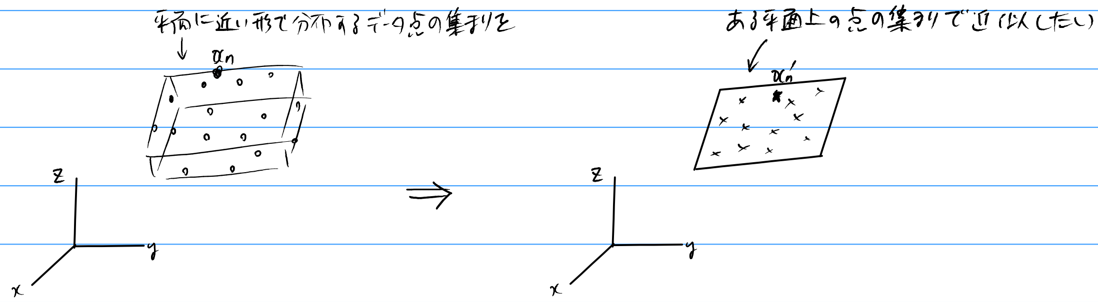
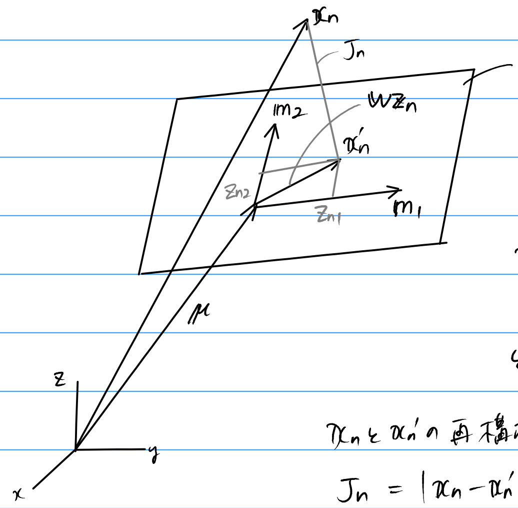
12.17

D=3, M=2 のときの状況



3次元データを ⇒ 2次元データで近似する

点 x_n についての近似はこんな感じになる



平面は $\boldsymbol{\mu}$ と $\mathbf{m}_1, \mathbf{m}_2$ で与えられる

平面上に x_n の近似点 x_n' を取る（正射影でなくてもよい）

x_n' は平面上にあるので

$$x_n' = \boldsymbol{\mu} + z_{n1}\mathbf{m}_1 + z_{n2}\mathbf{m}_2$$
$$= \boldsymbol{\mu} + \mathbf{W}\mathbf{z}_n$$

← $\mathbf{W} = (\mathbf{m}_1\ \mathbf{m}_2)$ なので

$\mathbf{W}\mathbf{z}_n = (\mathbf{m}_1\ \mathbf{m}_2)\begin{pmatrix} z_{n1} \\ z_{n2} \end{pmatrix}$

$= z_{n1}\mathbf{m}_1 + z_{n2}\mathbf{m}_2$

と表わされる。

x_n と x_n' の再構成コストは

$$J_n = |x_n - x_n'|^2 = |x_n - \boldsymbol{\mu} - \mathbf{W}\mathbf{z}_n|^2$$

で与えられるとする。

（$\boldsymbol{\mu}$, $\mathbf{W}$, $\mathbf{z}_n$ の推定方法）

① x_n, $\mathbf{z}_n$ の分布を考えて最尤推定で $\boldsymbol{\mu}$, $\mathbf{W}$, $E[\mathbf{z}_n]$ を推定する。
最尤推定を EMアルゴリズムで実行する。⇒ 12.2.2節

② x_n, $\mathbf{z}_n$ の分布は考えない。全データ点のトータルの再構成コストを最小にするように
$\boldsymbol{\mu}$, $\mathbf{W}$, $\mathbf{z}_n$ を推定する。⇒ この問題

