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.

$$\arg\max_{y} p(y \mid x)$$

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

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

- 従って，

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

$C = -\log p(x)$ ：定数

推定方法

- $\{(x_i, y_i)\}_{i=1}^{n}$: 訓練標本

$$x_i \in D\left(\subset \Re^d\right) \qquad y_i \in \{1,2,\ldots,m\}$$

- n_y : カテゴリ y に属する訓練標本数
- 事前確率 $p(y)$: カテゴリ y に含まれる標本の割合で推定

$$\hat{p}(y) = \frac{n_y}{n}$$

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

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

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 y の平均ベクトル μ_y
  - カテゴリ y の分散共分散行列 Σ_y

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

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

$$\hat{\mu}_y = \frac{1}{n_y} \sum_{i:y_i=y} x_i$$

$$\hat{\Sigma}_y = \frac{1}{n_y} \sum_{i:y_i=y} (x_i - \hat{\mu}_y)(x_i - \hat{\mu}_y)^T$$

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,

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

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

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

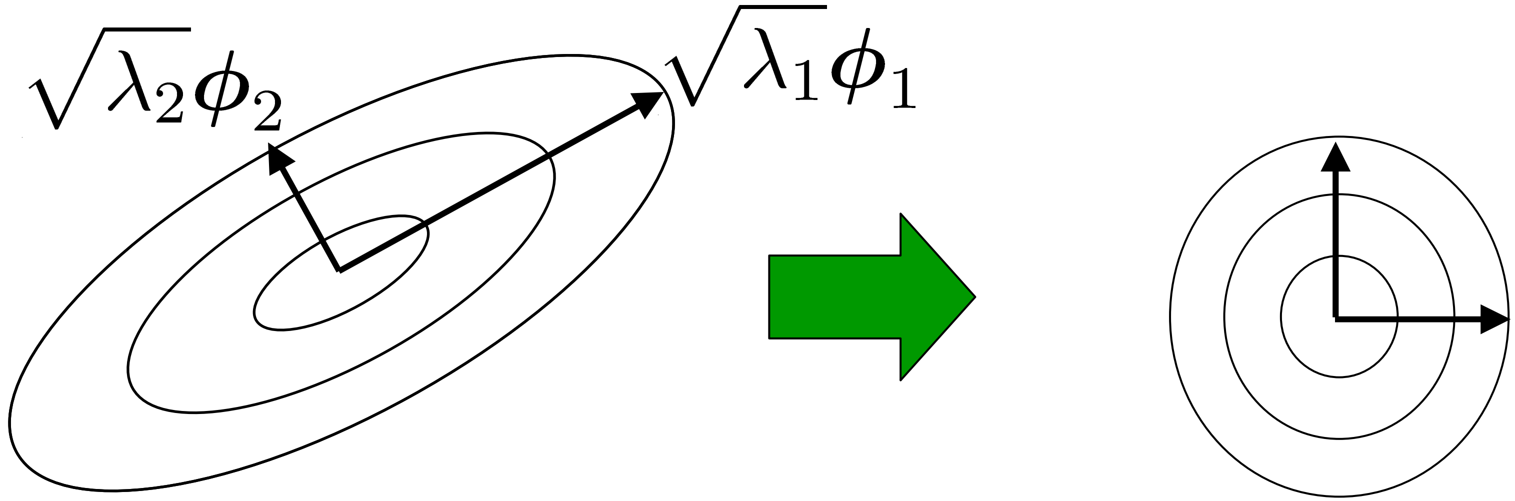
$$C' = -\frac{d}{2}\log 2\pi - \log n + C$$

- カテゴリの対数事後確率は x の二次形式

マハラノビス距離

- "楕円を正円に変換した距離"

$$(x-\mu)^T \Sigma^{-1}(x-\mu) = \left\|\Sigma^{-1/2}(x-\mu)\right\|^2$$



$$\mathbf{\Sigma} = \lambda_1 \phi_1 \phi_1^\top + \lambda_2 \phi_2 \phi_2^\top$$

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

$$\hat{\Sigma} = \frac{1}{n}\sum_{y=1}^{m}\sum_{i:y_i=y}(x_i - \hat{\mu}_y)^T(x_i - \hat{\mu}_y)$$

$$= \sum_{y=1}^{m}\frac{n_y}{n}\hat{\Sigma}_y$$

⇒ 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,

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

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

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

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

- 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

- カテゴリ数が2のとき，決定境界は

$$\hat{p}(y=1\,|\,x)=\hat{p}(y=2\,|\,x)$$

- 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)．

$$a^T x + b = 0$$

$$a = \hat{\Sigma}^{-1}(\hat{\mu}_1 - \hat{\mu}_2)$$

$$b = -\tfrac{1}{2}(\hat{\mu}_1{}^T \hat{\Sigma}^{-1}\hat{\mu}_1 - \hat{\mu}_2{}^T \hat{\Sigma}^{-1}\hat{\mu}_2) + \log n_1/n_2$$

