今日の話：回帰と判別

[午前]

- 回帰分析の多変量版(重回帰)
- 決定係数なども含めて具体的に計算をフォローしてみる
- 多変量版のまとめ

[午後]

- 判別と回帰
- 判別分析
- マハラノビス距離

回帰と判別

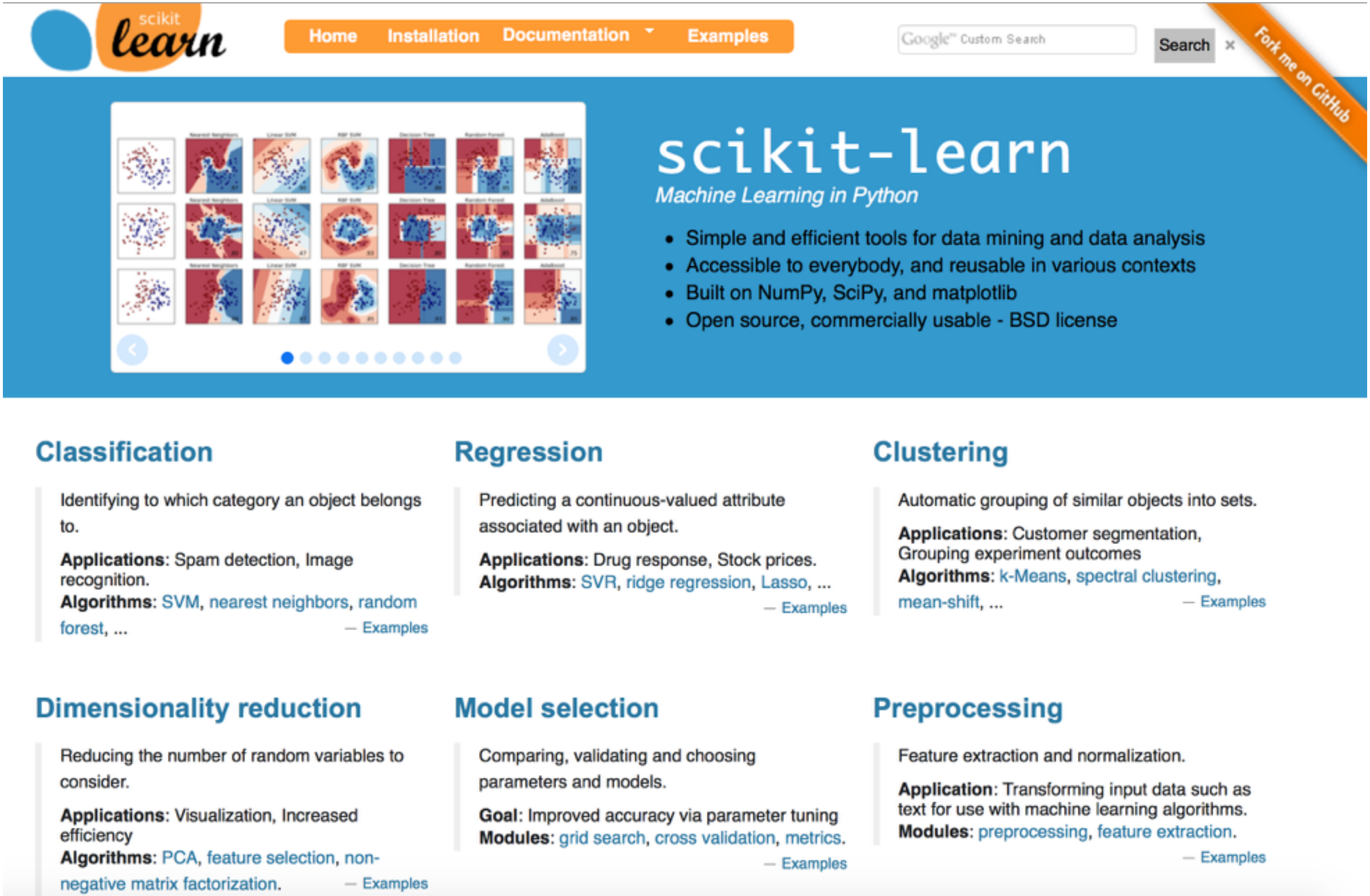
- 計算的な側面だけみると
  - 回帰 $(x_1, x_2, \ldots, x_p)' \mapsto y \in \mathbb{R}$
  - 判別/分類 $(x_1, x_2, \ldots, x_p)' \mapsto y \in \{1, 2, \ldots, C\}$

- 機械学習・パターン認識の用語でいうと
  - 教師付き学習の問題

すなわちこの問題を解く手法はたくさんある!

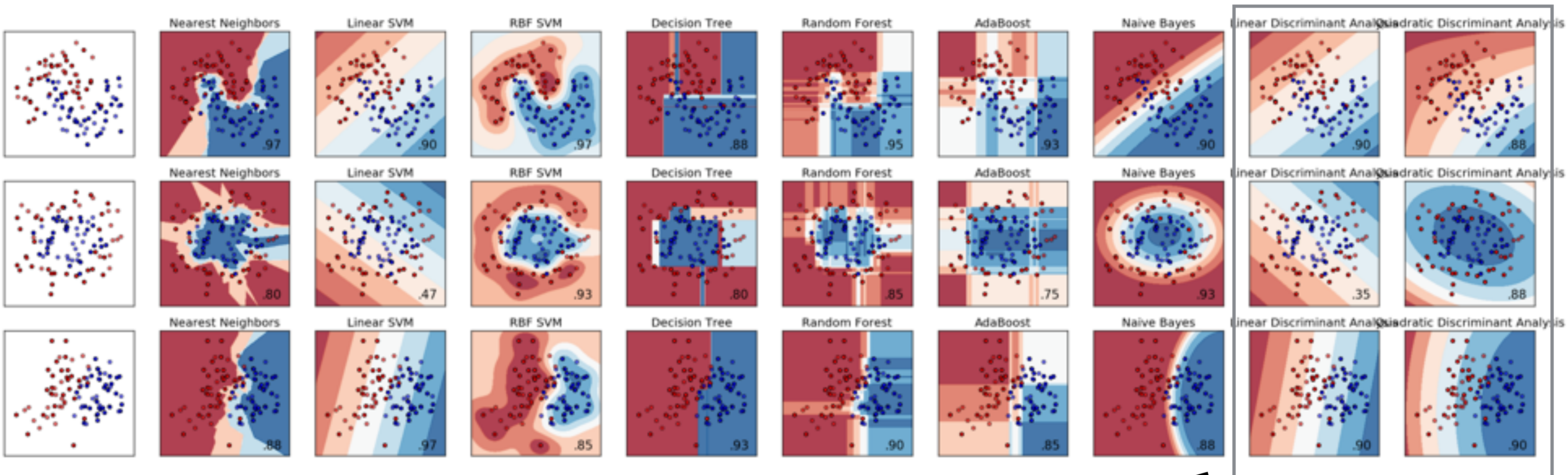
(DL/NN,SVM, SVR, RF, ExtraTrees, xgboost, FM, $\cdots$)

http://scikit-learn.org/stable/



すなわちこの問題を解く手法はたくさんある!

(DL/NN,SVM, SVR, RF, ExtraTrees, xgboost, FM, …)



