文字を用いて表し, その成分は a_{12} のように小文字を使って表わします. また第 i 行と第 j 列が交差する要素のことを (i, j) 成分と言います.

【例 2.2】 $A = \begin{pmatrix} 5 & 6 & 7 \\ 0 & 2 & 1 \end{pmatrix}$ における $(2, 1)$ 成分は 0.

問 2.1 例 2.2 の行列 A の $(1, 3)$ 成分を求めよ.

行列 A, B が同じ型を持ち, かつそれぞれの成分が等しいとき, A と B は等しいといい $A = B$ と書く.

【例 2.3】 $A = \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}, B = \begin{pmatrix} 1 & 2 \\ 3 & 5 \end{pmatrix}, C = \begin{pmatrix} 1 & 2 \\ 3 & 4 \\ 5 & 6 \end{pmatrix}, D = \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}$ とする. このとき $A \neq B, A \neq C$ (行列の型が違う), $A = D$ である.

2.1 行列の和, 差

同じ型を持つ行列 A, B のそれぞれの成分を足し合わせたものを, A と B の和と言いい, $A + B$ と表す. 例えば

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} + \begin{pmatrix} e & f \\ g & h \end{pmatrix} = \begin{pmatrix} a+e & b+f \\ c+g & d+h \end{pmatrix} \quad (2.1)$$

また同様に $A - B$ も定義する. 例えば

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} - \begin{pmatrix} e & f \\ g & h \end{pmatrix} = \begin{pmatrix} a-e & b-f \\ c-g & d-h \end{pmatrix} \quad (2.2)$$

また $\begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$ のように, 成分のすべてが 0 である行列を零行列といい, O を用いて表す.

【例 2.4】

$$\begin{pmatrix} -3 & 6 \\ 2 & 5 \end{pmatrix} + \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix} = \begin{pmatrix} -3+1 & 6+2 \\ 2+3 & 5+4 \end{pmatrix} = \begin{pmatrix} -2 & 8 \\ 5 & 9 \end{pmatrix}$$

2.2 行列の実数倍

実数 k に対して, 行列 A の各成分を k 倍した成分を持つ行列を kA と書く. 例えば

$$k \begin{pmatrix} a & b \\ c & d \end{pmatrix} = \begin{pmatrix} ka & kb \\ kc & kd \end{pmatrix} \quad (2.3)$$

【例 2.5】

$$5 \begin{pmatrix} -3 & 6 \\ 2 & 5 \end{pmatrix} = \begin{pmatrix} 5 \cdot (-3) & 5 \cdot 6 \\ 5 \cdot 2 & 5 \cdot 5 \end{pmatrix} = \begin{pmatrix} -15 & 30 \\ 10 & 25 \end{pmatrix}$$

問 2.2 $A = \begin{pmatrix} 5 & 3 \\ 6 & 4 \end{pmatrix}, B = \begin{pmatrix} 0 & 2 \\ 1 & 7 \end{pmatrix}$ とする. この時, 以下を計算せよ.

- ① $A + B$ ② $A - B$ ③ $A - 2B$

2.3 行列の積

次の 2 つの行列 A, B を考える.

$$A = \begin{pmatrix} a_1 & a_2 \\ c_1 & c_2 \end{pmatrix}, \quad B = \begin{pmatrix} b_1 & d_1 \\ b_2 & d_2 \end{pmatrix}.$$

このとき A と B の積 (かけ算) AB は以下のように定義します.

$$AB = \begin{pmatrix} a_1 & a_2 \\ c_1 & c_2 \end{pmatrix} \begin{pmatrix} b_1 & d_1 \\ b_2 & d_2 \end{pmatrix} = \begin{pmatrix} a_1b_1 + a_2b_2 & a_1d_1 + a_2d_2 \\ c_1b_1 + c_2b_2 & c_1d_1 + c_2d_2 \end{pmatrix}$$

上の定義において例えば, AB の $(2,1)$ 成分であれば, 左の行列 A の 2 行目と右の行列 B の 1 列目を上のように掛け合わせていることに注意してください

【例 2.6】 $A = \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}, B = \begin{pmatrix} -2 & 0 \\ 1 & 5 \end{pmatrix}$ とする.

$$\begin{aligned} AB &= \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix} \begin{pmatrix} -2 & 0 \\ 1 & 5 \end{pmatrix} = \begin{pmatrix} 1 \cdot (-2) + 2 \cdot 1 & 1 \cdot 0 + 2 \cdot 5 \\ 3 \cdot (-2) + 4 \cdot 1 & 3 \cdot 0 + 4 \cdot 5 \end{pmatrix} = \begin{pmatrix} 0 & 10 \\ -2 & 20 \end{pmatrix} \\ BA &= \begin{pmatrix} -2 & 0 \\ 1 & 5 \end{pmatrix} \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix} = \begin{pmatrix} -2 & -4 \\ 16 & 22 \end{pmatrix} \\ \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix} \begin{pmatrix} 5 \\ 6 \end{pmatrix} &= \begin{pmatrix} 1 \cdot 5 + 2 \cdot 6 \\ 3 \cdot 5 + 4 \cdot 6 \end{pmatrix} = \begin{pmatrix} 17 \\ 39 \end{pmatrix} \quad \text{よって} \quad 3 \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix} \begin{pmatrix} 5 \\ 6 \end{pmatrix} = \begin{pmatrix} 51 \\ 117 \end{pmatrix} \\ \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix} 3 \begin{pmatrix} 5 \\ 6 \end{pmatrix} &= \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix} \begin{pmatrix} 15 \\ 18 \end{pmatrix} = \begin{pmatrix} 1 \cdot 15 + 2 \cdot 18 \\ 3 \cdot 15 + 4 \cdot 18 \end{pmatrix} = \begin{pmatrix} 51 \\ 117 \end{pmatrix} \end{aligned}$$

行列の乗算における性質

- 行列の乗算において一般的には $AB \neq BA$ であることに注意 (上の例 2.6 を見よ).
- $A(B + C) = AB + AC$
- $(AB)C = A(BC)$
- 実数 k に対して $kAB = A(kB)$. (例 2.6 参照)

つまり行列の乗算は順番を交換できないこと以外は、普通の文字式 (例えば $x(y + z) = xy + xz, 3xy = x(3y)$) と同じように計算してよいということを意味している.

また 3×3 行列同士の乗算についても同様に定義します.

$$\begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix} \begin{pmatrix} b_{11} & b_{12} & b_{13} \\ b_{21} & b_{22} & b_{23} \\ b_{31} & b_{32} & b_{33} \end{pmatrix} = \begin{pmatrix} a_{11}b_{11} + a_{12}b_{21} + a_{13}b_{31} & a_{11}b_{12} + a_{12}b_{22} + a_{13}b_{32} & a_{11}b_{13} + a_{12}b_{23} + a_{13}b_{33} \\ a_{21}b_{11} + a_{22}b_{21} + a_{23}b_{31} & a_{21}b_{12} + a_{22}b_{22} + a_{23}b_{32} & a_{21}b_{13} + a_{22}b_{23} + a_{23}b_{33} \\ a_{31}b_{11} + a_{32}b_{21} + a_{33}b_{31} & a_{31}b_{12} + a_{32}b_{22} + a_{33}b_{32} & a_{31}b_{13} + a_{32}b_{23} + a_{33}b_{33} \end{pmatrix}$$

問 2.3 次を計算せよ.

$$\textcircled{1} \begin{pmatrix} 5 & -1 \\ 2 & 3 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 3 & -1 \end{pmatrix} \quad \textcircled{2} \begin{pmatrix} 0 & -2 & 1 \\ -1 & 1 & 1 \\ 2 & 5 & -5 \end{pmatrix} \begin{pmatrix} 10 & 5 & 3 \\ 3 & 2 & 1 \\ 7 & 4 & 2 \end{pmatrix} \quad \textcircled{3} \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix} \begin{pmatrix} 3 & 4 \end{pmatrix} \quad \textcircled{4} \begin{pmatrix} \sqrt{2} & 1 \end{pmatrix} \begin{pmatrix} 3 \\ -1 \end{pmatrix}$$

2.4 単位行列

左上から右下に向かった成分は全て 1 で, それ以外は 0 である行列を単位行列と言い, I で表す. ここで I は 2×2 の単位行列であれば

$$I = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \tag{2.4}$$

3×3 の単位行列であれば

$$I = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \tag{2.5}$$

となる. 任意の正方行列 A に対して, 同じ型の単位行列を I とすると $AI = IA = A$ が成り立つ. つまり単位行列は, 普通の数でいうところの 1 に相当する. ($5 \times 1 = 5$ のように 1 をかけても値は変わらない)

問 2.4 $A = \begin{pmatrix} 5 & -1 \\ 2 & 3 \end{pmatrix}$, $I = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$, $W = \begin{pmatrix} 5 \\ -2 \end{pmatrix}$ とする. この時, 以下を計算せよ.

- ① AI ② IW

2.5 転置行列

行列 A の行と列を入れ替えて得られる行列を A の転置行列と言い A^T と表す. 例えば

$$A = \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{pmatrix}, \quad A^T = \begin{pmatrix} 1 & 4 \\ 2 & 5 \\ 3 & 6 \end{pmatrix} \quad (2.6)$$

要は 1 行目の要素が 1 列目の要素となり, 2 行目の要素が 2 列目の要素となるということ.

性質

- 行列 A, B に対して $(AB)^T = B^T A^T$ が成り立つ.

- $n \times 1$ 行列 $\mathbf{a} = \begin{pmatrix} a_1 \\ \vdots \\ a_n \end{pmatrix}$, $\mathbf{b} = \begin{pmatrix} b_1 \\ \vdots \\ b_n \end{pmatrix}$ において, $\mathbf{a}^T \mathbf{b} = \mathbf{b}^T \mathbf{a}$.

証明: 例えば $A = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}$, $B = \begin{pmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{pmatrix}$ とおいて計算すれば等しいことがわかる.

問 2.5 $\mathbf{a} = \begin{pmatrix} 1 \\ 2 \end{pmatrix}$, $\mathbf{b} = \begin{pmatrix} 3 \\ 4 \end{pmatrix}$, $A = \begin{pmatrix} 1 & -2 \\ 0 & 5 \end{pmatrix}$ とする. この時, 以下を計算せよ.

- ① $\mathbf{a}^T \mathbf{b}$ ② $\mathbf{b}^T \mathbf{a}$ ③ $\mathbf{a}^T A \mathbf{b}$

3 行列式 (Determination)

今, 配列 $\begin{vmatrix} a & b \\ c & d \end{vmatrix}$ を考えます. これは行列式と呼ばれるもので, 以下のように計算します.

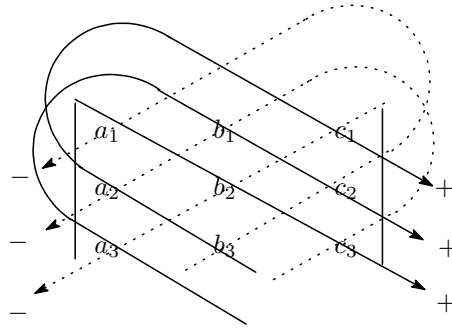
$$\begin{vmatrix} a & b \\ c & d \end{vmatrix} = ad - bc \quad (3.1)$$

これは行が 2 つで列が 2 つの行列式なので 2 次の行列式と言います. ここで括弧が丸ではなく, $|$ となっています. 上の計算方法からも分かるように行列と行列式はまったく違うものなので注意してください. さて話がそれましたが, 3 次の行列式は以下のように定義します.

$$\begin{vmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{vmatrix} = a_1 b_2 c_3 + a_2 b_3 c_1 + a_3 b_1 c_2 - a_1 b_3 c_2 - a_2 b_1 c_3 - a_3 b_2 c_1 \quad (3.2)$$

これは覚えづらいですが, サラスの方法と呼ばれる方法を用いると比較的楽に覚えることができます.

サラスの方法



実線の矢印のものは全て足して、点線の矢印のものは全てにマイナスをかける。

【例 3.1】 $\begin{vmatrix} -2 & 1 \\ 3 & 4 \end{vmatrix} = (-2) \cdot 4 - 1 \cdot 3 = -11.$

$$\begin{vmatrix} 6 & 2 & -5 \\ -2 & 1 & 0 \\ 0 & 4 & -3 \end{vmatrix} = 6 \cdot 1 \cdot (-3) + 4 \cdot (-2) \cdot (-5) + 0 \cdot 2 \cdot 0 - (-5) \cdot 1 \cdot 0 - 4 \cdot 0 \cdot 6 - (-3) \cdot 2 \cdot (-2) = 10$$

問 3.1 以下の行列式を計算せよ。

① $\begin{vmatrix} 6 & 2 & -5 \\ 0 & 4 & -3 \\ -2 & 1 & 0 \end{vmatrix}$ ② $\begin{vmatrix} 1 & 2 \\ 3 & 4 \end{vmatrix}$ ③ $\begin{vmatrix} 6 & 2 & -5 \\ 1 & 2 & 3 \\ 1 & 2 & 3 \end{vmatrix}$

3.1 高次の行列式

4 次の行列式を考えます。

$$A = \begin{vmatrix} a_1 & b_1 & c_1 & d_1 \\ a_2 & b_2 & c_2 & d_2 \\ a_3 & b_3 & c_3 & d_3 \\ a_4 & b_4 & c_4 & d_4 \end{vmatrix} \quad (3.3)$$

このとき、4 次の行列式は以下のように計算します。

$$A = a_1 \begin{vmatrix} b_2 & c_2 & d_2 \\ b_3 & c_3 & d_3 \\ b_4 & c_4 & d_4 \end{vmatrix} - b_1 \begin{vmatrix} a_2 & c_2 & d_2 \\ a_3 & c_3 & d_3 \\ a_4 & c_4 & d_4 \end{vmatrix} + c_1 \begin{vmatrix} a_2 & b_2 & d_2 \\ a_3 & b_3 & d_3 \\ a_4 & b_4 & d_4 \end{vmatrix} - d_1 \begin{vmatrix} a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \\ a_4 & b_4 & c_4 \end{vmatrix} \quad (3.4)$$

3.2 行列を用いた行列式の表し方

正方行列 $A = (a_{ij})$ の成分の配列と同じ配列を持つ行列式を行列 A の行列式と言い、記号で

$$|A| \text{ または } \det(A)$$

と表す。

【例 3.2】 $A = \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}$ とすると $|A| = \begin{vmatrix} 1 & 2 \\ 3 & 4 \end{vmatrix} = -2$ と計算します。

定理 3.1

A, B を正方行列とする。この時、 $|AB| = |A||B|$ が成り立つ

証明は省略するが、具体的な例をあげる。 $A = \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}$, $B = \begin{pmatrix} -2 & 0 \\ 1 & 5 \end{pmatrix}$ とする。ここで $AB = \begin{pmatrix} 0 & 10 \\ -2 & 20 \end{pmatrix}$ より $|AB| = 20$ となる。一方で $|A| = -2$, $|B| = -10$ より $|A||B| = (-2)(-10) = 20$ となり、 $|AB|$ と等しくなることがわかる。

定理 3.2

A を $n \times n$ 正方行列, かつ \mathbf{w} を $n \times 1$ 行列とする. この時, $A\mathbf{w} = \mathbf{0}$ かつ $\mathbf{w} \neq \mathbf{0}$ であるとき $|A| = 0$ が成り立つ.

証明 $n = 2$ の場合のみ示す. $A = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}$ と $\mathbf{w} = \begin{pmatrix} x \\ y \end{pmatrix}$ とする. ここで

$$\begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \iff \begin{cases} a_{11}x + a_{12}y = 0 & \cdots \textcircled{1} \\ a_{21}x + a_{22}y = 0 & \cdots \textcircled{2} \end{cases} \quad (3.5)$$

$\textcircled{1} \times a_{22}$, $\textcircled{2} \times a_{12}$ とすると

$$\begin{cases} a_{11}a_{22}x + a_{12}a_{22}y = 0 \\ a_{21}a_{12}x + a_{12}a_{22}y = 0 \end{cases} \quad \text{より } (a_{11}a_{22} - a_{21}a_{12})x = 0. \text{ これより } |A|x = 0 \text{ と書ける.}$$

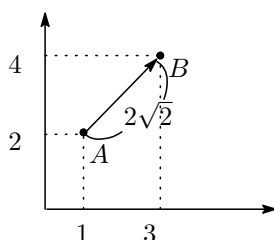
一方で $\textcircled{1} \times a_{21}$, $\textcircled{2} \times a_{11}$ とすると

$$\begin{cases} a_{11}a_{21}x + a_{12}a_{21}y = 0 \\ a_{11}a_{21}x + a_{11}a_{22}y = 0 \end{cases} \quad \text{より } (a_{11}a_{22} - a_{21}a_{12})y = 0. \text{ よって } |A|y = 0 \text{ と書ける. ここで } \mathbf{w} \neq \mathbf{0} \text{ から}$$

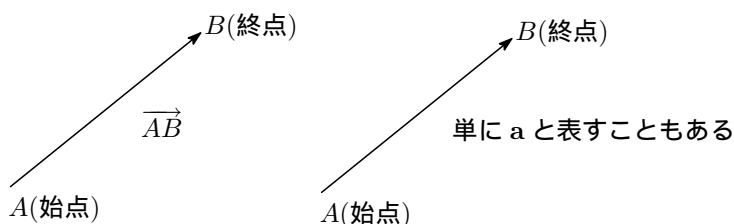
$x \neq 0$ もしくは $y \neq 0$ が成り立つ. よって $|A| = 0$ となる.

4 ベクトル

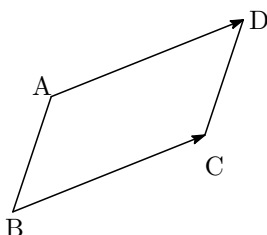
身長や体重は何 cm, 何 kg のように数値 (大きさ) で表すことができるが, 風向きは西の風秒速 10m のように向きと大きさで表わされる. また座標平面上においても点 $A(1, 2)$ から点 $B(3, 4)$ に移動した場合, 向きと距離によって表すことができる.



このように向きと大きさにより定まる線分¹をベクトルと言う. 以下の図のようにベクトルは, 始点と呼ばれる出発点と, 終点と呼ばれる到着点がある. ベクトルはこの始点 A と終点 B を用いて \overrightarrow{AB} と表すことにする. またしばしばベクトルは小文字を用いて \vec{a} と表したり, 太文字を用いて \mathbf{a} と表すこともある. またベクトルの長さのことをベクトルの大きさと言い, $|\overrightarrow{AB}|$, $|\vec{a}|$, $|\mathbf{a}|$ のように表す.



ただしベクトルはどこから出発するか (つまり, 始点がどこであるか) は問題にしないことに注意が必要である. 例えば, 以下の平行四辺形において $\overrightarrow{AD} = \overrightarrow{BC}$ である.



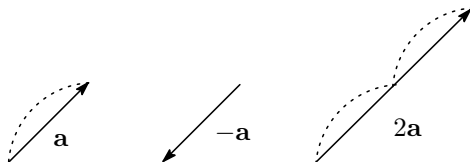
また大きさが 0 のベクトルを零ベクトルと言い, $\mathbf{0}$ と表す.

¹厳密には有向線分という

4.1 ベクトルの実数倍

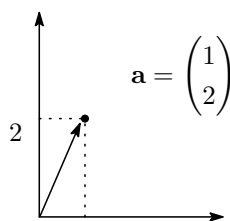
ベクトル \mathbf{a} と実数 k に対して $k\mathbf{a}$ を以下のように定義します.

- $k > 0$ の場合: \mathbf{a} と同じ向きで, 大きさが $|\mathbf{a}|$ の k 倍であるベクトル.
- $k < 0$ の場合: \mathbf{a} と逆の向きで, 大きさが $|\mathbf{a}|$ の k 倍であるベクトル. 例えば $-\mathbf{a}$ はベクトル \mathbf{a} と逆向きのベクトルを意味する.
- $k = 0$ の場合: $0\mathbf{a} = \mathbf{0}$ (零ベクトル) とする.



