情報認識
「ガウス最尤推定による
パターン識別器の構成」

- **講師：　杉山　将（計算工学専攻）**
- **居室：**　W8E-505
- **電子メール**:　sugi@cs.titech.ac.jp

最大事後確率則

- 入力パターンが与えられてそれがどのカテゴリに属するかを決めるとき，それが属する可能性が最も高いカテゴリを選ぶ．
- これは，事後確率が最大のカテゴリに分類することに対応．

$$x \in \omega_k : k = \arg\max_i p(\omega_i \mid x)$$

カテゴリの対数事後確率

- 対数をとっても大小関係は変わらないため，事後確率の対数をとったものを用いる．
- ベイズの定理より

$$p(\omega_i \mid x) = \frac{p(x \mid \omega_i)\, p(\omega_i)}{p(x)}$$

- 従って，

$$\log p(\omega_i \mid x) = \log p(x \mid \omega_i) + \log p(\omega_i) + C$$

$C = -\log p(x)$ ：定数

推定方法

- $\{x_i\}_{i=1}^{n}$ ：訓練標本
- n_i ：カテゴリ ω_i に属する訓練標本数
- 事前確率 $p(\omega_i)$ ：カテゴリ ω_i に含まれる標本の割合で推定

$$\hat{p}(\omega_i) = \frac{n_i}{n}$$

- 条件付き確率 $p(x \mid \omega_i)$ ：ガウスモデルに対する最尤推定

$$\hat{p}(x \mid \omega_i) = q(x; \hat{\mu}_i, \hat{\Sigma}_i)$$

ガウスモデル (正規分布)

- d 次元の確率ベクトル: $x = (x^{(1)}, x^{(2)}, \ldots, x^{(d)})^T$
- 2つのパラメータ:
  - カテゴリ ω_i の平均ベクトル　μ_i
  - カテゴリ ω_i の分散共分散行列　Σ_i

$$q(x; \mu_i, \Sigma_i) = \frac{1}{(2\pi)^{d/2} |\Sigma_i|^{1/2}} \exp\left(-\frac{1}{2} (x - \mu_i)^T \Sigma_i^{-1} (x - \mu_i) \right)$$

ガウスモデルにおける最尤推定

■ ガウスモデルの最尤推定量

$$\hat{\mu}_i = \frac{1}{n_i} \sum_{x \in \omega_i} x$$

$$\hat{\Sigma}_i = \frac{1}{n_i} \sum_{x \in \omega_i} (x - \hat{\mu}_i)(x - \hat{\mu}_i)^T$$

カテゴリの対数事後確率の計算

■ 分類したい入力パターンを x とすれば,

$$\log \hat{p}(\omega_i \mid x)$$

$$= -\frac{d}{2}\log 2\pi - \frac{1}{2}\log|\hat{\Sigma}_i| - \frac{1}{2}(x-\hat{\mu}_i)^T \hat{\Sigma}_i^{-1}(x-\hat{\mu}_i) + \log\frac{n_i}{n} + C$$

$$= -\frac{1}{2}\underbrace{(x-\hat{\mu}_i)^T \hat{\Sigma}_i^{-1}(x-\hat{\mu}_i)}_{\text{マハラノビス距離}} - \frac{1}{2}\log|\hat{\Sigma}_i| + \log n_i + C'$$

(Mahalanobis distance)

$$C' = -\frac{d}{2}\log 2\pi - \log n + C$$

■ カテゴリの対数事後確率は x の**二次形式**

共分散行列が共通のとき

- 各カテゴリの分散共分散行列が等しいという前提知識があるとき，

$$\Sigma_1 = \Sigma_2 = \cdots = \Sigma_m = \Sigma$$

共通の分散共分散行列 Σ の最尤推定量は

$$\hat{\Sigma} = \frac{1}{n}\sum_{i=1}^{m}\sum_{j:x_j \in \omega_i}(x_j - \hat{\mu}_i)(x_j - \hat{\mu}_i)^T$$

$$= \sum_{i=1}^{m}\frac{n_i}{n}\hat{\Sigma}_i$$

各カテゴリの分散共分散行列の重み付き平均

共分散行列が共通のとき (続き)

