独立成分分析
Independent Component Analysis
ICA

明治大学 理工学部 応用化学科
データ化学工学研究室　金子 弘昌

独立成分分析 (ICA) とは？

✓主成分分析 (Principal Component Analysis, PCA) は、
説明変数 X から互いに無相関な成分 (主成分) を計算する手法

- 主成分は、寄与率の大きい順に並べることが可能
  参考： http://datachemeng.com/principalcomponentanalysis/

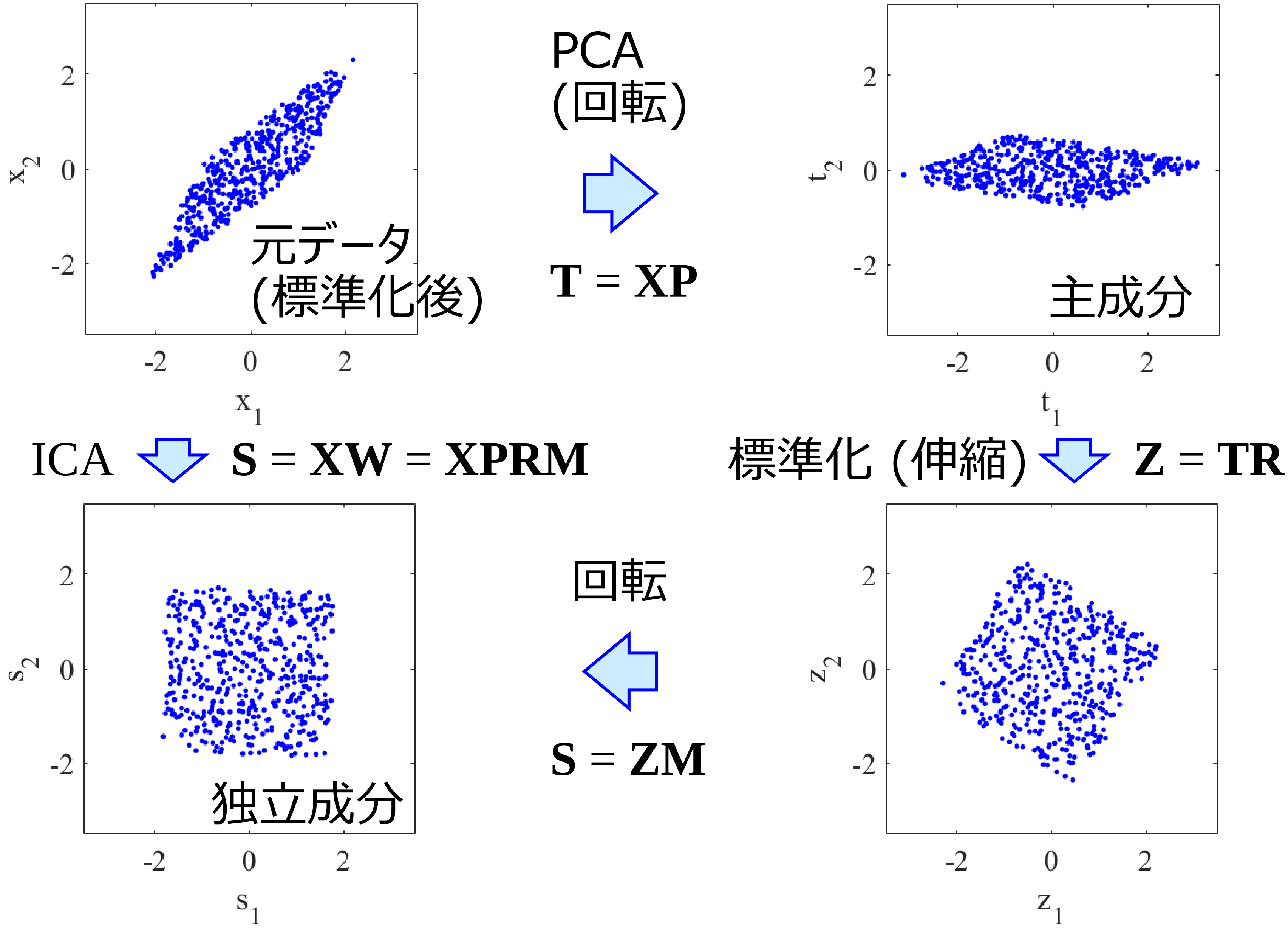
✓独立成分分析 (Independent Component Analysis, ICA) は、
説明変数 X から互いに独立な成分 (独立成分) を計算する手法