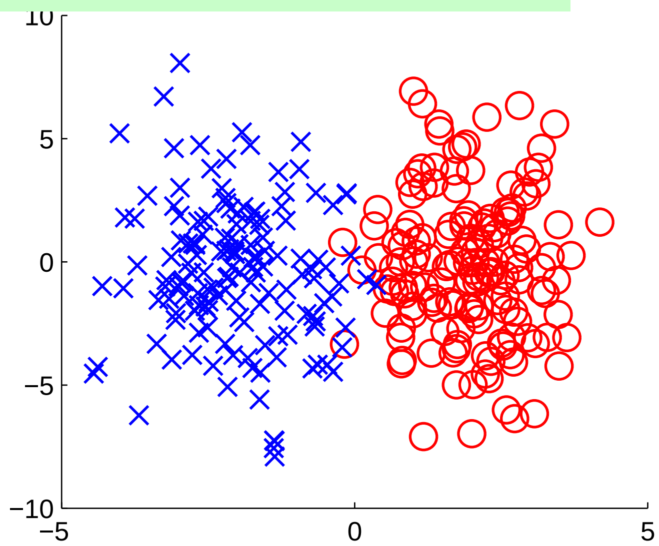
- この場合を特に，フィッシャーの線形判別分析(Fisher’s linear discriminant analysis)という.

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：$p(y=1)=p(y=2)=1/2$
- 各カテゴリの条件付き確率 $p(x\,|\,y)$ は正規分布．平均と分散共分散行列はそれぞれ

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

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



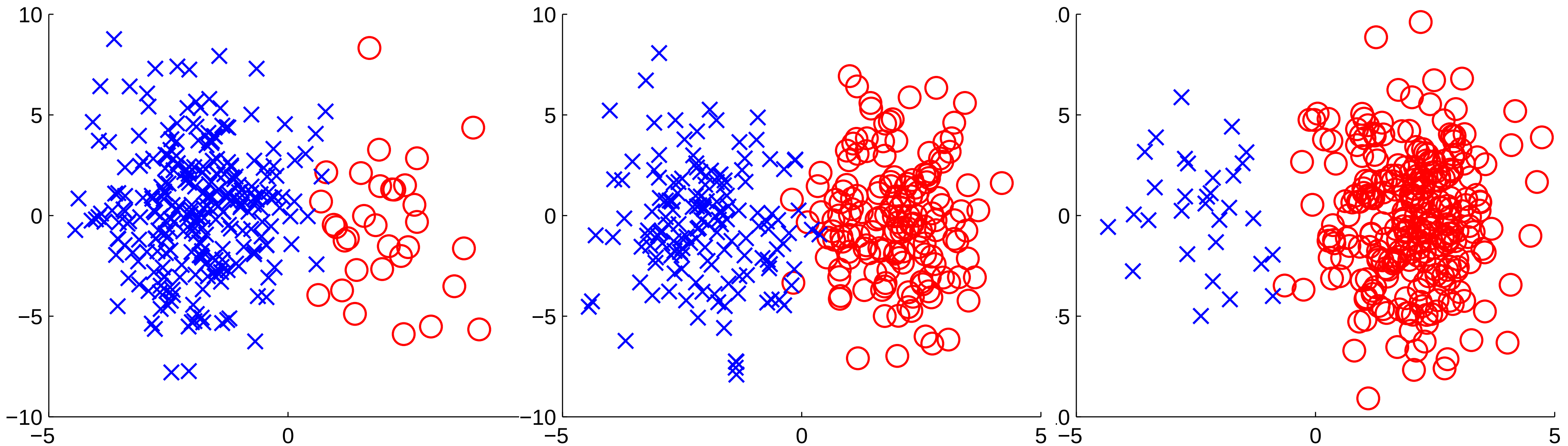
演習(続き)

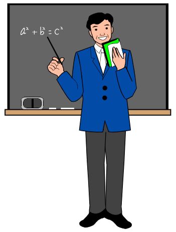
2. 演習1の問題で事前確率を次のように設定する.