- 各カテゴリの分散共分散行列が等しい時,

$$\log \hat{p}(\omega_i \mid x)$$

$$= -\frac{1}{2} x^T \hat{\Sigma}^{-1} x + \hat{\mu}_i^T \hat{\Sigma}^{-1} x - \frac{1}{2} \hat{\mu}_i^T \hat{\Sigma}^{-1} \hat{\mu}_i - \frac{1}{2} \log |\hat{\Sigma}| + \log \frac{n_i}{n} + C$$

$$= \hat{\mu}_i^T \hat{\Sigma}^{-1} x - \frac{1}{2} \hat{\mu}_i^T \hat{\Sigma}^{-1} \hat{\mu}_i + \log n_i + C'$$

$$C' = -\frac{1}{2} x^T \hat{\Sigma}^{-1} x - \frac{1}{2} \log |\hat{\Sigma}| - \log n + C$$

- カテゴリの対数事後確率は x の**一次形式**

決定境界の計算

- カテゴリ数 $m = 2$ の場合，決定境界は次の方程式で与えられる．

$$\hat{p}(\omega_1 \mid x) = \hat{p}(\omega_2 \mid x)$$

- ガウスモデルと最尤推定を用いたとき，決定境界は二次形式．

- 更に分散共分散行列が共通のとき，決定境界は一次形式，即ち，**超平面**(hyper-plane)．

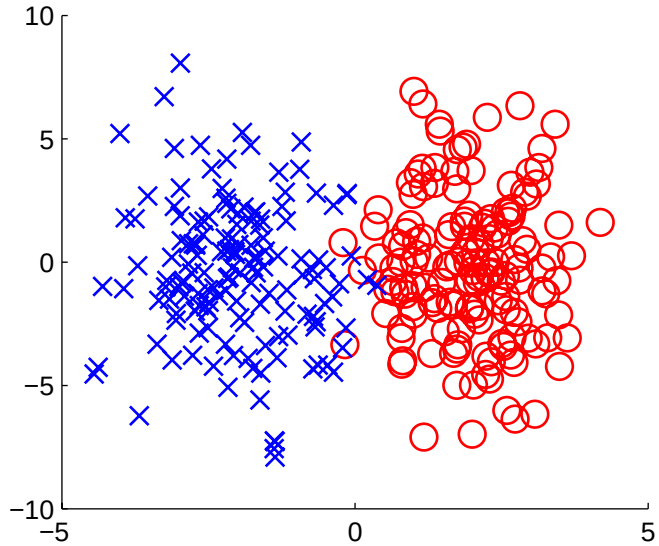
$$(\hat{\mu}_1 - \hat{\mu}_2)^T \hat{\Sigma}^{-1} x - \frac{1}{2}(\hat{\mu}_1{}^T \hat{\Sigma}^{-1} \hat{\mu}_1 - \hat{\mu}_2{}^T \hat{\Sigma}^{-1} \hat{\mu}_2) + \log \frac{n_1}{n_2} = 0$$

演習

- 入力次元 $d = 2$ ，カテゴリ数 $m = 2$
- 各カテゴリの事前確率： $p(\omega_1) = p(\omega_2) = 1/2$
- 各カテゴリの条件付き確率 $p(x \mid \omega_i)$ は正規分布．平均と分散共分散行列はそれぞれ

$$\mu_1 = \begin{pmatrix} 2 \\ 0 \end{pmatrix} \quad \mu_2 = \begin{pmatrix} -2 \\ 0 \end{pmatrix} \quad \Sigma_1 = \Sigma_2 = \Sigma = \begin{pmatrix} 1 & 0 \\ 0 & 9 \end{pmatrix}$$

1. 上記の設定のもと，最大事後確率則に基づいた決定境界を求めよ



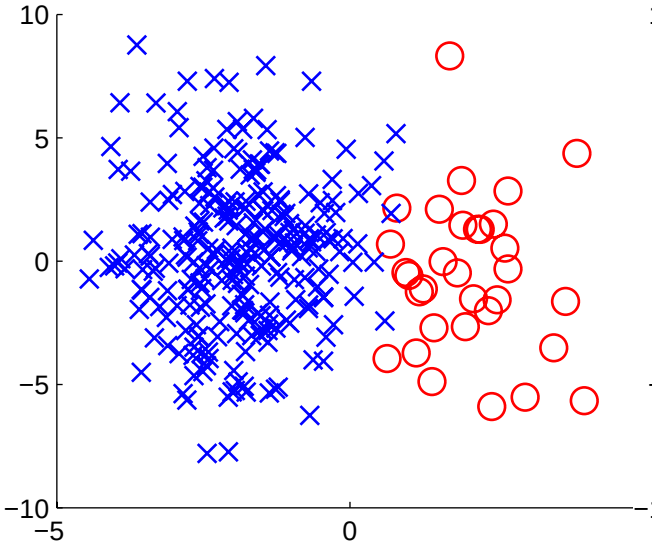
演習(続き)

2. 演習1の問題で事前確率を次のように設定する.