(問題文の「…と同様な表現に導くことを示せ」というのは
μについて最小化したJを示せということ。)

(反復法でJの最小解を得る)

Jの最小解は

$$
\begin{cases}
\dfrac{\partial J}{\partial \mu}=0 \\
\dfrac{\partial J}{\partial W}=0 \\
\dfrac{\partial J}{\partial z_n}=0 \quad (n=1\sim N)
\end{cases}
$$

という連立方程式で与えられる。
直接連立方程式を解いてもいいが、問題文はこれを反復法で解くことを示唆している。
反復法の手順は

(1) W, z_n を反復の前回の値 W^{old}, z_n^{old} に固定して $\frac{\partial J}{\partial \mu}=0$ となる μ^{new} を求める
(2) μ を μ^{new} に固定し、W を W^{old}、$z_{i\neq n}$ を z_i^{old} に固定して $\frac{\partial J}{\partial z_n}=0$ となる z_n^{new} を求める
(3) μ を μ^{new} に固定し、z_n を z_n^{old} に固定して $\frac{\partial J}{\partial W}=0$ となる W^{new} を求める
(4) μ, W, z_n が収束するまで(1)~(3)を反復する。

とする。μについてはヤコビ法、W, z_n についてはガウスザイデル法のような反復になっている。ちょっと変則のように見えるが、このような反復にすると問題で要求されている(12.58), (12.59)とよく似た更新式を導くことができる。

(収束値が解になっていること)

収束したとき $W^{old}=W^{new}$, $z_n^{old}=z_n^{new}$ となり

$$
\left(\frac{\partial J}{\partial \mu}\right)_{\mu^{new}, W^{new}, z_n^{new}}=0,\ \left(\frac{\partial J}{\partial W}\right)_{\mu^{new}, W^{new}, z_n^{new}}=0,\ \left(\frac{\partial J}{\partial z_n}\right)_{\mu^{new}, W^{new}, z_n^{new}}=0 \quad (n=1\sim N)
$$

が成立するので、収束値は元の連立方程式の解になっている。

(収束するかどうか)

ヤコビ法とザイデル法の混ざったようなアルゴリズムなので
Jが μ, W, z_n について凸関数ならば収束すると思われる。(証明は?)

($\boldsymbol{\mu}$の更新式)

$$J = \sum_{n=1}^{N} \| \mathbf{x}_n - \boldsymbol{\mu} - \mathbf{W}\mathbf{z}_n \|^2 \cdots (12.95)$$

より

$$J = \sum_{n=1}^{N} (\mathbf{x}_n - \boldsymbol{\mu} - \mathbf{W}\mathbf{z}_n)^T (\mathbf{x}_n - \boldsymbol{\mu} - \mathbf{W}\mathbf{z}_n)$$

$$= \sum_{n=1}^{N} \left(\mathbf{x}_n^T\mathbf{x}_n - \mathbf{x}_n^T\boldsymbol{\mu} - \mathbf{x}_n^T\mathbf{W}\mathbf{z}_n - \boldsymbol{\mu}^T\mathbf{x}_n + \boldsymbol{\mu}^T\boldsymbol{\mu} + \boldsymbol{\mu}^T\mathbf{W}\mathbf{z}_n - \mathbf{z}_n^T\mathbf{W}^T\mathbf{x}_n + \mathbf{z}_n^T\mathbf{W}^T\boldsymbol{\mu} + \mathbf{z}_n^T\mathbf{W}^T\mathbf{W}\mathbf{z}_n \right)$$

これより

$$0 = \frac{\partial J}{\partial \boldsymbol{\mu}} = \sum_{n=1}^{N} \left\{ -\mathbf{x}_n - \mathbf{x}_n + 2\boldsymbol{\mu} + (\mathbf{W}\mathbf{z}_n) + \mathbf{W}\mathbf{z}_n \right\}$$

$$= \sum_{n=1}^{N} \left\{ -2\mathbf{x}_n + 2\boldsymbol{\mu} + 2\mathbf{W}\mathbf{z}_n \right\}$$

