独立成分分析を用いた線形代数の学習傾向の考察

— 期末試験の得点分布を分析して
学生の理解度を調べる —

金沢工業大学
数理工基礎教育課程
高村　松三

研究の目的

試験のヒストグラムが正規分にならない

⬇

学生の理解度の個人差が大きい

⬇

学生がどのように授業内容を理解しているのか、もう少し詳しく調べたい。

⬇

試験の問題ごとの正解率のデータを独立成分分析により調べ、学生が独立に理解している内容を推定する。

カクテルパーティ効果

- 大勢の人が同時に談笑しているパーティ会場でも人間は特定の人と会話できる

⬇

ブラインド信号源分離

- 多くの信号が混在する中から必要な信号を分離して取り出す問題

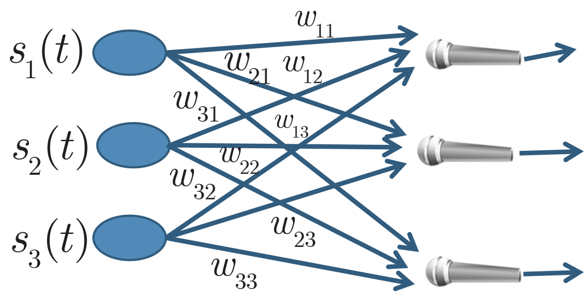
⬇

- 信号源の独立性を仮定するだけで解ける

独立成分分析

（測定不可能）　　（測定可能）

信号源 $\boldsymbol{s}(t)$　　観測信号 $\boldsymbol{x}(t)$



$x_1(t) = w_{11}s_1 + w_{12}s_2 + w_{13}s_3$

$x_2(t) = w_{21}s_1 + w_{22}s_2 + w_{23}s_3$

$x_3(t) = w_{31}s_1 + w_{32}s_2 + w_{33}s_3$

$$\boldsymbol{s}(t) = \begin{pmatrix} s_1(t) \\ s_2(t) \\ s_3(t) \end{pmatrix} \quad \boldsymbol{x}(t) = \begin{pmatrix} x_1(t) \\ x_2(t) \\ x_3(t) \end{pmatrix} \quad W = \begin{pmatrix} w_{11} & w_{12} & w_{13} \\ w_{21} & w_{22} & w_{23} \\ w_{31} & w_{32} & w_{33} \end{pmatrix}$$

$$\boldsymbol{x}(t)=\begin{pmatrix}x_1(t)\\x_2(t)\\x_3(t)\end{pmatrix}\quad W=\begin{pmatrix}w_{11}&w_{12}&w_{13}\\w_{21}&w_{22}&w_{23}\\w_{31}&w_{32}&w_{33}\end{pmatrix}\quad \boldsymbol{s}(t)=\begin{pmatrix}s_1(t)\\s_2(t)\\s_3(t)\end{pmatrix}$$

$$\boldsymbol{x}(t)=W\,\boldsymbol{s}(t)$$

ブラインド信号源分離の問題

- 測定した $\boldsymbol{x}(t)$ から未知の W と $\boldsymbol{s}(t)$ を推定する。
- 信号源 $\boldsymbol{s}(t)$ は互いに独立であると仮定する。