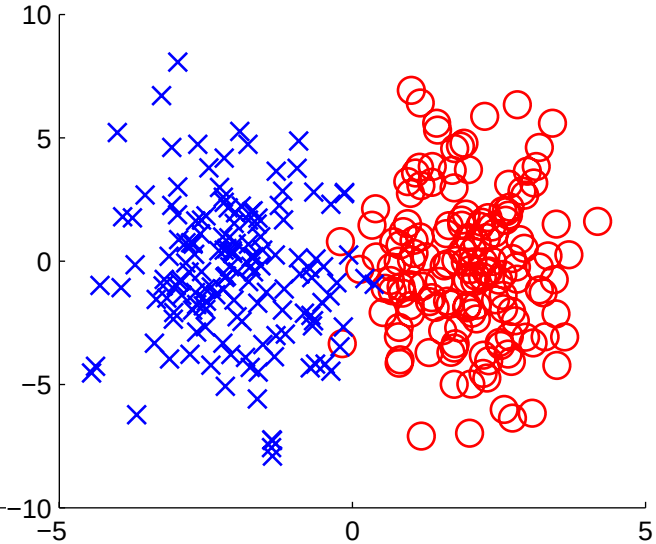
$$p(\omega_1) = \alpha \qquad\qquad p(\omega_2) = 1 - \alpha$$

- $\alpha = \frac{1}{10}, \frac{1}{3}, \frac{1}{2}, \frac{2}{3}, \frac{9}{10}$ に対する決定境界を求めよ
- α を $0 < \alpha < 1$ の間で変化させれば決定境界はどのように変化するか？

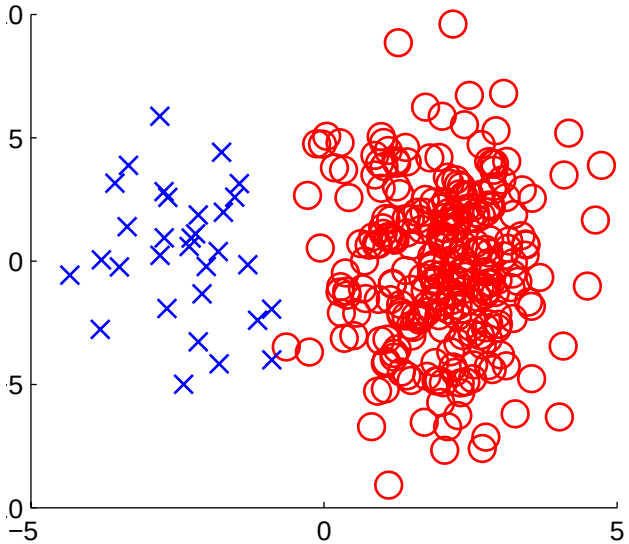
$\alpha = 1/10$



$\alpha = 1/2$



$\alpha = 9/10$



まとめ

- 各カテゴリの標本が正規分布に従うとき，
  - 対数事後確率は二次形式
  - 2カテゴリのとき，決定境界は二次形式
- 更に各カテゴリの共分散行列が共通のとき
  - 対数事後確率の主要項は一次形式
  - 2カテゴリのとき，決定境界は一次形式（超平面）

小レポート (第6回)