- (C.19) $\dfrac{\partial \mathbf{x}^T\mathbf{a}}{\partial \mathbf{x}} = \dfrac{\partial \mathbf{a}^T\mathbf{x}}{\partial \mathbf{x}} = \mathbf{a}$
- ベクトル微分のchain rule
  $\mathbf{u} = \mathbf{u}(\mathbf{x})$, $\mathbf{v} = \mathbf{v}(\mathbf{x})$ のとき
  $\dfrac{\partial \mathbf{u}^T\mathbf{v}}{\partial \mathbf{x}} = \dfrac{\partial \mathbf{u}}{\partial \mathbf{x}}\mathbf{u} + \dfrac{\partial \mathbf{v}}{\partial \mathbf{x}}\mathbf{v}$ (denominator layout)
  なので
  $\dfrac{\partial \mathbf{x}^T\mathbf{x}}{\partial \mathbf{x}} = 2\mathbf{x}$ (denominator layout)

注 (C.18)はi, jが逆になっている

× $\left(\dfrac{\partial \mathbf{a}}{\partial \mathbf{b}}\right)_{ij} = \dfrac{\partial a_i}{\partial b_j}$ (分子レイアウト)

○ $\left(\dfrac{\partial \mathbf{a}}{\partial \mathbf{b}}\right)_{ij} = \dfrac{\partial a_j}{\partial b_i}$ (分母レイアウト)

他の公式が分母レイアウトになってるので

(スカラー微分以外)

よって、$\boldsymbol{\mu}$の更新式は

$$\boldsymbol{\mu}^{new} = \frac{1}{N} \sum_{n=1}^{N} (\mathbf{x}_n - \mathbf{W}\mathbf{z}_n)$$

$$\therefore \boldsymbol{\mu}^{new} = \bar{\mathbf{x}} - \mathbf{W}\bar{\mathbf{z}}$$

となる。ただし右辺の$\mathbf{W}, \bar{\mathbf{z}}$は反復の前回値である。$\boldsymbol{\mu}^{new}$を(12.95)に入れて

$$J_{\boldsymbol{\mu}^{new}} = \sum_{n=1}^{N} \| \mathbf{x}_n - (\bar{\mathbf{x}} - \mathbf{W}\bar{\mathbf{z}}) - \mathbf{W}\mathbf{z}_n \|^2$$

$$= \sum_{n=1}^{N} \| (\mathbf{x}_n - \bar{\mathbf{x}}) - \mathbf{W}(\mathbf{z}_n - \bar{\mathbf{z}}) \|^2$$

を得る。

(z_n の更新式)

z_n の更新値 z_n^{new} は $\mu=\mu^{new}$ のときの J を最小にする z_n である。

$$0=\frac{\partial J}{\partial z_n}=\frac{\partial}{\partial z_n}\sum_{i=1}^{N}\|x_i-\mu-Wz_i\|^2$$

$$=\frac{\partial}{\partial z_n}\|x_n-\mu-Wz_n\|^2=\frac{\partial}{\partial z_n}(x_n-\mu-Wz_n)^T(x_n-\mu-Wz_n)$$

$$=\frac{\partial}{\partial z_n}\left\{(x_n-\mu)^T(x_n-\mu)-2(Wz_n)^T(x_n-\mu)+(Wz_n)^T(Wz_n)\right\}$$

$$=-2W^T(x_n-\mu)+2W^TWz_n$$

$$\frac{\partial z_n^T W^T(x_n-\mu)}{\partial z_n}=W^T(x_n-\mu) \leftarrow (C.19)\text{より}$$

$$\frac{\partial z_n^T W^T W z_n}{\partial z_n}=\frac{\partial}{\partial z_n}\mathrm{Tr}(z_n^T W^T W z_n) \leftarrow \text{スカラーの}\mathrm{Tr}$$

$$=\left(\frac{\partial}{\partial z_n^T}\mathrm{Tr}(z_n^T W^T W z_n)\right)^T \leftarrow \text{成分比較より}$$