4.2 ベクトルの成分表示

2次元平面においてベクトル \mathbf{a} は $\mathbf{a} = \begin{pmatrix} a_1 \\ a_2 \end{pmatrix}$ と表すことがある. ここで a_1 は x 軸方向へどれだけ進んだかを表し, a_2 は y 軸方向へどれだけ進んだかを表している. 例えば $\mathbf{a} = \begin{pmatrix} 1 \\ 2 \end{pmatrix}$ とすると, 以下の図のようになる.



同様にして3次元ベクトルであれば $\mathbf{a} = \begin{pmatrix} a_1 \\ a_2 \\ a_3 \end{pmatrix}$ と n 次元ベクトルであれば $\mathbf{a} = \begin{pmatrix} a_1 \\ \vdots \\ a_n \end{pmatrix}$ と表すことが

できる². この時, ベクトルの大きさ, 和, 実数倍を以下のように定義します.

ベクトルの大きさ, 和, 実数倍

ベクトル $\mathbf{a} = \begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{pmatrix}$, $\mathbf{b} = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{pmatrix}$ とする.

- \mathbf{a} の大きさは $|\mathbf{a}| = \sqrt{a_1^2 + a_2^2 + \cdots + a_n^2}$ と定義する. (つまりベクトルの大きさ = ベクトルの長さ)
- 大きさ1のベクトルを単位ベクトルと言う.

• \mathbf{a} を k 倍したものを $k\mathbf{a} = \begin{pmatrix} ka_1 \\ ka_2 \\ \vdots \\ ka_n \end{pmatrix}$, • ベクトルの和を $\mathbf{a} + \mathbf{b} = \begin{pmatrix} a_1 + b_1 \\ a_2 + b_2 \\ \vdots \\ a_n + b_n \end{pmatrix}$ とする.

【例 4.2】 $\mathbf{a} = \begin{pmatrix} -1 \\ 2 \end{pmatrix}$, $\mathbf{b} = \begin{pmatrix} 5 \\ 6 \end{pmatrix}$ とする. この時, $3\mathbf{a} = 3 \begin{pmatrix} -1 \\ 2 \end{pmatrix} = \begin{pmatrix} -3 \\ 6 \end{pmatrix}$ で, その大きさは $|3\mathbf{a}| = \sqrt{(-3)^2 + 6^2} = 3\sqrt{5}$.

²ただしイメージできるは3次元までです

$\mathbf{a} - 2\mathbf{b} = \begin{pmatrix} -1 \\ 2 \end{pmatrix} - 2 \begin{pmatrix} 5 \\ 6 \end{pmatrix} = \begin{pmatrix} -1 \\ 2 \end{pmatrix} + \begin{pmatrix} -10 \\ -12 \end{pmatrix} = \begin{pmatrix} -11 \\ -10 \end{pmatrix}$. $\mathbf{d} = \begin{pmatrix} 1/\sqrt{2} \\ 1/\sqrt{2} \end{pmatrix}$ は, $|\mathbf{d}| = \sqrt{(1/\sqrt{2})^2 + (1/\sqrt{2})^2} = 1$ より単位ベクトルである.

問 4.1 $\mathbf{a} = \begin{pmatrix} -1 \\ 2 \end{pmatrix}$, $\mathbf{b} = \begin{pmatrix} 5 \\ 6 \end{pmatrix}$ とする. この時 $\mathbf{a} + 4\mathbf{b}$, $|\mathbf{a}|$ を求めよ.

問 4.2 ベクトル $\mathbf{a} = k \begin{pmatrix} 1 \\ 2 \end{pmatrix}$ が単位ベクトルとなるように k の値を求めよ.

4.3 方向ベクトルを持つ直線

ここでは, 話を簡単にするため 2 次元の場合のみ言います. 原点 O を通り, 方向ベクトル $\mathbf{a} = \begin{pmatrix} a_1 \\ a_2 \end{pmatrix}$ を持つ直線を考えます. もし, 方向ベクトル $\mathbf{a} = \begin{pmatrix} 2 \\ 4 \end{pmatrix}$ のとき, $x_2 = \frac{4}{2}x_1 = 2x_1$ となります. しかし方向ベクトル $\mathbf{a} = \begin{pmatrix} 10 \\ 20 \end{pmatrix}$ のときも, $x_2 = \frac{20}{10}x_1 = 2x_1$ となります. つまりこれは直線を表す方向ベクトル \mathbf{a} は無数にあることを意味します. これは式を求めるのには困った性質でして, これを避けるために直線を表す方向ベクトルは単位ベクトル (つまり大きさが 1 のベクトル) に限定します.

【例 4.3】 $x_2 = 2x_1$ の方向を表す単位ベクトルは $\mathbf{a} = \pm \begin{pmatrix} \frac{1}{\sqrt{5}} \\ \frac{2}{\sqrt{5}} \end{pmatrix}$ となります.

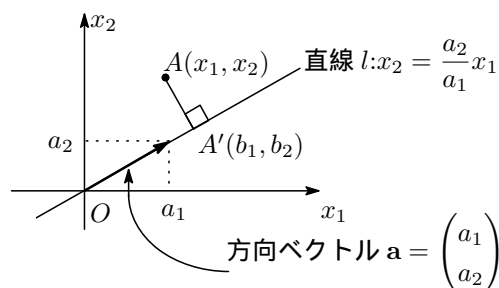
問 4.4 直線 $x_2 = \frac{1}{2}x_1$ の方向を表す単位ベクトルを求めよ.

定理 4.1(重要!!!)

原点 O を通り, 単位ベクトル $\mathbf{a} = \begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_p \end{pmatrix}$ を方向ベクトルにもつ直線を l とする. この時, 点 $\mathbf{x} = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_p \end{pmatrix}$ から直線 l に降ろした垂線の足を A' とする. この時, 直線 l を新しい軸とする A' の座標は $\mathbf{a}^T \mathbf{x}$ で与えられる.

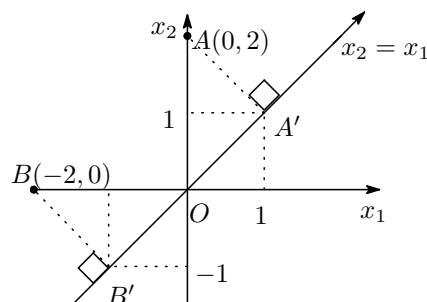
^a垂線の足とは, 垂線と直線の交点のことで, ここでは点 A' を表します

これは若干イメージが沸きづらいので, 2 次元 ($p = 2$) の場合の上の定理の意味を紹介します.



直線の方向を表す単位ベクトルを $\mathbf{a} = \begin{pmatrix} a_1 \\ a_2 \end{pmatrix}$ (ここで $a_1^2 + a_2^2 = 1$) とし, 点 A の座標を $A(x_1, x_2)$ とします. ここでこの定理は $OA' = \mathbf{a}^T \mathbf{x} = \begin{pmatrix} a_1 & a_2 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = a_1x_1 + a_2x_2$ となることを言っています. もちろん 3 次元の場合も同様に定義します. ただし 4 次元以上はイメージできないので, 数式による表現だけとなります.

【例 4.5】 直線 l (ここでは $x_2 = x_1$) を新しい軸とするときの A' の座標というのは, A' のある方向によって正になるときに負になるときがある長さ OA' と考えられます. これを符号付き距離と言います.



ここで直線 $x_2 = x_1$ の方向ベクトルは $\mathbf{a} = \begin{pmatrix} 1/\sqrt{2} \\ 1/\sqrt{2} \end{pmatrix}$ より $OA' = 0 \cdot \frac{1}{\sqrt{2}} + 2 \cdot \frac{1}{\sqrt{2}} = \sqrt{2}$ ですが, $OB' = (-2) \cdot \frac{1}{\sqrt{2}} + 0 \cdot \frac{1}{\sqrt{2}} = -\sqrt{2}$ となります.

定理 4.1 の証明: A' の座標を $A'(b_1, b_2)$ とすると, 直線 AA' と直線 l は直角に交わるので

$$\frac{b_2 - x_2}{b_1 - x_1} \cdot \frac{a_2}{a_1} = -1 \quad (4.1)$$

となる. この式を整理すると,

$$a_2(b_2 - x_2) + a_1(b_1 - x_1) = 0 \quad (4.2)$$

となる. 一方で, ベクトル \mathbf{a} を何倍かすれば $\overrightarrow{OA'}$ となるので, $ka_1 = b_1$, $ka_2 = b_2$ となる. これを上 (4.1) 式に代入すると

$$\begin{aligned} a_2(ka_2 - x_2) + a_1(ka_1 - x_1) &= 0 \\ (a_1^2 + a_2^2)k &= a_1x_1 + a_2x_2 \end{aligned}$$

ここで $a_1^2 + a_2^2 = 1$ より $k = a_1x_1 + a_2x_2$ となる. \square

問 4.5 点 $A(2, -3)$ とし, 直線 l を方向ベクトル $\mathbf{a} = \begin{pmatrix} \frac{1}{\sqrt{5}} \\ \frac{2}{\sqrt{5}} \end{pmatrix}$ を持ち, 原点 O を通る直線とする. A' を点 A から直線 l に降ろした垂線の足とすると, 軸 l 上での A' の座標を求めよ.

5 固有値 (Eigenvalue)

$n \times n$ 正方行列 A に対して, 0 でないベクトル \mathbf{w} と実数 λ が

$$A\mathbf{w} = \lambda\mathbf{w} \quad (5.1)$$

を満たすとき, λ を固有値, \mathbf{w} を固有ベクトルという.

【例題 5.1】 $n = 2$ の場合の正方行列 A を考える. ここで $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$, $\mathbf{w} = \begin{pmatrix} x \\ y \end{pmatrix}$ とすると式 (5.1) は

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

となる. 例えば, $A = \begin{pmatrix} 5 & 1 \\ 3 & 3 \end{pmatrix}$, $\mathbf{w} = \begin{pmatrix} -1 \\ 3 \end{pmatrix}$, $\lambda = 2$ とすると $\begin{pmatrix} 5 & 1 \\ 3 & 3 \end{pmatrix} \begin{pmatrix} -1 \\ 3 \end{pmatrix} = 2 \begin{pmatrix} -1 \\ 3 \end{pmatrix}$ が成り立つことがわかる.

定理 5.1

A の固有ベクトルを \mathbf{w} とすると, $k\mathbf{w}$ (k は 0 でない実数) も固有ベクトルになる

証明: $A\mathbf{w} = \lambda\mathbf{w}$ の両辺に実数 k をかけると $kA\mathbf{w} = k\lambda\mathbf{w}$. よって 2.3 章の行列の性質より $A(k\mathbf{w}) = \lambda(k\mathbf{w})$.

固有値の求め方

$|A - \lambda I| = 0$ を満たす λ を求める.

【例題 5.2】 $A = \begin{pmatrix} 5 & 1 \\ 3 & 3 \end{pmatrix}$ の固有値を求める.

$$A - \lambda I = \begin{pmatrix} 5 & 1 \\ 3 & 3 \end{pmatrix} - \lambda \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 5 - \lambda & 1 \\ 3 & 3 - \lambda \end{pmatrix} \text{ より}$$

$$|A - \lambda I| = \begin{vmatrix} 5 - \lambda & 1 \\ 3 & 3 - \lambda \end{vmatrix} = \lambda^2 - 8\lambda + 12 = (\lambda - 2)(\lambda - 6) = 0$$

よって固有値は $\lambda = 2, 6$.

問 5.1 以下の行列の固有値を求めよ ① $A = \begin{pmatrix} 5 & -1 \\ -1 & 5 \end{pmatrix}$ ② $B = \begin{pmatrix} 1 & r \\ r & 1 \end{pmatrix}$, ただし $r > 0$.

固有ベクトルの求め方

A の固有値 $\lambda = \lambda^*$ がわかっているとき, $(A - \lambda^* I)\mathbf{w} = \mathbf{0}$ を満たす \mathbf{w} を求める

$n = 2$ の場合は

$$\begin{cases} (a - \lambda^*)x + by = 0 \\ cx + (d - \lambda^*)y = 0 \end{cases} \quad (5.2)$$

を満たすような x と y を求めればよい. ただし (5.2) を満たす x と y は無数にあることに注意.

【例題 5.3】 【例題 5.2】で求めた固有値に対する固有ベクトルを求める.

2×2 の行列の場合, (5.2) に値を代入すると

$$\begin{cases} 3x + y = 0 \\ 3x + y = 0 \end{cases} \quad (5.3)$$

よって $3x + y = 0$ を得る. ここで $\frac{y}{x} = \frac{-3}{1}$ より $x : y = 1 : (-3)$ となる. よって固有ベクトルは $\mathbf{u} = k \begin{pmatrix} 1 \\ -3 \end{pmatrix}$.

ここで $k \neq 0$.

同様にして $\lambda = 6$ の時,

$$\begin{cases} -x + y = 0 \\ 3x - 3y = 0 \end{cases} \quad (5.4)$$

よって $x - y = 0$ を得る. ここで $\frac{y}{x} = \frac{1}{1}$ より $x : y = 1 : 1$ となる. よって固有ベクトルは $\mathbf{v} = l \begin{pmatrix} 1 \\ 1 \end{pmatrix}$. ここで $l \neq 0$.

問 5.2 問 5.1 における A, B の固有ベクトルを求めよ.

5.1 3 次の正方行列に対する固有値, 固有ベクトル

ここで 3×3 の行列 $A = \begin{pmatrix} 3 & 1 & 1 \\ 1 & 2 & 0 \\ 1 & 0 & 2 \end{pmatrix}$ の固有値, 固有ベクトルを求めることを考える.

① 固有値を求める. $|A - \lambda I| = 0$ を解くと, $\begin{vmatrix} 3 - \lambda & 1 & 1 \\ 1 & 2 - \lambda & 0 \\ 1 & 0 & 2 - \lambda \end{vmatrix} = 0$ となる. これより左辺の行列式を

展開すると

$$\begin{aligned}(3-\lambda)(2-\lambda)(2-\lambda) - (2-\lambda) - (2-\lambda) &= 0 \\ (2-\lambda)(\lambda^2 - 5\lambda + 6 - 2) &= 0 \\ (\lambda-2)(\lambda-1)(\lambda-4) &= 0\end{aligned}$$

よって固有値は $\lambda = 1, 2, 4$.

② 固有ベクトルを求める. $(A - \lambda I)\mathbf{w} = \mathbf{0}$ を式で表すと

$$\begin{cases} (3-\lambda)x + y + z = 0 \\ x + (2-\lambda)y = 0 \\ x + (2-\lambda)z = 0 \end{cases} \quad (5.5)$$

となる.

$$\bullet \lambda = 1 \text{ のとき } \begin{cases} 2x + y + z = 0 \\ x + y = 0 \\ x + z = 0 \end{cases} \quad \text{から } y = -x, z = -x \text{ を得る. ここで } x = k \text{ とおくと, } y = -k, z = -k$$

となる. よって求める固有ベクトルは $\mathbf{u} = k \begin{pmatrix} 1 \\ -1 \\ -1 \end{pmatrix}$, $k \neq 0$ となる.

$$\bullet \lambda = 2 \text{ のとき } \begin{cases} x + y + z = 0 \\ x = 0 \end{cases} \quad \text{から } x = 0, y = -z \text{ を得る. よって求める固有ベクトルは } \mathbf{v} = l \begin{pmatrix} 0 \\ 1 \\ -1 \end{pmatrix},$$

$l \neq 0$ となる.

$$\bullet \lambda = 4 \text{ のとき } \begin{cases} -x + y + z = 0 \\ x - 2y = 0 \\ x - 2z = 0 \end{cases} \quad \text{から } y = \frac{1}{2}x, z = \frac{1}{2}x \text{ を得る. ここで } x = m \text{ とおくと, } y = \frac{m}{2}, z = \frac{m}{2}$$

となる. よって求める固有ベクトルは $\mathbf{w} = m \begin{pmatrix} 1 \\ \frac{1}{2} \\ \frac{1}{2} \end{pmatrix}$, $m \neq 0$ となる.

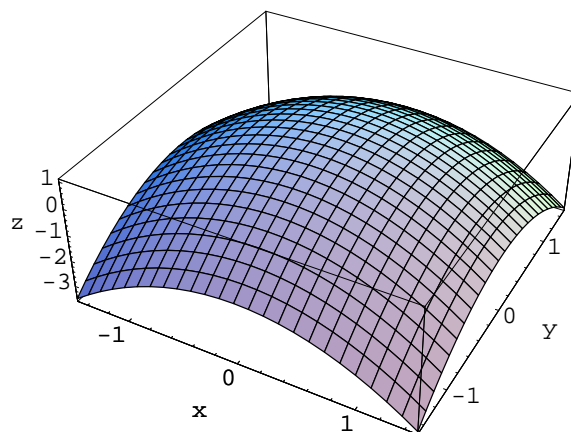
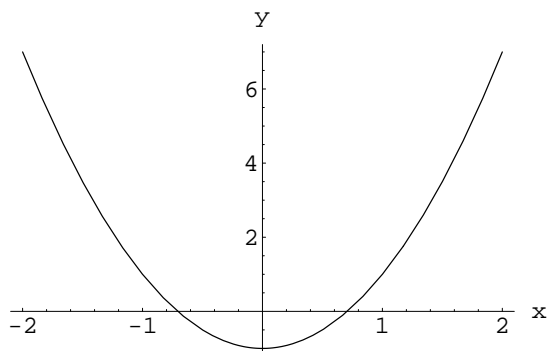
問 5.3 次の行列の固有値と固有ベクトルを求めよ.

$$\textcircled{1} \begin{pmatrix} 1 & 0 & -1 \\ 0 & 2 & 2 \\ -1 & 2 & 3 \end{pmatrix} \quad \textcircled{2} \begin{pmatrix} 1 & -1 & 1 \\ -1 & 1 & 1 \\ 1 & 1 & -1 \end{pmatrix}$$

6 偏導関数

6.1 2変数関数

これまでは左下のグラフ $y = 2x^2 - 1$ のように y を x を用いて表し, 一般的には $y = f(x)$ として表しました. 一方で z を x, y という2つの変数で表すことを考えてみます. 例えば $z = 1 - x^2 - y^2$ とすると, 右下のグラフのように, 3次元空間における曲面になることがわかります. これは一般的には $z = f(x, y)$ と表します. また $f(x, y) = 1 - x^2 - y^2$ のように表すこともあります.



6.2 偏微分

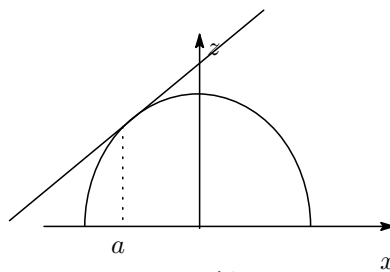
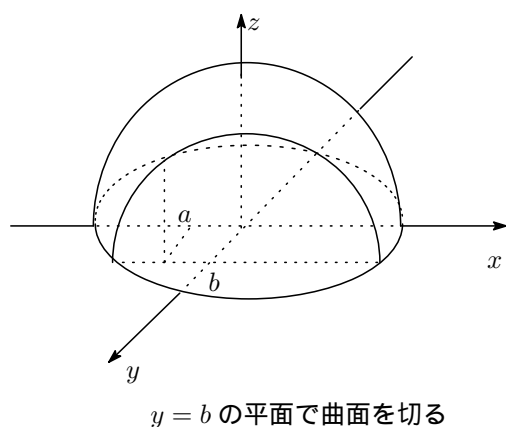
関数 $z = f(x, y)$ の点 (a, b) における x に関する偏微分は $\frac{\partial}{\partial x} f(a, b)$ もしくは $f_x(a, b)$ のように表し,

$$\frac{\partial}{\partial x} f(a, b) = \lim_{h \rightarrow 0} \frac{f(a + h, b) - f(a, b)}{h} \quad (6.1)$$

と定義します. 同様にして関数 $z = f(x, y)$ の点 (a, b) における y に関する偏微分も $\frac{\partial}{\partial y} f(a, b)$ もしくは $f_y(a, b)$ と表し

$$\frac{\partial}{\partial y} f(a, b) = \lim_{h \rightarrow 0} \frac{f(a, b + h) - f(a, b)}{h} \quad (6.2)$$

と定義します. ここで y に関する偏微分のイメージを説明します.



