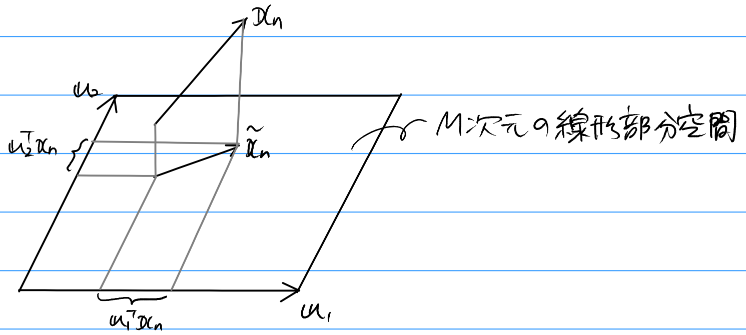
12.1

(M次元の線形部分空間へのデータの射影の分散)



x_n のM次元線形部分空間への射影 $\tilde{x}_n$ は成分表示で

$$\tilde{x}_n = (u_1^T x_n, \cdots, u_M^T x_n)$$

となる。同じように $\bar{x}$ のM次元線形部分空間への射影 $\tilde{\bar{x}}$ は

$$\tilde{\bar{x}} = (u_1^T \bar{x}, \cdots, u_M^T \bar{x})$$

これより

$$\tilde{x}_n - \tilde{\bar{x}} = (u_1^T x_n - u_1^T \bar{x}, \cdots, u_M^T x_n - u_M^T \bar{x}) = (u_1^T (x_n - \bar{x}), \cdots, u_M^T (x_n - \bar{x}))$$

よって

$$(\tilde{x}_n - \tilde{\bar{x}})^2 = \sum_{i=1}^{M} \{u_i^T (x_n - \bar{x})\}^2 = \sum_{i=1}^{M} \{u_i^T (x_n - \bar{x})\}\{u_i^T (x_n - \bar{x})\}$$

$$= \sum_{i=1}^{M} \{u_i^T (x_n - \bar{x})\}\{(x_n - \bar{x})^T u_i\}$$ スカラーなので転置しても同じ

$$= \sum_{i=1}^{M} \{u_i^T (x_n - \bar{x})(x_n - \bar{x})^T u_i\}$$

よってM次元線形部分空間へ射影されたデータの分散は

$$\frac{1}{N} \sum_{n=1}^{N} (\tilde{x}_n - \tilde{\bar{x}})^2 = \frac{1}{N} \sum_{n=1}^{N} \sum_{i=1}^{M} \{u_i^T (x_n - \bar{x})(x_n - \bar{x})^T u_i\}$$

$$= \sum_{i=1}^{M} u_i^T \left\{ \frac{1}{N} \sum_{n=1}^{N} (x_n - \bar{x})(x_n - \bar{x})^T \right\} u_i$$

$$= \sum_{i=1}^{M} u_i^T S u_i$$ ← $S = \frac{1}{N} \sum_{n=1}^{N} (x_n - \bar{x})(x_n - \bar{x})^T$ (12.3)

となる。

(逐次的に求めた解が一括して求める場合の解になっていることの確認)

第1主成分、第2主成分、…、第M主成分と逐次的に求めた u_i $(Su_i=\lambda_i u_i)$ が、M次元空間を一括して求める場合の解になっていることを確認する

M次元空間を一括して求める場合のラグランジアンは

$$L=\sum_{i=1}^{M} u_i^T S u_i+\sum_{i=1}^{M}\lambda_i(1-u_i^T u_i)+\sum_{i=1}^{M}\sum_{j=i+1}^{M}\mu_{ij}u_i^T u_j$$

これより

$$0=\frac{\partial L}{\partial u_i}=2(Su_i-\lambda_i u_i)+\sum_{j=i+1}^{M}\mu_{ij}u_j \quad \cdots ①$$

という条件を得る。

u_iが逐次的に求めた解 $Su_i=\lambda_i u_i$であるとすると

①の両辺に u_j^T $(j=i+1\sim M)$をかけて

$$0=2u_j^T S u_i+\mu_{ij} \quad \leftarrow u_j^T u_i=\delta_{ji}$$

$Su_i=\lambda_i u_i$より

$$=2\lambda_i u_j^T u_i+\mu_{ij}$$

$$=\mu_{ij}$$

を得る。このμ_{ij}を①に入れると

$$0=Su_i-\lambda_i u_i$$

となり、$Su_i=\lambda_i u_i$のとき

①は

$$0=0$$

となり、$Su_i=\lambda_i u_i$は①の解であることが確認できる。

(問題の答)

第 M 主成分まで逐次的に u_i を求めたとき

u_i は $Su_i = \lambda_i u_i,\ i = 1 \sim M$, λ_i は大きい方の固有値

であると仮定する。

この仮定の下で第 $M+1$ 主成分 u_{M+1} を求める。

分散最大、$u_{M+1}^T u_{M+1} = 1$, $u_{M+1}^T u_j = 0$ という条件

に対するラグランジアンは、

$$L = \sum_{i=1}^{M+1} u_i^T S u_i + \lambda_{M+1}(1 - u_{M+1}^T u_{M+1}) + \sum_{j=1}^{M} \mu_j u_{M+1}^T u_j$$

両辺を u_{M+1} で微分して 0 とおくと

$\left\{ \begin{array}{l} \cdot\ \dfrac{\partial x^T A x}{\partial x} = (A^T + A)x = 2Ax \quad \leftarrow A \text{ 対称のとき} \\ \cdot\ \dfrac{\partial x^T a}{\partial x} = a \quad \leftarrow \text{但し denominator layout} \end{array} \right.$

$$0 = \frac{\partial L}{\partial u_{M+1}} = 2(Su_{M+1} - \lambda_{M+1} u_{M+1}) + \sum_{j=1}^{M} \mu_j u_j \quad \cdots ①$$

を得る。両辺に左から u_j^T をかけると

$$0 = 2u_j^T S u_{M+1} + \mu_j \quad \leftarrow u_j^T u_i = \delta_{ji}$$

$$= \mu_j$$

スカラーなので転置しても同じ

$$u_j^T S u_{M+1} = u_{M+1}^T S u_j$$

$$= u_{M+1}^T \lambda_j u_j \quad \leftarrow \text{仮定 } Su_j = \lambda_j u_j \text{ より}$$

$$= 0 \quad \leftarrow u_{M+1}^T u_j = 0 \text{ なので}$$

となる。これを①に戻して

$$Su_{M+1} = \lambda_{M+1} u_{M+1} \quad \cdots ②$$

を得る。

このとき分散は

$$\sum_{i=1}^{M+1} u_i^T S u_i = \sum_{i=1}^{M} u_i^T S u_i + u_{M+1}^T S u_{M+1}$$

$$= \sum_{i=1}^{M} \lambda_i + \lambda_{M+1}$$

仮定より $\sum_{i=1}^{M} u_i^T S u_i = \sum_{i=1}^{M} u_i^T \lambda_i u_i = \sum \lambda_i$

②より $u_{M+1}^T S u_{M+1} = u_{M+1}^T \lambda_{M+1} u_{M+1} = \lambda_{M+1}$

となる。

仮定より λ_i は大きい方の M 個の固有値なので、

分散を最大にするには λ_{M+1} は $M+1$ 番目に大きい固有値になる。

よって第 $M+1$ 主成分は $Su_{M+1} = \lambda_{M+1} u_{M+1}$, λ_{M+1} は $M+1$ 番目に大きい固有値

となる。