- 独立成分は、どれも平等

✓独立は無相関より強力

✓データセット内に外れ値があると、外れ値が強調されたような
独立成分が抽出される

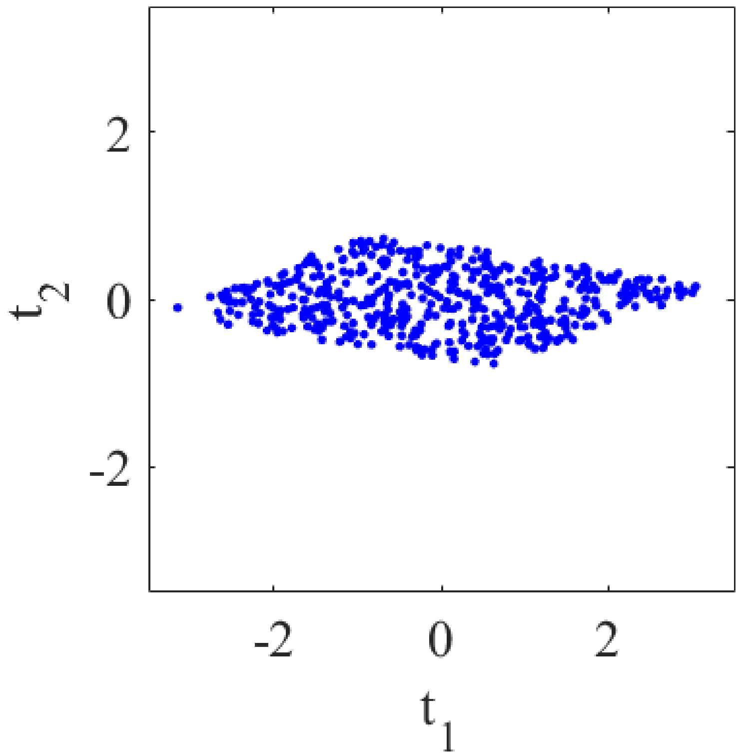
ICAの図解



無相関な変数 (成分) [数学的には p.25]

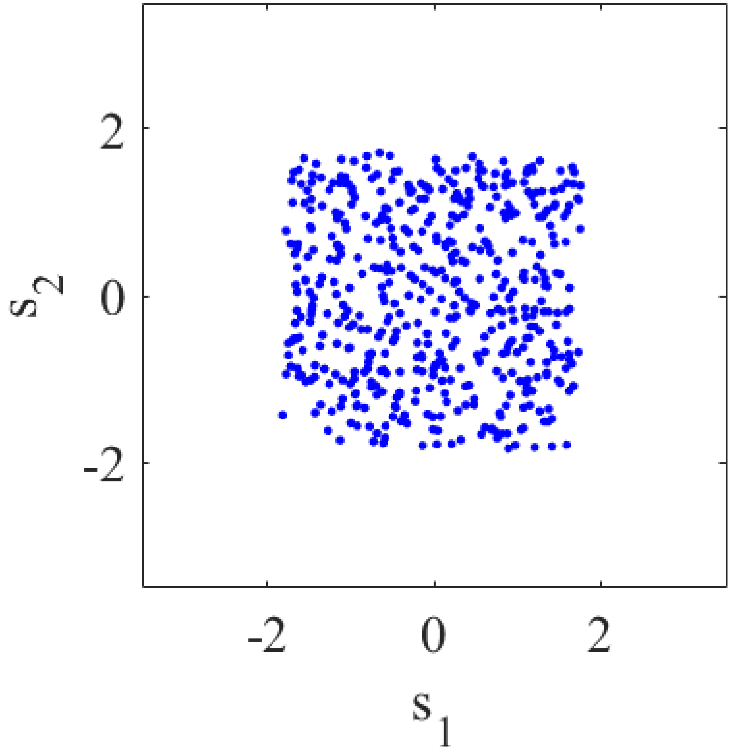
✓相互に関連がない変数

✓共分散が0である変数



独立な変数 (成分) [数学的には p.25]