このあたりは統計学じゃなく
機械学習の本で

- 教科書の判別分析はいわゆる「Fisherの判別分析」
- 双璧は回帰っぽくやる「ロジスティック回帰/Softmax回帰」

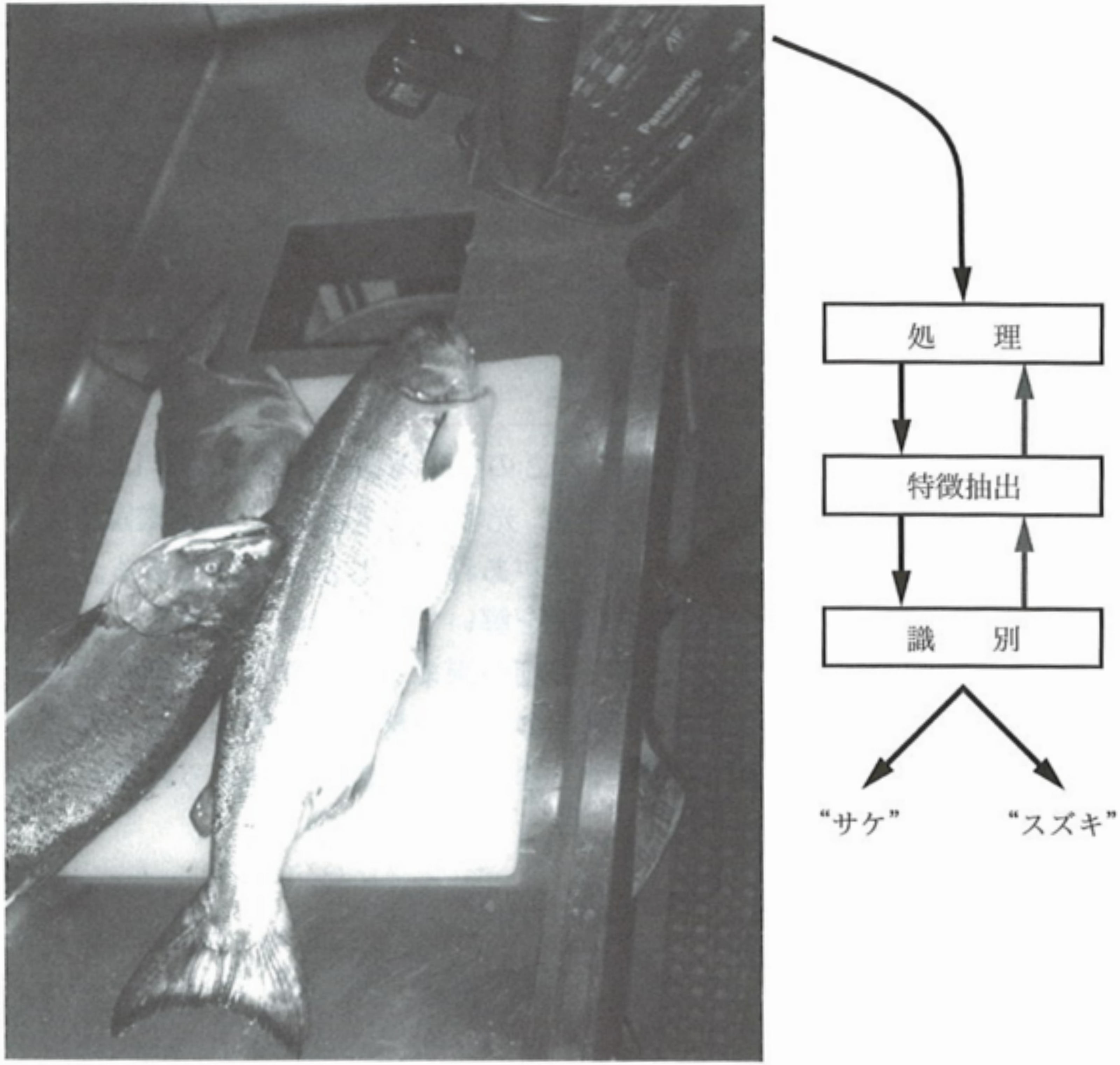
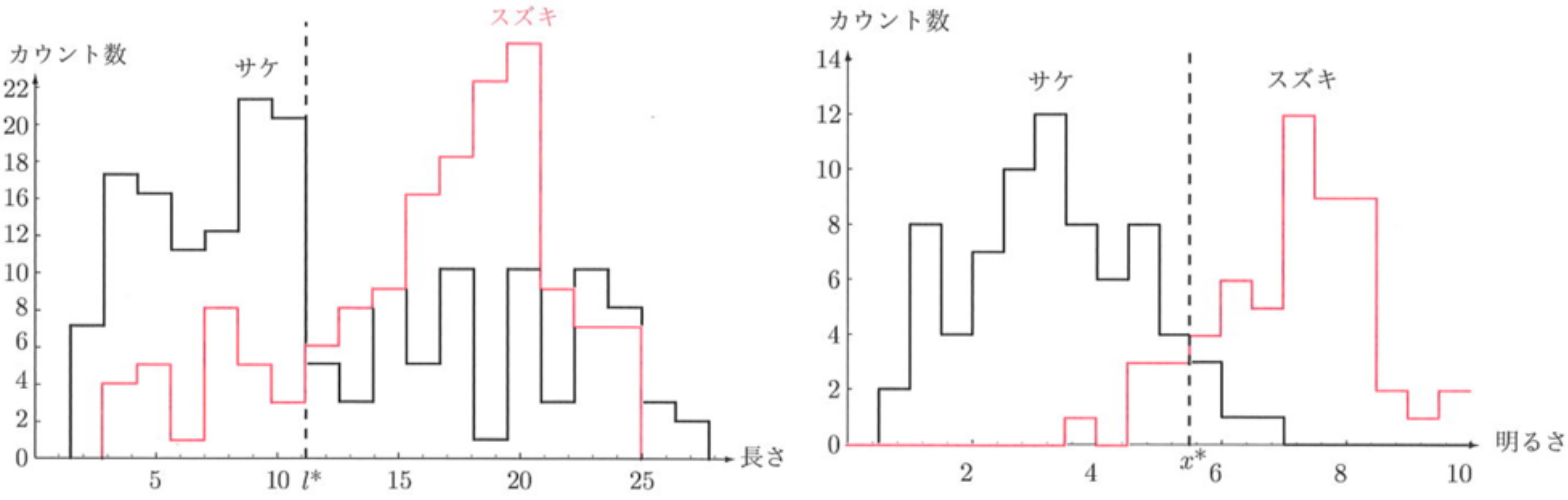
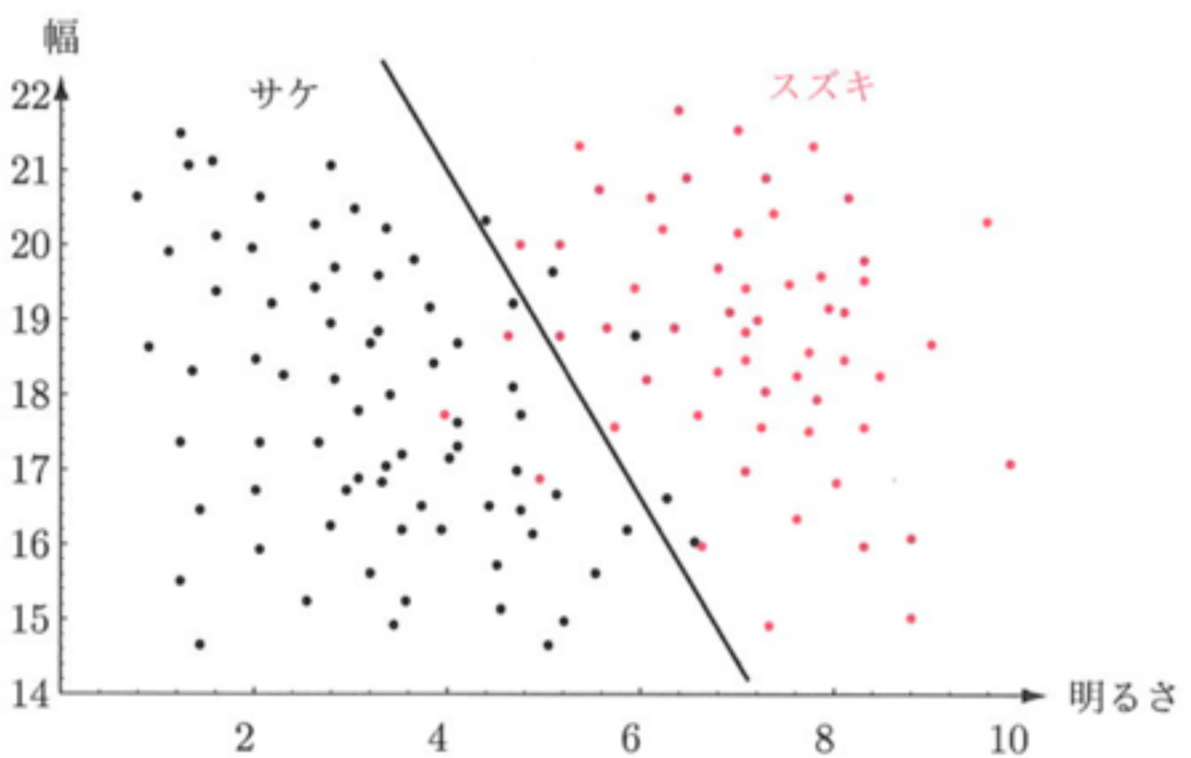


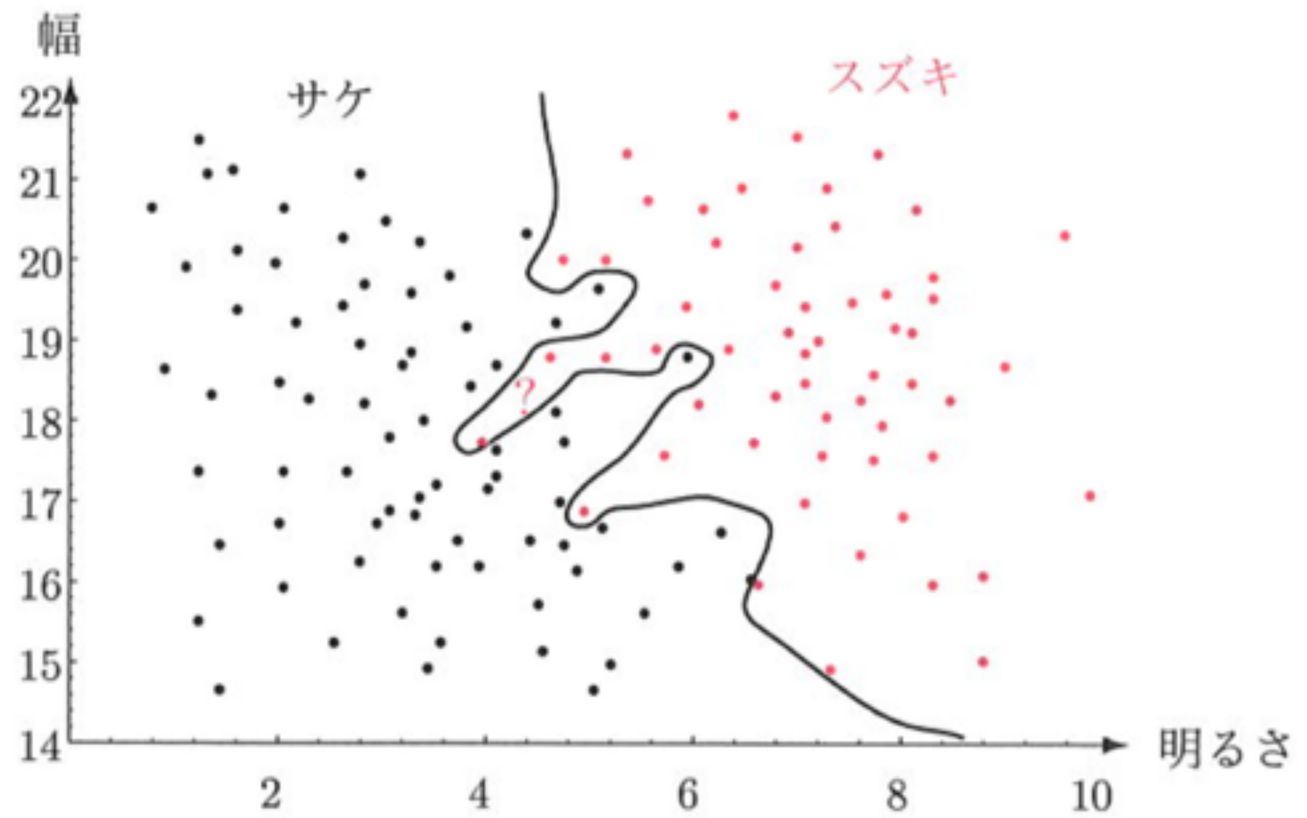
図 1.1　分類すべき対象物は，まず変換器（カメラ）で撮影され，その映像が前処理される。次にいくつかの特徴が抽出され，最終的に“サケ”か“スズキ”かの分類が出力される。この情報の流れは多くの場合，起点から識別器へと向かっているが，細い矢印で示したように後の段階での試験的または予備的応答にもとづいて，初期の段階の処理を変えるシステムもある。また他のシステムでは，たとえば分割処理と特徴抽出を同時に行うなど，複数の段階を1つの段階にまとめることがある。



