

「パターン認識と機械学習」

2.3 ガウス分布

小野寺喜行

ガウス分布(正規分布)

変数が x の1つの場合のガウス分布

$$N(x | \mu, \sigma^2) = \frac{1}{(2\pi\sigma^2)^{1/2}} \exp\left\{-\frac{1}{2\sigma^2}(x - \mu)^2\right\}$$

μ : 平均 σ^2 : 分散

D 次元ベクトル \mathbf{x} に対する多変量ガウス分布

$$N(\mathbf{x} | \boldsymbol{\mu}, \boldsymbol{\Sigma}) = \frac{1}{(2\pi)^{D/2}} \frac{1}{|\boldsymbol{\Sigma}|^{1/2}} \exp\left\{-\frac{1}{2}(\mathbf{x} - \boldsymbol{\mu})^T \boldsymbol{\Sigma}^{-1} (\mathbf{x} - \boldsymbol{\mu})\right\}$$

$\boldsymbol{\mu}$: D 次元の平均ベクトル $\boldsymbol{\Sigma}$: $D \times D$ の共分散行列

$|\boldsymbol{\Sigma}|$: $\boldsymbol{\Sigma}$ の行列式

中心極限定理

複数の確率変数の和の確率変数



足し合わされる変数の数が増えていく

徐々にガウス分布に従うようになる

ガウス分布の性質

指数分布部分

$$\Delta^2 = (\mathbf{x} - \boldsymbol{\mu})^T \boldsymbol{\Sigma}^{-1} (\mathbf{x} - \boldsymbol{\mu})$$

を通して \mathbf{x} に依存。 Δ^2 が \mathbf{x} 空間中で定数となる面で、ガウス分布の密度は一定。

(Δ : $\boldsymbol{\mu}$ から \mathbf{x} までのマハラノビス距離

(Mahalanobis distance))

ガウス分布の性質

$i=1, \dots, D$ について、共分散行列に対する固有ベクトルの方程式

$$\Sigma \mathbf{u}_i = \lambda_i \mathbf{u}_i$$

次の様に互いに正規直行するようなものを選ぶ事が可能

$$\mathbf{u}_i^T \mathbf{u}_j = I_{ij}$$

(I_{ij} は単位行列の i,j 要素)