✓変数 x_1 の値の情報が x_2 の値のどのような情報に対しても寄与しないならば、x_1 と x_2 は独立



✓s_1 の値が決まったとしても、s_2 の値には影響を与えていない

無相関と独立 [数学的には p.26, 27]

✓独立な変数ならば、無相関な変数

✓無相関な変数であっても、独立とは限らない

ICAの問題設定 1/2

変数

互いに独立な成分

サンプル

X から **Y** を推定する

X = **YA** となる混合行列 **A** を推定する

ICAの問題設定 2/2

X = **YA** となる混合行列 $\boxed{\mathbf{A}}$ を推定する

⇒ 分解行列 $\boxed{\mathbf{W}}$ を用いて、**S** = **XW** で計算できる **S** の

各成分が互いに独立となる **W** を求める

前処理としての主成分分析 (PCA)

✓変数の数が独立成分の数より多いならば、**W** は低次元化を行う行列

✓互いに独立な成分ならば、それらは無相関でもあるため、
W は無相関化する行列

ICA の前処理として PCA が利用される

PCA を行うとどうなる？

PCA $\mathbf{T} = \mathbf{XP}$

主成分スコアを標準化 (オートスケーリング) した行列を $\mathbf{Z}$ とすると、

$\mathbf{Z} = \mathbf{XM}$ $\mathbf{Z}$: 標準偏差が 1 で、互いに無相関な変数の行列

$\mathbf{Z} = \mathbf{XM} = \mathbf{YAM} = \mathbf{YB}$ とすると、

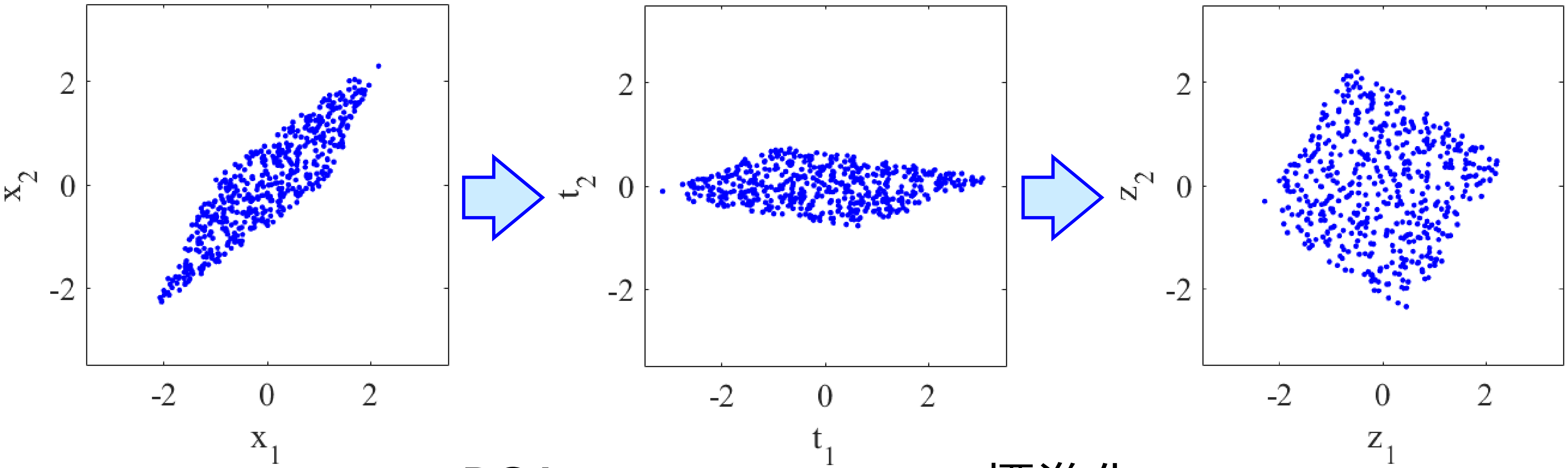
$\mathbf{Z}^{\mathrm{T}}\mathbf{Z} = \mathbf{E}$ $\mathbf{E}$: 単位行列 また、