$$=\left[z_n^T\{W^TW+(W^TW)^T\}\right]^T \leftarrow (C.27)$$

$$=2W^TWz_n$$

よって

$$W^TWz_n=W^T(x_n-\mu)$$

ここで、$\mu=\mu^{new}=\bar{x}-W\bar{z}$ とすると

$$W^TWz_n=W^T(x_n-\bar{x}+W\bar{z})$$

$$\therefore\ W^TW(z_n-\bar{z})=W^T(x_n-\bar{x})$$

よって z_n の更新式は

$$(z_n-\bar{z})^{new}=(W^TW)^{-1}W^T(x_n-\bar{x})$$

となる。ただし

$$(z_n-\bar{z})^{new}=z_n^{new}-\bar{z}^{new}$$

$$\bar{z}^{new}=\frac{1}{N}\left(z_n^{new}+\sum_{\substack{i=1\\ i\neq n}}^{N}z_i\right)$$

である。が、反復計算において z_n, $\bar{z}$ を個別に扱うことはなく必ず $(z_n-\bar{z})$ の形で処理される。

また式中の $z_{i\neq n}$, W は反復の前回値である。

$n=1 \sim N$ について並べて行列表記にすると

$$\left((z_1-\bar{z})^{new} \cdots (z_N-\bar{z})^{new}\right) = \left((W^TW)^{-1}W^T(x_1-\bar{x}) \cdots (W^TW)^{-1}W^T(x_N-\bar{x})\right)$$

$$= (W^TW)^{-1}W^T\left((x_1-\bar{x}) \cdots (x_N-\bar{x})\right)$$

ここで

$$\Omega = \left((z_1-\bar{z}) \cdots (z_N-\bar{z})\right), \quad \tilde{X} = \begin{pmatrix} (x_1-\bar{x})^T \\ \vdots \\ (x_N-\bar{x})^T \end{pmatrix}$$

とおくと

$$\Omega^{new} = (W^TW)^{-1}W^T\tilde{X}^T$$

となる。ただし Ω^{new} は Ω の $(z_n-\bar{z})$ を $(z_n-\bar{z})^{new}$ にしたもの。

これは (12.58) で $E[z_n]$ を $(z_n-\bar{z})^{new}$ に置き換えたものと同じになっている。

(Wの更新式)

Wの更新値 W^{new} は $\mu=\mu^{new}$ のときの J を最小にする W である。

$$0=\frac{\partial J}{\partial W}=\frac{\partial}{\partial W}\sum_{i=1}^{N}\|x_i-\mu-Wz_i\|^2$$

$$=\sum_{i=1}^{N}\frac{\partial}{\partial W}(x_n-\mu-Wz_n)^T(x_n-\mu-Wz_n)$$

$$=\sum_{i=1}^{N}\frac{\partial}{\partial W}\left\{(x_n-\mu)^T(x_n-\mu)-2(Wz_n)^T(x_n-\mu)+(Wz_n)^T(Wz_n)\right\}$$

$$=\sum_{n=1}^{N}\left\{-2(x_n-\mu)z_n^T+2Wz_nz_n^T\right\}$$

$$\frac{\partial}{\partial W}z_n^TW^T(x_n-\mu)=\frac{\partial}{\partial W}\mathrm{Tr}\left(z_n^TW^T(x_n-\mu)\right)\leftarrow \text{スカラーのTr}$$

$$=\frac{\partial}{\partial W}\mathrm{Tr}\left(W^T(x_n-\mu)z_n^T\right)\leftarrow(C.9)$$

$$=(x_n-\mu)z_n^T\leftarrow(C.25)$$

$$\frac{\partial}{\partial W}z_n^TW^TWz_n=\frac{\partial}{\partial W}\mathrm{Tr}(z_n^TW^TWz_n)\leftarrow \text{スカラーのTr}$$