ガウス分布の性質

固有ベクトルを用いて共分散行列 Σ は以下の様に表せる

$$\Sigma = \sum_{i=1}^D \lambda_i \mathbf{u}_i \mathbf{u}_i^T$$

同様に、共分散行列の逆行列 Σ^{-1} も以下の様に表せる

$$\Sigma^{-1} = \sum_{i=1}^D \frac{1}{\lambda_i} \mathbf{u}_i \mathbf{u}_i^T$$

ガウス分布の性質

共分散行列の逆行列の式より

$$\Delta^2 = \sum_{I=1}^D \frac{y_i^2}{\lambda_i}$$

ただし

$$y_i = \mathbf{u}_i^T (\mathbf{x} - \boldsymbol{\mu})$$

ガウス分布の性質

$y_1 \sim y_D$ をベクトル $\mathbf{y} = (y_1, \dots, y_D)^T$ にまとめると

$$\mathbf{y} = \mathbf{U}(\mathbf{x} - \boldsymbol{\mu})$$

ただし、 \mathbf{U} はその i 行が \mathbf{u}_i^T である行列。また、直交行列となる。

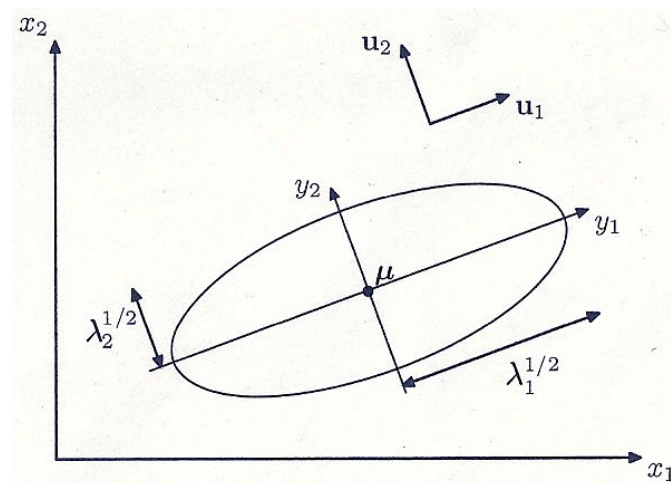
ガウス分布の性質

y_i が定数となる面では、ガウス分布の密度も一定になる



その面が

もし全ての固有値 λ が正ならば、これらの面は図のような楕円体になる



ガウス分布の性質

共分散行列の全ての固有値 λ_i が正でなくてはならない

→ 分布を適切に正規化できるようにし、定義する為

正定値 (positive definite) :

固有値が厳密に正である行列

半正定値 (positive semidefinite) :

すべての固有値が非負である行列

y_i で定義された座標系でのガウス分布

x から y の座標系に写すために、ヤコビ行列 J を用いる

ヤコビ行列の要素: $J_{ij} = \frac{\delta x_i}{\delta y_j} = U_{ji}$

(U_{ji} は行列 U^T の要素)

y_i で定義された座標系でのガウス分布

ヤコビ行列の行列式の2乗

$$|\mathbf{J}|^2 = |\mathbf{U}^T|^2 = |\mathbf{U}^T| |\mathbf{U}| = |\mathbf{U}^T \mathbf{U}| = |\mathbf{I}| = 1$$

共分散行列の行列式 $|\Sigma|$ も固有ベクトルの積で書けるので、

$$|\Sigma|^{1/2} = \prod_{j=1}^D \lambda_j^{1/2}$$

y_i で定義された座標系でのガウス分布

y_i 座標系では、多変量ガウス分布は D 個の独立な1変数ガウス分布の積となる

$$p(\mathbf{y}) = p(\mathbf{x})|\mathbf{J}| = \prod_{j=1}^D \frac{1}{(2\pi\lambda_j)^{1/2}} \exp\left\{-\frac{y_j^2}{2\lambda_j}\right\}$$

このとき

$$\int p(\mathbf{y}) d\mathbf{y} = \prod_{j=1}^D \int_{-\infty}^{\infty} \frac{1}{(2\pi\lambda_j)^{1/2}} \exp\left\{-\frac{y_j^2}{2\lambda_j}\right\} dy_j = 1$$

ガウス分布のモーメント

ガウス分布の下での \mathbf{x} の期待値

$$\mathbf{E}[\mathbf{x}] = \frac{1}{(2\pi)^{D/2}} \frac{1}{|\Sigma|^{1/2}} \int \exp\left\{-\frac{1}{2}(\mathbf{x} - \boldsymbol{\mu})^T \Sigma^{-1}(\mathbf{x} - \boldsymbol{\mu})\right\} \mathbf{x} d\mathbf{x}$$

$$= \frac{1}{(2\pi)^{D/2}} \frac{1}{|\Sigma|^{1/2}} \int \exp\left\{-\frac{1}{2}\mathbf{z}^T \Sigma^{-1}\mathbf{z}\right\} (\mathbf{z} + \boldsymbol{\mu}) d\mathbf{z}$$

$$= \boldsymbol{\mu} \quad (\text{因子}(\mathbf{z} + \boldsymbol{\mu})\text{中の}\mathbf{z}\text{の項は対称性で消える})$$

(ただし $\mathbf{z} = \mathbf{x} - \boldsymbol{\mu}$)