$\mathbf{Z}^{\mathrm{T}}\mathbf{Z} = (\mathbf{YB})^{\mathrm{T}}\mathbf{YB} = \mathbf{B}^{\mathrm{T}}\mathbf{Y}^{\mathrm{T}}\mathbf{YB} = \mathbf{B}^{\mathrm{T}}\mathbf{B}$ より、 $\mathbf{B}^{\mathrm{T}}\mathbf{B} = \mathbf{E}$

⇨ PCAで前処理することで、任意の行列 $\mathbf{A}$ を推定する問題から、
直交行列 (かけると単位行列になる行列) $\mathbf{B}$ を推定する問題になった！

前処理を2次元の図で表す

元データ (標準化後) 主成分



PCA 標準化

$\mathbf{T} = \mathbf{XP}$ $\mathbf{Z} = \mathbf{TR}$

4 次キュムラント

✓平均・分散などの統計量の一つ

✓ある変数 x の 4次キュムラントを *kurt*(x) とすると、

$$kurt(\mathrm{x}) = E\left(\mathrm{x}^4\right) - 3E\left(\mathrm{x}^2\right)^2$$

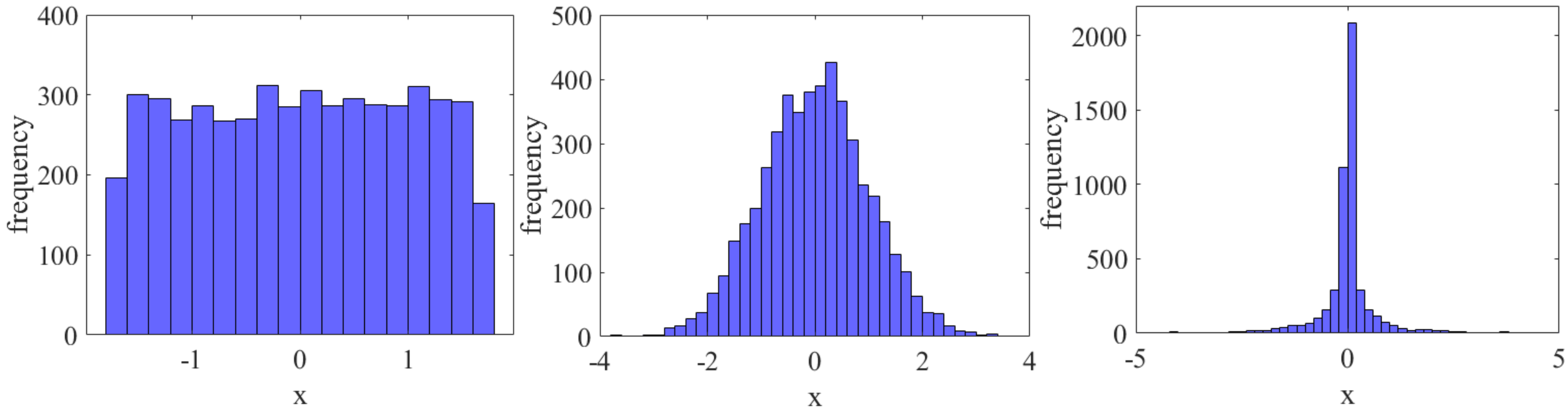
E(x) は x の期待値 (平均)

✓ある変数について、平均が 0、分散が 1 のとき、
尖度 (kurtosis) と呼ばれる

✓正規分布のとき、0 となる

- 尖った分布のとき、大きくなる
- 平べったい分布のとき、小さくなる

4 次キュムラントの例



$kurt(\mathrm{x}) = -1.1934$ $kurt(\mathrm{x}) = 0.0018$ $kurt(\mathrm{x}) = 52.9501$

外れ値があると、4 次キュムラントが大きくなる

4次キュムラントの性質

✓$kurt(\, a\mathrm{x}\,) = a^4\, kurt(\,\mathrm{x}\,)$

✓x_1 と x_2 が独立ならば、$kurt(\,\mathrm{x}_1 + \mathrm{x}_2\,) = kurt(\,\mathrm{x}_1\,) + kurt(\,\mathrm{x}_2\,)$