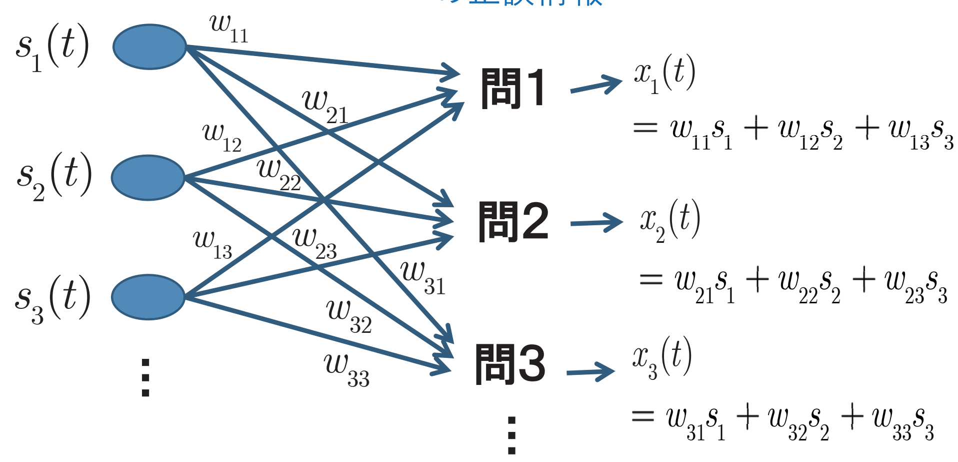
期末試験に適用する場合

信号源 $\boldsymbol{s}(t)$

学生が独立に理解している内容

観測信号 $\boldsymbol{x}(t)$

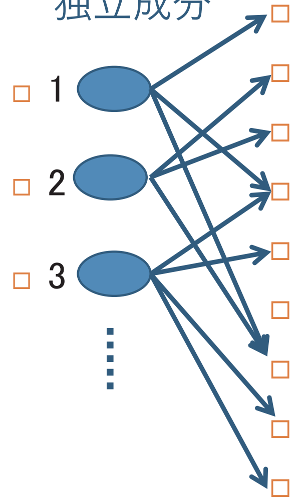
試験の問題ごとの正誤情報



線形代数の学習内容

	学生001	学生002	…	学生538
1 3～5次の行列式の計算	0.8	0.9	…	0.6
2 変数を含む行列式と因数分解	0.6	0.8	…	0.3
3 3次正方行列の逆行列	0.5	0.6	…	0.2
4 クラメルの公式	0.7	0.9	…	0.4
5 直交行列	0.3	0.7	…	0.1
6 ベクトルの1次独立	0.7	1.0	…	0.5
7 (線形空間の)基底と次元	0.6	0.8	…	0.6
8 固有値	0.7	1.0	…	0.8
9 固有ベクトル	0.6	1.0	…	0.7

学生が理解している内容の独立成分



線形代数の学習内容

	学生001
1 3～5次の行列式の計算	0.8
2 変数を含む行列式と因数分解	0.6
3 3次正方行列の逆行列	0.5
4 クラメルの公式	0.7
5 直交行列	0.3
6 ベクトルの1次独立	0.7
7 (線形空間の)基底と次元	0.6
8 固有値	0.7
9 固有ベクトル	0.6

独立成分分析（fastICA）

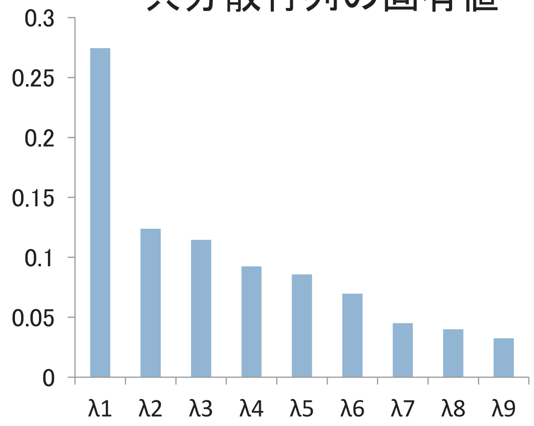
- ステップ1　$p \leftarrow 1$ とする
- ステップ2　$\boldsymbol{w}_p$ の初期値を乱数で設定.4へ飛ぶ
- ステップ3　$\boldsymbol{w}_p \leftarrow E\left[\boldsymbol{z}(k)g(\boldsymbol{w}_p^T\boldsymbol{z}(k))\right] - E\left[g'(\boldsymbol{w}_p^T\boldsymbol{z}(k))\right]$
  $g(x) = \tanh(x)$ とした
- ステップ4　$\boldsymbol{w}_p \leftarrow \boldsymbol{w}_p - \sum_{j=1}^{p-1} \boldsymbol{w}_p^T\boldsymbol{w}_j\ \boldsymbol{w}_j$　（直交化）
- ステップ5　$\boldsymbol{w}_p \leftarrow \boldsymbol{w}_p / \|\boldsymbol{w}_p\|$　（大きさを1に正規化）
- ステップ6　$\boldsymbol{w}_p$ が収束していなければ3に戻る.
- ステップ7　$p \leftarrow p+1$　$p \leqq 4$ であれば2に戻る.