単一の計測量では２つの種類をきっぱり分けることができない

→ 多変量化



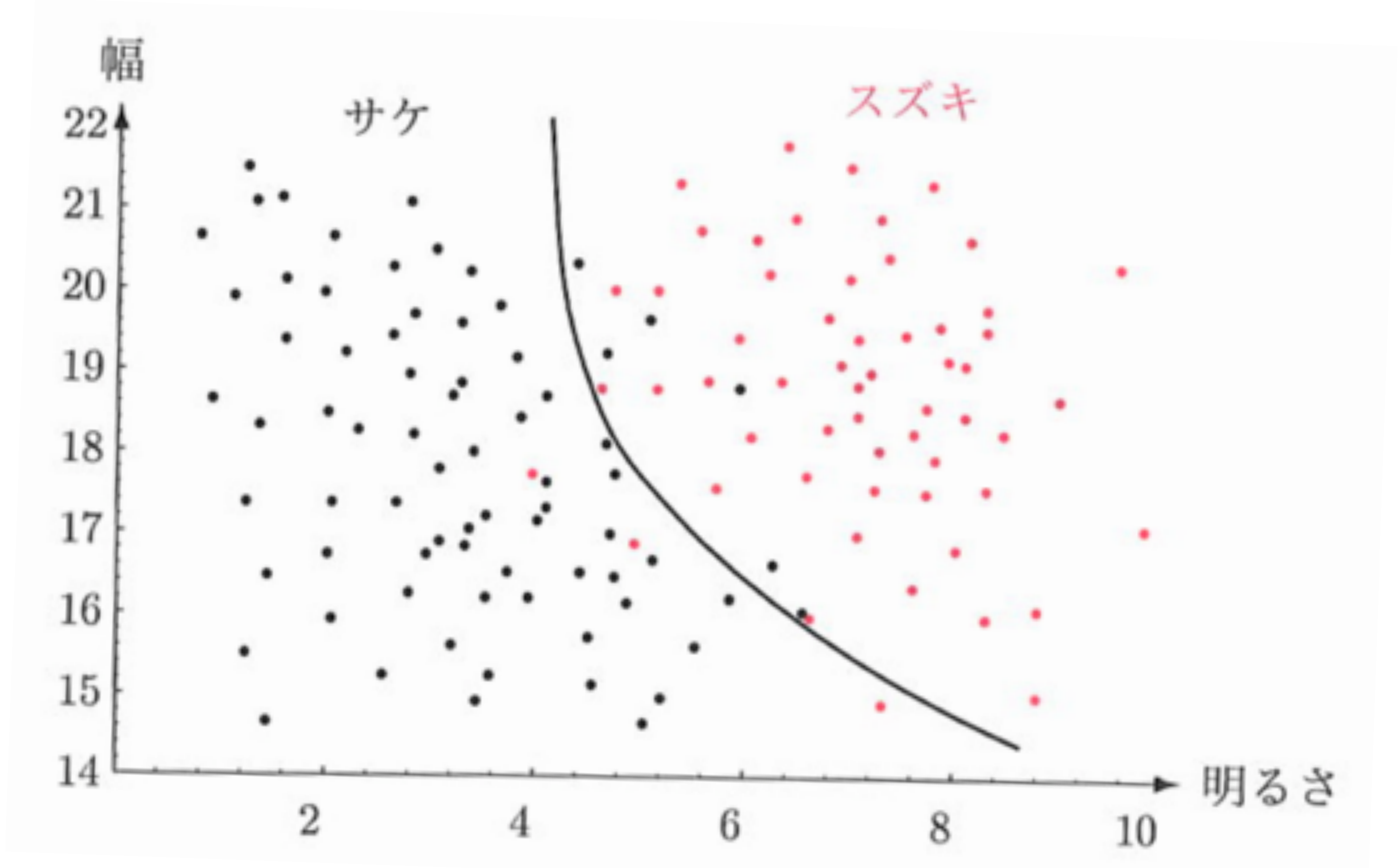


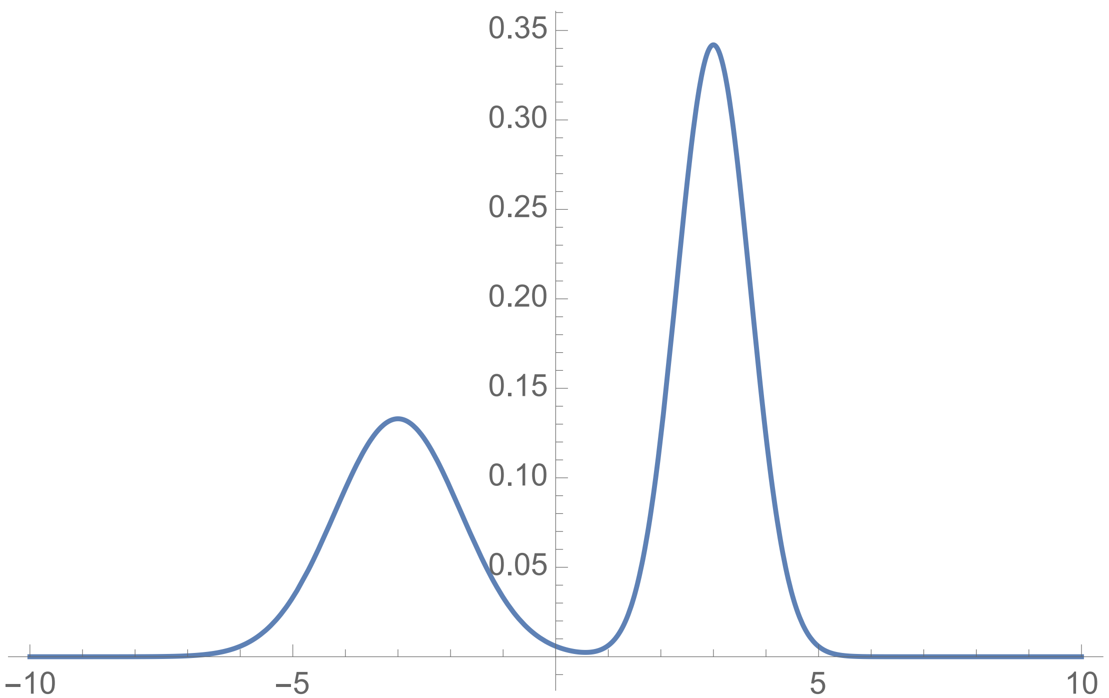
判別は幾何的には

境界線作成問題

or

領土塗り分け問題





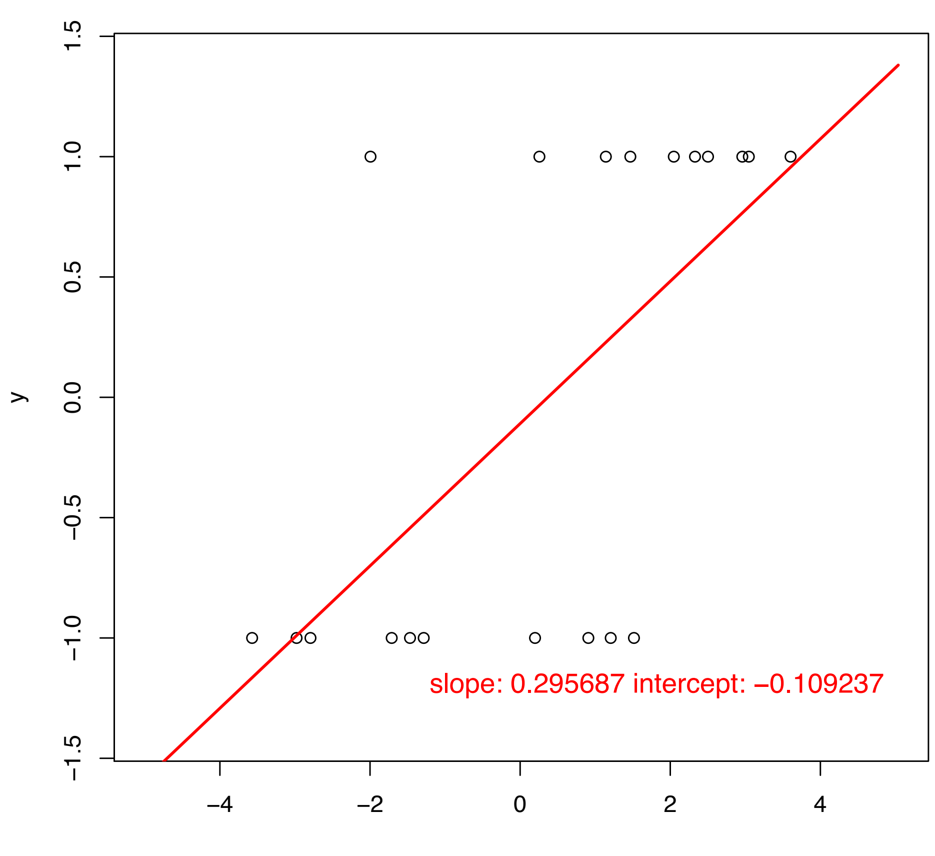
y=1 (xがサケのとき)
y=-1 (xがスズキのとき)

と置いて (x,y) の回帰

として解くとまずいの？

↓

OKなときもあるが、変になり得る(左例)のでよろしくはない



判別の場合、+1/-1の値自体を正確に当てる必要がなく、回帰より簡単な問題

正規線形モデル

$$\boldsymbol{y} = \boldsymbol{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon}$$

$\boldsymbol{X}$ ：計画行列

$$\boldsymbol{\varepsilon} \sim N(\boldsymbol{0}, \sigma^2 \boldsymbol{I})$$

$$\begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix} = \begin{bmatrix} 1 & x_{11} & \cdots & x_{1p} \\ 1 & x_{21} & \cdots & x_{2p} \\ \vdots & & & \vdots \\ 1 & x_{p1} & \cdots & x_{pp} \end{bmatrix} \begin{bmatrix} \beta_0 \\ \beta_1 \\ \vdots \\ \beta_p \end{bmatrix} + \begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \vdots \\ \varepsilon_n \end{bmatrix}$$

$$\varepsilon_1, \varepsilon_2, \ldots, \varepsilon_n \sim N(0, \sigma^2)$$

$$\Leftrightarrow \begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \vdots \\ \varepsilon_n \end{bmatrix} \sim N\left(\begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix}, \begin{bmatrix} \sigma^2 & 0 & \cdots & 0 \\ 0 & \sigma^2 & \cdots & 0 \\ \vdots & & & \\ 0 & 0 & \cdots & \sigma^2 \end{bmatrix} \right)$$

