知能機械と自然言語処理

知能機械部　第7回

ソフトウェア情報学部

Goutam Chakraborty

d次元の正規分布（2）

教師データ

	長さ	重さ	ラベル
$k=1$	46.1	180	さんま
:	:	:	:
$k=2$	31.6	241	にしん
:	:	:	:

例えばさんまのデータについて長さと重さの平均がそれぞれ
$\mu_1 = 50,\ \mu_2 = 200$ であったとする。
これをベクトルを用いて表すとこのようになる $\vec{\mu} = \begin{bmatrix} \mu_1 \\ \mu_2 \end{bmatrix}$

ここで魚 $\vec{x} = (52,\ 189)$ が与えられた $\vec{x} = \begin{bmatrix} 52 \\ 189 \end{bmatrix}$

これで $\vec{x} - \vec{\mu}$ を求める

$$\vec{x} - \vec{\mu} = \begin{bmatrix} 52 \\ 189 \end{bmatrix} - \begin{bmatrix} 50 \\ 200 \end{bmatrix} = \begin{bmatrix} 2 \\ -11 \end{bmatrix}$$

相関行列（6）

相関行列の表し方

$$\Sigma_2 = \begin{bmatrix} \sigma_1{}^2 & \sigma_{12} \\ \sigma_{21} & \sigma_2{}^2 \end{bmatrix}$$

相関行列の計算結果が

$$\Sigma = \begin{bmatrix} 3 & 1 \\ 1 & 5 \end{bmatrix}$$ であったとき

$$|\Sigma| = 3 \cdot 5 - 1 \cdot 1 = 14$$

$$\Sigma^{-1} = \frac{1}{14}\begin{bmatrix} 5 & -1 \\ -1 & 3 \end{bmatrix}$$

行列の復習

- 逆行列

$$\begin{bmatrix} 2 & 1 \\ 3 & 5 \end{bmatrix}^{-1} = \frac{1}{2 \cdot 5 - 1 \cdot 3}\begin{bmatrix} 5 & -1 \\ -3 & 2 \end{bmatrix} = \frac{1}{7}\begin{bmatrix} 5 & -1 \\ -3 & 2 \end{bmatrix}$$

- 検算

$$\begin{bmatrix} 2 & 1 \\ 3 & 5 \end{bmatrix} \times \frac{1}{7}\begin{bmatrix} 5 & -1 \\ -3 & 2 \end{bmatrix} = \frac{1}{7}\begin{bmatrix} 7 & 0 \\ 0 & 7 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

単位行列になればOK！

d次元の正規分布の確率密度関数

2次元の正規分布

$$f(\vec{x}) = \frac{1}{2\pi\sqrt{|\Sigma|}} \exp\left(-\frac{1}{2}(\vec{x}-\vec{\mu})^T \Sigma^{-1} (\vec{x}-\vec{\mu})\right)$$

$$f(x) = \frac{1}{2\pi\sqrt{14}} \exp\left(-\frac{1}{2}\underline{[2 \quad -11]\ \frac{1}{14}\begin{bmatrix} 5 & -1 \\ -1 & 3 \end{bmatrix}}\begin{bmatrix} 2 \\ -11 \end{bmatrix}\right)$$

$$\frac{1}{14}[21 \quad -35]\begin{bmatrix} 2 \\ -11 \end{bmatrix} = \frac{1}{14}(42+385) = 30.5$$

$$\therefore\ f(x) = \frac{1}{2\pi\sqrt{14}}\exp\left(-\frac{30.5}{2}\right) \fallingdotseq 1.0 \times 10^{-8}$$

2次元の正規分布の確率を求める簡単な方法

確率を求めるには、

$$\int_{51.5}^{52.5}\int_{188}^{190} f(l, w)\, dl\, dw$$

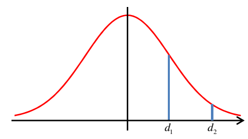
むずかしい・・・

Numerical Analysis（数値計算）ではできる。

マハラノビス距離の求め方-
データの平均の位置から該当するデータまでのマハラノビス距離を求める事で確率が分かる

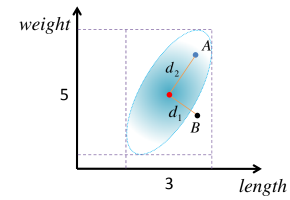
- Mahalanobis Distance（マハラノビス距離）
  - $d_m = \sqrt{(\vec{x}-\vec{\mu})^T \Sigma^{-1} (\vec{x}-\vec{\mu})}$
- Euclidean Distance（ユークリッド距離）
  - $d_e = \sqrt{\sum (x_i^2 - x_i^1)^2}$ （いわゆる普通の距離）
- Manhattan Distance（マンハッタン距離）
  - $d_M = |y_2 - y_1| + |x_2 - x_1|$

マハラノビス距離



$d_1 < d_2$
$P(d_1) > P(d_2)$

相関行列は $\begin{bmatrix} 3 & 1 \\ 1 & 5 \end{bmatrix}$



サンプルAとBの位置は図で示されている。
$d_1 < d_2$ だけど、サンプル Aはサンプル Bより青い分布のクラスである確率が高い

データの平均値からサンプルまでのMahalanobis距離を求める事でクラスの確率が求められる。マハラノビス距離が近い程確率が高い。

マハラノビス距離

$$P(C_1 \mid t) = P(t \mid C_1)\, \boxed{P(C_1)}$$
$$P(C_2 \mid t) = P(t \mid C_2)\, \boxed{P(C_2)}$$

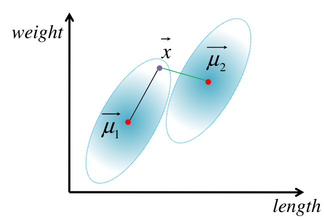
$n(C_1) = n(C_2)$
のときはいらない。

$C_1 \rightarrow \Sigma_1, \mu_1$
$\quad\ \ \|\quad \nparallel$
$C_2 \rightarrow \Sigma_2, \mu_2$ のとき、

$\dfrac{1}{2\pi\sqrt{|\Sigma|}}$ は同じ値になる

Sample $\vec{x}$ の場合、

$d_{e_1} > d_{e_2}$ だけど、
$d_{m_1} < d_{m_2}$ となる。



Sample $\vec{x}$ は
Class1のものである。

マハラノビス距離とクラス決定

$$f(\vec{x}) = \frac{1}{\underbrace{2\pi\sqrt{|\Sigma|}}_{\text{一定}}} \exp\left(-\frac{1}{2}\left(\vec{x}-\vec{\mu}\right)^T \Sigma^{-1} \left(\vec{x}-\vec{\mu}\right)\right)$$

$\rightarrow e^{-\frac{1}{2}{d_m}^2}$

d_m の値が低いと、$e^{-\frac{1}{2}{d_m}^2}$ が高くなり、確率も高くなる。

課題

C_1	C_1とC_2の相関行列は同じ $\Sigma = \begin{bmatrix} 1.1 & 0.3 \\ 0.3 & 1.9 \end{bmatrix}$	$\vec{\mu}_1 = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$
C_2		$\vec{\mu}_2 = \begin{bmatrix} 3 \\ 3 \end{bmatrix}$

$\vec{x} = \begin{bmatrix} 1.0 \\ 2.2 \end{bmatrix}$

このデータは C_1, C_2 のどっちのクラスでしょうか？

$\vec{x}$から $\vec{\mu}_1$ の ${d_m}^2 = ?$
$\vec{x}$から $\vec{\mu}_2$ の ${d_m}^2 = ?$

距離が近いほうが $\vec{x}$ のClassである