

# パターン認識と機械学習

## 2.3 ガウス分布

05T4074A

久保田 敦

# 目標

- ガウス分布の様々な解析的性質に触れる。その際に使う技法に慣れることで、後のさらに複雑なモデルの理解に役立てるようになる。

# 目次

- ガウス分布
- 中心極限定理
- 座標系の変換
- 正定値
- $y$ 座標系
- ガウス分布の問題点

# ガウス分布

$x$ : 変数      ( $\mu$ : 平均    $\sigma$ : 分散)

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

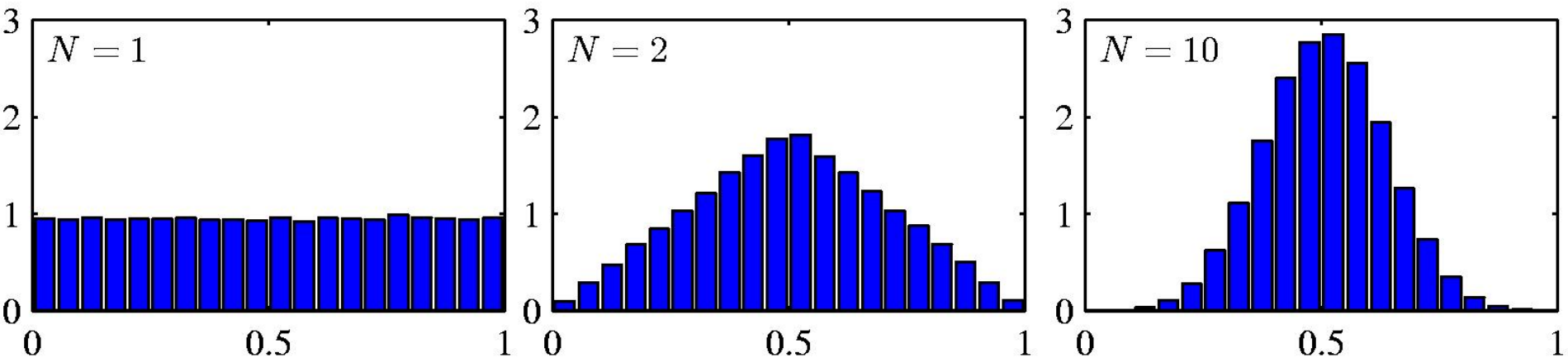
# 多変量ガウス分布

$\mathbf{x}$ :  $D$ 次元ベクトル      ( $\Sigma$ :  $D \times D$ の共分散行列)

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

# 中心極限定理

- どんな確立分布の確立変数でもそれらの和の極限をとるとその平均はガウス分布になる。



# 座標系の変換

- マハラノビス距離:  $\Delta$

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

$\boldsymbol{\Sigma}$ が単位行列ならユークリッド距離  
(異なる変数の間に相関がない)

- ◆この二次形式の値が $\mathbf{x}$ 空間で定数となる面で  
ガウス分布の密度は一定である。

$\Sigma$ は対称行列であるとして、共分散に対する固有ベクトルの方程式は

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

固有値が実数なので、固有ベクトルも正規直交するものが選べる

$$\mathbf{u}_i^T \mathbf{u}_j = I_{ij} \quad I_{ij} = \begin{cases} 1 & i=j \text{ のとき} \\ 0 & \text{それ以外} \end{cases}$$

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

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

マハラノビス距離の式に代入

$$\Delta^2 = \sum_{i=1}^D \frac{y_i^2}{\lambda_i} \quad [y_i = \mathbf{u}_i^T (\mathbf{x} - \boldsymbol{\mu})]$$

$x_i$  の座標系  $\Rightarrow$   $\mathbf{u}_i$  の座標系

$$\mathbf{y} = (y_1, \dots, y_D)^T = \mathbf{U}(\mathbf{x} - \boldsymbol{\mu})$$

(U : i行が  $\mathbf{u}_i^T$  である直交行列)

# 正定値

- ガウス分布が適切に定義されるには共分散行列の固有値  $\lambda_i$  が正でなくては分布を正規化できない。

正定値：固有値が正である行列

反正定値：固有値が非負の行列

# y 座標系

多変量ガウス分布:

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

同時確率分布を独立な分布の積に分解できた。  
(1.48)を用いると正規化されていることがわかる

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

# ガウス分布の問題点

パラメータ数が多いので計算が困難



共分散行列の形式の制限



適切に表現できる分布の範囲が制限される



潜在変数の導入(2.3.9)