$y = b$ での断面
その $y = b$ における切り口を 2 次元のグラフで表し, $z = f(x, b)$ を $x = a$ の点で微分したのが $\frac{\partial}{\partial x} f(a, b)$ となる.

これは言い換えると, 曲面 $z = f(x, y)$ において $y = b$ と固定したときの, $x = a$ における接線の傾きが x に関する偏微分だということが出来ます. 次の偏微分の計算の仕方を説明します.

【例 6.1】 $f(x, y) = 1 - x^2 - y^2$ の点 $(1, 2)$ における x に関する偏微分を求めます.

- (1) $f(x, y) = 1 - x^2 - y^2$ に $y = 2$ を代入.
- (2) $f(x, 2) = 1 - x^2 - 2^2 = 3 - x^2$ を点 x で微分すると $\frac{df(x, 2)}{dx} = -2x$.
- (3) $x = 1$ を代入すると, 偏微分 $\frac{\partial f(1, 2)}{\partial x} = -2$ を得る.

問 6.1 $f(x, y) = 1 - 2x^2 - y^2 + xy$ とする.

- ① 点 $(1, 0)$ における x に関する偏微分を求めよ.
- ② 点 $(1, 1)$ における x に関する偏微分を求めよ.
- ③ 点 $(1, 2)$ における x に関する偏微分を求めよ.

6.3 偏導関数

偏微分では、問 6.1 のように y のそれぞれの値 b に対して偏微分 $\frac{\partial}{\partial x} f(a, b)$ を対応させると、1 つの新しい関数を得られます。この新しい関数を $z = f(x, y)$ の x に関する偏導関数と言い、 $\frac{\partial}{\partial x} f(x, y)$, $f_x(x, y)$ で表します。関数 $z = f(x, y)$ の x に関する偏導関数は

$$\frac{\partial}{\partial x} f(x, y) = \lim_{h \rightarrow 0} \frac{f(x+h, y) - f(x, y)}{h} \quad (6.3)$$

と定義します。同様にして関数 $z = f(x, y)$ の y に関する偏導関数も

$$\frac{\partial}{\partial y} f(x, y) = \lim_{h \rightarrow 0} \frac{f(x, y+h) - f(x, y)}{h} \quad (6.4)$$

と定義します。ただし実際の計算では、上の定義は使いません。計算はいたって簡単で次のようにします。

【例 6.2】ここでは $f(x, y) = 1 - x^2 - y^2$ の x に関する偏導関数を求めます。

これは y を定数だとみなして、 x で微分すればよいということです。よって $\frac{\partial}{\partial x} f(x, y) = -2x$ となります。

同様にして $f(x, y) = 1 - x^2 - y^2$ の y に関する偏導関数も、 x を定数だとみなして、 y で微分すればよいので $\frac{\partial}{\partial y} f(x, y) = -2y$ となります。

問 6.2 $f(x, y) = 1 - 2x^2 - y^2 + xy$ とする時、 $\frac{\partial}{\partial x} f(x, y)$, $\frac{\partial}{\partial y} f(x, y)$ を求めよ。

今、 $z = f(x, y)$ という 2 つの変数 x, y の関数に対する偏微分を考えたが、3 変数の関数 $z = f(x_1, x_2, x_3)$ や、より一般的な n 個の変数の関数 $z = f(x_1, x_2, \dots, x_n)$ に対する偏微分も同様に考えることができます。

【例 6.3】 $f(x_1, x_2, x_3, x_4) = x_1^2 x_2 + x_3 x_4 + 1$ とする。このとき、 $\frac{\partial}{\partial x_1} f(x_1, x_2, x_3, x_4) = 2x_1 x_2$ 。
 $\frac{\partial}{\partial x_3} f(x_1, x_2, x_3, x_4) = x_4$ 。

6.4 極大値、極小値

$z = f(x, y)$ が点 (a, b) の近くの任意の点 (x, y) において

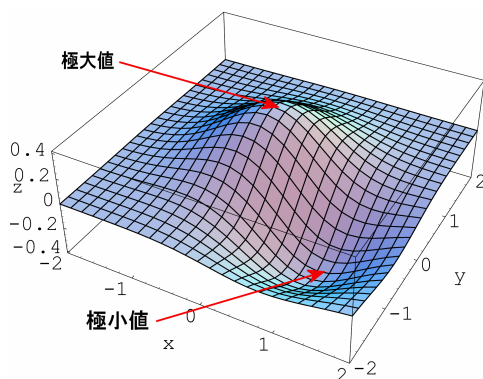
$$f(x, y) \leq f(a, b) \quad (6.5)$$

を満たすとき $f(a, b)$ は極大値であると言う。一方で点 (a^*, b^*) の近くの任意の点 (x, y) において

$$f(x, y) \geq f(a^*, b^*) \quad (6.6)$$

を満たすとき $f(a^*, b^*)$ は極小値であると言う。また極大値と極小値をあわせたものを極値と言う。

何かわかりづらい定義の仕方だと思いますが、要は周辺と比べて山があってその山の頂上を極大値、凹み（へこみ）があってその一番底の部分が極小値と考えてもらえばよいと思います。



ここで偏微分と極値の関係を一つ紹介します。

極値の性質

$z = f(x, y)$ が点 (a, b) で極値を持つとき、

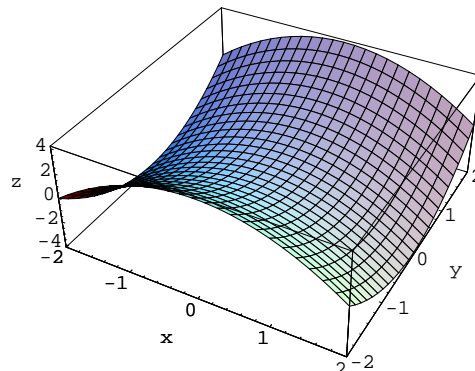
$$\frac{\partial}{\partial x} f(a, b) = 0, \quad \frac{\partial}{\partial y} f(a, b) = 0 \quad (6.7)$$

が成り立つ。

ここで $\frac{\partial}{\partial x} f(a, b)$ は $f(x, y)$ を x で偏微分した後で、点 (a, b) を代入したものを表す。 $\frac{\partial}{\partial y} f(a, b)$ も同様に解釈する。また式 (6.7) を満たす点を停留点と言う。

注意 1 この定理は式 (6.7) より求められる点 (a, b) は極値の候補を含んでいるが、それらが常に極値になるかどうかはわからない。例えば、 $f(x, y) = y^2 - x^2$ は $\frac{\partial}{\partial x} f(x, y) = -2x = 0$, $\frac{\partial}{\partial y} f(x, y) = 2y = 0$ より $x = 0, y = 0$ は停留点ではあるが、極値ではないことは右のグラフから分かる。ちょうど点 $(0, 0)$ は乗馬で使う鞍の真ん中の点^aとなっている。

^aこのような点を鞍点と言う。



注意 2 極値の判定方法はあるが、主成分分析では、常に極値を持つ関数しか扱わないのでここでは省略する。

【例 6.4】 $f(x, y) = 2x^2 + 3y^2 - y$ の極値を求めよ。 $0 = \frac{\partial}{\partial x} f(x, y) = 4x$, $0 = \frac{\partial}{\partial y} f(x, y) = 6y - 1$ より $x = 0, y = \frac{1}{6}$. これより $(0, 1/6)$ の時に $f(x, y)$ は極値 (実は最小値) を持つ。

問 6.3 以下の関数に対する最小値を求めよ。

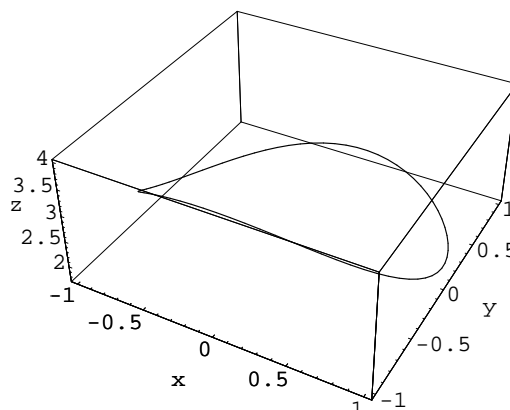
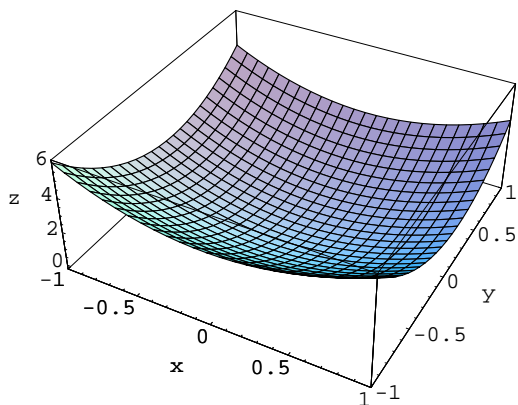
- ① $f(x, y) = x^2 - 2x + y^2 - y$.
- ② $f(x, y) = 2x^2 - 4xy - 2x + 4y^2 + 1$

7 ラグランジュの未定乗数法

以下の問題を考えてみましょう。

条件 $x^2 + y^2 = 1$ の下で $f(x, y) = 2x^2 + 3y^2 - y$ の極値を求めよ。

$f(x, y) = 2x^2 + 3y^2 - y$ は左下のグラフなのですが、その曲面上で $x^2 + y^2 = 1$ を満たす点の集まり (右下のグラフ) の中から極値を求めよという問題です。



この問題を解くためには、次のラグランジュの未定乗数法を使います。

ラグランジュの未定乗数法

制約条件 $g(x_1, x_2, \dots, x_n) = 0$ の下で関数 $f(x_1, x_2, \dots, x_n)$ の極値をとる点は

$$F(x_1, x_2, \dots, x_n, \lambda) = f(x_1, x_2, \dots, x_n) - \lambda g(x_1, x_2, \dots, x_n) \quad (7.1)$$

とおくとき, $\frac{\partial}{\partial x_i} F(x_1, x_2, \dots, x_n, \lambda) = 0$, $(i = 1, 2, \dots, n)$, $\frac{\partial}{\partial \lambda} F(x_1, x_2, \dots, x_n, \lambda) = 0$ を満たす.

証明:ここでは省略するが, 金谷 (2005) [6] に分かりやすい説明が載っている.

【例 7.1】条件 $x^2 + y^2 = 1$ の下で $f(x, y) = 2x^2 + 3y^2 - y$ の極値を求める.

$g(x, y) = x^2 + y^2 - 1$, $f(x, y) = 2x^2 + 3y^2 - y$ とおき,

$$F(x, y, \lambda) = 2x^2 + 3y^2 - y - \lambda(x^2 + y^2 - 1) \quad (7.2)$$

とする. ここで

$$\frac{\partial}{\partial x} F(x, y, \lambda) = 4x - 2\lambda x = 0 \quad \dots \textcircled{1}$$

$$\frac{\partial}{\partial y} F(x, y, \lambda) = 6y - 1 - 2\lambda y = 0 \quad \dots \textcircled{2}$$

$$\frac{\partial}{\partial \lambda} F(x, y, \lambda) = -(x^2 + y^2 - 1) = 0 \quad \dots \textcircled{3}$$

ここで ①より $2x(2 - \lambda) = 0$ より $\lambda = 2$ または $x = 0$ となる.

$\lambda = 2$ の時 これを ②に代入すると $y = \frac{1}{2}$. これを ③に代入すると $x = \pm \frac{\sqrt{3}}{2}$ を得る. これより点 $\left(\pm \frac{\sqrt{3}}{2}, \frac{1}{2}\right)$ において $f\left(\pm \frac{\sqrt{3}}{2}, \frac{1}{2}\right) = 2 \cdot \frac{3}{4} + 3 \cdot \frac{1}{4} - \frac{1}{2} = \frac{7}{4}$ となる.

$x = 0$ の時 ③に代入すると $y = \pm 1$ となる. 点 $(0, 1)$ において $f(0, 1) = 2 \cdot 0^2 + 3 \cdot 1 - 1 = 2$.

点 $(0, -1)$ において $f(0, -1) = 2 \cdot 0^2 + 3 \cdot 1 + 1 = 4$ となる.

これより点 $(x, y) = (0, -1)$ の時, 最大値 4. 点 $(x, y) = \left(\pm \frac{\sqrt{3}}{2}, \frac{1}{2}\right)$ の時, 最小値 $\frac{7}{4}$ となる.

注意 このようにある制限の下で, 関数 $f(x, y)$ の極値を求める問題を条件付極値問題という. ラグランジュ乗数法は極値の候補を与えるが, それらが全て極値になるとは限らない. 極値かどうかの判定法はあるが, 主成分分析では条件付き極値を持つものだけ扱うので, ここでは省略する.

問 7.1

① $x + y = 1$ の下で $f(x, y) = 2x^2 + 3y^2 + xy$ は最小値を持つことが分かっている. $f(x, y)$ が最小になる (x, y) を求めよ.

② $x^2 + y^2 + z^2 = 1$ の下で $f(x, y, z) = xy + yz + zx$ の最大値を求めよ. ただし $x > 0, y > 0, z > 0$ とする.

8 標本平均 (Sample mean, Average)

この章では, 数値を使ってデータを表すことを考えます. データの特徴を数値を使って表わしたものを代表値と言います.

データ x_1, x_2, \dots, x_n に対して, 全てのデータを足して, それをデータの個数で割ったもの

$$\bar{x} = \frac{x_1 + x_2 + \dots + x_n}{n} \quad (8.1)$$

を標本平均という. ここで \sum 記号を用いると $\bar{x} = \frac{1}{n} \sum_{i=1}^n x_i$ と表せる.

【例 8.1】次の表は美香子さんの過去 10 回の数学と英語の 10 点満点の小テストの結果である.

回数	1	2	3	4	5	6	7	8	9	10
数学 (点)	9	5	8	8	10	6	9	9	7	9
英語 (点)	8	9	7	9	8	9	8	8	7	7

表 8.1

ここで数学のデータを $x_1 = 9, x_2 = 5, \dots, x_{10} = 9$ とすると、その 標本平均は

$$\bar{x} = \frac{x_1 + x_2 + \dots + x_{10}}{10} = \frac{9 + 5 + 8 + 8 + 10 + 6 + 9 + 9 + 7 + 9}{10} = 8$$

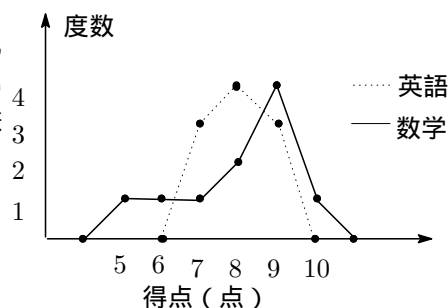
となります。

問 8.1 表 8.1 における美香子さんの英語の平均を求めよ。

【学生】 標本平均はデータの何を表しているのですか？

【先生】 標本平均は与えられたデータが大体どこを中心に分布しているかを表しています。上の例ですと数学の標本平均は 8 点ですが、9 点、10 点のように 8 点より大きな点数と 7, 6 点のように 8 点より小さな点数の度合いが大体同じことがわかります。

実は数学と英語の平均点が 8 点になるのですが、表 1.1 のデータのそれぞれの度数折れ線を書いてみると、数学と英語では点数の散らばりの様子が異なっていますよね。ここでデータの散らばり具合を表す指標として、標本分散というものをご紹介します。



9 標本分散 (Sample Variance)

データ x_1, \dots, x_n に対して、

$$s^2 = \frac{(x_1 - \bar{x})^2 + (x_2 - \bar{x})^2 + \dots + (x_n - \bar{x})^2}{n} \quad (9.1)$$

を標本分散 s^2 と言う。ここで $\bar{x} = \frac{1}{n} \sum_{i=1}^n x_i$ 。

また Σ 記号を用いると $s^2 = \frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})^2$ と表せる。

標本分散は単に分散ということもあるので注意してください。また

$$s = \sqrt{s^2} = \sqrt{\frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})^2}$$

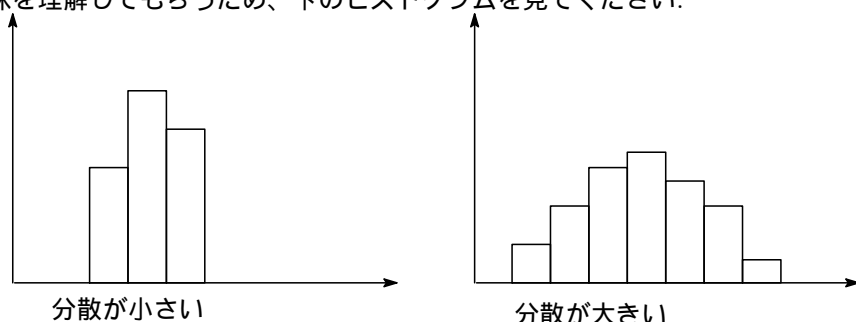
を標準偏差と呼びます。分散と標準偏差はデータの散らばり具合を表す 1 つの指標です。

9.1 標本分散の意味

式 (9.1) の標本分散において、 $(x_1 - \bar{x})^2$ という項は \bar{x} と x_1 がどれだけ離れているかを表しています。

例えば $\bar{x} = 8$ で、 $x_1 = 9$ のときには $(9 - 8)^2 = 1$ となりますが、 $x_1 = 10$ の時には $(10 - 8)^2 = 4$ となります。つまり \bar{x} から離れていればいるほど大きな値をとるということです。同様に $(x_2 - \bar{x})^2, \dots$,

$(x_n - \bar{x})^2$ も、それぞれ x_2, \dots, x_n が \bar{x} からどれだけ離れているかを表しているものとなります。このため、標本分散は標本平均から平均的に見てどれくらい離れているかを表しています。ここで標本分散の意味を理解してもらうため、下のヒストグラムを見てください。



【例 9.1】表 2.4 における数学のデータに対する標本分散 s_x^2 は、 $x_1 = 9, x_2 = 5, \dots, x_{10} = 9$ より、

$$\begin{aligned} s_x^2 &= \frac{(x_1 - \bar{x})^2 + (x_2 - \bar{x})^2 + \dots + (x_{10} - \bar{x})^2}{10} \\ &= \frac{(9-8)^2 + (5-8)^2 + \dots + (9-8)^2}{10} \\ &= \frac{1^2 + 3^2 + 2^2 + 2^2 + 1^2 + 1^2 + 1^2 + 1^2}{10} = \frac{22}{10} = 2.2 \end{aligned}$$

となります。一方、英語の標本分散 s_y^2 も同様にして

$$s_y^2 = \frac{(8-8)^2 + (9-8)^2 + \dots + (7-8)^2}{10} = \frac{3}{5} = 0.6$$

となります。これより、榎本君の数学の成績と英語の成績の標本平均は同じであるが、数学の方が英語より成績にばらつきが大きいことがわかります。

問 9.1 5 人の生徒に駅から学校までの通学時間 (分) を聞いたところ以下のようなデータが得られた。

15, 10, 15, 14, 16

- ① データの標本平均を求めよ。
- ② データの標本分散を求めよ。
- ③ データの標準偏差を求めよ。ただし $\sqrt{110} \doteq 10.5$ とする。

平均値が整数の値を取る場合は上の例のように計算すればよいですが、もし平均値が小数の値をとると、分散の計算は面倒くさくなります。この煩わしさをなくすために以下の公式が知られています。

分散の計算公式

(9.1) の標本分散 s^2 は以下の形で表せる。

$$s^2 = \frac{1}{n} \sum_{i=1}^n x_i^2 - (\bar{x})^2 \quad (9.2)$$

これより分散は (データを 2 乗したものの平均) - (データの平均)² で計算できることがわかります。

証明

$$\begin{aligned}s^2 &= \frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})^2 \\&= \frac{1}{n} \sum_{i=1}^n (x_i^2 - 2\bar{x}x_i + \bar{x}^2) \\&= \frac{1}{n} \sum_{i=1}^n x_i^2 - 2\bar{x} \frac{1}{n} \sum_{i=1}^n x_i + \frac{1}{n} \sum_{i=1}^n (\bar{x})^2 \\&= \frac{1}{n} \sum_{i=1}^n x_i^2 - 2(\bar{x})^2 + (\bar{x})^2 \\&= \frac{1}{n} \sum_{i=1}^n x_i^2 - (\bar{x})^2.\end{aligned}$$

【例 9.2】 6 人に対して携帯電話の 1 日あたりの通話時間 (分) を調べたところ 10, 13, 8, 5, 12, 2 (分) というデータを得た. このとき, 平均値は $\bar{x} = \frac{25}{3}$ となるので

$$s^2 = \frac{10^2 + 13^2 + 5^2 + 8^2 + 12^2 + 2^2}{6} - \left(\frac{25}{3}\right)^2 = \frac{253}{3} - \frac{625}{9} = \frac{134}{9}$$

となります.