$$p(y=1)=\alpha \qquad p(y=2)=1-\alpha$$

- $\alpha=\frac{1}{10},\frac{1}{3},\frac{1}{2},\frac{2}{3},\frac{9}{10}$ 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？

$\alpha=1/10$ $\qquad$ $\alpha=1/2$ $\qquad$ $\alpha=9/10$





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4回)

1. 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？

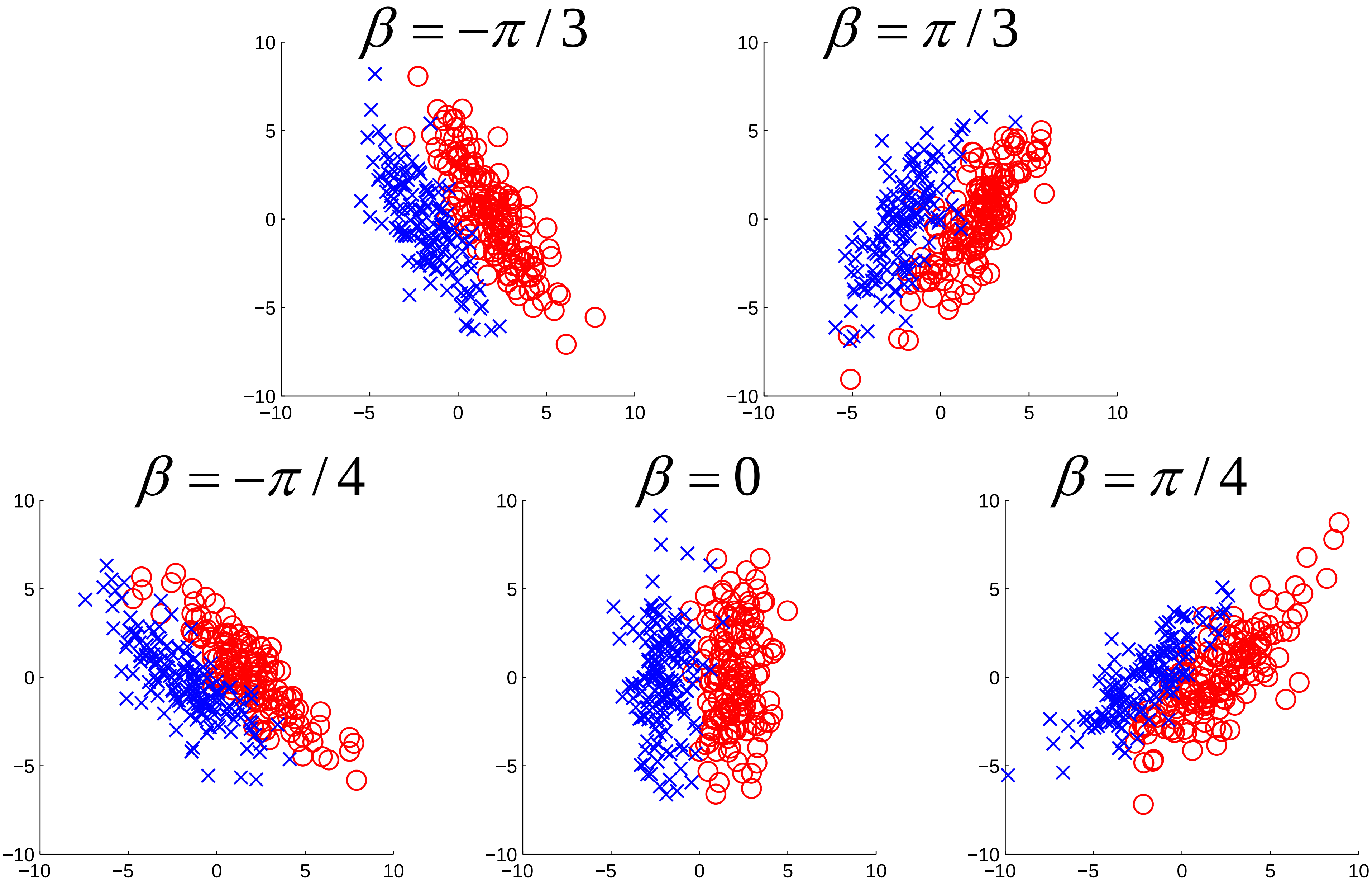
■ヒント：

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

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

$$\sin\frac{1}{6}\pi = 1/2$$

小レポート(続き)



小レポート(続き)

2. ガウスモデルと最尤推定法を用いて,
   手書き文字認識装置を作成せよ.

- 詳細は添付の資料を参照せよ.