これはX自体には確率変動をなんら仮定していない！

(Xとyの関係が線形で正規乱数分ゆらぐと仮定)

正規線形モデルと回帰

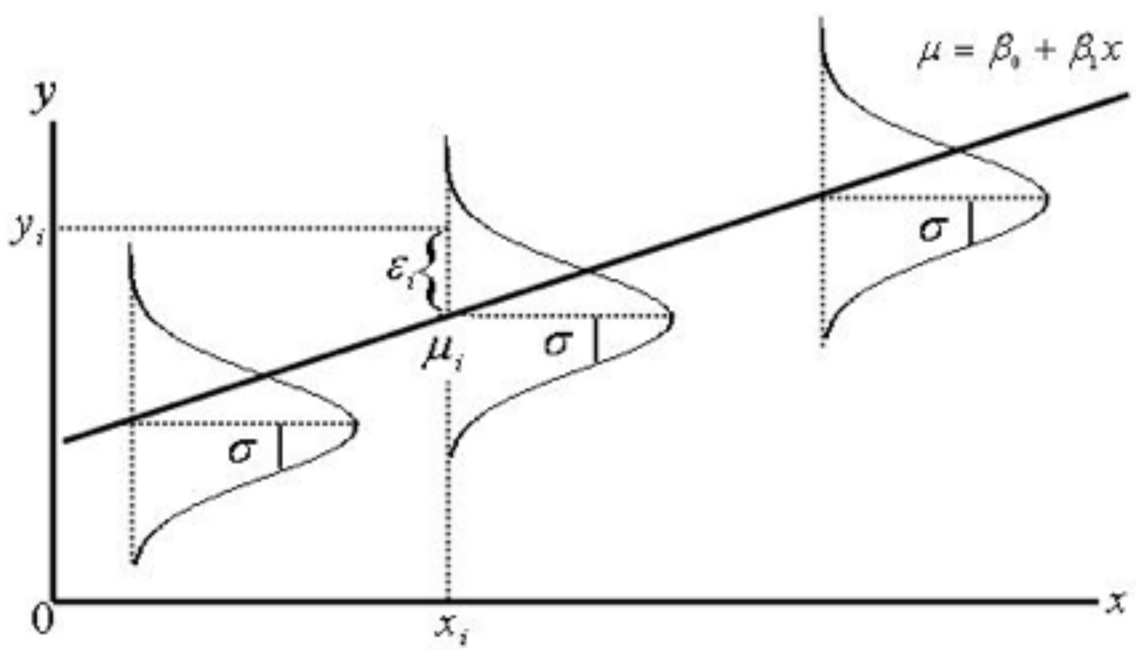
Given: $x_1, x_2, \ldots, x_n$

$$Y_i = a\,x_i + b + Z_i$$

正規乱数 $Z_i \overset{\text{iid}}{\sim} N(0, \sigma^2)$

Observe: $Y_1, Y_2, \cdots, Y_n$

Q: 観測された $x_1, x_2, \ldots, x_n$ および $Y_1, Y_2, \cdots, Y_n$ から、その背後にある未知パラメタ a, b, σ^2 をどのくらいの精度で当てられる？？



未知なる真の直線y=ax+bがあり
あるxでの観測値yは真の値まわりでN(0,σ²)に従ってずれて得られたと考える。分布はxに依らず常に同じもの。

1.2 重回帰分析とは

表 1.3 は東京のある駅の徒歩圏内の中古マンションに関するデータである.

表 1.3　中古マンションのデータ

サンプル No.	広さ x_1 (m^2)	築年数 x_2 (年数)	価格 y (千万円)
1	51	16	3.0
2	38	4	3.2
3	57	16	3.3
4	51	11	3.9
5	53	4	4.4
6	77	22	4.5
7	63	5	4.5
8	69	5	5.4
9	72	2	5.4
10	73	1	6.0

このデータに基づいて知りたいことは次の通りである.

(1) 価格は広さと築年数とによって予測できるだろうか.

(2) 予測できるとすればその精度はどのくらいか.

(3) 同じ地区で $x_1 = 70$, $x_2 = 10$, $y = 5.8$ を提示された. 価格は妥当か.

1.4　判別分析とは

表 1.5 は健常者とある疾病にかかっている患者に対する 2 種類の検査値 x_1（量的変数）と x_2（量的変数）のデータである．

表 1.5　健常者・患者の検査値のデータ

サンプル No.	健常者・患者	検査値 1 x_1	検査値 2 x_2
1	健常者	50	15.5
2	健常者	69	18.4
3	健常者	93	26.4
4	健常者	76	22.9
5	健常者	88	18.6
6	患者	43	16.9
7	患者	56	21.6
8	患者	38	12.2
9	患者	21	16.0
10	患者	25	10.5

このデータに基づいて知りたいことは次の通りである．

(1) 疾病にかかっているか否かを検査値 1 と検査値 2 より判別できるか．

(2) 判別できるとすればその精度はどのくらいか．

(3) 例えば，$x_1 = 70$，$x_2 = 19.0$ ならどのように判別されるか．

（2）判別分析の解析ストーリー

判別分析の解析の流れは以下の通りである．

判別分析の解析ストーリー

（1）健常者を母集団 [1]，患者を母集団 [2] とする．母集団 [1] と母集団 [2] における変数の確率分布を母平均（ベクトル）が異なる正規分布と想定する．