独立と関係ありそう！

4次キュムラントを最大もしくは最小にしてみる

PCA で前処理した２変数 z_1, z_2 について考える　　$\mathbf{Z} = [\mathbf{z}_1 \quad \mathbf{z}_2]$

互いに独立な成分 y_1, y_2 として、 $[\mathbf{z}_1 \quad \mathbf{z}_2] = [\mathbf{y}_1 \quad \mathbf{y}_2]\mathbf{B}$　　とする

このとき、$\mathbf{w}$ を単位ベクトルとして　　$\mathbf{s} = [\mathbf{z}_1 \quad \mathbf{z}_2]\mathbf{w}$

とした $\mathbf{s}$ の４次キュムラントを最大もしくは最小にすると、どうなるか？

ちなみに、　$\mathbf{s} = [\mathbf{z}_1 \quad \mathbf{z}_2]\mathbf{w} = [\mathbf{y}_1 \quad \mathbf{y}_2]\mathbf{B}\mathbf{w}$

4次キュムラントと独立性 式変形

$$\mathbf{s} = \begin{bmatrix} \mathbf{z}_1 & \mathbf{z}_2 \end{bmatrix} \mathbf{w} = \begin{bmatrix} \mathbf{y}_1 & \mathbf{y}_2 \end{bmatrix} \mathbf{Bw}$$

$$(\mathbf{Bw})^{\mathrm{T}} \mathbf{Bw} = \mathbf{w}^{\mathrm{T}} \mathbf{B}^{\mathrm{T}} \mathbf{Bw} = \mathbf{w}^{\mathrm{T}} \mathbf{Ew} = \mathbf{w}^{\mathrm{T}} \mathbf{w} = 1$$ より、

$\because \mathbf{B}^{\mathrm{T}} \mathbf{B} = \mathbf{E}$　　$\because$ $\mathbf{w}$ は単位ベクトル

$\mathbf{Bw}$ の大きさは 1 であるため、 $\mathbf{Bw} = \begin{bmatrix} \sin\theta & \cos\theta \end{bmatrix}^{\mathrm{T}}$ とおける

よって、 $\mathbf{s} = \sin\theta \mathbf{y}_1 + \cos\theta \mathbf{y}_2$

4次キュムラントと独立性 性質

$$\mathbf{s} = \sin\theta \mathbf{y}_1 + \cos\theta \mathbf{y}_2$$

y_1, y_2 は互いに独立であることから、4次キュムラントの性質より、

$$\begin{aligned} kurt(\mathbf{s}) &= kurt(\sin\theta \mathbf{y}_1) + kurt(\cos\theta \mathbf{y}_2) \\ &= \sin^4\theta kurt(\mathbf{y}_1) + \cos^4\theta kurt(\mathbf{y}_2) \end{aligned}$$

4次キュムラントと独立性 微分

$$kurt(\mathbf{s}) = \sin^4\theta kurt(\mathbf{y}_1) + \cos^4\theta kurt(\mathbf{y}_2)$$

θ で微分すると、

$$\begin{aligned}\frac{dkurt(\mathbf{s})}{d\theta} &= 4\sin^3\theta\cos\theta kurt(\mathbf{y}_1) - 4\cos^3\theta\sin\theta kurt(\mathbf{y}_2)\\ &= 4\sin\theta\cos\theta\left\{\sin^2\theta kurt(\mathbf{y}_1) - \cos^2\theta kurt(\mathbf{y}_2)\right\}\end{aligned}$$

4次キュムラントと独立性

$$\frac{dkurt(\mathbf{s})}{d\theta} = 4\sin\theta\cos\theta\left\{\sin^2\theta\, kurt(\mathbf{y}_1) - \cos^2\theta\, kurt(\mathbf{y}_2)\right\}$$

$\mathbf{s}$ の４次キュムラントが最大もしくは最小 $\Leftrightarrow$ $kurt(\mathbf{s})$ が極値をもつ

$\Leftrightarrow$ $\dfrac{dkurt(\mathbf{s})}{d\theta} = 0$ $\Leftrightarrow$ $\sin\theta = 0$ もしくは $\cos\theta = 0$