- 演習1の問題で共分散行列を次のように設定する

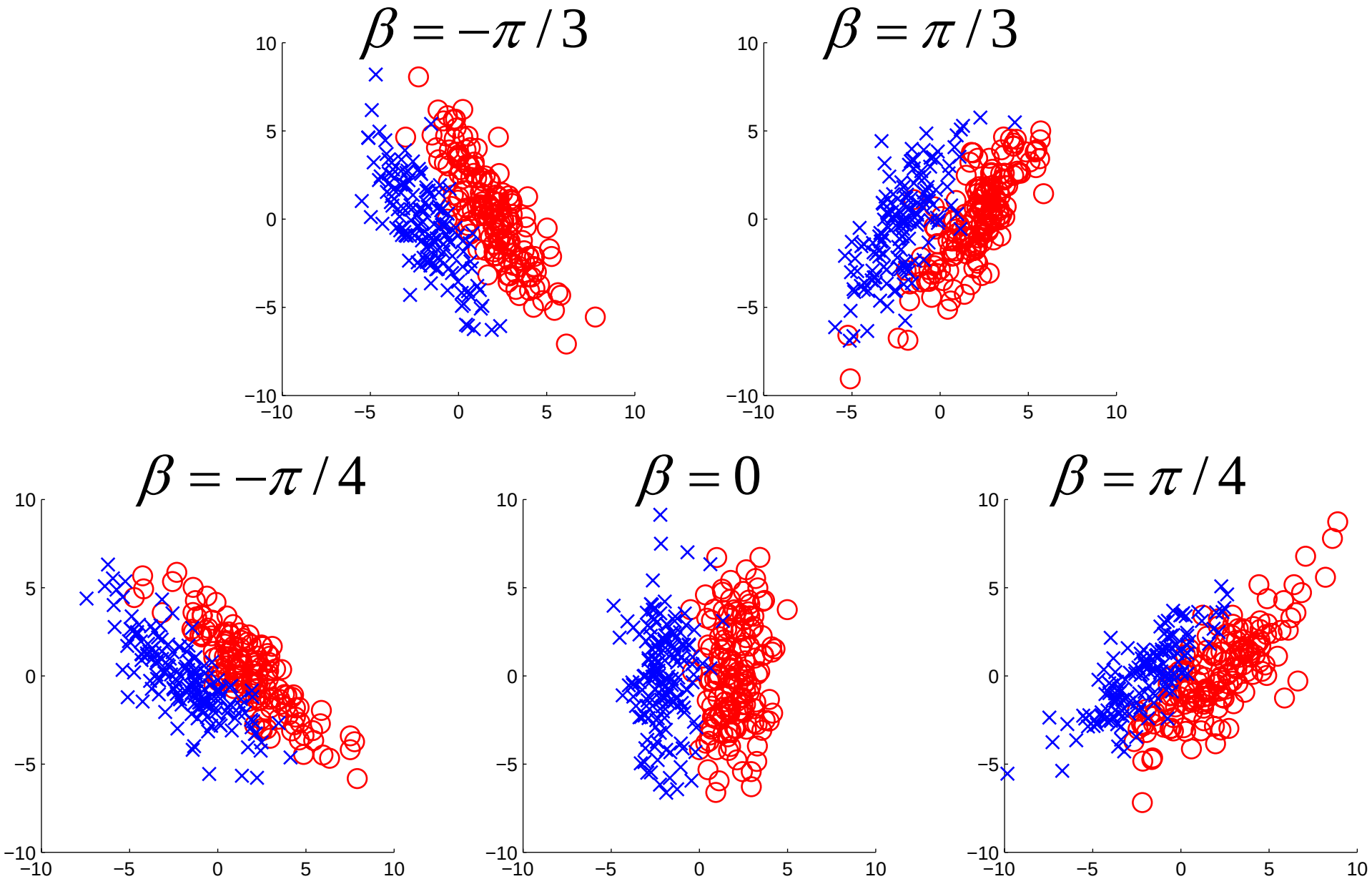
$$\Sigma = \begin{pmatrix} 9-8\cos^2\beta & 8\sin\beta\cos\beta \\ 8\sin\beta\cos\beta & 9-8\sin^2\beta \end{pmatrix}$$

これは，各カテゴリのデータを β だけ回転させたものに対応している (次ページの図参照)

- $\beta = -\frac{1}{4}\pi, -\frac{1}{6}\pi, 0, \frac{1}{6}\pi, \frac{1}{4}\pi$ に対する決定境界を求めよ．
- β を $-\pi/4 \le \beta \le \pi/4$ の間で変化させれば，決定境界はどのように変化するか？

- ヒント：

$$\Sigma^{-1} = \begin{pmatrix} 1-\frac{8}{9}\sin^2\beta & -\frac{8}{9}\sin\beta\cos\beta \\ -\frac{8}{9}\sin\beta\cos\beta & 1-\frac{8}{9}\cos^2\beta \end{pmatrix}$$

$$\sin\tfrac{1}{4}\pi = 1/\sqrt{2}$$
$$\sin\tfrac{1}{6}\pi = 1/2$$

小レポート(続き)