変数の値からそれぞれの母集団への距離としてマハラノビスの距離の2乗を求める．マハラノビスの距離の2乗値の小さい母集団へ判別するという判別方式を定める．

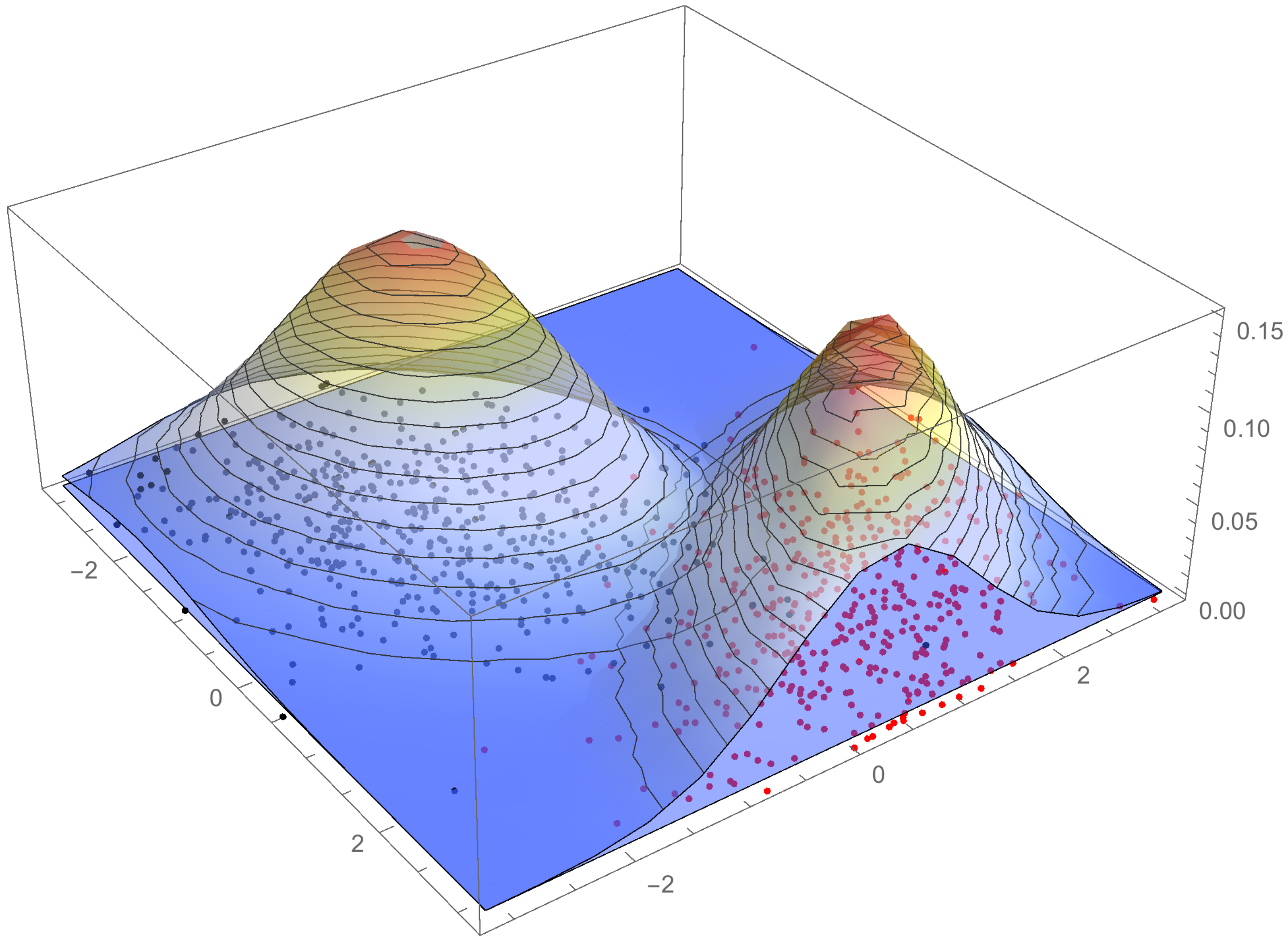
（2）誤判別の確率を求め，得られた判別方式の精度を評価する．

（3）変数選択を行い，有用な変数を選択する．

（4）得られた判別方式を利用して，どちらの母集団に属するのか不明なサンプルの判別を行う．

判別分析

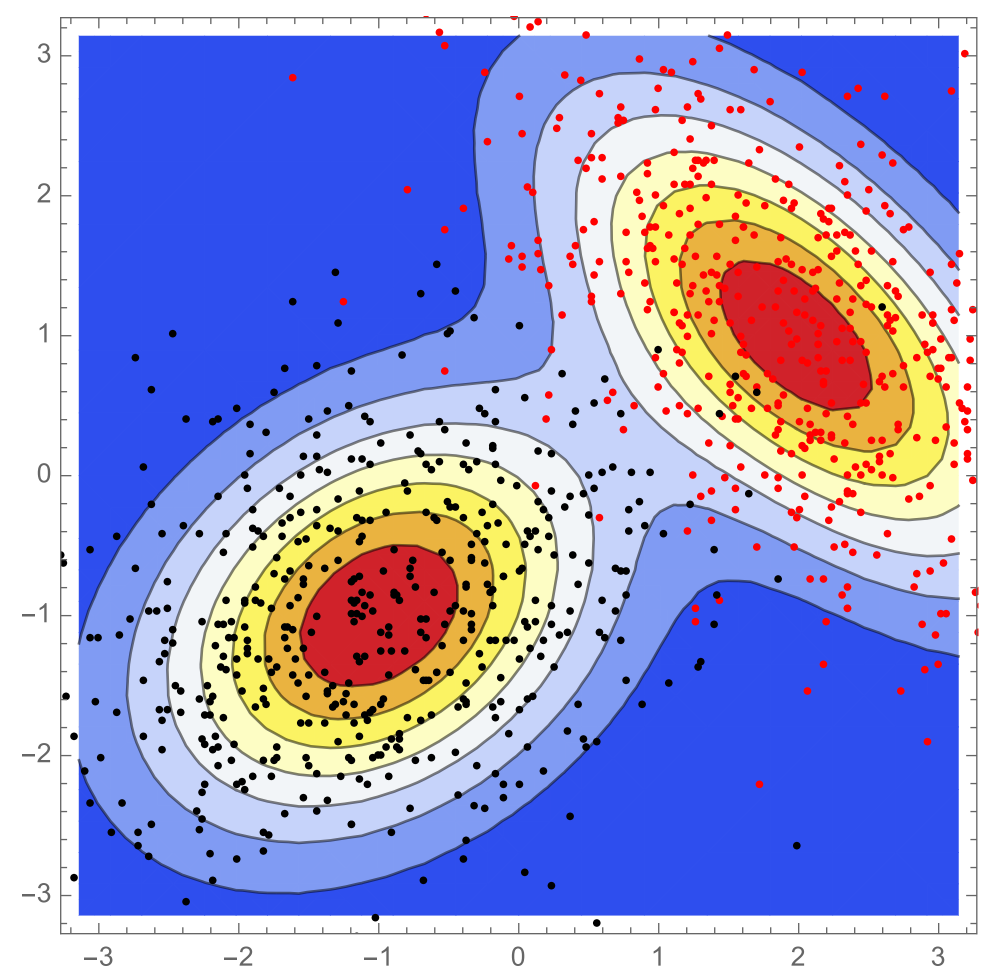
discriminant analysis



事前に与えられた二群のデータについて、各々平均と分散共分散を推定し所属が未知な検査点xがどちらの群に由来するかの判別を行う

検査点xと各々の群の平均ベクトルとのマハラノビス距離を計算し、最も距離が近いものを結論とする

(パターン認識で、クラスの事前確率が等しいケースでのpluginベイズの二次判別)



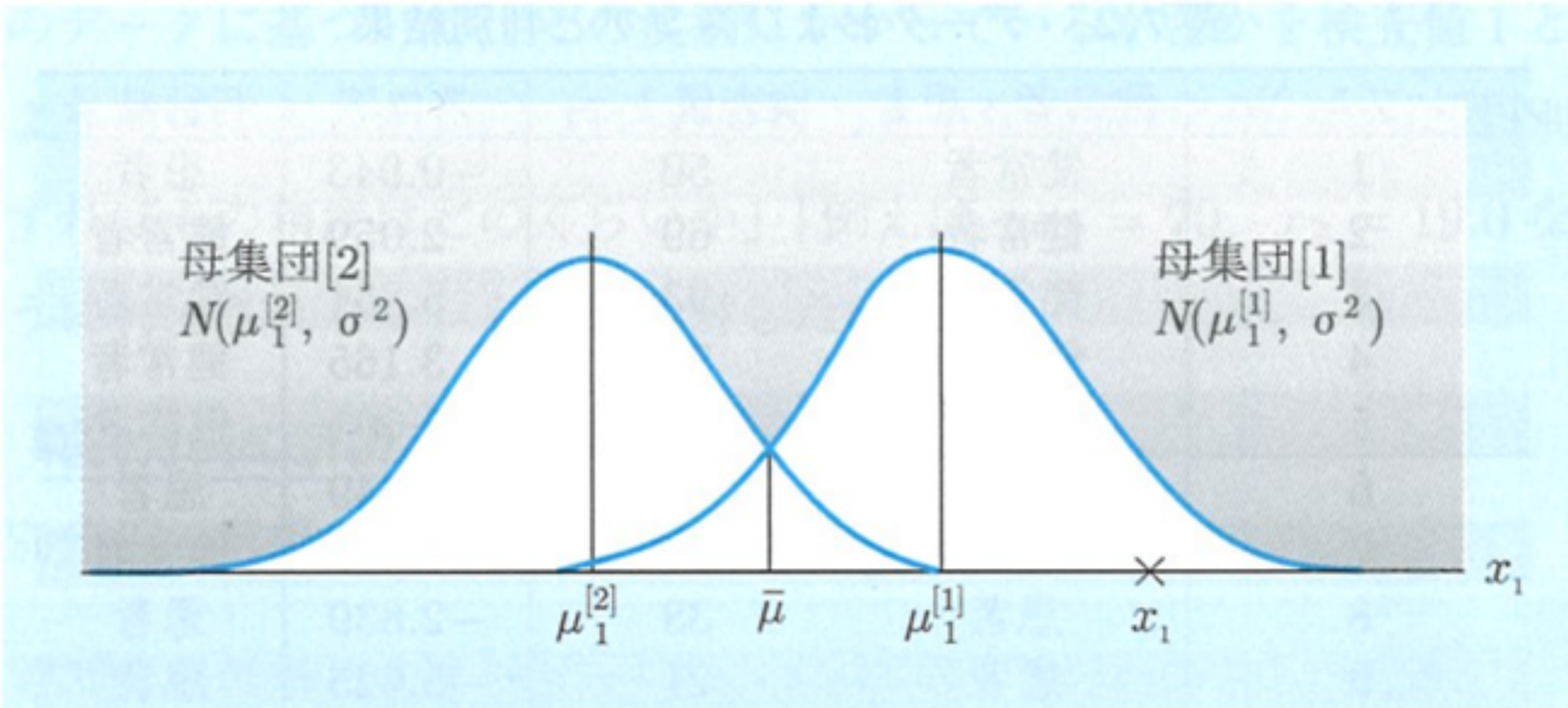


図 7.1　変数が 1 個の場合の判別

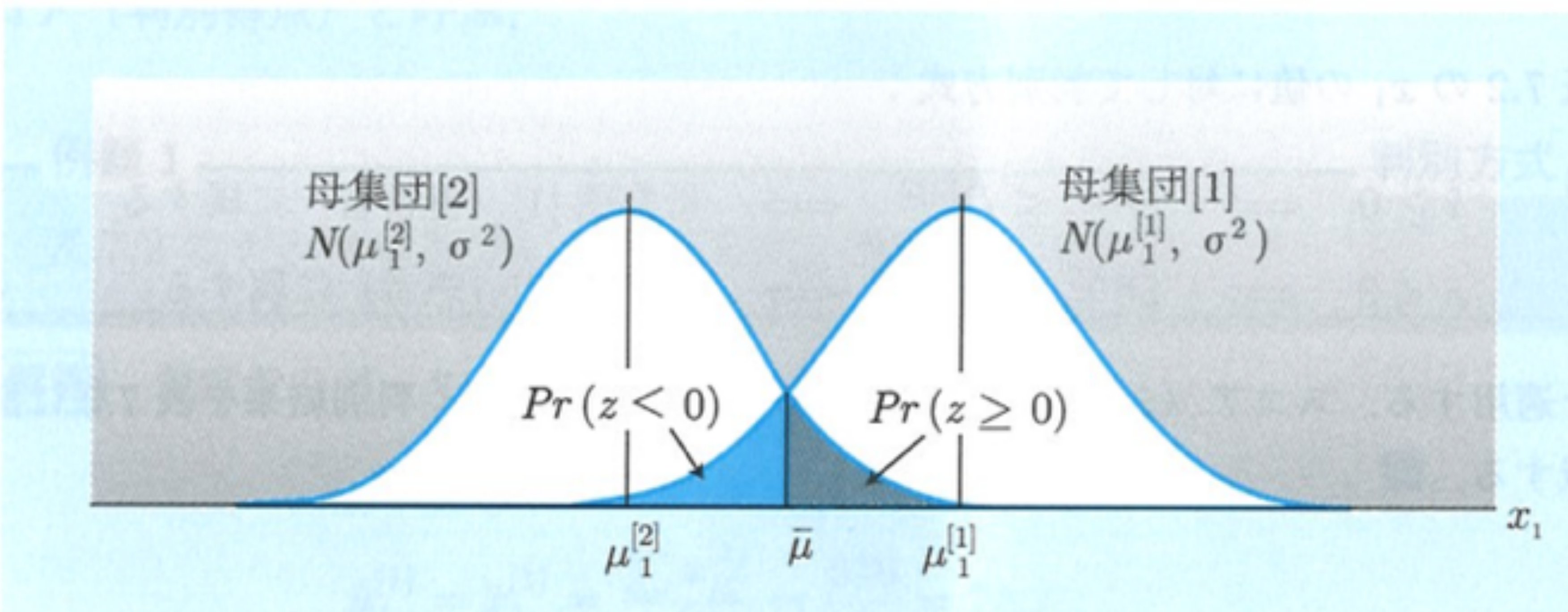
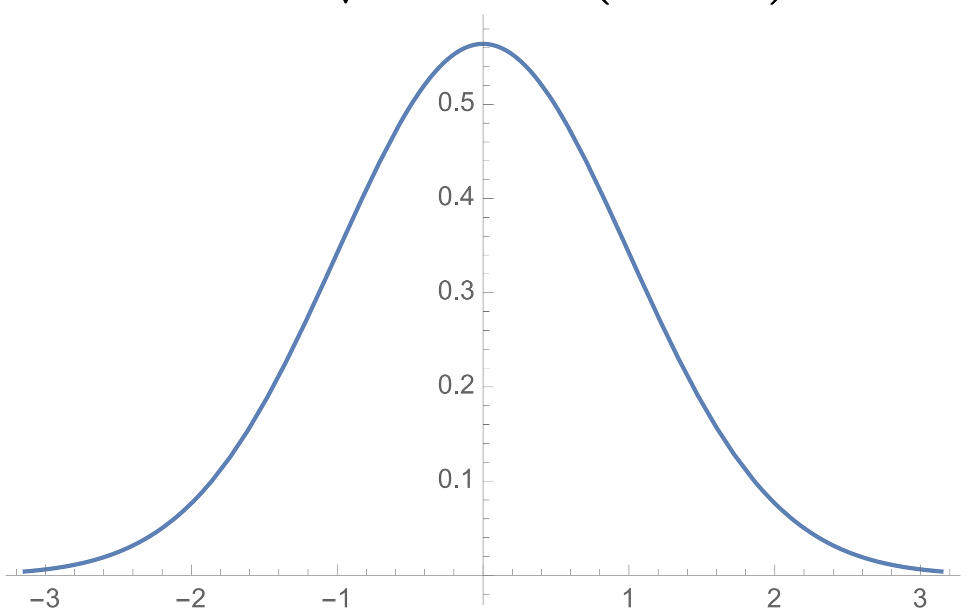