$\Leftrightarrow$ $\mathbf{s} = \pm\,\mathbf{y}_1$ もしくは $\mathbf{s} = \pm\,\mathbf{y}_2$ $\mathbf{y}_1, \mathbf{y}_2$ は独立な成分より、

４次キュムラントを最大もしくは最小にすれば、独立成分を計算できる

独立成分の計算

$$\mathbf{s} = \mathbf{Z}\mathbf{w}$$

$$\begin{aligned}
kurt(\mathbf{s}) = kurt(\mathbf{Z}\mathbf{w}) &= E\left\{(\mathbf{Z}\mathbf{w})^4\right\} - 3E\left\{(\mathbf{Z}\mathbf{w})^2\right\}^2 \\
&= E\left\{(\mathbf{Z}\mathbf{w})^4\right\} - 3E\left\{(\mathbf{Z}\mathbf{w})^{\mathrm{T}}\,\mathbf{Z}\mathbf{w}\right\}^2 \\
&= E\left\{(\mathbf{Z}\mathbf{w})^4\right\} - 3E\left(\mathbf{w}^{\mathrm{T}}\mathbf{Z}^{\mathrm{T}}\mathbf{Z}\mathbf{w}\right)^2 \\
&= E\left\{(\mathbf{Z}\mathbf{w})^4\right\} - 3E\left(\mathbf{w}^{\mathrm{T}}\mathbf{w}\right)^2 \qquad \because \mathbf{Z}^{\mathrm{T}}\mathbf{Z} = \mathbf{E} \\
&= E\left\{(\mathbf{Z}\mathbf{w})^4\right\} - 3E\left(\|\mathbf{w}\|^2\right)^2 \\
&= E\left\{(\mathbf{Z}\mathbf{w})^4\right\} - 3\|\mathbf{w}\|^4
\end{aligned}$$

Lagrangeの未定乗数法

$\mathbf{w}$ は単位ベクトル ($\|\mathbf{w}\|^2 = 1$) という制約のもとで、

$$kurt(\mathbf{s}) = E\left\{(\mathbf{Zw})^4\right\} - 3\|\mathbf{w}\|^4$$ を最大化もしくは最小化する

Lagrange の未定乗数法より、Lagrange 定数を λ とすると、

$$G = E\left\{(\mathbf{Zw})^4\right\} - 3\|\mathbf{w}\|^4 + \lambda\left(\|\mathbf{w}\|^2 - 1\right)$$

$$\frac{\partial G}{\partial \mathbf{w}} = E\left\{4(\mathbf{Zw})^3\,\mathbf{Z}\right\} - 12\|\mathbf{w}\|^3\,\mathbf{w} + 2\lambda\mathbf{w} = 0$$

$$\frac{\partial G}{\partial \lambda} = \|\mathbf{w}\|^2 - 1 = 0$$

w の決め方～不動点法～

$$E\left\{4(\mathbf{Zw})^3\mathbf{Z}\right\}-12\|\mathbf{w}\|^3\mathbf{w}+2\lambda\mathbf{w}=0$$

$$\Rightarrow\quad \mathbf{w}=-\frac{1}{2\lambda}E\left\{4(\mathbf{Zw})^3\mathbf{Z}\right\}-12\|\mathbf{w}\|^3\mathbf{w}$$

$\mathbf{w}$ の大きさは 1 ($\|\mathbf{w}\|^2=1$) なので、重要なのは $\mathbf{w}$ の方向のみ

$$\mathbf{w}_{\mathrm{new}}=-\frac{1}{2\lambda}E\left\{4(\mathbf{Zw}_{\mathrm{old}})^3\mathbf{Z}\right\}-12\|\mathbf{w}_{\mathrm{old}}\|^3\mathbf{w}_{\mathrm{old}}$$

上の式で $\mathbf{w}_{\mathrm{old}}\rightarrow\mathbf{w}_{\mathrm{new}}$ と更新し、 $\mathbf{w}_{\mathrm{old}}=\dfrac{\mathbf{w}_{\mathrm{new}}}{\|\mathbf{w}_{\mathrm{new}}\|}$