ガウス分布のモーメント

ガウス分布の2次モーメント

$$\begin{aligned}\mathbf{E}[\mathbf{xx}^T] &= \frac{1}{(2\pi)^{D/2}} \frac{1}{|\Sigma|^{1/2}} \int \exp\left\{-\frac{1}{2}(\mathbf{x} - \boldsymbol{\mu})^T \Sigma^{-1}(\mathbf{x} - \boldsymbol{\mu})\right\} \mathbf{xx}^T d\mathbf{x} \\ &= \frac{1}{(2\pi)^{D/2}} \frac{1}{|\Sigma|^{1/2}} \int \exp\left\{-\frac{1}{2}\mathbf{z}^T \Sigma^{-1}\mathbf{z}\right\} (\mathbf{z} + \boldsymbol{\mu})(\mathbf{z} + \boldsymbol{\mu})^T d\mathbf{z}\end{aligned}$$

- $\boldsymbol{\mu}\mathbf{z}^T$ と $\mathbf{z}\boldsymbol{\mu}^T$ の交差項は指数部分の対称性で消える
- $\boldsymbol{\mu}\boldsymbol{\mu}^T$ の項の積分部分は1となる

ガウス分布のモーメント

ガウス分布の2次モーメント

- $\mathbf{z}\mathbf{z}^T$ を含む項について

共分散行列の固有ベクトル展開と、固有ベクトル集合の完全性より $y_j = \mathbf{u}_j^T \mathbf{z}$ なので

$$\mathbf{z} = \sum_{j=1}^D y_j \mathbf{u}_j$$

ガウス分布のモーメント

ガウス分布の2次モーメント

- $\mathbf{z}\mathbf{z}^T$ を含む項について

$$\begin{aligned} & \frac{1}{(2\pi)^{D/2}} \frac{1}{|\Sigma|^{1/2}} \int \exp\left\{-\frac{1}{2}\mathbf{z}^T \Sigma^{-1} \mathbf{z}\right\} \mathbf{z}\mathbf{z}^T d\mathbf{z} \\ &= \frac{1}{(2\pi)^{D/2}} \frac{1}{|\Sigma|^{1/2}} \sum_{i=1}^D \sum_{j=1}^D \mathbf{u}_i \mathbf{u}_j^T \int \exp\left\{-\sum_{k=1}^D \frac{y_k^2}{2\lambda_k}\right\} y_i y_j d\mathbf{y} \\ &= \sum_{i=1}^D \mathbf{u}_i \mathbf{u}_i^T \lambda_i = \Sigma \end{aligned}$$

ガウス分布のモーメント

ガウス分布の2次モーメント

$$\mathbf{E}[\mathbf{x}\mathbf{x}^T] = \boldsymbol{\mu}\boldsymbol{\mu}^T + \boldsymbol{\Sigma}$$

- 共分散

$$\text{cov}[\mathbf{x}] = \mathbf{E}\left[(\mathbf{x} - \mathbf{E}[\mathbf{x}])(\mathbf{x} - \mathbf{E}[\mathbf{x}])^T\right] = \boldsymbol{\Sigma}$$



共分散行列と呼ぶ

ガウス分布の自由パラメータの数

一般対称共分散行列 Σ : $D(D+1)/2$ 個

μ : D 個

合計: $D(D+3)/2$ 個



パラメータの総数が D に対して2乗の割合で増加して計算が困難に

ガウス分布の自由パラメータの数

共分散行列が対角 (diagonal) ($\Sigma = \text{diag}(\sigma_i^2)$) であると制限を加える



パラメータの総数が $2D$ になる

さらに $\Sigma = \sigma^2 \mathbf{I}$ のように共分散行列が単位行列に比例するように制限を強める (等方共分散行列 (isotropic covariance matrix))



パラメータの総数が $D+1$ 個になる

ガウス分布の自由パラメータの数

共分散行列の形式に制限を加える事で、分布の自由度が制限される。



- 共分散行列の逆行列をより高速に求められる
- データ中の興味深い相関を捉える能力を制限してしまう

ガウス分布の問題

- パラメータが多すぎる
- 本質的に単峰系という制限があり、多峰系の分布をうまく近似できない



潜在変数 (latent variable) (**隠れ変数** (hidden variable) / **非観測変数** (unobserved variable)) を導入することで対処する。