図 7.2　変数が 1 個の場合の誤判別の確率

標準正規分布 → 多変量標準正規分布

標準正規分布

$$p(x) := \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{1}{2}x^2\right)$$



平均 0
分散 1

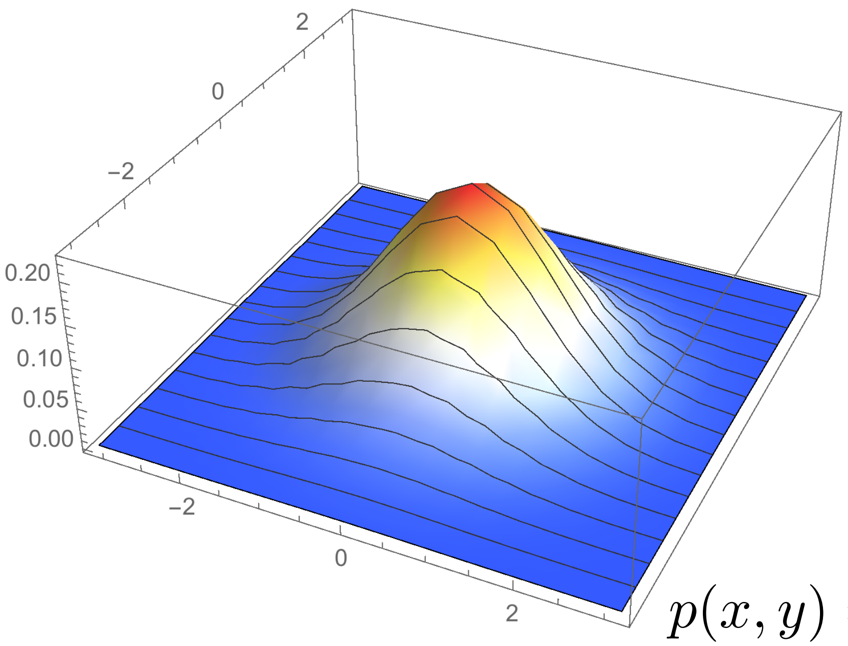
$$\boldsymbol{x} := (x_1, x_2, \ldots, x_n)'$$

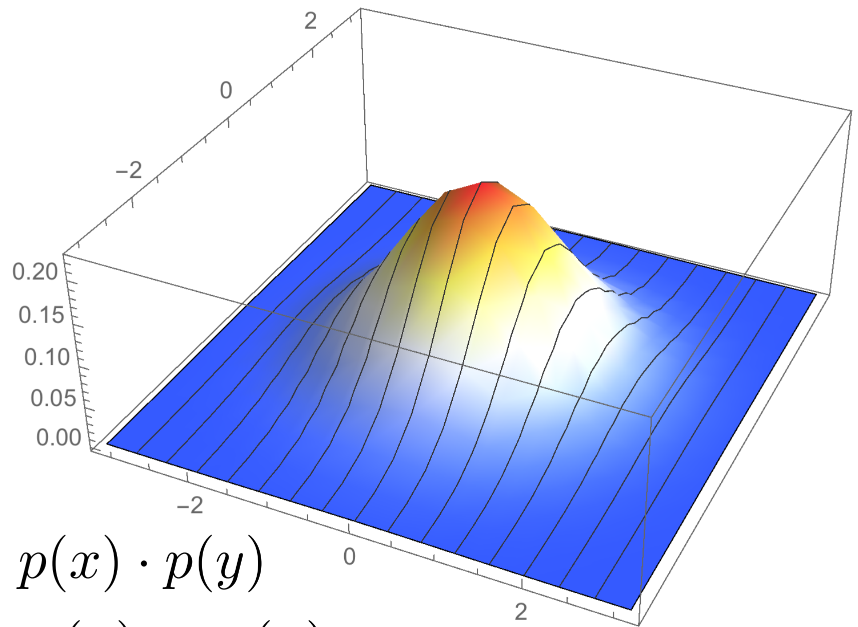
各々の変数が独立としてかけるだけ。

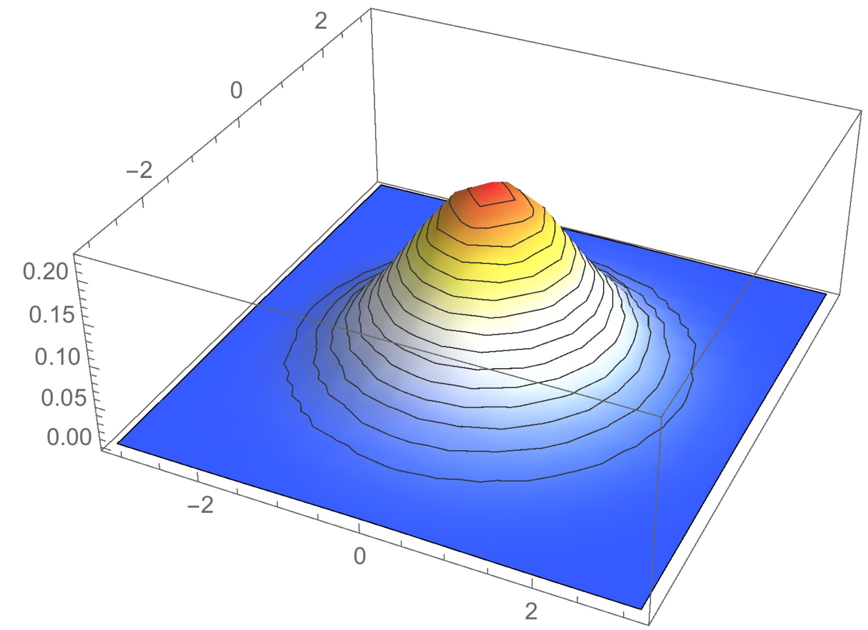
$$\begin{aligned} p(\boldsymbol{x}) &= p(x_1)p(x_2)\cdots p(x_n) \\ &= \frac{1}{(\sqrt{2\pi})^n} \exp\left(-\frac{1}{2}\sum_{i=1}^{n} x_i^2\right) \\ &= \frac{1}{(\sqrt{2\pi})^n} \exp\left(-\frac{1}{2}\boldsymbol{x}'\boldsymbol{x}\right) \end{aligned}$$

$$\boxed{p(\boldsymbol{x}) := \frac{1}{(\sqrt{2\pi})^n} \exp\left(-\frac{1}{2}\boldsymbol{x}'\boldsymbol{x}\right)}$$

２変数の標準正規分布の例

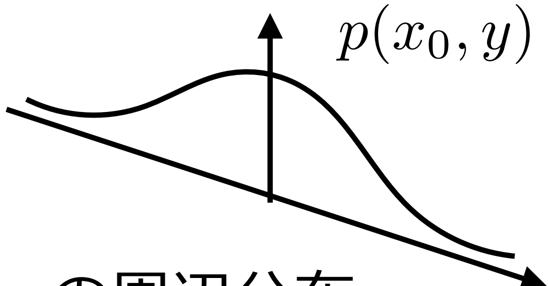




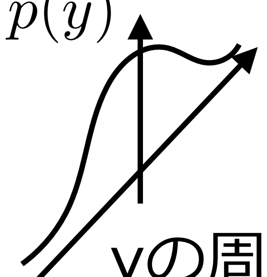


$$p(x, y) = p(x) \cdot p(y)$$

$$p(x_0, y) = p(x_0) \cdot p(y) \propto p(y)$$



xの周辺分布

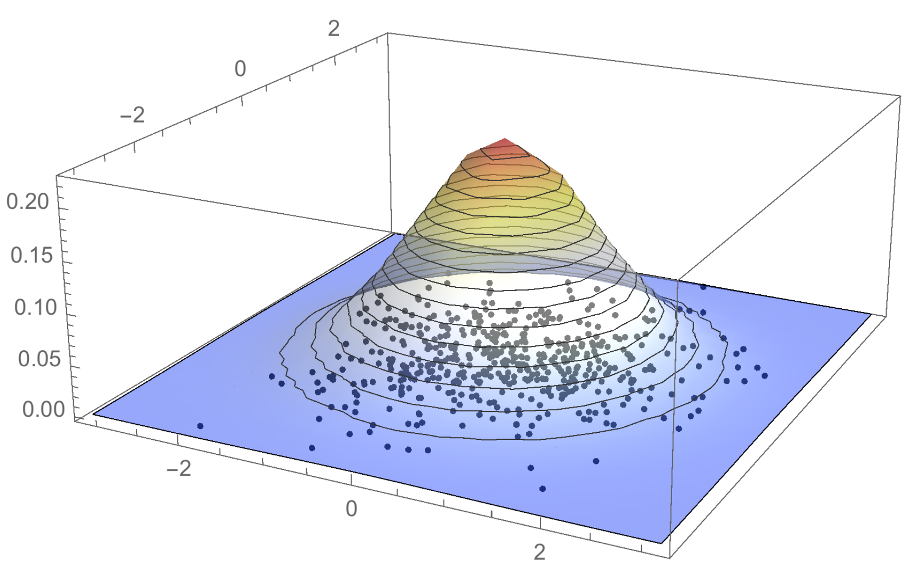


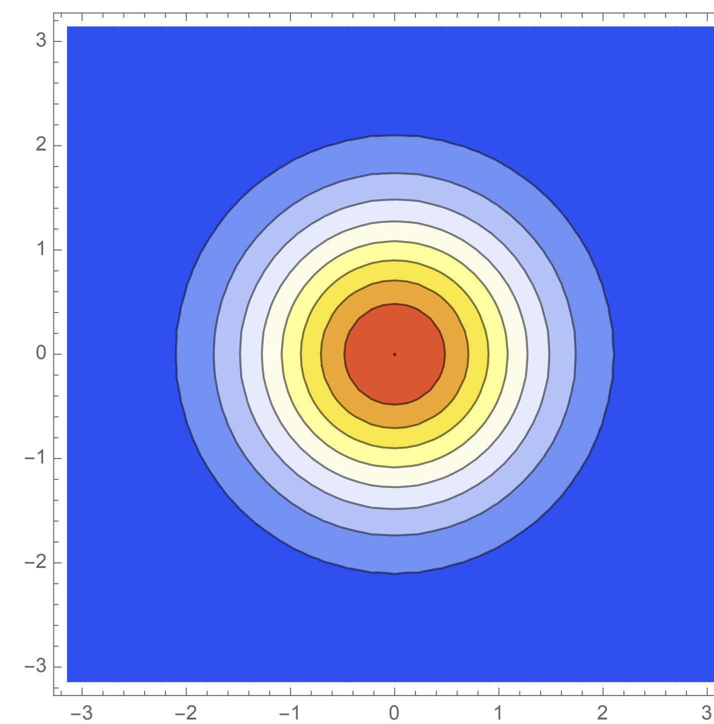
yの周辺分布

等確率面(等高線)

xとyは独立なので、絡みはなく
等高線(等確率面)は球形になる

(2Dでは真円形→)