問 9.2 2001 年にある高校の男子と女子生徒それぞれ 10 人に対して 1 年に約何枚の CD を買うのかアンケートをとったところ以下の結果が得られた.

男子	20, 3, 25, 11, 0, 5, 7, 5, 5, 5
女子	10, 15, 6, 4, 9, 6, 8, 4, 7, 12

- ① 男子と女子それぞれの CD を買う枚数の平均値を求めよ.
- ② それぞれの高校の標本分散と標準偏差を分散の計算公式を用いて計算せよ. (要電卓)
- ③ ②の結果からどのようなことが分かるか.

9.2 標本分散の性質... 単位が変われば分散の値も変わる!

標本分散は小さいほど, ばらつきが少ないということですが, その数字自体の意味はほとんどありません. 例えば, ある 4 本のヒマワリの高さをメートルという単位で測った時, $0.8m, 0.7m, 0.6m, 1.1m$ であったとしましょう. このとき標本平均は $\bar{x} = 0.8$ となり, 標本分散は

$$s_x^2 = \frac{0.8^2 + 0.7^2 + 0.6^2 + (1.1)^2}{4} - (0.8)^2 = \frac{7}{200} = 0.035$$

となります. 次にヒマワリの長さをセンチメートル単位で表すと $80cm, 70cm, 60cm, 110cm$ と表せます. これから標本分散 s_w^2 を計算すると $s_w^2 = 350$ となります. つまり, 単位が変われば, 分散の値も変わるということです.

実は, 単位の数値が 100 倍になれば (例えばメートル法で $1.1m$ のものはセンチメートル法ではその 100 倍の数値にあたる $110cm$ として表される) 分散の値は 10000 倍となることが知られています. これは, 分散を考える上で注意しなくてはならないことです.

10 標準化

$\bar{x} = \frac{1}{n} \sum_{i=1}^n x_i, s = \sqrt{\frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})^2}$ をデータ x_1, \dots, x_n の標本平均、標準偏差とします. この時, 次の変換を考えましょう.

$$z_1 = \frac{x_1 - \bar{x}}{s}, z_2 = \frac{x_2 - \bar{x}}{s}, \dots, z_n = \frac{x_n - \bar{x}}{s}$$

この変換を標準化と呼び、 z_1, \dots, z_n のことを z -スコアと言います。この時

$$\bar{z} = \frac{1}{n} \sum_{i=1}^n z_i = 0, \quad s_z^2 = \frac{1}{n} \sum_{i=1}^n (z_i - \bar{z})^2 = 1 \quad (10.1)$$

が成り立ちます。この変換を行うと、どんなデータであっても (標本) 平均 0, (標本) 分散 1 となるようなデータに変換できることを意味します。

【例 10.1】 【例 9.2】のデータにおいて、6 人のうちの 1 人である B 君の通話時間は 2 分であった。この時、

$$\text{B 君の } z\text{-スコア} = \frac{2 - \frac{25}{3}}{\sqrt{\frac{134}{9}}} \div -1.64$$

となる。

問 10.1 【例 9.2】のデータにおいて、6 人のうちの 1 人である A 君の通話時間は 10 分であった。A 君の z -スコアを求めよ。

問 10.2 式 (10.1) を証明せよ。

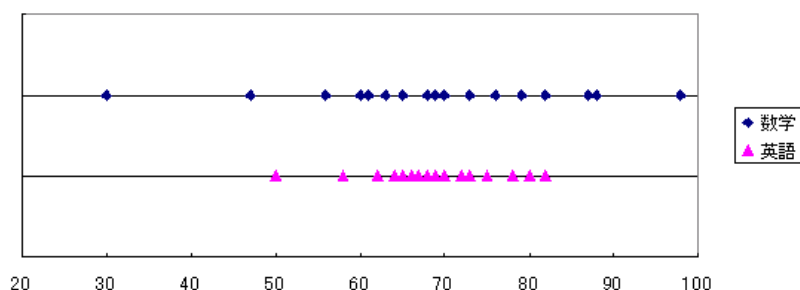
10.1 なぜ標準偏差 s で割る必要があるのですか？

20 人のクラスに対して、数学と英語のテストを行った時、以下の結果が得られたとしましょう。

氏名	数学	英語	氏名	数学	英語	氏名	数学	英語
A	47	58	H	87	82	O	82	78
B	73	70	I	76	72	P	88	75
C	68	65	J	69	68	Q	65	68
D	79	73	K	56	62	R	69	70
E	70	69	L	30	50	S	98	80
F	61	64	M	63	67	T	56	58
G	63	66	N	60	65			

表 10.1

ここで数学と英語の平均は共に 68 点となります。この時 O さんの数学 82 点と H さんの英語 82 点どちらが価値があるのかを考えて見ましょう。実は数学と英語の平均点を直線上に点をうっていくと、以下のようになります。



つまり英語の方はたいていの生徒が 68 点付近の点数を取っている中での 82 点であって、一方で数学の方は 68 点からの散らばり具合が大きいの、80 点台をとった人も何人かいるなかでの 82 点ということになります。つまりは英語の 82 点の方が価値があるということになります。

z-スコアは、データの散らばり具合も考慮にいたした上での平均値 \bar{x} からのずれを測っていることとなります。実際、数学の標準偏差は 14.9、英語の標準偏差は 7.5 です。

$$z = \frac{82 - 68}{14.9} = 0.93 \quad (\text{O さんの数学の z-スコア})$$

$$z = \frac{82 - 68}{7.5} = 1.86 \quad (\text{H さんの英語の z-スコア})$$

となり、英語の 82 点の方が価値が高いことがわかります。

10.2 標準化のいいところは何ですか？

例えば、アサガオの丈の長さを測ったところ 20cm, 60cm, 80cm, 100cm, 140cm だったとしましょう。この時、標本平均 $\bar{x} = 80$ 、標本分散 $s_x^2 = 1600$ より最初のデータ 20 は

$$z_1 = \frac{20 - 80}{\sqrt{1600}} = -\frac{3}{2} = -1.5$$

となります。残りの 60, 80, 100, 140 も同様にして変換すると次の表のようになります。

元のデータ (cm)	標準化されたデータ
20	-1.5
60	-0.5
80	0
100	0.5
140	1.5

一方で、これらのデータをメートルを使って表すと 0.2m, 0.6m, 0.8m, 1m, 1.4m となりますよね。この時、標本平均 $\bar{y} = 0.8$ 、標本分散 $s_y^2 = 0.16$ から最初のデータ 0.2 は

$$z_1 = \frac{0.2 - 0.8}{\sqrt{0.16}} = -\frac{3}{2} = -1.5$$

となります。残りの 0.6, 0.8, 1, 1.4 も同様にして変換すると次のようになります。

元のデータ (m)	標準化されたデータ
0.2	-1.5
0.6	-0.5
0.8	0
1	0.5
1.4	1.5

結局、何が分かったかというと、標準化されたデータは単位の影響を受けないということです。

11 2つの変数の関係を調べる。

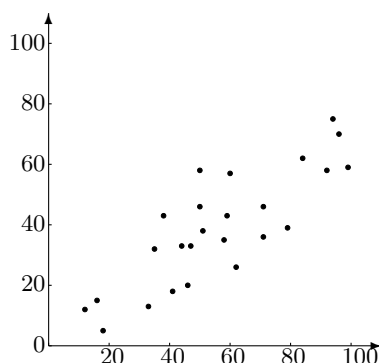
11.1 散布図

ある高校の選択科目で、1 学期の成績を x 、2 学期の成績を y とし、2 つの変数の組 (x, y) を 25 人の受講者について調べたところ、次のようになりました。

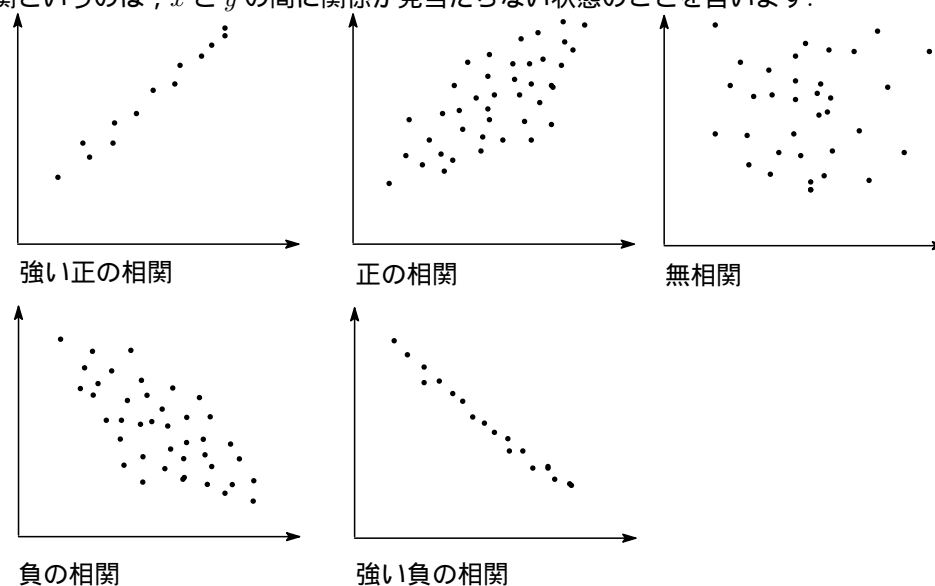
(99,59)	(84,62)	(62,26)	(92,58)	(59,43)	(35,32)	(71,36)	(33,13)	(47,33)
(18,5)	(50,46)	(51,38)	(58,35)	(60,57)	(71,46)	(46,20)	(38,43)	(12,12)
(16,15)	(79,39)	(50,58)	(96,70)	(94,75)	(44,33)	(41,18)		

表 11.1

これらの 25 個の資料のそれぞれを xy 平面上の点として以下のように図示してみると、 x の値が大きいと、 y の値も大きくなる傾向があることがわかります。



このように、2 つの変量からなるデータを平面上に図示して、その関係を目で見てわかるようにしたものを散布図といいます。散布図には大きく分けて次の 5 つの場合に区別することができます。正の相関があるとは、 x の値が大きくなると、それに対応する y の値も大きくなる傾向があるものを言います。例えば表 11.1 のように国語の成績がいいものは数学の成績もよいという関係があることです。負の相関があるとは x の値が大きくなると、逆にそれに対応する y の値は小さくなる傾向があるものを言います。無相関というのは、 x と y の間に関係が見当たらない状態のことを言います。



11.2 (標本) 共分散

表 11.1 のような対のデータ $(x_1, y_1), (x_2, y_2), \dots, (x_n, y_n)$ に対して、どのような関係があるのかを数値で表すために、共分散と呼ばれるものを紹介します。共分散は s_{xy} という記号で、以下のように計算します。

共分散

$$s_{xy} = \frac{(x_1 - \bar{x})(y_1 - \bar{y}) + (x_2 - \bar{x})(y_2 - \bar{y}) + \cdots + (x_n - \bar{x})(y_n - \bar{y})}{n}$$

これは \sum 記号を使うと $s_{xy} = \frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})$ と表せます。

共分散には以下のような特徴があります。

- 散布図に正の相関がある $\iff s_{xy} > 0$
- 散布図に負の相関がある $\iff s_{xy} < 0$
- 散布図は無相関である $\iff s_{xy} \div 0$

【例 11.1】 ある 5 人に対して英語 x と数学 y の小テストを行ったところ、以下の対のデータ (x, y) を得た。この時、共分散を求めよ。

x	9	5	8	8	10
y	5	9	7	9	8

これは表を使って計算するとよい。

x	y	$x - \bar{x}$	$y - \bar{y}$	$(x - \bar{x})(y - \bar{y})$	
9	5	1	-2.6	-2.6	
5	9	-3	1.4	-4.2	
8	7	0	-0.6	0	
8	9	0	1.4	0	
10	8	2	0.4	0.8	
計	40	38	0	0	-6

これより共分散 $s_{xy} = \frac{-6}{5} = -1.2$ となります。

問 11.1 以下の対のデータ (x, y) に対する共分散を求めよ。

x	1	4	4	1	5
y	2	6	5	3	6

Excel には共分散を求めてくれる機能がありますが、手計算や電卓を使うときには以下の公式を使うと計算しやすくなることが多いです。

共分散の計算公式

$\overline{xy} = \frac{x_1y_1 + x_2y_2 + \cdots + x_ny_n}{n}$ とするとき、共分散 s_{xy} は以下の形で表せる。

$$s_{xy} = \overline{xy} - \bar{x} \cdot \bar{y} \quad (11.1)$$

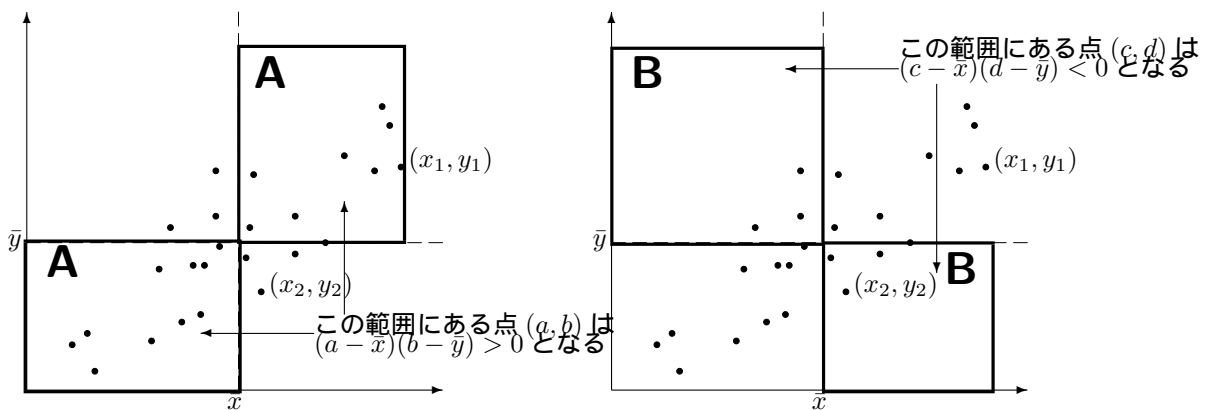
また \sum 記号を用いると $\overline{xy} = \frac{1}{n} \sum_{i=1}^n x_iy_i$ と表せる。

問 11.2 以下の対のデータ (x, y) に対する共分散を計算公式を用いて求めよ。

x	2	5	3	9	12
y	3	7	4	10	12

11.3 共分散はなぜ正か負で 2 つの変数の関係を表せるの？

これは表 11.1 の散布図を用いて説明しましょう。



この2つの図より，点の集まりがBよりAの領域にたくさんある場合には，

$$s_{xy} = \frac{1}{n} \left(\underbrace{(x_1 - \bar{x})(y_1 - \bar{y}) + \dots + (x_n - \bar{x})(y_n - \bar{y})}_{\text{正の値をとる}} \right) \quad (11.2)$$

において正の値をとる項が多くなりますので，これらの n 個の項を足し合わせた場合 s_{xy} は正となります．一方で，AよりBの領域にたくさん点がある場合には，負の値をとる項が多くなりますので，これらの n 個の項を足し合わせた場合 s_{xy} は負となります．

11.4 共分散の弱点

実は共分散には弱点があります．これを示すために以下の例を考えてみましょう．

No.	国語 (100点満点)	数学 (100点満点)
1	80	70
2	50	60
3	70	50
4	90	30
5	70	90

表 11.2

No.	国語 (100点満点)	数学 (10点満点)
1	80	7
2	50	6
3	70	5
4	90	3
5	70	9

表 11.3

ここで共分散を計算すると，左の表の共分散は -100 となり，右の表の共分散は -10 となります．これより国語と数学の関係は同じなのに評価する点数の単位によって共分散の値が変わってしまいます．つまり

- 共分散 s_{xy} は2つ変数 x と y の単位によって値が変わる．

ということなのです．ですから共分散が正の大きい値をとっているからといって，正の相関が強いということはいえません．ここで2つの変数の関係を表す数値として望まれるのは単位の影響を受けないということです．実は単位の影響を受けずに2つの変数の関係を表す数値として，相関係数と呼ばれるものがあります．

11.5 相関係数

相関係数はある量 x とある量 y との直線的³な関係の強さを表す指標です。式で表すと、対のデータ $(x_1, y_1), \dots, (x_n, y_n)$ が与えられたとき、 x と y の相関係数 r は

$$r = \frac{\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^n (x_i - \bar{x})^2} \sqrt{\sum_{i=1}^n (y_i - \bar{y})^2}} \quad (11.3)$$

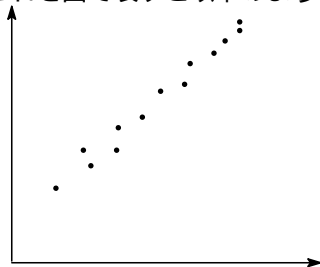
となります。ここで $\bar{x} = \frac{1}{n} \sum_{i=1}^n x_i$ (x の平均), $\bar{y} = \frac{1}{n} \sum_{i=1}^n y_i$ (y の平均) です。相関係数 r は

- $-1 \leq r \leq 1$
- 1 に近いほど強い正の相関がある。
- -1 に近いほど強い負の相関がある。
- 0 に近いほど相関がない。

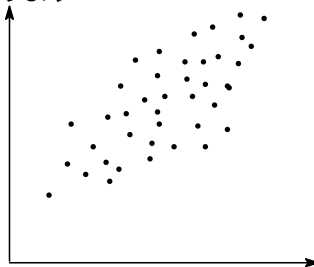
という性質をもっています。また相関の強さの目安としては以下ようになります。

$0 \sim \pm 0.2$	ほとんど相関なし
$\pm 0.2 \sim \pm 0.4$	やや相関あり
$\pm 0.4 \sim \pm 0.7$	中程度の相関あり
$\pm 0.7 \sim \pm 0.9$	高い相関あり
$\pm 0.9 \sim \pm 1$	非常に高い相関あり

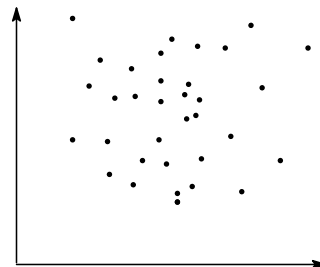
これを図で表すと以下ようになります。



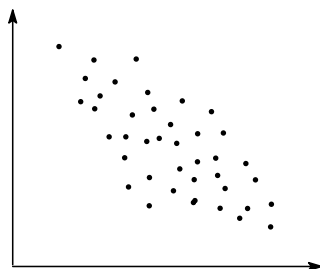
強い正の相関 ($r \doteq 1$)



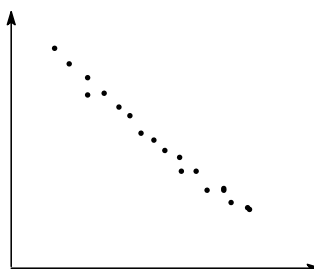
正の相関 ($r > 0$)



無相関 ($r = 0$)



負の相関 ($r < 0$)



強い負の相関 ($r \doteq -1$)

【例 11.2】 【例 11.1】のデータに対する相関係数を求めてみましょう。ここも、以下のように表を用いて計算すると、比較的楽に計算できます。

³直線的な関係とは線形的な関係と言うこともあります。詳しくは 3.8 章を見てください。

	x	y	$x - \bar{x}$	$y - \bar{y}$	$(x - \bar{x})^2$	$(y - \bar{y})^2$	$(x - \bar{x})(y - \bar{y})$
	9	5	1	-2.6	1	6.76	-2.6
	5	9	-3	1.4	9	1.96	-4.2
	8	7	0	-0.6	0	0.36	0
	8	9	0	1.4	0	1.96	0
	10	8	2	0.4	4	0.16	0.8
計	40	38	0	0	14	11.2	-6

これより

$$r = \frac{-6}{\sqrt{14}\sqrt{11.2}} \doteq -0.479.$$

問 11.3 11.4 章の表 11.2 のデータの相関係数を求めよ.