として $\mathbf{w}$ の大きさを 1 に修正すればよい
→ 再度 $\mathbf{w}_{\mathrm{new}}=\cdots$ の式で $\mathbf{w}$ を更新、大きさの修正、を繰り返す

不動点法のアルゴリズム

1. $\mathbf{w}_0$ を $\|\mathbf{w}_0\| = 1$ となるように乱数で初期化

2. $k = 0$ とする

3. $$\mathbf{w}_{k+1} = -\frac{1}{2\lambda} E\left\{4\left(\mathbf{Z}\mathbf{w}_k\right)^3 \mathbf{Z}\right\} - 12\|\mathbf{w}_k\|^3 \mathbf{w}_k$$

4. $$\mathbf{w}_{k+1} = \frac{\mathbf{w}_{k+1}}{\|\mathbf{w}_{k+1}\|}$$

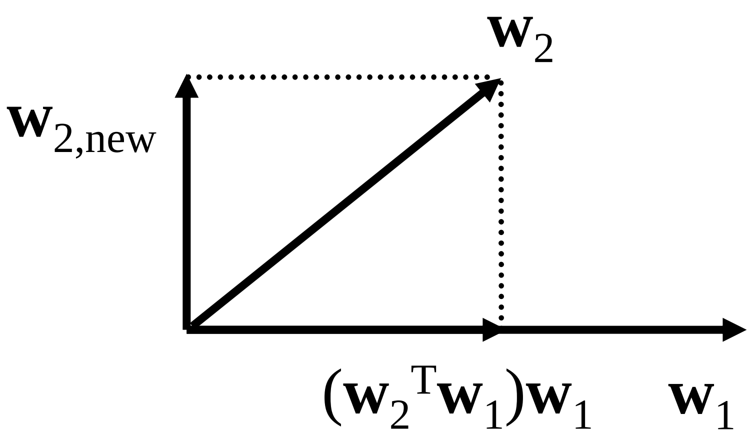
5. $|\,\mathbf{w}_k{}^{\mathrm{T}}\mathbf{w}_{k+1}\,|$ が十分 1 に近ければ終了、
そうでなければ $\mathrm{k} = \mathrm{k} + 1$ として 3. へ

次の独立成分を計算する

✓毎回異なる独立成分を求めるために、$\mathbf{w}_i^{\mathrm{T}}\mathbf{w}_j = 0$ ($i \neq j$) であることを利用する

✓新しい $\mathbf{w}_k$ を求めるときに、更新された $\mathbf{w}_k$ からすでに計算したすべての $\mathbf{w}$ の射影 ($\mathbf{w}_k^{\mathrm{T}}\mathbf{w}_i$) $\mathbf{w}_i$ ($i = 1, 2, \ldots, k-1$) を引く

$$\mathbf{w}_k = \mathbf{w}_k - \sum_{i=1}^{k-1}\left[\left\{\mathbf{w}_k^{\mathrm{T}}\mathbf{w}_i\right\}\mathbf{w}_i\right]$$



$$\mathbf{w}_2 = \mathbf{w}_{2,new} + \left(\mathbf{w}_2^{\mathrm{T}}\mathbf{w}_1\right)\mathbf{w}_1$$

$$\mathbf{w}_{2,new} = \mathbf{w}_2 - (\mathbf{w}_2^{\mathrm{T}}\mathbf{w}_1)\mathbf{w}_1$$

[参考] 数学的なこと

✓離散分布：確率関数 f(x) $$\sum_x f(x) = 1$$

✓連続分布：確率密度関数 p(x) $$\int_{-\infty}^{\infty} p(x)\,dx = 1$$

✓期待値：

$$E(\mathrm{x}) = \sum_x x f(x)$$

$$E(\mathrm{x}) = \int_{-\infty}^{\infty} x p(x)\,dx$$

✓共分散：

$$\begin{aligned}\mathrm{cov}(\mathrm{x}_1, \mathrm{x}_2) &= E\{(\mathrm{x}_1 - E(\mathrm{x}_1))(\mathrm{x}_2 - E(\mathrm{x}_2))\} \\ &= E(\mathrm{x}_1\mathrm{x}_2) - E(\mathrm{x}_1)E(\mathrm{x}_2) - E(\mathrm{x}_2)E(\mathrm{x}_1) + E\{E(\mathrm{x}_1)E(\mathrm{x}_2)\} \\ &= E(\mathrm{x}_1\mathrm{x}_2) - E(\mathrm{x}_1)E(\mathrm{x}_2)\end{aligned}$$