$$=\frac{\partial}{\partial W}\mathrm{Tr}(Wz_nz_n^TW^T)\leftarrow(C.9)$$

$$=W\left(z_nz_n^T+(z_nz_n^T)^T\right)\leftarrow(C.27)$$

$$=2Wz_nz_n^T$$

よって

$$W\sum_{n=1}^{N}z_nz_n^T=\sum_{n=1}^{N}(x_n-\mu)z_n^T$$

ここで、$\mu=\mu^{new}=\bar{x}-W\bar{z}$ とすると

$$W\sum_{n=1}^{N}z_nz_n^T=\sum_{n=1}^{N}(x_n-\bar{x}+W\bar{z})z_n^T$$

$$\therefore W\sum_{n=1}^{N}(z_n-\bar{z})z_n^T=\sum_{n=1}^{N}(x_n-\bar{x})z_n^T$$

よって

$$W\left\{\sum_{n=1}^{N}(z_n-\bar{z})z_n^T-\sum_{n=1}^{N}(z_n-\bar{z})\bar{z}^T\right\}=\sum_{n=1}^{N}(x_n-\bar{x})z_n^T-\sum_{n=1}^{N}(x_n-\bar{x})\bar{z}^T$$

$$\sum_{n=1}^{N}(z_n-\bar{z})\bar{z}^T=\left(\sum_{n=1}^{N}z_n\right)\bar{z}^T-N\bar{z}\bar{z}^T=N\bar{z}\bar{z}^T-N\bar{z}\bar{z}^T=0$$

$$\sum_{n=1}^{N}(x_n-\bar{x})\bar{z}^T=\left(\sum_{n=1}^{N}x_n\right)\bar{z}^T-N\bar{x}\bar{z}^T=N\bar{x}\bar{z}^T-N\bar{x}\bar{z}^T=0$$

$$\therefore W\left\{\sum_{n=1}^{N}(z_n-\bar{z})(z_n-\bar{z})^T\right\}=\sum_{n=1}^{N}(x_n-\bar{x})(z_n-\bar{z})^T$$

これよりWの更新式は

$$W^{new}=\left\{\sum_{n=1}^{N}(x_n-\bar{x})(z_n-\bar{z})^T\right\}\left\{\sum_{n=1}^{N}(z_n-\bar{z})(z_n-\bar{z})^T\right\}^{-1}$$

となる。ただし右辺の $z_n,\bar{z}$ は反復の前回値である。

右辺の{…}の中は行列の積で表わせる

$$W^{new}=\left\{\left((x_1-\bar{x})\cdots(x_N-\bar{x})\right)\begin{pmatrix}(z_1-\bar{z})^T\\ \vdots\\ (z_N-\bar{z})^T\end{pmatrix}\right\}\left\{\left((z_1-\bar{z})\cdots(z_N-\bar{z})\right)\begin{pmatrix}(z_1-\bar{z})^T\\ \vdots\\ (z_N-\bar{z})^T\end{pmatrix}\right\}^{-1}$$

なので上記 Ω, $\tilde{X}$ を使うと

$$W^{new}=\tilde{X}^T\Omega^T(\Omega\Omega^T)^{-1}$$

となる。これは(12.59)で $E[z_n]$ を $z_n-\bar{z}$ に置き換えたものと同じになっている。

(更新値の与え方について)

z_nの更新値を $\frac{\partial J}{\partial z_n}=0$ を与える z_n の $\mu=\mu^{new}$ での値としたが、

これは $\frac{\partial J}{\partial z_n}=0$ と $\mu=\mu^{new}(z_n, w)$ の交点を更新値としたことになる。

他にも $\mu=\mu^{new}(z_n, w)$ 空間において $\frac{\partial J}{\partial z_n}=0$ を与える z_n を探索しこれを z_n の更新とすることもできる。

この場合 z_n の偏微分において μ^{new} と z_n は独立ではなくなる。

w についても同様

このような更新値であっても収束値は元の方程式の解になっていると思われるが

$\mu=\mu^{new}(z_n, w)$ 空間に解が入ることがなければ収束しなかったりすると思われる。