また x の標本分散 $s_x^2 = \frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})^2$, y の標本分散 $s_y^2 = \frac{1}{n} \sum_{i=1}^n (y_i - \bar{y})^2$, x と y の標本共分散 $s_{xy} = \frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})$ を用いると, 相関係数は $r = \frac{s_{xy}}{s_x s_y}$ と表すことができます.

【生徒】 相関係数はどうして (11.3) のような形になるの?

【先生】 以前, 標準化という単位の影響を受けない数値の出し方を学びました. ここではこれを利用して, 単位の影響を受けないで, かつ 2 つの関連性を表す指標を考えます.

① 対のデータ $(x_1, y_1), (x_2, y_2), \dots, (x_n, y_n)$ に対して標準化をおこなう. つまり $a_i = \frac{x_i - \bar{x}}{s_x}$, $b_i = \frac{y_i - \bar{y}}{s_y}$ ($i = 1, \dots, n$) とすることです.

② ①の操作で得られた標準化された対のデータ $(a_1, b_1), (a_2, b_2), \dots, (a_n, b_n)$ の共分散 r を求める. つまり

$$r = \frac{1}{n} \sum_{i=1}^n (a_i - \bar{a})(b_i - \bar{b}) \quad (11.4)$$

ということです. ここで $\bar{a} = \frac{1}{n} \sum_{i=1}^n a_i$, $\bar{b} = \frac{1}{n} \sum_{i=1}^n b_i$ ですが, (10.1) 式より $\bar{a} = 0$, $\bar{b} = 0$ がわかりますので

$$\begin{aligned} r &= \frac{1}{n} \sum_{i=1}^n (a_i - \bar{a})(b_i - \bar{b}) = \frac{1}{n} \sum_{i=1}^n a_i b_i = \frac{1}{n} \sum_{i=1}^n \frac{(x_i - \bar{x})}{s_x} \frac{(y_i - \bar{y})}{s_y} \\ &= \frac{1}{n s_x s_y} \sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y}) \\ &= \frac{\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^n (x_i - \bar{x})^2} \sqrt{\sum_{i=1}^n (y_i - \bar{y})^2}} \end{aligned}$$

となります. つまり相関係数は標準化されたデータの共分散ということがわかりました.

問 11.4 ある高校生 10 人の 1 日平均の勉強時間 (x) と年間欠席数 (y) を調査し, それらを組 (x, y) にしたところ以下のような結果が得られた.

$(2, 3), (3, 1), (4, 5), (2, 2), (3, 6), (1, 1), (2, 3), (1, 2), (1, 3), (3, 1)$

ここで $\sqrt{261} = 16.15$, $\sqrt{6} = 2.45$ として計算せよ.

(1) x と y の平均 \bar{x} と \bar{y} を求めよ.

(2) x と y の標準偏差 s_x, s_y を求めよ.

(3) x と y の相関係数 r を求めよ.

12 主成分分析

12.1 分散共分散行列

次の章で、主成分分析という統計手法を説明する訳ですが、その前に分散共分散行列を紹介します。今、2変量データ (x_i, y_i) , $(i = 1, \dots, n)$ が得られたとしましょう。この時、分散共分散行列は

$$\Sigma = \begin{pmatrix} x \text{ の分散} & x \text{ と } y \text{ の共分散} \\ x \text{ と } y \text{ の共分散} & y \text{ の分散} \end{pmatrix} \quad (12.1)$$

と定義します。

【例 12.1】 【例 11.2】 (p28) のデータにおいて $s_x^2 = \frac{14}{5} = 2.8$, $s_y^2 = \frac{11.2}{5} = 2.24$, $s_{xy} = \frac{-6}{5} = -1.2$ となる。よって分散共分散行列は

$$\Sigma = \begin{pmatrix} 2.8 & -1.2 \\ -1.2 & 2.24 \end{pmatrix} \quad (12.2)$$

となります。また 3 次元データ (x_i, y_i, z_i) , $(i = 1, \dots, n)$ についての分散共分散行列も同様にして

$$\Sigma = \begin{pmatrix} x \text{ の分散} & x \text{ と } y \text{ の共分散} & x \text{ と } z \text{ の共分散} \\ x \text{ と } y \text{ の共分散} & y \text{ の分散} & y \text{ と } z \text{ の共分散} \\ x \text{ と } z \text{ の共分散} & y \text{ と } z \text{ の共分散} & z \text{ の分散} \end{pmatrix} \quad (12.3)$$

と定義します。以下 n 変数の分散共分散行列も同様に定義します。

次に相関行列を説明します。今、2 変量データ (x_i, y_i) , $(i = 1, \dots, n)$ に対する相関行列は

$$R = \begin{pmatrix} 1 & x \text{ と } y \text{ の相関係数} \\ x \text{ と } y \text{ の相関係数} & 1 \end{pmatrix} \quad (12.4)$$

と定義します。3 変量の場合も同様にして

$$R = \begin{pmatrix} 1 & x \text{ と } y \text{ の相関係数} & x \text{ と } z \text{ の相関係数} \\ x \text{ と } y \text{ の相関係数} & 1 & y \text{ と } z \text{ の相関係数} \\ x \text{ と } z \text{ の相関係数} & y \text{ と } z \text{ の相関係数} & 1 \end{pmatrix} \quad (12.5)$$

と定義します。

【例 12.2】 【例 11.2】 のデータにおいて、相関係数 $r = -0.479$ より相関行列 R は

$$R = \begin{pmatrix} 1 & -0.479 \\ -0.479 & 1 \end{pmatrix} \quad (12.6)$$

となる。

11.5 章 (p27-28) で述べたように、対のデータ (x_i, y_i) , $(i = 1, \dots, n)$ を標準化した変数を $a_i = \frac{x_i - \bar{x}}{s_x}$, $b_i = \frac{y_i - \bar{y}}{s_y}$ ($i = 1, \dots, n$) とする時

- a_i ($i = 1, \dots, n$) の標本分散は $s_a^2 = 1$
- b_i ($i = 1, \dots, n$) の標本分散は $s_b^2 = 1$
- (a_i, b_i) ($i = 1, \dots, n$) の共分散は $s_{ab} = \frac{1}{n} \sum_{i=1}^n (a_i - \bar{a})(b_i - \bar{b}) = r$ (x と y の相関係数)

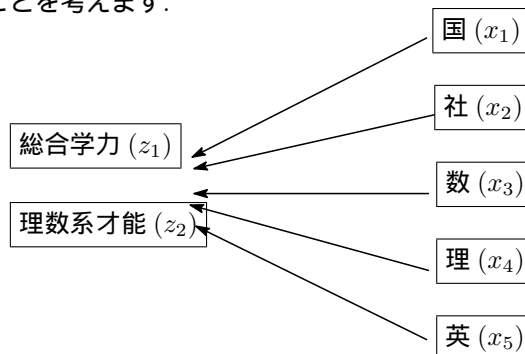
が分かりました。つまり相関行列は標準化したデータの分散共分散行列であることを思い出してください。

問 12.1 11.4 章の表 11.2(p26) のデータにおける分散共分散行列と相関行列を求めよ。

12.2 主成分分析

まず主成分分析がどのようなものであるかを説明します。ある学校の 50 人分の国語 x_1 , 社会 x_2 , 数学 x_3 , 理科 x_4 , 英語 x_5 のデータが得られたとします。主成分分析とは, 5 次元のデータを”総合的な学力”や”理系の能力”といった情報にまとめる技術です。数学的に書きますと, p 次元の変数 x_1, x_2, \dots, x_p を以下のように組み合わせて, 圧縮した情報 z_1, \dots, z_m ($m < p$) を取り出すことを考えます。

$$\begin{aligned} z_1 &= a_{11}x_1 + a_{12}x_2 + \cdots + a_{1p}x_p \\ z_2 &= a_{21}x_1 + a_{22}x_2 + \cdots + a_{2p}x_p \\ &\vdots \\ z_m &= a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mp}x_p \end{aligned}$$



まず簡単な例を説明します。以下のデータは 10 人の身長 w_1 と体重 w_2 とします。身長を \bar{w}_1 , 身長の標準偏差 s_{w_1} とし, 体重の平均を \bar{w}_2 , 体重の標準偏差 s_{w_2} として, 標準化 $x_{1i} = \frac{w_{1i} - \bar{w}_1}{s_{w_1}}$, $x_{2i} = \frac{w_{2i} - \bar{w}_2}{s_{w_2}}$, ($i = 1, 2, \dots, 10$) を行います。

出席番号	w_1	w_2	出席番号	x_1 (身長)	x_2 (体重)
1	172	63	1	0.5021	0.2717
2	173	64	2	0.7663	0.4981
3	172	61	3	0.5021	-0.1811
4	171	62	4	0.2378	0.0453
5	171	64	5	0.2378	0.4981
6	166	56	6	-1.0834	-1.3131
7	162	54	7	-2.1403	-1.7658
8	168	63	8	-0.5549	0.2717
9	171	61	9	0.2378	-0.1811
10	175	70	10	1.2948	1.8564

表 12.1: 身長, 体重の z-score

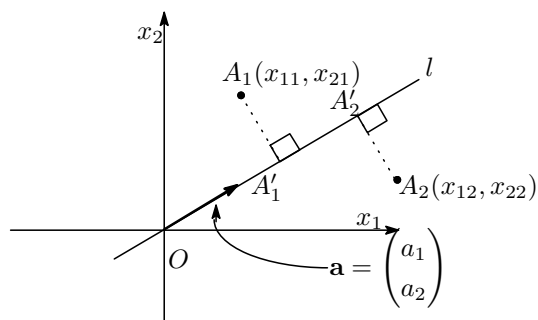
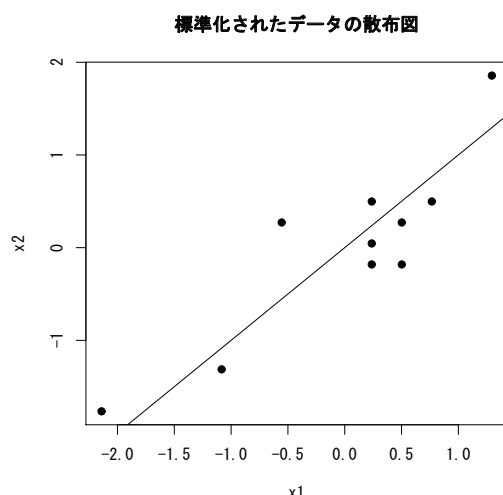
ここで $a_1^2 + a_2^2 = 1$ とし, 身長 x_1 と体重 x_2 を 1 つの指標

$$z_1 = a_1 \times \text{身長}(x_1) + a_2 \times \text{体重}(x_2) \quad (12.7)$$

で表すことを考えます。さて z_1 はどのように決めればよいのでしょうか? この時, 以下の手順で 2 次元データ (x_{1i}, x_{2i}) , ($i = 1, 2, \dots, n$) を 1 次元の数で表します。

- ① 散布図をうまく表す軸 (直線) を引く
- ② 各点から軸 l に降ろした垂線の足と原点 O の 符号付きの距離⁴を考える。

⁴原点からの向きにより, 正の値と負の値をとる距離。【例 4.3】参照。



今, 右上の図の点 $A_1(x_{11}, x_{21})$ における⁵量 $z_{11} = a_1x_{11} + a_2x_{21}$ を考えてみましょう. ここで, 散布図をうまく表している軸の方向ベクトルを $\mathbf{a} = \begin{pmatrix} a_1 \\ a_2 \end{pmatrix}$ とすると, OA'_1 は定理 4.1 から

$$OA'_1 = a_1x_{11} + a_2x_{21} \quad (12.8)$$

と表すことが出来ます. これは (12.7) の形をしていることがわかります. 同様に点 $A_2(x_{12}, x_{22})$ においても $z_{12} = a_1x_{12} + a_2x_{22} = OA'_2$ となります. このように OA'_1 や OA'_2 のことを第 1 主成分スコアと言い, 方向ベクトル \mathbf{a} を主成分もしくは主成分ベクトルと言います.

次にデータからどのようにして第 1 主成分の方向ベクトル $\mathbf{a} = \begin{pmatrix} a_1 \\ a_2 \end{pmatrix}$ を求めるかを説明します. ここで原点とそれぞれの点の垂線の足までの符号付きの長さ

$$OA'_1, OA'_2, \dots, OA'_n,$$

を考え,

$$f(a_1, a_2) = \frac{1}{n} \left((OA'_1)^2 + (OA'_2)^2 + \dots + (OA'_n)^2 \right) = \frac{1}{n} \sum_{i=1}^n (a_1x_{1i} + a_2x_{2i})^2 \quad (12.9)$$

を最大にするような a_1, a_2 を求めます.

ただし 4.3 章で説明されたように, 直線を表す a_1, a_2 が無数に存在しますので制限 $a_1^2 + a_2^2 = 1$ を入れます. これはラグランジュ未定乗数法より

$$F(a_1, a_2, \lambda) = \frac{1}{n} \sum_{i=1}^n (a_1x_{1i} + a_2x_{2i})^2 - \lambda(a_1^2 + a_2^2 - 1) \quad (12.10)$$

とおきます. ここで $\bar{x}_1 = 0, \bar{x}_2 = 0$ ですので, x_{1i} の分散を $s_{11} = \frac{1}{n} \sum_{i=1}^n x_{1i}^2$, x_{2i} の分散を $s_{22} = \frac{1}{n} \sum_{i=1}^n x_{2i}^2$,

x_{1i} と x_{2i} の共分散を $s_{12} = \frac{1}{n} \sum_{i=1}^n x_{1i}x_{2i}$ とおくと,

$$F(a_1, a_2, \lambda) = s_{11}a_1^2 + 2s_{12}a_1a_2 + s_{22}a_2^2 - \lambda(a_1^2 + a_2^2 - 1) \quad (12.11)$$

と表すことができます. これを偏微分すると

⁵このような具体的な $A_1(x_{11}, x_{21})$ における主成分は厳密には主成分得点と言います

$$\frac{\partial}{\partial a_1} F(a_1, a_2, \lambda) = 2a_1 s_{11} + 2a_2 s_{12} - 2\lambda a_1 = 0 \quad (12.12)$$

$$\frac{\partial}{\partial a_2} F(a_1, a_2, \lambda) = 2a_1 s_{12} + 2a_2 s_{22} - 2\lambda a_2 = 0 \quad (12.13)$$

$$a_1^2 + a_2^2 = 1 \quad (12.14)$$

となります。ここで式 (12.12), (12.13) は行列を用いると

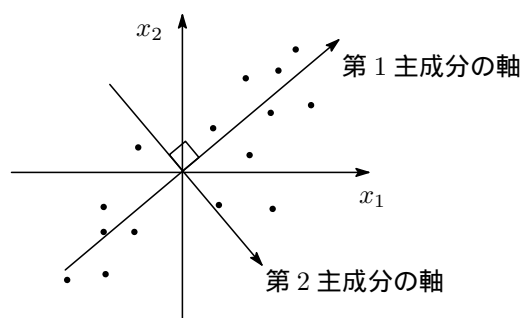
$$\begin{pmatrix} s_{11} & s_{12} \\ s_{12} & s_{22} \end{pmatrix} \begin{pmatrix} a_1 \\ a_2 \end{pmatrix} = \lambda \begin{pmatrix} a_1 \\ a_2 \end{pmatrix} \quad (12.15)$$

と表すことが出来ます。実は、 (x_{1i}, x_{2i}) , $(i = 1, 2, \dots, n)$ は標準化されたデータなので、 $s_{11} = s_{22} = 1$, $s_{12} = r$ (w_1 と w_2 の相関係数) となるので、主成分の係数 a_1, a_2 は相関行列 R の大きさ 1 の固有ベクトルを求めればよいことが分かります。

データが 2 次元の時、(12.15) の固有値、固有ベクトルは 2 つ出てきます。ここで

- 最も大きな固有値に対する固有ベクトルを第 1 主成分
- 2 番目に大きな固有値に対する固有ベクトルを第 2 主成分

と定義します。主成分は軸の方向を表すベクトルであることに注意してください。また第 2 主成分は、以下のグラフのように、第 1 主成分を表す軸に直交する形で現れます。



またデータが 3 次元の場合は、第 1 主成分から第 3 主成分まで現れ、全ての軸が直交している形となります。

【例 12.3】表 12.1 の標準化されたデータ (x_{1i}, x_{2i}) , $(i = 1, 2, \dots, 10)$ の分散共分散行列は $R = \begin{pmatrix} 1 & 0.8854 \\ 0.8854 & 1 \end{pmatrix}$

であった。ここで R の固有値は 1.8854, 0.1146 となり、それに対応する第 1 主成分は $\begin{pmatrix} \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \end{pmatrix}$ 、第 2 主成分は $\begin{pmatrix} \frac{1}{\sqrt{2}} \\ -\frac{1}{\sqrt{2}} \end{pmatrix}$ となる。

注意

相関行列 $R = \begin{pmatrix} 1 & r \\ r & 1 \end{pmatrix}$ に対する固有値は $1+r, 1-r$ になり、それに対応する大きさ 1 の固有ベクトルは $\begin{pmatrix} \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \end{pmatrix}, \begin{pmatrix} \frac{1}{\sqrt{2}} \\ -\frac{1}{\sqrt{2}} \end{pmatrix}$ となることが知られている。

【学生】 あれ？ 1 つの固有値に対応する大きさ 1 の固有ベクトルは 2 つ出てくるんじゃないですか？

【先生】 そうですね。例えば、上の注意においても固有値 $1+r$ に対する大きさ 1 の固有ベクトルは $\begin{pmatrix} \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \end{pmatrix}$ と $\begin{pmatrix} -\frac{1}{\sqrt{2}} \\ -\frac{1}{\sqrt{2}} \end{pmatrix}$ が出てきます。しかし実際 $\begin{pmatrix} \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \end{pmatrix}$ で表わされる直線は $y = x$ ですし、 $\begin{pmatrix} -\frac{1}{\sqrt{2}} \\ -\frac{1}{\sqrt{2}} \end{pmatrix}$ で表わ

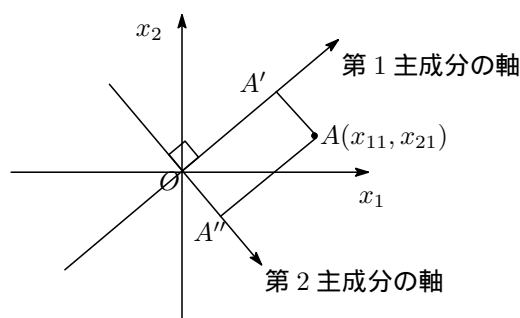
される直線も $y = x$ です。つまりどちらに向いている直線 (軸) を採用するかということになりますので、どちらを採用しても構わないということになります。⁶

問 12.2 問 11.4 の相関行列 $R = \begin{pmatrix} 1 & 0.42 \\ 0.42 & 1 \end{pmatrix}$ に対する主成分と固有値を求めよ。

12.3 主成分得点 (主成分スコア)

ここでは 2 次元の情報を 1 次元の数値を用いて表すことを目的に説明します。

前の章で、主成分を求めました。この時、 (x_{1i}, x_{2i}) から第 1 主成分の軸に降ろした垂線の足と原点 O の符号付きの距離を用いれば、2 次元の情報を 1 次元の数値で表すことができることがわかると思います。実は、この符号付きの距離のことを (x_{1i}, x_{2i}) の第 1 主成分得点 (スコア) と言います。



OA' が A の第 1 主成分スコア

OA'' が A の第 2 主成分スコア

【例 12.4】 【例 12.3】 のデータにおいて第 1 主成分は $\mathbf{a}^* = \begin{pmatrix} 1/\sqrt{2} \\ 1/\sqrt{2} \end{pmatrix}$ ですので、出席番号 1 番の生徒 $(0.5021, 0.2717)$ の第 1 主成分スコアは、表 12.1 と定理 4.1 を用いると

$$\frac{1}{\sqrt{2}} \times 0.5021 + \frac{1}{\sqrt{2}} \times 0.2717 = 0.5472$$

となります。出席番号 1 番の第 2 主成分は $\mathbf{a}^{**} = \begin{pmatrix} 1/\sqrt{2} \\ -1/\sqrt{2} \end{pmatrix}$ ですので、第 2 主成分スコアは

$$\frac{1}{\sqrt{2}} \times 0.5021 + \left(-\frac{1}{\sqrt{2}}\right) \times 0.2717 = 0.1629$$