[参考] 数学的な無相関・独立な変数

✓無相関な変数

- 相互に関連がない変数
- 共分散が0である変数

$$\mathrm{cov}\left(\mathrm{x}_1, \mathrm{x}_2\right) = E\left(\mathrm{x}_1 \mathrm{x}_2\right) - E\left(\mathrm{x}_1\right) E\left(\mathrm{x}_2\right)$$
$$= 0$$

✓独立な変数

- 変数 x_1 の値の情報が x_2 の値のどのような情報に対しても寄与しないならば、x_1 と x_2 は独立
- x_1 と x_2 の結合の確率密度関数がそれぞれの確率密度関数に分解可能であるならば、x_1 と x_2 は独立

あらゆる x_1 と x_2 に対して、

$$f\left(\mathrm{x}_1, \mathrm{x}_2\right) = f_1\left(\mathrm{x}_1\right) f_2\left(\mathrm{x}_2\right) \qquad \left(f_1\left(x_1\right) = \sum_{x_2} f\left(x_1, x_2\right) \right)$$

$$p\left(\mathrm{x}_1, \mathrm{x}_2\right) = p_1\left(\mathrm{x}_1\right) p_2\left(\mathrm{x}_2\right) \qquad \left(p_1\left(x_1\right) = \int_{-\infty}^{\infty} p\left(x_1, x_2\right) dx_2 \right)$$

[参考] 数学的な無相関と独立の関係

✓無相関な変数

$$E(\mathrm{x}_1\mathrm{x}_2) - E(\mathrm{x}_1)E(\mathrm{x}_2) = 0$$

$$E(\mathrm{x}_1\mathrm{x}_2) = E(\mathrm{x}_1)E(\mathrm{x}_2)$$

✓独立な変数

あらゆる x_1 と x_2 に対して、

$$f(\mathrm{x}_1, \mathrm{x}_2) = f_1(\mathrm{x}_1) f_2(\mathrm{x}_2)$$

ある関数 h_1 と h_2 に対して、

$$\begin{aligned}
E\left\{h_1(X_1)h_2(X_2)\right\} &= \sum_{x_1}\sum_{x_2} h_1(x_1)h_2(x_2)f(x_1, x_2) \\
&= \sum_{x_1}\sum_{x_2} h_1(x_1)h_2(x_2)f_1(x_1) f_2(x_2) \\
&= \sum_{x_1}\sum_{x_2} h_1(x_1) f_1(x_1)h_2(x_2)f_2(x_2)
\end{aligned}$$

[参考] 数学的な無相関と独立の関係

✓無相関な変数

$$E(\mathrm{x}_1\mathrm{x}_2) - E(\mathrm{x}_1)E(\mathrm{x}_2) = 0$$

$$E(\mathrm{x}_1\mathrm{x}_2) = E(\mathrm{x}_1)E(\mathrm{x}_2)$$

✓独立な変数

あらゆる x_1 と x_2 に対して、

$$f(\mathrm{x}_1, \mathrm{x}_2) = f_1(\mathrm{x}_1) f_2(\mathrm{x}_2)$$

ある関数 h_1 と h_2 に対して、

$$\begin{aligned} E\{h_1(X_1)h_2(X_2)\} &= \sum_{x_1}\sum_{x_2} h_1(x_1) f_1(x_1) h_2(x_2) f_2(x_2) \\ &= \sum_{x_1} h_1(x_1) f_1(x_1) \sum_{x_2} h_2(x_2) f_2(x_2) \\ &= E\{h_1(X_1)\} E\{h_2(X_2)\} \end{aligned}$$

⇨ 独立な変数ならば、無相関な変数 (逆はいえない)

[参考] 独立性の指標

✓４次キュムラント

✓エントロピー

✓相互情報量