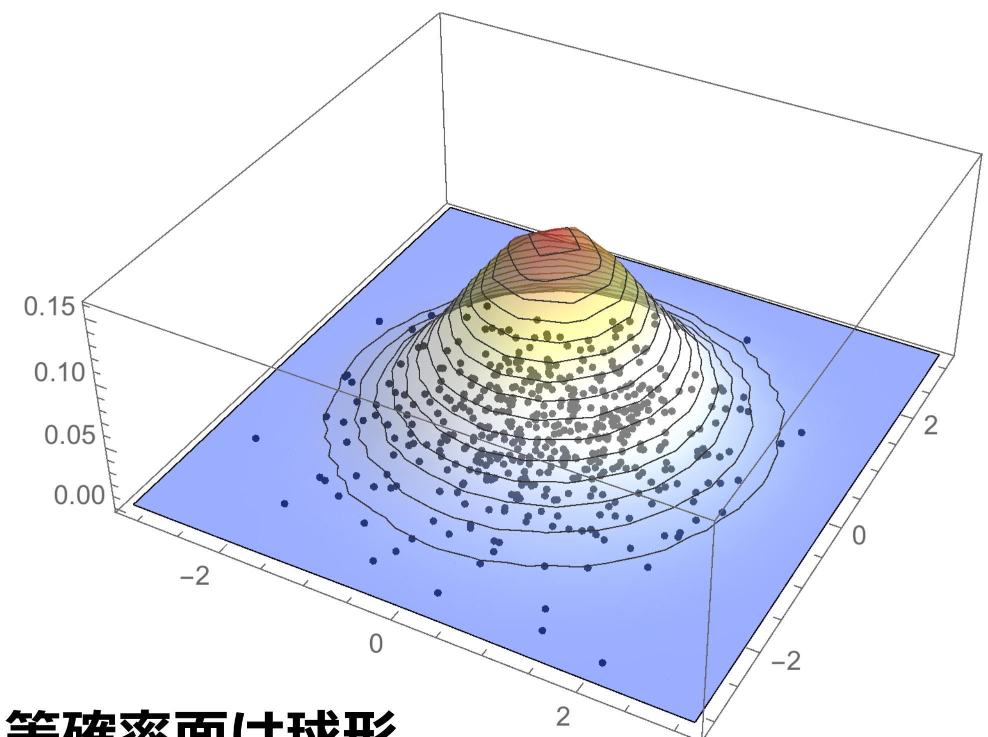
標準正規分布から一般の多変量正規分布へ

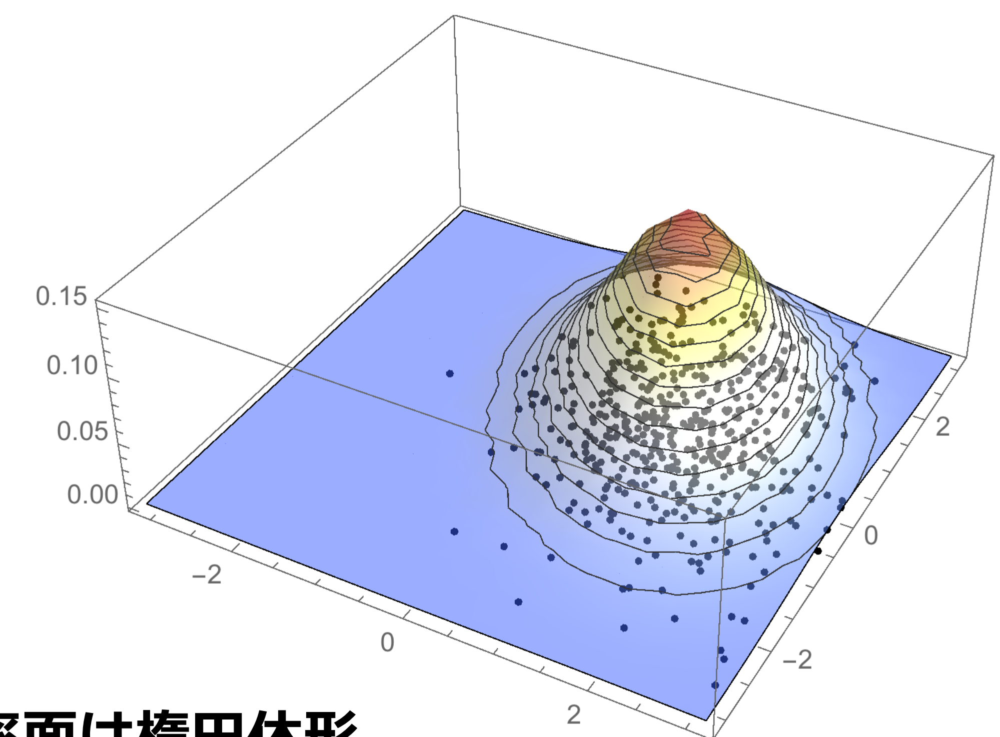
$N(\mathbf{0}, \boldsymbol{I})$

$$\frac{1}{(\sqrt{2\pi})^n} \exp\left(-\frac{1}{2}\boldsymbol{x}'\boldsymbol{x}\right)$$

$N(\boldsymbol{\mu}, \boldsymbol{\Sigma})$

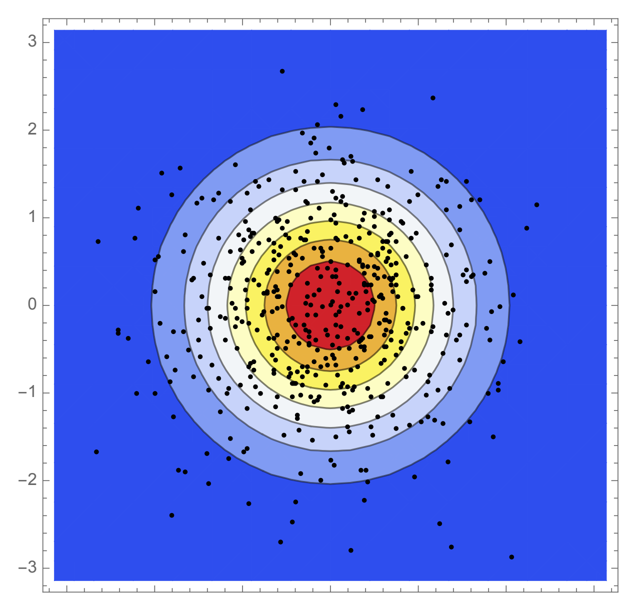
$$\frac{1}{(\sqrt{2\pi})^n \sqrt{\det(\boldsymbol{\Sigma})}} \exp\left(-\frac{1}{2}(\boldsymbol{x}-\boldsymbol{\mu})'\boldsymbol{\Sigma}^{-1}(\boldsymbol{x}-\boldsymbol{\mu})\right)$$





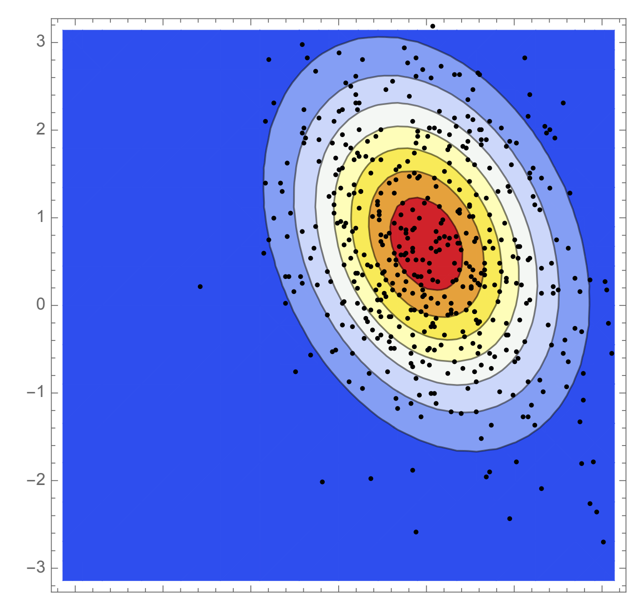
等確率面は球形

2D:等高線が真円形



等確率面は楕円体形

2D:等高線が楕円形



標本分散共分散行列

$$\hat{\Sigma} = \frac{1}{n-1}\sum_{i=1}^{n}(\boldsymbol{x}_i - \bar{\boldsymbol{x}})(\boldsymbol{x}_i - \bar{\boldsymbol{x}})' \quad \text{ただし} \ \bar{\boldsymbol{x}} := \frac{1}{n}\sum_{i=1}^{n}\boldsymbol{x}_i$$

単変量の標本分散のときと同じ理屈で不偏推定量になる

$$\begin{aligned}
\mathbb{E}\{\hat{\Sigma}\} &= \frac{1}{n-1}\mathbb{E}\left\{\sum_{i=1}^{n}(\boldsymbol{x}_i - \boldsymbol{\mu})(\boldsymbol{x}_i - \boldsymbol{\mu})' - 2n(\bar{\boldsymbol{x}} - \boldsymbol{\mu})(\bar{\boldsymbol{x}} - \boldsymbol{\mu})' + n(\bar{\boldsymbol{x}} - \boldsymbol{\mu})(\bar{\boldsymbol{x}} - \boldsymbol{\mu})'\right\} \\
&= \frac{1}{n-1}\left(\sum_{i=1}^{n}\mathbb{E}\left\{(\boldsymbol{x}_i - \boldsymbol{\mu})(\boldsymbol{x}_i - \boldsymbol{\mu})'\right\} - n\mathbb{E}\left\{(\bar{\boldsymbol{x}} - \boldsymbol{\mu})(\bar{\boldsymbol{x}} - \boldsymbol{\mu})'\right\}\right) \\
&= \frac{1}{n-1}\left(\sum_{i=1}^{n}\Sigma - n\frac{1}{n}\Sigma\right) = \Sigma
\end{aligned}$$