となります。これを出席番号 2 番から 10 番まで同様の計算を行うと以下の表を得ることができます。

出席番号	x_1 (身長)	x_2 (体重)	第 1 主成分得点	第 2 主成分得点
1	0.5021	0.2717	0.5472	0.1629
2	0.7663	0.4981	0.8941	0.1896
3	0.5021	-0.1811	0.227	0.4831
4	0.2378	0.0453	0.2002	0.1361
5	0.2378	0.4981	0.5204	-0.1841
6	-1.0834	-1.3131	-1.6946	0.1624
7	-2.1403	-1.7658	-2.762	-0.2648
8	-0.5549	0.2717	-0.2003	-0.5845
9	0.2378	-0.1811	0.0401	0.2962
10	1.2948	1.8564	2.2282	-0.3971

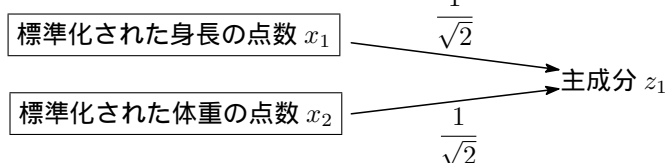
主成分得点から分かること 各個体 (ここでは生徒) の主成分の大きさの度合いがわかります。よって第 1 主成分スコアは $z_1 = \frac{1}{\sqrt{2}}x_1 + \frac{1}{\sqrt{2}}x_2$ と表せ、第 2 主成分スコアは $z_2 = \frac{1}{\sqrt{2}}x_1 - \frac{1}{\sqrt{2}}x_2$ と表せます。⁷

⁶本質的なものはまったく同じですが、後で紹介する主成分得点はどちらの主成分を採用するかによって符号だけが変わりますので注意してください

⁷しばしばこの式を主成分と呼ぶことがあるので注意

12.4 主成分の解釈

さて主成分は何を表しているのでしょうか？ 実は主成分の解釈は解析者にまかされているのです。例えば、【例 12.3】における第 1 主成分であれば、以下の図のようになります。



この場合、身長 z_{score} が高くても、体重の z_{score} が高くても主成分 z_1 は増えるので、 z_1 は「体格の良さ」と解釈することができます。これより最も第 1 主成分得点が高いのは 10 番の生徒なので、この生徒が最も体格が良いと言えます。

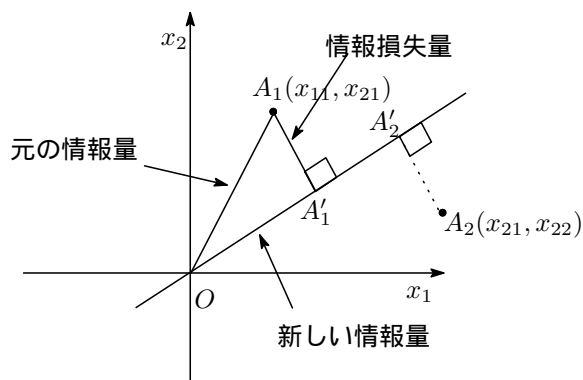
【学生】 ということは主成分の解釈というのは、自分で勝手に決めていいということでしょうか？

【先生】 まあ、平たく言えばそういうことです。ですので、いかにうまい解釈を与えるかが解析者の腕の見せ所となります。⁸

12.5 寄与率

寄与率は、主成分が元のデータの情報をどのくらい保持しているかを考えます。これは情報の損失という観点から定義します。

第 1 主成分に対応する固有ベクトルを $\mathbf{a}^* = \begin{pmatrix} a_1^* \\ a_2^* \end{pmatrix}$ ，その固有値を λ^* とします。ここで新しい情報量を定義します。これは主成分によって説明できる量ということで、点 A_1 の新しい情報量であれば長さ $OA'_1 = a_1^*x_{11} + a_2^*x_{21}$ で定義します。同様に元の情報量は長さ $OA_1 = \sqrt{x_{11}^2 + x_{21}^2}$ によって定義します。



この時、第 1 主成分の寄与率は

$$\text{主成分 } \mathbf{a}^* \text{ の寄与率} = \frac{\text{それぞれの点の”新しい情報量の長さの 2 乗”を全て足したもの}}{\text{それぞれの点の”元の情報量の長さの 2 乗”を全て足したもの}}$$

と定義します。これより

$$\begin{aligned} \text{主成分 } \mathbf{a}^* \text{ の寄与率} &= \frac{(OA'_1)^2 + \cdots + (OA'_n)^2}{(OA_1)^2 + \cdots + (OA_n)^2} = \frac{\sum_{i=1}^n (a_1^*x_{1i} + a_2^*x_{2i})^2}{\sum_{i=1}^n (x_{1i}^2 + x_{2i}^2)} \\ &= \frac{\begin{pmatrix} a_1^* & a_2^* \end{pmatrix} \begin{pmatrix} s_{11} & s_{12} \\ s_{12} & s_{22} \end{pmatrix} \begin{pmatrix} a_1^* \\ a_2^* \end{pmatrix}}{\frac{1}{n} \sum_{i=1}^n (x_{1i}^2 + x_{2i}^2)} \quad (R\mathbf{a}^* = \lambda^* \mathbf{a}^* \text{ より}) \\ &= \frac{\begin{pmatrix} a_1^* & a_2^* \end{pmatrix} \lambda^* \begin{pmatrix} a_1^* \\ a_2^* \end{pmatrix}}{\frac{1}{n} \sum_{i=1}^n (x_{1i}^2 + x_{2i}^2)} \quad (a_1^* + a_2^* = 1 \text{ より}) \\ &= \frac{\lambda^*}{2} \end{aligned} \tag{12.16}$$

⁸ただし主成分の意味付けが難しいときもある。そのようなときは無理やり意味を付けない方がよい

となります。つまり

大きな固有値を持つ主成分 \iff 大きな寄与率を持つ主成分

がわかります。これより最も寄与率が高い主成分のことを第 1 主成分、2 番目に大きい主成分のことを第 2 主成分と考えられます。

また今後、相関行列 R の対角成分を全て足し合わせたものを $tr(R)$ と表します。即ち

$$\frac{1}{n} \sum_{i=1}^n (x_{1i}^2 + x_{2i}^2) = \frac{1}{n} \sum_{i=1}^n x_{1i}^2 + \frac{1}{n} \sum_{i=1}^n x_{2i}^2 = s_{11} + s_{22} = 2 = tr(R)$$

となります。これより、寄与率 (12.16) の分母の部分は $tr(R)$ と表せます。

【例 12.5】 【例 12.3】において

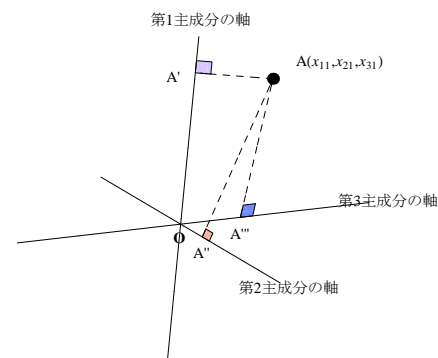
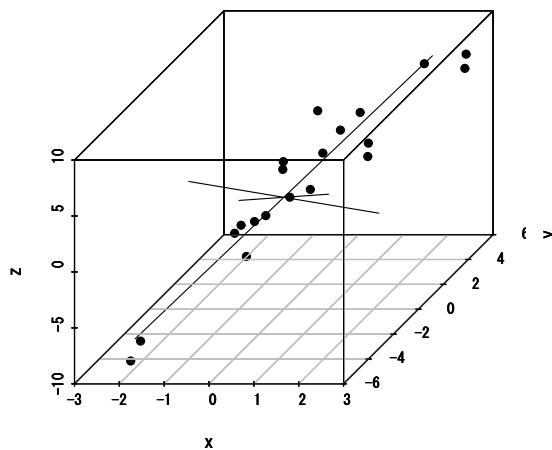
$$\text{第 1 主成分の寄与率} = \frac{1.8854}{1+1} = 0.9427 \quad (12.17)$$

これより第 1 主成分で、データの約 94% が説明されていることが分かる。

問 12.3 問 12.1 のデータにおける第 1 主成分の寄与率を求めよ。

13 3 変数以上の主成分分析

今までは 2 変数のデータを扱ってきましたが、下図のように 3 次元の場合であれば、散布図をうまく表す軸は 3 つ現れるので、主成分も 3 つ出てきます。



右上の図から分かるように、点 A からそれぞれの主成分を表す軸に垂線を降ろすことができます。ここで OA' を A の第 1 主成分スコア、 OA'' を A の第 2 主成分スコア、 OA''' を A の第 3 主成分スコアと考えればよいということになります。数学的な説明は 2 次元データの場合と同様に行えますが、ここではより一般的な p 次元データに対する主成分分析について説明します。⁹

① 与えられたデータに対して標準化を行う

② 相関行列 R を求める。

③ R の固有値 $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_p$ とそれに対応する主成分 (固有ベクトル) a_1, \dots, a_p を求める。この時、 λ_1 に対する a_1 を第 1 主成分、 λ_2 に対する a_2 を第 2 主成分、 λ_k に対する a_k を第 k 主成分となる

④ 主成分の寄与率を求める。

第 1 主成分の寄与率は $\frac{\lambda_1}{tr(R)}$ であり、第 2 主成分の寄与率は $\frac{\lambda_2}{tr(R)}$ 、第 3 主成分の寄与率は $\frac{\lambda_3}{tr(R)}$ となる

⁹ グラフによる説明は 3 次元までが限界なので、後は数式による説明となる。

が、第 1 主成分から第 k 主成分までで、データのどのくらいを説明できたかを見るために累積寄与率というのを考えます。

$$\text{第 1 主成分から第 } k \text{ 主成分までの累積寄与率} = \frac{\lambda_1 + \cdots + \lambda_k}{\text{tr}(R)} = \frac{\lambda_1 + \cdots + \lambda_k}{p}$$

実は $\text{tr}(R) = \lambda_1 + \cdots + \lambda_p$ (この証明は省略) となることが知られているので、全ての主成分の累積寄与率は $\frac{\lambda_1 + \cdots + \lambda_p}{\text{tr}(R)} = 1$ となります。

【例 13.1】第 1 主成分と第 2 主成分による累積寄与率は $\frac{\lambda_1 + \lambda_2}{\text{tr}(R)}$ 。

さてデータを圧縮する場合、何個の主成分を採用すればよいのでしょうか？ いくつかの方法がありますが、以下の 2 つのどちらかがよく使われます。¹⁰

方法 1 累積寄与率が 80% 以上大きくなるようにする。

方法 2 固有値が 1 以上の主成分 (固有ベクトル) を採用 (これは標準化したデータに対してのみ使用できる)。

【例 13.2】20 人の生徒に対する国語、社会、数学、物理、英語の 5 次元データが得られたとする。

	国語	社会	数学	物理	英語
相川 英之	91	98	63	72	88
吾妻 博	57	70	83	79	63
阿部 美穂	63	64	78	82	66
飯田 卓也	66	67	67	40	82
上原 広幸	55	66	70	59	66
内田 弓子	95	95	81	76	91
遠藤 恵	64	79	73	69	69
岡田 啓吾	91	96	79	74	87
加藤 玲子	85	89	69	57	69
川端 洋二	91	95	85	79	86
北川 俊夫	66	70	63	67	71
久米 浩二	54	63	77	74	77
後藤 健一	75	72	66	50	73
斉藤 正志	84	76	80	79	83
佐藤 隆志	74	67	66	52	78
佐藤 愛	86	84	76	70	73
桜井 幸信	80	84	84	66	84
島貫 寛一朗	61	61	60	44	79
菅原 美紀	76	76	79	85	85
瀬川 潤	50	54	55	58	81

ここで相関行列は

$$R = \begin{pmatrix} 1 & 0.9066 & 0.3867 & 0.2855 & 0.5878 \\ 0.9066 & 1 & 0.4564 & 0.3909 & 0.4695 \\ 0.3867 & 0.4564 & 1 & 0.7086 & 0.1093 \\ 0.2855 & 0.3909 & 0.7086 & 1 & 0.1046 \\ 0.5878 & 0.4695 & 0.1093 & 0.1046 & 1 \end{pmatrix} \quad (13.1)$$

となるので、統計処理ソフトを用いて固有値、固有ベクトルを求めると以下のようになります。

¹⁰ただし理論的な根拠はありません。

	第 1 主成分	第 2 主成分	第 3 主成分	第 4 主成分	第 5 主成分
国語	0.526	0.309	0.326	-0.013	0.722
社会	0.535	0.177	0.451	-0.166	-0.672
数学	0.412	-0.531	-0.059	0.737	-0.032
物理	0.377	-0.577	-0.31	-0.647	0.1
英語	0.353	0.508	-0.769	0.102	-0.126
固有値	2.8266	1.2878	0.526	0.2829	0.0768
寄与率	0.565	0.258	0.105	0.0566	0.0154
累積寄与率	0.565	0.823	0.928	0.9847	1.0000

表 13.1

ここで第 1 主成分はすべての教科で正の値となっているので”総合的な学力”と考えることができ、第 2 主成分は文系科目が正の値を値となっているので、これは”文系の能力”と考えることができる。

また第 2 主成分までの累積寄与率は

$$\frac{\lambda_1 + \lambda_2}{tr(R)} = \frac{2.8266 + 1.2878}{5} \div 0.823$$

より 80% 以上になっていることがわかる。

次に主成分得点（スコア）を求めてみましょう。これは

	国語	社会	数学	物理	英語
平均	73.20	76.30	72.70	66.60	77.55
標準偏差	13.87	12.81	8.52	12.67	8.13

より相川君の第 1 主成分得点 z_{11} は

$$\begin{aligned}
z_{11} &= 0.526 \times \underbrace{\frac{91 - 73.2}{13.87}}_{\text{国語の zscore}} + 0.535 \times \underbrace{\frac{98 - 76.3}{12.81}}_{\text{社会の zscore}} + 0.412 \times \underbrace{\frac{63 - 72.70}{8.52}}_{\text{数学の zscore}} \\
&\quad + 0.377 \times \underbrace{\frac{72 - 66.60}{12.67}}_{\text{物理の zscore}} + 0.353 \times \underbrace{\frac{88 - 77.55}{8.13}}_{\text{英語の zscore}} \\
&\div 1.727
\end{aligned}$$

となります。ここで上の 0.526 や 0.535 のような zscore の前にある係数は、第 1 主成分から計算します。上の表 13.1 からわかるように、

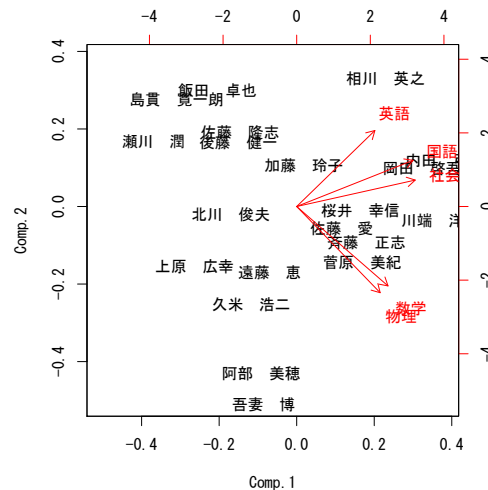
第 1 主成分の国語の部分×相川君の国語の zscore+…+第 1 主成分の英語の部分×相川君の英語の zscore

という対応をさせながら計算します。また相川君の第 2 主成分得点 z_{12} も同様に計算します。

$$\begin{aligned}
z_{12} &= 0.309 \times \underbrace{\frac{91 - 73.2}{13.87}}_{\text{国語の zscore}} + 0.177 \times \underbrace{\frac{98 - 76.3}{12.81}}_{\text{社会の zscore}} - 0.531 \times \underbrace{\frac{63 - 72.70}{8.52}}_{\text{数学の zscore}} \\
&\quad - 0.577 \times \underbrace{\frac{72 - 66.60}{12.67}}_{\text{物理の zscore}} + 0.508 \times \underbrace{\frac{88 - 77.55}{8.13}}_{\text{英語の zscore}} \\
&\div 1.709
\end{aligned}$$

13.1 Biplot(バイプロット)

それぞれの生徒の第 1 主成分得点と第 2 主成分得点を座標であると考えて、散布図を描くと生徒それぞれの特徴が分かる。これを biplot(バイプロット)と言う。



ここで Comp.1 は第 1 主成分スコア, Comp.2 は第 2 主成分スコアを表しています。

Biplot で表示されるベクトルは、大きく分けると 2 種類の表し方があります。

表し方 1 第 1 主成分、第 2 主成分を $P = (a_1, a_2)$ とした時のそれぞれの行の要素をベクトルとして表す。例えば 5 教科データの主成分 (表 7.1) であれば、

$$P = \begin{matrix} & x \text{ 成分} & y \text{ 成分} \\ \text{国語} \rightarrow & \begin{pmatrix} 0.526 & 0.309 \end{pmatrix} \\ \text{社会} \rightarrow & \begin{pmatrix} 0.535 & 0.177 \end{pmatrix} \\ \text{数学} \rightarrow & \begin{pmatrix} 0.412 & -0.531 \end{pmatrix} \\ \text{物理} \rightarrow & \begin{pmatrix} 0.377 & -0.577 \end{pmatrix} \\ \text{英語} \rightarrow & \begin{pmatrix} 0.353 & 0.508 \end{pmatrix} \end{matrix}$$

となりますので国語を表すベクトルは $b_1 = \begin{pmatrix} 0.526 \\ 0.309 \end{pmatrix}$, 社会を表すベクトルは $b_2 = \begin{pmatrix} 0.535 \\ 0.177 \end{pmatrix}$ となります。

表し方 2 次に紹介するの統計ソフト R の biplot においてデフォルト¹¹として設定されている方法です。第 1 主成分、第 2 主成分を $P = (\sqrt{\lambda_1}na_1, \sqrt{\lambda_2}na_2)$ とした時のそれぞれの行の要素をベクトルとして表す。**【例 13.2】** の 5 教科データの主成分であれば、

$$P = \begin{matrix} & x \text{ 成分} & y \text{ 成分} \\ \text{国語} \rightarrow & \begin{pmatrix} 3.952 & 1.568 \end{pmatrix} \\ \text{社会} \rightarrow & \begin{pmatrix} 4.026 & 0.897 \end{pmatrix} \\ \text{数学} \rightarrow & \begin{pmatrix} 3.101 & -2.696 \end{pmatrix} \\ \text{物理} \rightarrow & \begin{pmatrix} 2.835 & -2.927 \end{pmatrix} \\ \text{英語} \rightarrow & \begin{pmatrix} 2.655 & 2.58 \end{pmatrix} \end{matrix}$$

となるので国語を表すベクトルは $b_1 = \begin{pmatrix} 3.952 \\ 1.568 \end{pmatrix}$, 社会を表すベクトルは $b_2 = \begin{pmatrix} 4.026 \\ 0.897 \end{pmatrix}$ となります。

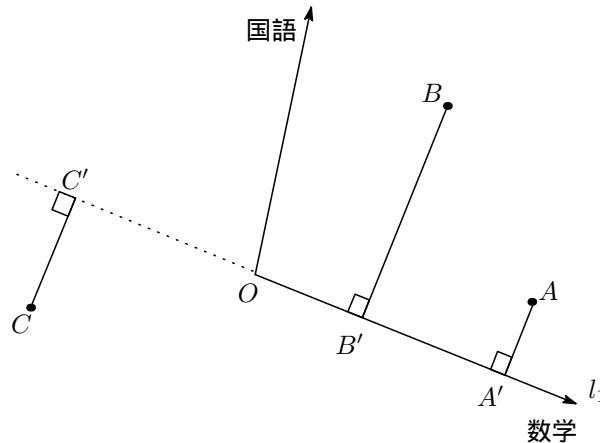
ただし方法 1 でも方法 2 でもそれほど気に留める必要はないと思います。

13.2 Biplot の解釈

さて Biplot において、このベクトルはどのような解釈すればよいのでしょうか？

今、 A, B, C という生徒があり、数学と国語を表すベクトルが以下のように与えられたとします。

¹¹何も指示されていない時には、自動的に選択される設定のことをデフォルトと言います



ここで点 A, B から数学を表すベクトルの直線に垂線を降ろし、その交点を A', B' とします。この時、原点から垂線の足までの符号付きの長さ OA', OB' が 2 人の能力を表していると考えます。例えば、上の図において $OA' > OB'$ ですので、 A 君の方が B 君よりも数学が出来ると解釈します。