独立成分の個数の求め方（１）

期末試験のデータ(538名分)

X(1)	X(2)	…	X(538)
0.8	0.9	…	0.6
0.6	0.8	…	0.3
0.5	0.6	…	0.2
0.7	0.9	…	0.4
0.3	0.7	…	0.1
0.7	1.0	…	0.5
0.6	0.8	…	0.6
0.7	1.0	…	0.8
0.6	1.0	…	0.7

共分散行列の固有値



独立成分の個数の求め方（２）

従来法

データの共分散行列の小さい固有値に対応する次元を削除することで低次元化

提案法

低次元化せずに独立成分分析を適用

得られた独立成分ごとの尖度に着目して低次元化

独立成分分析で利用している非ガウス性の信号の性質

ガウス性の信号 ＋ ガウス性の信号　ガウス性の信号

分離不可能

非ガウス性の信号 ＋ 非ガウス性の信号 ＋ …　ガウス性の信号に近づく

独立成分分析では・・・
独立性に着目して分離

ガウス性の信号の性質を利用した独立成分の個数の決定法

実際のデータ(例:期末試験の採点データ)

非ガウス性の信号 ＋ 非ガウス性の信号 ＋ …

＋ ガウス性の信号1＋ ガウス性の信号2＋ …

…

分離不可能

非ガウス性の信号の個数 ＝ 独立成分の個数

尖度 κ の定義 (非ガウス性の尺度として使用)

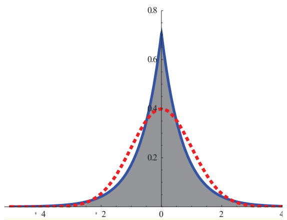
重みベクトルに射影して平均値を 0 に正規化

$$p_i(k) = \boldsymbol{w}_i^T \boldsymbol{z}(k)$$

$p_i(k) \ (k = 1, 2, \cdots, N)$ **の尖度**

$$\kappa(p_i) = \frac{E[p_i^4]}{E[p_i^2]^2} - 3$$

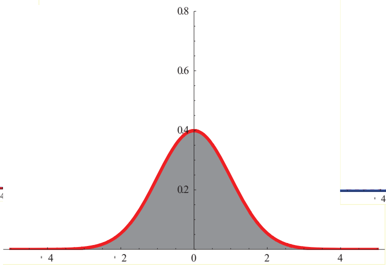
尖度の例 (分散 1 に正規化した確率密度関数)



優ガウス性

$\kappa(p_i) > 0$

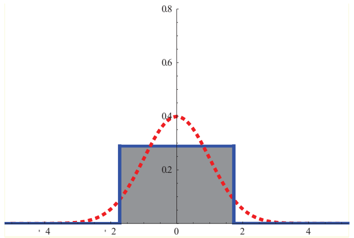
尖度が正値



ガウス分布

$\kappa(p_i) = 0$

尖度が0



劣ガウス性

$\kappa(p_i) < 0$

尖度が負値

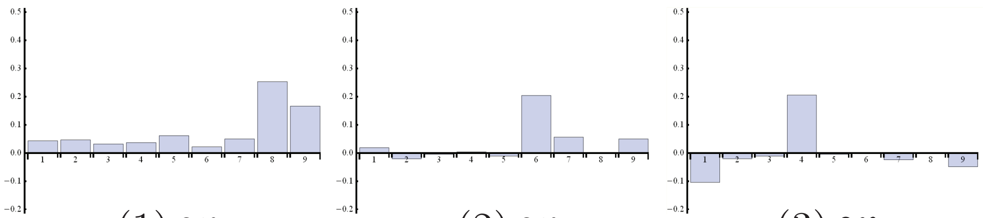
表1 重み重み $\boldsymbol{w}_i$ に射影したデータの尖度 $\kappa(p_i)$ (i=1,…,9)

$\kappa(p_1)$	$\kappa(p_2)$	$\kappa(p_3)$	$\kappa(p_4)$	$\kappa(p_5)$	$\kappa(p_6)$	$\kappa(p_7)$	$\kappa(p_8)$	$\kappa(p_9)$
6.32	2.83	1.71	−1.25	1.23	0.49	0.43	−0.35	−0.23