注意 1 ただし " $OA' > OB' \implies$ 必ず A 君の方が B 君より数学ができる" とはなりませんので注意してください。これは多次元のデータを 2 次元に圧縮しており、誤差があるからです。

注意 2 この OA' や OB' は数学を表すベクトルと同じ向きであれば正、反対方向であれば負の値をとるものとします。例えば、上の図において生徒 C から直線 l_1 に垂線を降ろし、その交点を C' とします。この時、 OC' は OB' より長さとしては大きいですが、 OB' は正の値、 OC' は負の値となります。これより $OB' > OC'$ となります。つまり B 君の方が C 君より数学が出来るということになります。

さてこのように考えると、13.1 章の Biplot より以下のことがわかります。

- それぞれの個体がどの個体と特徴が似ているかがわかる。(例えば、内田弓子と岡田啓吾は似た特徴を持っている)
- この例では一番右側にある生徒が最も総合力が高いことがわかる
- biplot のベクトルの向きより、縦軸は文系向きかどうかを表している。即ち上に名前があるほど純粋に文系向きで、下に名前があるほど純粋に理系向きということである。
- 吾妻君は理系科目はよくできるが、文系科目はとても悪い etc

問 13.1 100 個の 3 次元データ (x, y, z) から相関行列を作成したところ

$$\begin{pmatrix} 1 & \frac{1}{2} & -\frac{1}{2} \\ \frac{1}{2} & 1 & 0 \\ -\frac{1}{2} & 0 & 1 \end{pmatrix} \quad (13.2)$$

となった。この時、このデータの主成分と寄与率を全て求めよ。また第 2 主成分までの累積寄与率を求めよ。

問 13.2 各国の国民一人当たりの石炭 (kg), 原油 (kg), 天然ガス (100 万ジュール), 電力 (kwh) の使用料を調べたところ以下の結果を得た。(総務省統計局 [2] より引用)

	石炭 (kg)	原油 (kg)	天然ガス	電力 (kwh)
日本	1416	1576	26334	8459
イラン	25	1101	49258	2460
インド	375	120	934	618
インドネシア	65	232	5944	476
韓国	1708	2348	24031	7716
タイ	118	733	15957	2020
中国	1480	223	1119	1684
トルコ	267	368	12312	2122
バングラデシュ	5	9	3360	154
フィリピン	116	127	1185	686
マレーシア	564	1117	49842	3476
ミャンマー	2	21	1288	129
アメリカ合衆国	1710	2723	81341	14240
カナダ	565	2561	113674	18408
キューバ	1	428	2423	1380
メキシコ	17	693	19416	2130
南アメリカ	80	583	11064	2149
アルゼンチン	25	711	39507	2714
ブラジル	109	473	4038	2340
イギリス	1019	1382	68095	6756
イタリア	419	1584	53054	6029
デンマーク	1357	1486	39937	6967
ドイツ	823	1355	44378	7442
フランス	335	1390	30156	8231
ブルガリア	547	678	14864	4582
ロシア	1008	1330	112108	6425
アルジェリア	17	522	23819	889
エジプト	27	481	16699	1465

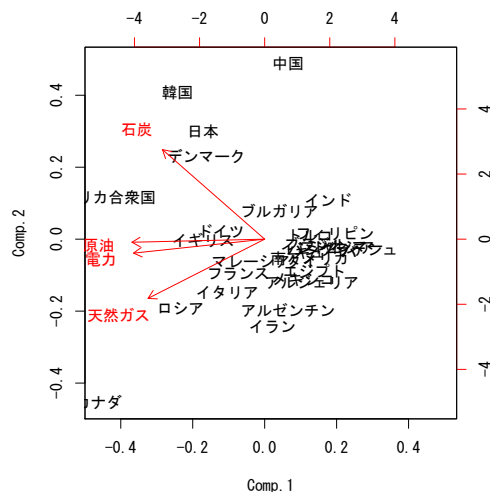
このデータに対して主成分分析を行ったところ以下の結論を得た.

	Comp.1	Comp.2	Comp.3	Comp.4
石炭	-0.42	0.83	-0.36	0.09
原油	-0.55	-0.03	0.39	-0.74
天然ガス	-0.48	-0.55	-0.69	0.02
電力	-0.54	-0.13	0.50	0.67
固有値	3.1	0.62	0.21	0.07

- ① 第 1 主成分は何を表しているか?
- ② 第 2 主成分は何を表しているか?
- ③ 第 2 主成分までの累積寄与率を求めよ
- ④ 以下の表は石炭, 原油, 天然ガス, 電力の平均と標準偏差である. この時, 日本とアメリカの第 1 主成分得点を求め, 考察を述べよ.

	石炭	原油	天然ガス	電力
平均	507.14	941.25	30933.46	4362.39
標準偏差	563.70	739.0	31069.17	4281.69

- ⑤ 以下の biplot から見て、日本とアメリカどちらが石炭エネルギーを用いている傾向があるか？
 ⑥ 以下の biplot から、日本と最も特徴が似ている国を挙げよ。



14 R を用いた統計処理

ここではフリーの統計処理ソフト R を用いた主成分分析を説明します。尚, R のダウンロードとインストールの仕方は <http://www1.tcu.ac.jp/home1/ymiyatagbt/statR.html> に置いておくので自由に印刷してください。

14.1 データ (csv ファイル) を準備する

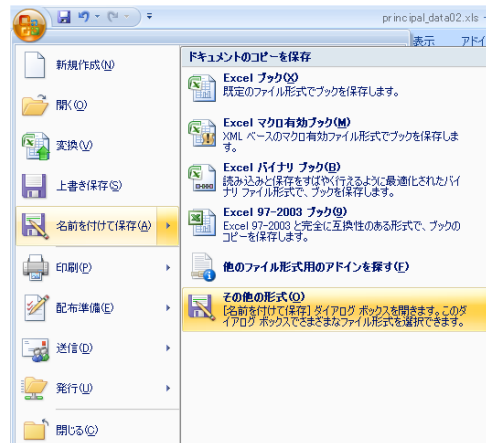
まず, 主成分分析をするためのデータを準備しなければなりません。今, 宮田のホームページからダウンロードできる Excel ファイル principal_data02.xls を使います。これを次の操作をすることにより, csv 形式¹²で c:\Program Files\R に保存します。

- ① Excel でデータを作成する, もしくは Excel ファイルを開く。

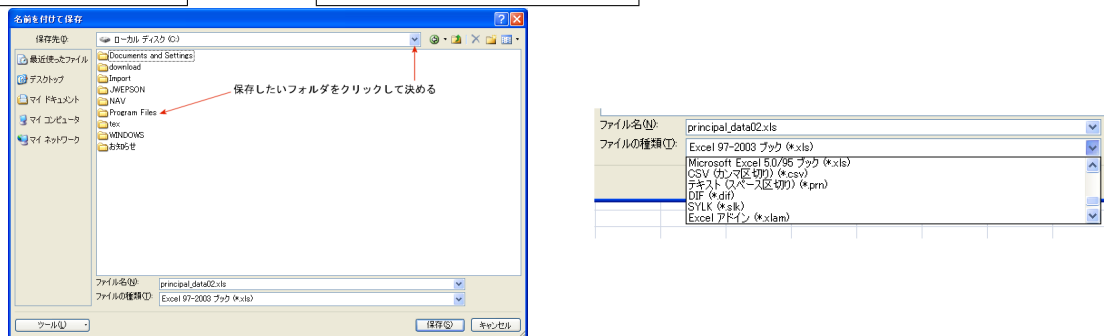
	A	B	C	D	E	F
1	name	国語	社会	数学	物理	英語
2	相川 英之	91	98	63	72	88
3	吾妻 博	57	70	83	79	63
4	阿部 美穂	63	64	78	82	66
5	飯田 卓也	66	67	67	40	82
6	上原 広幸	55	66	70	59	66
7	内田 弓子	95	95	81	76	91
8	遠藤 恵	64	79	73	69	69
9	岡田 啓吾	91	96	79	74	87
10	加藤 玲子	85	89	69	57	69
11	川端 洋二	91	95	85	79	86

- ② **スタートボタン** (Excel2007 では左上の模様のついた丸いボタン) を押し, **名前を付けて保存 (A)** を選び, **その他の形式 (O)** をクリックする。

¹²csv ファイルとは comma separated values という名前のファイルで, テキスト形式でも Excel ファイルでも開くことができます。



③ 画面の上の方で「保存先 (I)」を以下の図のいずれかの矢印をクリックすることで指定する。(今回の場合は c:\Program Files\R に保存するつもりで操作を行っている。) その後、ボックスの下の方にある「ファイルの種類 (I)」のところで「CSV(カンマ区切り)(*.csv)」という項目があるので、それをクリックする。



それで「保存 (S)」ボタンを押せば、csv ファイルが作られる。

この時「選択したファイルの種類は複数のシートを含むブックをサポートしていません」や「～」には、csv (カンマ区切り) と互換性のない機能が含まれている可能性があります。～」などのコメントが現れるが、気にせず「OK」もしくは「はい (Y)」を押せば良い。

14.2 データのインポート

まず R を起動させた後で、データをインポート (要はデータを取り込むこと) します。今、インポートするデータは csv ファイルで c:\Program Files\R にあるものとします。

```
data02<-read.csv("c:\\Program Files\\R\\principal_data02.csv",header=T,row.name="name")
```

と入力し、エンターキーを押します。csv ファイルに行の名前があるときには、row.name= "行ラベルの名前" と入力しますが、行ラベルがない時はこの部分は削除します。今、行ラベルの名前は 14.1 章の ①で name としているので row.name="name"としています。一方で列ラベルに名前がある場合は header=T とします。

注意 \ はパソコンでは ｘ を入力します。

14.3 主成分分析を行う

一旦データをインポートできたら R での主成分分析は至って簡単です。

```
res2<-princomp(data02,cor=TRUE)
```

と入力し、エンターキーを押すだけです。これで data02 を主成分分析した結果が res2 に入っています。ここで cor というのは、相関行列を使うかどうかを尋ねていますが、ここでは標準化するので cor=TRUE と入力します。

① 主成分を求める res2\$loadings と入力しエンターキーを押すと、主成分 (固有ベクトル) が分かります。

Loadings:

	Comp.1	Comp.2	Comp.3	Comp.4	Comp.5
国語	0.526	0.309	0.326		0.722
社会	0.535	0.177	0.451	-0.166	-0.672
数学	0.412	-0.531		0.737	
物理	0.377	-0.577	-0.310	-0.647	0.100
英語	0.353	0.508	-0.769	0.102	-0.126

	Comp.1	Comp.2	Comp.3	Comp.4	Comp.5
SS loadings	1.0	1.0	1.0	1.0	1.0
Proportion Var	0.2	0.2	0.2	0.2	0.2
Cumulative Var	0.2	0.4	0.6	0.8	1.0

ここで Comp.1 は第 1 主成分, Comp.2 は第 2 主成分を表します。尚、空白のところは、0 とみなしますが、実際には 0.1 より小さな値をとっています。全ての主成分を表示させたい場合は、`print(res2$loadings, cutoff=0.01)` と入力します。

② 固有値を求める

一方で `summary(res2)` と入力し、エンターキーを押すと

Importance of components:

	Comp.1	Comp.2	Comp.3	Comp.4	Comp.5
Standard deviation	1.6812356	1.1348212	0.7252369	0.53187620	0.27706841
Proportion of Variance	0.5653106	0.2575638	0.1051937	0.05657846	0.01535338
Cumulative Proportion	0.5653106	0.8228744	0.9280682	0.98464662	1.00000000

となります。ここで Proportion of Variance は主成分の寄与率を表し、Cumulative Proportion は主成分の累積寄与率を表します。

単に固有値の値を知りたいときには `(res2$sdev)^2` と入力すれば良い。

③ 主成分得点を求める

主成分得点は `res2$scores` と入力し、エンターキーを押すと

	Comp.1	Comp.2	Comp.3	Comp.4	Comp.5
相川 英之	1.7265903	1.70928067	0.12905325	-1.28251248	-0.29260303
吾妻 博	-0.6418345	-2.56489434	0.39875516	0.17339481	-0.22968921
阿部 美穂	-0.6877820	-2.15089741	0.00616816	-0.30326796	0.39388629
飯田 卓也	-1.5357077	1.55551911	-0.22704070	1.04815146	-0.14485306
上原 広幸	-1.9791737	-0.75587299	0.50688889	0.16104098	-0.27907826
.....					
佐藤 愛	0.8704208	-0.25362987	0.89604976	-0.05687978	0.34778942
桜井 幸信	1.3894134	-0.01650842	-0.24309887	0.98294896	-0.19720267
島貫 寛一郎	-2.3264977	1.42838075	-0.32118612	0.28303354	0.01405411
菅原 美紀	1.2698013	-0.70612494	-1.14335365	-0.30035825	0.16772268
瀬川 潤	-2.7749767	0.88661525	-1.32299000	-0.73855918	-0.09314850

となります。ここで、第 1 主成分と第 2 主成分をデータの散布図として、グラフに描きたいときには `biplot(res2)` と入力し、エンターキーを押します。R における biplot の説明は W.N. ウエナブルズ, B.D. リブリー [5] で見ることができる。

14.4 R での biplot の目盛りについて

R の biplot のグラフの目盛はスケールが変更されているため、出力された主成分得点と主成分の値と一致しません。相関行列 R の大きい 2 つの固有値を $\lambda_1 \geq \lambda_2$ とし、対応する固有ベクトルを $\mathbf{a}_1, \mathbf{a}_2$ とする。

この時、例えば p36 の相川君のデータを標準化したものを $\tilde{\mathbf{x}}_1 = \begin{pmatrix} 1.2834 \\ 1.6939 \\ -1.1391 \\ 0.4261 \\ 1.2858 \end{pmatrix}$ とすると、相川君の第 1 と第

2 主成分得点は $\mathbf{a}_1^T \tilde{\mathbf{x}}_1, \mathbf{a}_2^T \tilde{\mathbf{x}}_1$ となる。

R の biplot のデフォルトでは、第 1 と第 2 主成分得点はそれぞれ $\frac{1}{\sqrt{\lambda_1 n}} \mathbf{a}_1^T \tilde{\mathbf{x}}_1, \frac{1}{\sqrt{\lambda_2 n}} \mathbf{a}_2^T \tilde{\mathbf{x}}_1$ となり、主成分ベクトルを $\sqrt{\lambda_1 n} \mathbf{a}_1, \sqrt{\lambda_2 n} \mathbf{a}_2$ とスケール調整してある。もし一致させたい場合、`biplot(res2, scale=0)` と入力すれば良い。(これは表 2.1 のデータで確認済み)

14.5 R での注意点

主成分ベクトルは相関行列の大きさ 1 の固有ベクトルであったため、しばしば期待している向きと逆向きに出てくることがしばしばある。例えば、5 教科データでの biplot も左に行くほど、(つまり第 1 主成分得点が小さいほど) 総合力が高くなる場合があります。しかし実際には、第 1 主成分得点が高いほど、総合力が高いと考える方が自然である。このことより biplot において、元のデータにも注意しながら軸の解釈を与える必要がある。

14.6 解析結果のエクスポート

R の統計処理で得た、解析結果を Excel で利用したいということもあるかと思います。ここで `write.csv` というコマンドを使い、主成分得点の結果を csv ファイルに書き直します。

```
output01<-res2$loadings (これは res2$loadings の結果を output01 に代入するという操作)
write.csv(output01, file = "c:\\Program Files\\R\\output.csv")
```

と入力すれば ok です。これで `c:\Program Files\R` に `output.csv` というファイルが作られます。もし、行の名前を入れたくなければ

```
write.csv(output01, file = "c:\\Program Files\\R\\output.csv", row.names=FALSE)
```

とすれば ok です。

【Windows VISTA での注意点】 VISTA では、管理者権限で入らないとデータをエクスポートできません。このため R を起動する時に、アイコンの上で右クリックをし、管理者として実行 (A)... を左クリックして、R を起動させてください。

14.7 princomp がうまくいかない時

`princomp` 関数は、データ数 n が変数の数 p より小さいと、主成分分析ができなくなります。この場合は `prcomp` 関数を用いるうまく行きますが、細かいことは省略します。(インターネットでいくつか公開されている)

参考文献

- [1] <http://aoki2.si.gunma-u.ac.jp/lecture/mb-arc/arc035/08319.html> かの有名な青木先生 (群馬大) のサイト. 主成分が2つ出てくるが, どちらを選べば良いのかを回答している.
- [2] 総務省統計局, <http://www.stat.go.jp/data/sekai/06.htm>, 出典 UN, Energy Statistics Yearbook 2004
- [3] 石村 貞夫, 石村光資郎 (2007) 入門はじめての多変量解析, 東京図書. このプリントとは別のアプローチで主成分分析を説明している. 分散が不偏分散を用いていることに注意が必要.
- [4] 高橋 信, 井上いろは, トレンド・プロ (2006). マンガでわかる統計学 因子分析編, オーム社. 主成分分析のイメージがつかめない人はこの本を読むとよい.
- [5] W.N. ウエナブルズ, B.D. リプリー, S-PLUS による統計解析, シュプリンガーフェアラーク東京
- [6] 金谷健一 (2005) これなら分かる最適化数学, 共立出版

15 解答

問 1.1 $x = \frac{3\sqrt{3} \pm \sqrt{27-24}}{4} = \frac{3\sqrt{3} \pm \sqrt{3}}{4}$ より $x = \sqrt{3}, \frac{\sqrt{3}}{2}$.

問 1.2 ① $x^3 - 7x + 6 = (x+3)(x-1)(x-2)$. よって $x = -3, 1, 2$. ② $2x^3 + 5x^2 + x - 2 = (2x - 1)(x+1)(x+2)$. よって $x = \frac{1}{2}, -1, -2$. 問 2.1 (1,3) 成分は 7. 問 2.2 ① $\begin{pmatrix} 5 & 3 \\ 6 & 4 \end{pmatrix} + \begin{pmatrix} 0 & 2 \\ 1 & 7 \end{pmatrix} = \begin{pmatrix} 5 & 5 \\ 7 & 11 \end{pmatrix}$.

② $\begin{pmatrix} 5 & 3 \\ 6 & 4 \end{pmatrix} - \begin{pmatrix} 0 & 2 \\ 1 & 7 \end{pmatrix} = \begin{pmatrix} 5 & 1 \\ 5 & -3 \end{pmatrix}$. ③ $\begin{pmatrix} 5 & 3 \\ 6 & 4 \end{pmatrix} - 2 \begin{pmatrix} 0 & 2 \\ 1 & 7 \end{pmatrix} = \begin{pmatrix} 5 & -1 \\ 4 & -10 \end{pmatrix}$.

問 2.3 ① $\begin{pmatrix} 5 & -1 \\ 2 & 3 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 3 & -1 \end{pmatrix} = \begin{pmatrix} 2 & 1 \\ 11 & -3 \end{pmatrix}$. ② $\begin{pmatrix} 0 & -2 & 1 \\ -1 & 1 & 1 \\ 2 & 5 & -5 \end{pmatrix} \begin{pmatrix} 10 & 5 & 3 \\ 3 & 2 & 1 \\ 7 & 4 & 2 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$ ③

$\begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix} \begin{pmatrix} 3 & 4 \end{pmatrix} = \begin{pmatrix} 3 & 4 \\ 6 & 8 \\ 9 & 12 \end{pmatrix}$ ④ $\begin{pmatrix} \sqrt{2} & 1 \end{pmatrix} \begin{pmatrix} 3 \\ -1 \end{pmatrix} = 3\sqrt{2} - 1$. 1×1 行列の場合は括弧をつけない.

問 2.4 ① $AI = \begin{pmatrix} 5 & -1 \\ 2 & 3 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 5 & -1 \\ 2 & 3 \end{pmatrix}$. ② $IW = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 5 \\ -2 \end{pmatrix} = \begin{pmatrix} 5 \\ -2 \end{pmatrix}$.

問 2.5 ① 11 ② 11 ③ $\begin{pmatrix} 1 & 2 \\ 0 & 5 \end{pmatrix} \begin{pmatrix} 3 \\ 4 \end{pmatrix} = \begin{pmatrix} 1 & 8 \\ 3 & 20 \end{pmatrix} = 3 + 32 = 35$.

問 3.1 ① -10 ② -2 ③ 0. 問 4.1 $\mathbf{a} + 4\mathbf{b} = \begin{pmatrix} 19 \\ 26 \end{pmatrix}$, $|\mathbf{a}| = \sqrt{(-1)^2 + 2^2} = \sqrt{5}$. 問 4.2 $k = \pm \frac{1}{\sqrt{5}}$.

問 4.4 傾き $\frac{1}{2}$ は x が 1 増えると y が $\frac{1}{2}$ 増えることを意味している. これより求める単位ベクトルは $\mathbf{a} = k \begin{pmatrix} 1 \\ \frac{1}{2} \end{pmatrix} = \begin{pmatrix} k \\ \frac{k}{2} \end{pmatrix}$. よって $1 = |\mathbf{a}| = \sqrt{k^2 + (\frac{k}{2})^2} = \sqrt{\frac{5}{4}k^2} = \frac{\sqrt{5}}{2}|k|$. $|k| = \frac{2}{\sqrt{5}}$ より $k = \pm \frac{2}{\sqrt{5}}$.

問 4.5 $\mathbf{a}^T \mathbf{x} = \begin{pmatrix} \frac{1}{\sqrt{5}} & \frac{2}{\sqrt{5}} \end{pmatrix} \begin{pmatrix} 2 \\ -3 \end{pmatrix} = \frac{2-6}{\sqrt{5}} = -\frac{4}{\sqrt{5}}$.

問 5.1 ①

$$\left| \begin{pmatrix} 5 & -1 \\ -1 & 5 \end{pmatrix} - \lambda \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \right| = \begin{vmatrix} 5-\lambda & -1 \\ -1 & 5-\lambda \end{vmatrix} = (5-\lambda)^2 - (-1)^2 = \lambda^2 - 10\lambda + 24 = (\lambda-6)(\lambda-4) = 0$$

よって $\lambda = 4, 6$.

$$\textcircled{2} \left| \begin{pmatrix} 1 & r \\ r & 1 \end{pmatrix} - \lambda \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \right| = \begin{vmatrix} 1-\lambda & r \\ r & 1-\lambda \end{vmatrix} = (1-\lambda)^2 - r^2 = (\lambda-1+r)(\lambda-1-r) = 0. \text{ よって } \lambda = 1 \pm r.$$

問 5.2 A の固有ベクトル

$$\begin{cases} (5-\lambda)x - y = 0 \\ -x + (5-\lambda)y = 0 \end{cases}$$

$\lambda = 4$ の時, $\begin{cases} x - y = 0 \\ -x + y = 0 \end{cases}$ より $x = y$. よって $\frac{y}{x} = \frac{1}{1}$ から固有ベクトルは $\mathbf{u} = k \begin{pmatrix} 1 \\ 1 \end{pmatrix}$, ($k \neq 0$) となる.

$\lambda = 6$ の時, $\begin{cases} -x - y = 0 \\ -x - y = 0 \end{cases}$ より $x = -y$. よって $\frac{y}{x} = \frac{-1}{1}$ から固有ベクトルは $\mathbf{v} = l \begin{pmatrix} 1 \\ -1 \end{pmatrix}$, ($l \neq 0$) となる.

B の固有ベクトル

$$\begin{cases} (1-\lambda)x + ry = 0 \\ rx + (1-\lambda)y = 0 \end{cases}$$

$\lambda = 1+r$ の時, $\begin{cases} -rx + ry = 0 \\ rx - ry = 0 \end{cases}$ より $x = y$. よって $\frac{y}{x} = \frac{1}{1}$ から固有ベクトルは $\mathbf{u} = k \begin{pmatrix} 1 \\ 1 \end{pmatrix}$, ($k \neq 0$) となる.

$\lambda = 1-r$ の時, $\begin{cases} rx + ry = 0 \\ rx + ry = 0 \end{cases}$ より $x = -y$. よって $\frac{y}{x} = \frac{-1}{1}$ から固有ベクトルは $\mathbf{v} = l \begin{pmatrix} 1 \\ -1 \end{pmatrix}$, ($l \neq 0$) となる.

問 5.3 ① 固有値は $\lambda = 3 + \sqrt{3}, 3 - \sqrt{3}, 0$. 固有ベクトルは $k \begin{pmatrix} -2 + \sqrt{3} \\ -1 + \sqrt{3} \\ 1 \end{pmatrix}, l \begin{pmatrix} -2 - \sqrt{3} \\ -1 - \sqrt{3} \\ 1 \end{pmatrix}, m \begin{pmatrix} 1 \\ -1 \\ 1 \end{pmatrix}$.

ただし k, l, m は 0 でない実数.

注意 ここで k, l, m は任意の数であるので, 固有ベクトルは色々な表し方がある. 例えば ①では $\lambda = 3 + \sqrt{3}$ の固有ベクトルは $k \begin{pmatrix} 1 \\ -1 - \sqrt{3} \\ -2 - \sqrt{3} \end{pmatrix}$, $\lambda = 3 - \sqrt{3}$ の固有ベクトルは $l \begin{pmatrix} 1 \\ -1 + \sqrt{3} \\ -2 + \sqrt{3} \end{pmatrix}$, と表すこともできる.

② 固有値は $\lambda = -2, 2, 1$. 固有ベクトルは $k \begin{pmatrix} -1 \\ -1 \\ 2 \end{pmatrix}, l \begin{pmatrix} -1 \\ 1 \\ 0 \end{pmatrix}, m \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$. ただし k, l, m は 0 でない実数.

問 6.1 $\frac{\partial f(1,0)}{\partial x} = -4, \frac{\partial f(1,1)}{\partial x} = -3, \frac{\partial f(1,2)}{\partial x} = -2$.

問 6.2 $\frac{\partial}{\partial x} f(x, y) = -4x + y, \frac{\partial}{\partial y} f(x, y) = -2y + x$. 問 6.3 ① $\frac{\partial}{\partial x} f(x, y) = 2x - 2 = 0, \frac{\partial}{\partial y} f(x, y) = 2y - 1 = 0$. これより $x = 1, y = \frac{1}{2}$. $(x, y) = (1, \frac{1}{2})$ の時, 最小値 $f(1, \frac{1}{2}) = 1 - 2 + (\frac{1}{2})^2 - \frac{1}{2} = -\frac{5}{4}$.

② $\frac{\partial}{\partial x} f(x, y) = 4x - 4y - 2 = 0, \frac{\partial}{\partial y} f(x, y) = -4x + 8y = 0$. これより $x = 1, y = \frac{1}{2}$. $(x, y) = (1, \frac{1}{2})$ の時, 最小値 $f(1, \frac{1}{2}) = 2 - 4 \cdot \frac{1}{2} - 2 \cdot 1 + 4(\frac{1}{2})^2 + 1 = 0$.

問 7.1 ① $F(x, y, \lambda) = 2x^2 + 3y^2 + xy - \lambda(x + y - 1)$ とおくと

$$\frac{\partial}{\partial x} F(x, y, \lambda) = 4x + y - \lambda = 0 \quad \cdots \textcircled{1}$$

$$\frac{\partial}{\partial y} F(x, y, \lambda) = x + 6y - \lambda = 0 \quad \cdots \textcircled{2}$$

$$\frac{\partial}{\partial \lambda} F(x, y, \lambda) = x + y - 1 = 0 \quad \cdots \textcircled{3}$$

よって ①=②より $3x - 5y = 0$. これと ③との連立方程式を解くと $x = \frac{5}{8}, y = \frac{3}{8}$.

② $F(x, y, z, \lambda) = xy + yz + zx - \lambda(x^2 + y^2 + z^2 - 1)$ とおくと

$$\frac{\partial}{\partial x} F(x, y, z, \lambda) = y + x - 2\lambda x = 0 \quad \cdots \text{①}$$

$$\frac{\partial}{\partial y} F(x, y, z, \lambda) = x + z - 2\lambda y = 0 \quad \cdots \text{②}$$

$$\frac{\partial}{\partial z} F(x, y, z, \lambda) = y + z - 2\lambda z = 0 \quad \cdots \text{③}$$

$$\frac{\partial}{\partial \lambda} F(x, y, z, \lambda) = x^2 + y^2 + z^2 - 1 = 0 \quad \cdots \text{④}$$

よって ①+②+③より $2(x + y + z) - 2\lambda(x + y + z) = 0$. よって $\lambda = 1$. これと ①, ②, ③に代入すると $x = y = z$ を得る. これを ④に代入すると $x = y = z = \frac{1}{\sqrt{3}}$ を得る.

問 8.1 $\bar{y} = \frac{8+9+7+9+8+9+8+8+7+7}{10} = 8$ となる. 問 9.1: (1) ① $\bar{x} = 14$.

$$\text{② } s^2 = \frac{(15-14)^2 + (10-14)^2 + (15-14)^2 + \cdots + (16-14)^2}{5} = \frac{22}{5}. \quad \text{③ } \sqrt{\frac{22}{5}} = \frac{\sqrt{110}}{5} = 2.1$$

問 9.2 男子の平均値 8.6、女子の平均値 8.1、男子の標本分散 56.441、女子の標本分散 11.09、男子の標準偏差 7.52、女子の標準偏差 3.33. この結果から男子の方が女子よりも CD を持っている人と持っていない人の差が多い.

問 10.1 A 君の z スコアは $\frac{10 - 25/3}{\sqrt{134/9}} \div 0.432$.

$$\text{問 10.2 } \bar{z} = \frac{1}{n} \sum_{i=1}^n z_i = \frac{1}{n} \sum_{i=1}^n \frac{x_i - \bar{x}}{s} = \frac{1}{ns} \sum_{i=1}^n (x_i - \bar{x}) = \frac{1}{s} \left(\frac{1}{n} \sum_{i=1}^n x_i - \frac{1}{n} \sum_{i=1}^n \bar{x} \right) = \frac{1}{s} (\bar{x} - \bar{x}) = 0.$$

$$\frac{1}{n} \sum_{i=1}^n (z_i - \bar{z})^2 = \frac{1}{n} \sum_{i=1}^n z_i^2 = \frac{1}{n} \sum_{i=1}^n \left(\frac{x_i - \bar{x}}{s} \right)^2 = \frac{1}{n} \sum_{i=1}^n \frac{(x_i - \bar{x})^2}{s^2} = \frac{1}{s^2} \frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})^2 = \frac{1}{s^2} \cdot s^2 = 1.$$

問 11.1 $\bar{x} = 3, \bar{y} = \frac{22}{5}$. 下の表より $s_{xy} = 13/5$.

x	y	$x - \bar{x}$	$y - \bar{y}$	$(x - \bar{x})(y - \bar{y})$	
1	2	-2	-2.4	4.8	
4	6	1	1.6	1.6	
4	5	1	0.6	0.6	
1	3	-2	-1.4	2.8	
5	6	2	1.6	3.2	
計	15	22	0	0	13

問 11.2 $\bar{x} = 31/5, \bar{y} = 36/5, \bar{xy} = 287/5$ より $s_{xy} = \frac{287}{5} - \frac{31}{5} \cdot \frac{36}{5} = \frac{319}{25} = 12.76$.

x	y	$x - \bar{x}$	$y - \bar{y}$	$(x - \bar{x})^2$	$(y - \bar{y})^2$	$(x - \bar{x})(y - \bar{y})$
80	70	8	10	64	100	80
50	60	-22	0	484	0	0
70	50	-2	-10	4	100	20
90	30	18	-30	324	900	-540
70	90	-2	30	4	900	-60
360	300	0	0	880	2000	-500

これより相関係数は約 $r = -0.38$.

問 11.4 右の表より平均 $\bar{x} = 22/10$, $\bar{y} = 27/10$. 分散 $s_x^2 = \frac{58}{10} - \left(\frac{22}{10}\right)^2 = \frac{96}{100}$, $s_y^2 = \frac{261}{100}$. よって $s_x = \frac{\sqrt{96}}{10}$. $s_y = \frac{\sqrt{261}}{10}$. 共分散 $s_{xy} = \frac{66}{10} - \frac{22}{10} \cdot \frac{27}{10} = \frac{66}{100}$. よって相関係数は $r = \frac{66/100}{\sqrt{96/100}\sqrt{261/100}} \doteq 0.42$.

x	y	x^2	y^2	xy
2	3	4	9	6
3	1	9	1	3
4	5	16	25	20
2	2	4	4	4
3	6	9	36	18
1	1	1	1	1
2	3	4	9	6
1	2	1	4	2
1	3	1	9	3
3	1	9	1	3
22	27	58	99	66

問 12.1 分散共分散行列は $\begin{pmatrix} 176 & -100 \\ -100 & 400 \end{pmatrix}$. 相関行列は $\begin{pmatrix} 1 & -0.38 \\ -0.38 & 1 \end{pmatrix}$.

問 12.2 $R = \begin{pmatrix} 1 & 0.42 \\ 0.42 & 1 \end{pmatrix}$ は【例 12.3】の注意を用いると固有値は 1.42, 0.58 となる. 一方で第 1 主成分は $\mathbf{a}^* = \begin{pmatrix} \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \end{pmatrix}$, 第 2 主成分は $\mathbf{a}^{**} = \begin{pmatrix} \frac{1}{\sqrt{2}} \\ -\frac{1}{\sqrt{2}} \end{pmatrix}$.

問 12.3 $\frac{1.42}{1.42 + 0.58} = \frac{1.42}{2} = 0.71$. よって寄与率は 71%

問 13.1 ① 固有値を求める. $|A - \lambda I| = 0$ を解くと, $\begin{vmatrix} 1 - \lambda & \frac{1}{2} & -\frac{1}{2} \\ \frac{1}{2} & 1 - \lambda & 0 \\ -\frac{1}{2} & 0 & 1 - \lambda \end{vmatrix} = 0$ となる. これより左辺の行列式を展開すると

$$\begin{aligned} (1 - \lambda)^3 - \frac{1}{4}(1 - \lambda) - \frac{1}{4}(1 - \lambda) &= 0 \\ (1 - \lambda) \left((1 - \lambda)^2 - \frac{1}{2} \right) &= 0 \\ -(\lambda - 1) \left(\lambda - 1 + \frac{1}{\sqrt{2}} \right) \left(\lambda - 1 - \frac{1}{\sqrt{2}} \right) &= 0 \end{aligned}$$

よって固有値は $\lambda = 1, 1 - \frac{1}{\sqrt{2}}, 1 + \frac{1}{\sqrt{2}}$.

② 固有ベクトル (主成分) を求める. $(A - \lambda I)\mathbf{w} = \mathbf{0}$ を式で表すと

$$\begin{cases} (1 - \lambda)x + \frac{1}{2}y - \frac{1}{2}z = 0 \\ \frac{1}{2}x + (1 - \lambda)y = 0 \\ -\frac{1}{2}x + (1 - \lambda)z = 0 \end{cases} \quad (15.1)$$

となる.

• $\lambda = 1 + \frac{1}{\sqrt{2}}$ のとき $\begin{cases} -\frac{1}{\sqrt{2}}x + \frac{1}{2}y - \frac{1}{2}z = 0 \\ \frac{1}{2}x - \frac{1}{\sqrt{2}}y = 0 \\ -\frac{1}{2}x - \frac{1}{\sqrt{2}}z = 0 \end{cases}$ から $x = \sqrt{2}y$, $x = -\sqrt{2}z$ を得る. よって求める固有ベクトルは $\mathbf{v} = l \begin{pmatrix} 1 \\ \frac{1}{\sqrt{2}} \\ -\frac{1}{\sqrt{2}} \end{pmatrix}$, $l \neq 0$ となる. ここで主成分の大きさは 1 であるので, $|\mathbf{v}| = \sqrt{l^2 + l^2/2 + l^2/2} = \sqrt{2}|l|$.

より $l = \pm \frac{1}{\sqrt{2}}$ となる. よって第 1 主成分は $f_2 = \frac{1}{\sqrt{2}}x + \frac{1}{2}y - \frac{1}{2}z$.

• $\lambda = 1$ のとき $\begin{cases} \frac{1}{2}y - \frac{1}{2}z = 0 \\ \frac{1}{2}x = 0 \\ -\frac{1}{2}x = 0 \end{cases}$ から $x = 0, y = z$ を得る. よって求める固有ベクトルは $\mathbf{u} = k \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix}$,

$k \neq 0$ となる. ここで主成分の大きさは 1 であるので, $|\mathbf{u}| = \sqrt{k^2 + k^2} = \sqrt{2}|k|$. より $k = \pm \frac{1}{\sqrt{2}}$ となる.

よって第 2 主成分は $f_1 = \frac{1}{\sqrt{2}}y + \frac{1}{\sqrt{2}}z$.

• $\lambda = 1 - \frac{1}{\sqrt{2}}$ のとき $\begin{cases} -\frac{1}{\sqrt{2}}x + \frac{1}{\sqrt{2}}y + \frac{1}{\sqrt{2}}z = 0 \\ \frac{1}{2}x + \frac{1}{\sqrt{2}}y = 0 \\ -\frac{1}{2}x + \frac{1}{\sqrt{2}}z = 0 \end{cases}$ から $x = -\sqrt{2}y, x = \sqrt{2}z$ を得る. よって固有ベクトル

は $\mathbf{w} = m \begin{pmatrix} 1 \\ -\frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \end{pmatrix}$, $m \neq 0$ となる. ここで主成分の大きさは 1 であるので, $|\mathbf{w}| = \sqrt{m^2 + m^2/2 + m^2/2} =$

$\sqrt{2}|m|$. より $m = \pm \frac{1}{\sqrt{2}}$ となる. よって第 3 主成分は $f_3 = \frac{1}{\sqrt{2}}x - \frac{1}{2}y + \frac{1}{2}z$.

また第 2 主成分までの累積寄与率は $\frac{1 + 1/\sqrt{3} + 1}{3} \div 0.90$. これより約 90% となる. 問 13.2 ① 総合的にエネルギーを消費している量を表す. 即ち左側にある程, エネルギーの消費量が多い. ② 天然ガスは他のエネルギーと比較して, 二酸化炭素, 二酸化硫黄の排出が少なく, また電力のように原子力による核廃棄物を出さない. 即ちエネルギーのクリーンさの度合いと考えられる. (他にももっとうまい意味付けがあるかもしれない)

③ $\frac{3.1 + 0.62}{3.1 + 0.62 + 0.21 + 0.07} = 0.93$. 第 2 主成分まででデータの 93% が説明できる.

④ まず日本の石炭, 原油, 天然ガス, 電力の z-score を求める.

	石炭	原油	天然ガス	電力
日本	1416	1576	26334	8459
平均	507.14	941.25	30933.46	4362.39
標準偏差	563.70	739.0	31069.17	4281.69
zscore	1.61	0.86	-0.15	0.96

これより,

$$\text{日本の第 1 主成分得点} = -0.42 \times 1.61 - 0.55 \times 0.86 - 0.48 \times (-0.15) - 0.54 \times 0.96 = -1.5956$$

	石炭	原油	天然ガス	電力
アメリカ	1710	2723	81341	14240
平均	507.14	941.25	30933.46	4362.39
標準偏差	563.70	739.0	31069.17	4281.69
zscore	2.13	2.41	1.62	2.31

これより,

$$\text{アメリカの第 1 主成分得点} = -0.42 \times 2.13 - 0.55 \times 2.41 - 0.48 \times 1.62 - 0.54 \times 2.31 = -4.2451$$

これよりアメリカの一人当たりのエネルギーの総合的な使用量は日本より多いことがわかる. ⑤ アメリカ

⑥ デンマーク