絶対値が大きく
正規分布に従わない
独立成分とみなせる

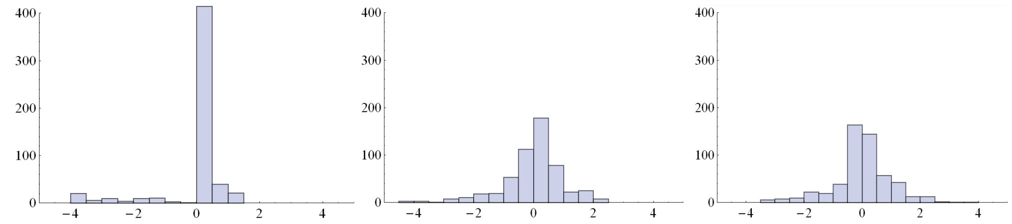
□ 絶対値が小さく
□ 正規分布とみなせる

図１ 独立成分分析による 重みベクトル $\boldsymbol{w}_i$（i＝1,⋯,5）



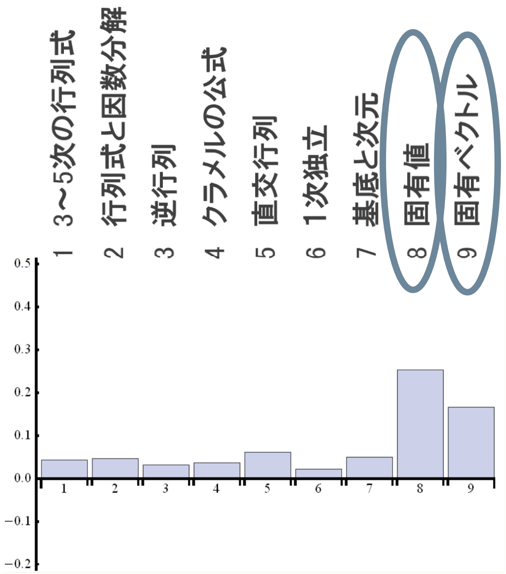
(1) $\boldsymbol{w}_1$　(2) $\boldsymbol{w}_2$　(3) $\boldsymbol{w}_3$　(4) $\boldsymbol{w}_4$　(5) $\boldsymbol{w}_5$

図２全採点データを重みベクトル $\boldsymbol{w}_i$ に射影した $p_i(k)$ (i=1,⋯,5)のヒストグラム



(1) $\kappa(p_1) = 6.32$　(2) $\kappa(p_2) = 2.83$　(3) $\kappa(p_3) = 1.71$　(4) $\kappa(p_4) = -1.25$　(5) $\kappa(p_5) = 1.23$

独立成分の重みベクトル w_1 と尖度 $\kappa(p_1)$

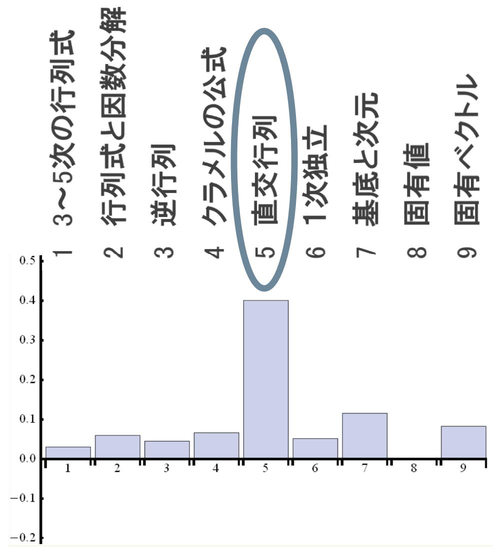


尖度 $\kappa(p_1) = 6.32$

学習内容

- □ 1　3〜5次の行列式
- □ 2　行列式と因数分解
- □ 3　逆行列
- □ 4　クラメルの公式
- □ 5　直交行列
- □ 6　1次独立
- □ 7　基底と次元
- □ 8　固有値
- □ 9　固有ベクトル

独立成分の重みベクトル w_4 と尖度 $\kappa(p_4)$



尖度 $\kappa(p_4) = -1.25$

学習内容

- □ 1　3〜5次の行列式
- □ 2　行列式と因数分解
- □ 3　逆行列
- □ 4　クラメルの公式
- □ 5　直交行列
- □ 6　1次独立
- □ 7　基底と次元
- □ 8　固有値
- □ 9　固有ベクトル

独立成分の重みベクトル w_i の解釈

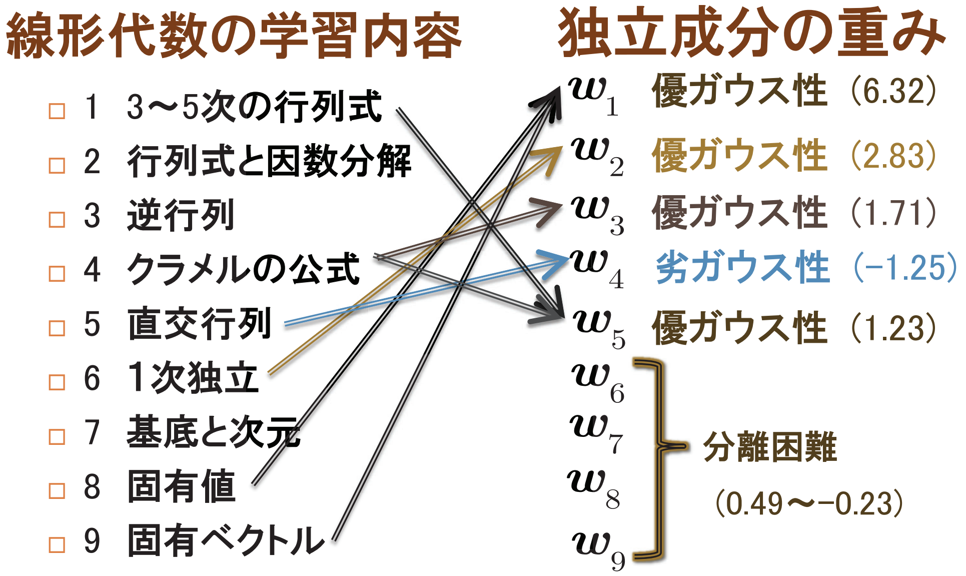
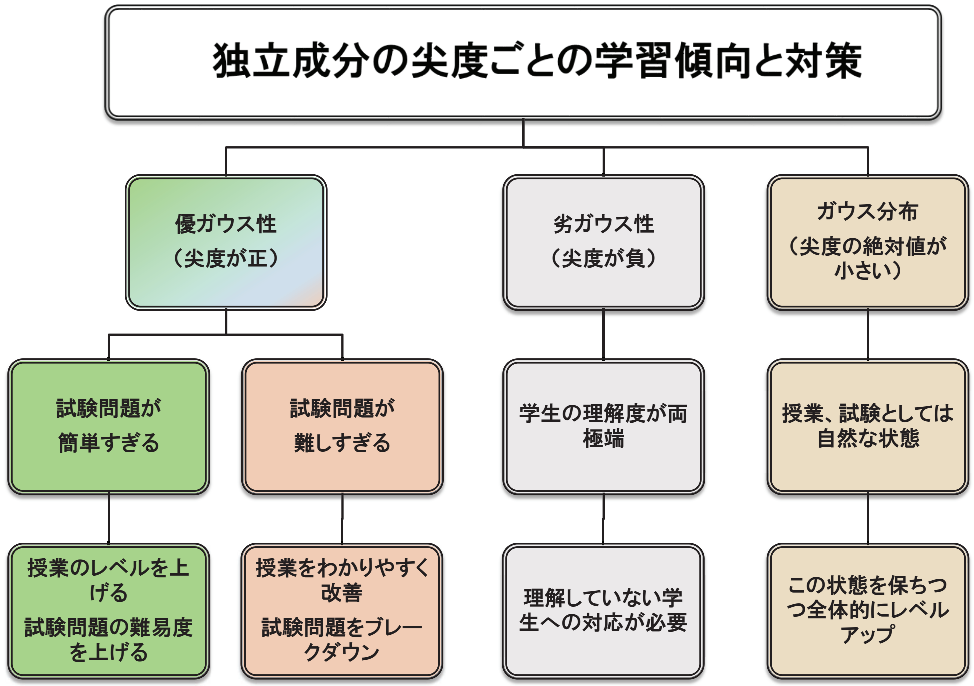


表２ 重み w_i (i=1,⋯,5)に射影したクラス毎のデータの尖度

クラス番号	$\kappa(p_1)$	$\kappa(p_2)$	$\kappa(p_3)$	$\kappa(p_4)$	$\kappa(p_5)$
1	7.35	5.91	4.35	−1.62	0.12
2	25.17	7.33	0.59	−1.84	1.51
3	9.92	3.97	−0.23	−0.33	−0.07
4	1.88	5.54	1.72	−1.60	0.46
5	−0.92	0.09	−0.10	−0.72	−0.65
6	6.43	1.79	0.98	−0.35	1.54
7	23.83	0.82	3.69	−1.28	1.27
8	21.58	10.61	1.24	−1.39	1.73
9	8.50	−0.06	−0.71	−1.39	1.40
10	6.98	−0.69	−0.84	−1.34	2.33



まとめ

- 独立な学習内容ごと学習傾向と対策を示した。
- クラスごとの独立成分の分析結果を授業にフィードバックし、進度や教え方をわかりやすく修正する目安を得られる。
- 解答の独立性を意識することで、学生の理解度を把握するのに効果的な試験問題を作成できる可能性がある。