- 参考) 正規分布する変数を線形変換しても正規分布する

証明「数理統計学ー基礎から学ぶデータ解析ー(鈴木武・山田作太郎著)」p.121 定理4.2

定理 4.2 $\boldsymbol{X}$ が k 次元正規分布 $N_k(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ に従うとき

$$\boldsymbol{Y}=\boldsymbol{AX}+\boldsymbol{a},$$

$\boldsymbol{A}$ は $p\times k$ の定数行列，$\boldsymbol{a}$ は p 次元定数ベクトル

とおくと，$\boldsymbol{Y}$ は $N_p(\boldsymbol{A\mu}+\boldsymbol{a}, \boldsymbol{A\Sigma A}')$ に従う．

証明 $\boldsymbol{t}=(t_1, \cdots, t_p)'$ に対して

$$\begin{aligned} M_Y(\boldsymbol{t}) &= E(e^{\boldsymbol{t}'(\boldsymbol{AX}+\boldsymbol{a})})=e^{\boldsymbol{t}'\boldsymbol{a}}E(e^{\boldsymbol{t}'\boldsymbol{AX}}) \\ &= e^{\boldsymbol{t}'\boldsymbol{a}}\exp\left\{\boldsymbol{t}'\boldsymbol{A\mu}+\frac{1}{2}\boldsymbol{t}'\boldsymbol{A\Sigma A}'\boldsymbol{t}\right\} \\ &= \exp\left\{\boldsymbol{t}'(\boldsymbol{A\mu}+\boldsymbol{a})+\frac{1}{2}\boldsymbol{t}'\boldsymbol{A\Sigma A}'\boldsymbol{t}\right\} \end{aligned}$$

を得る．積率母関数は確率分布と1対1に対応するので，(4.51)より $\boldsymbol{Y}$ は p 次元正規分布 $N_p(\boldsymbol{A\mu}+\boldsymbol{a}, \boldsymbol{A\Sigma A}')$ に従う． □

定理 4.2 では $\boldsymbol{A\Sigma A}'$ は対称行列ではあるが，必ずしも正値定符号ではない．$p\leq k$ のとき $\boldsymbol{Y}$ の分布が(4.45)の意味での p 次元正規分布であるためには，$\boldsymbol{A}$ のランク(rank)が p であればよい．

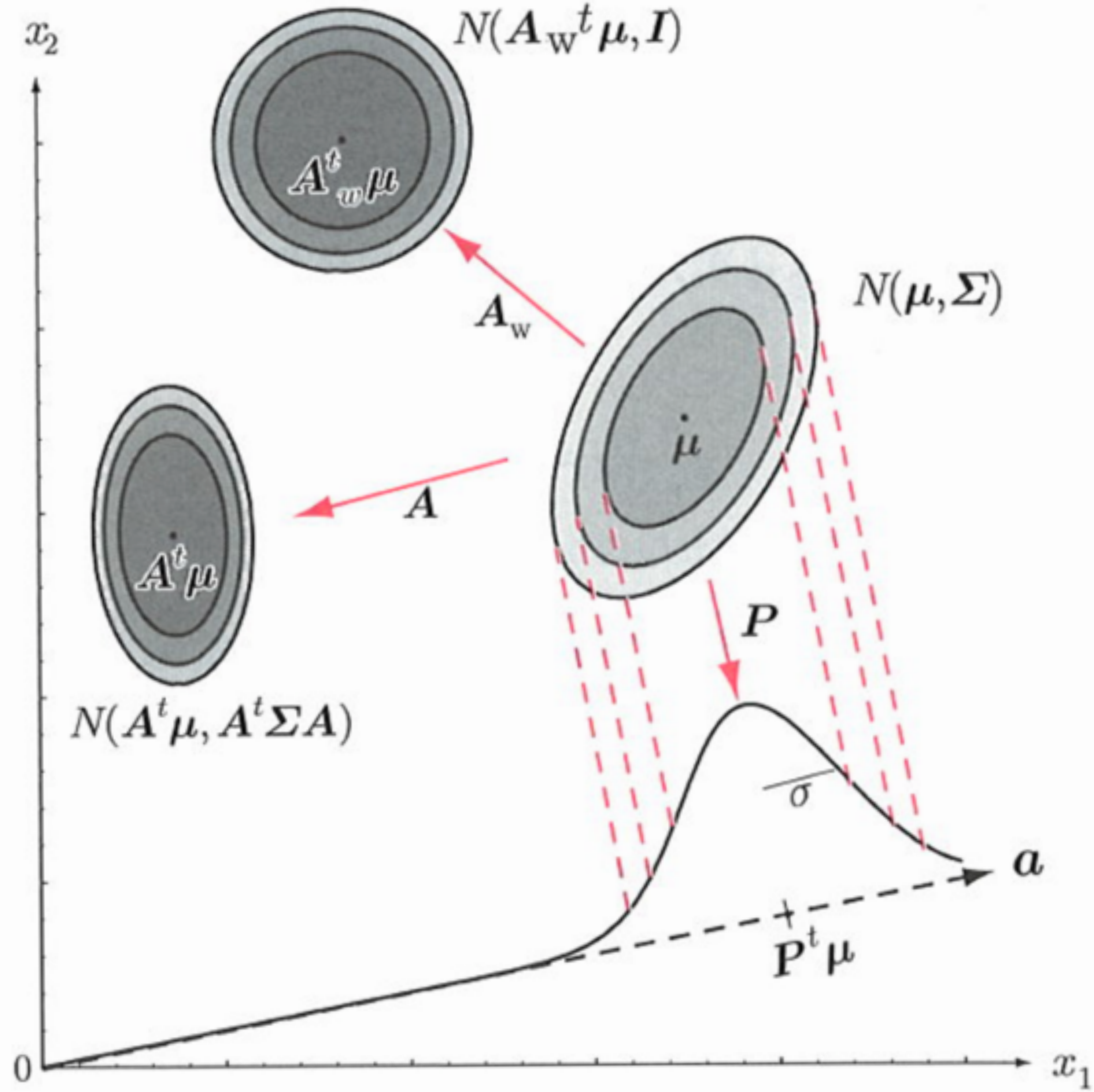
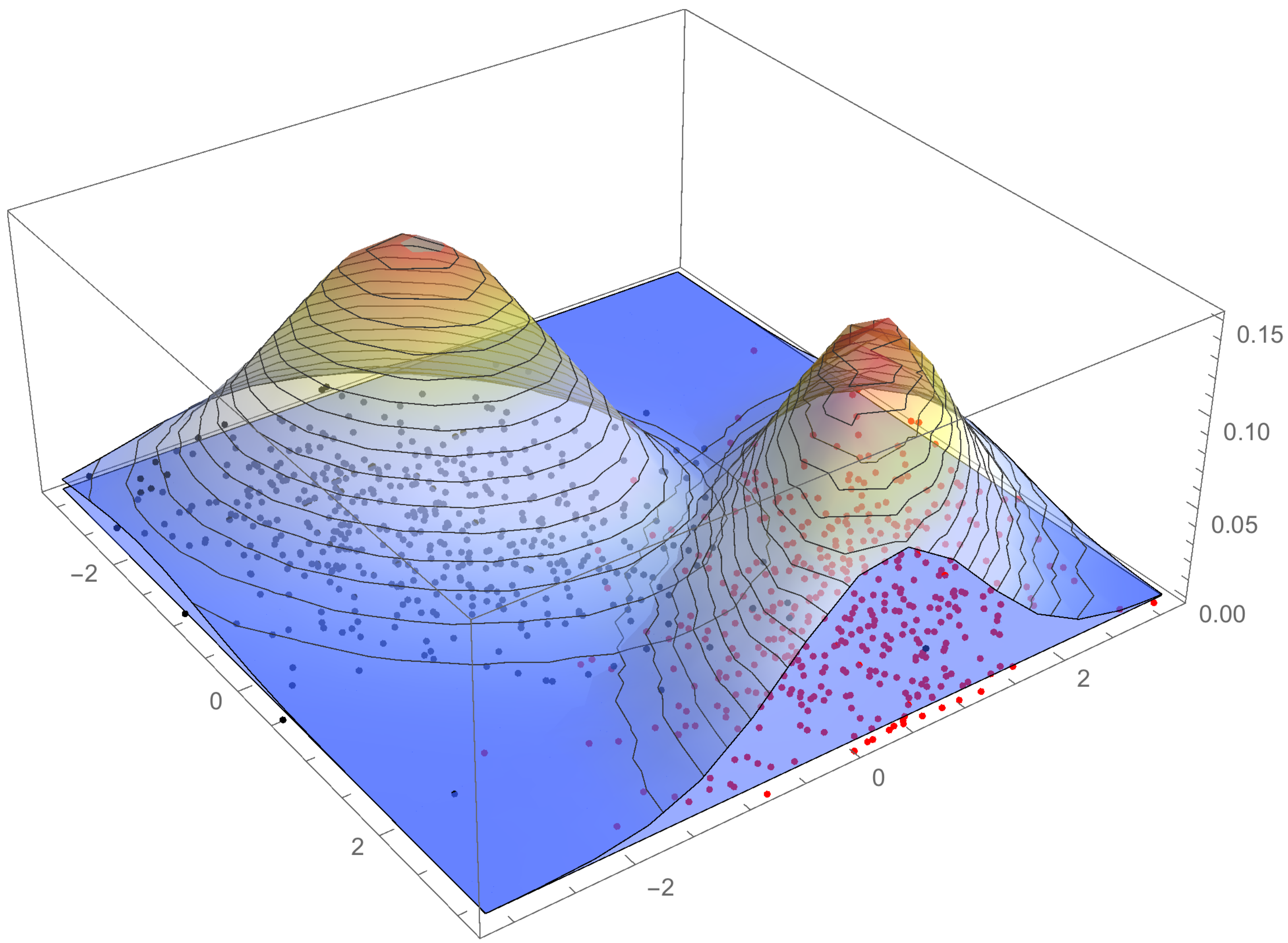


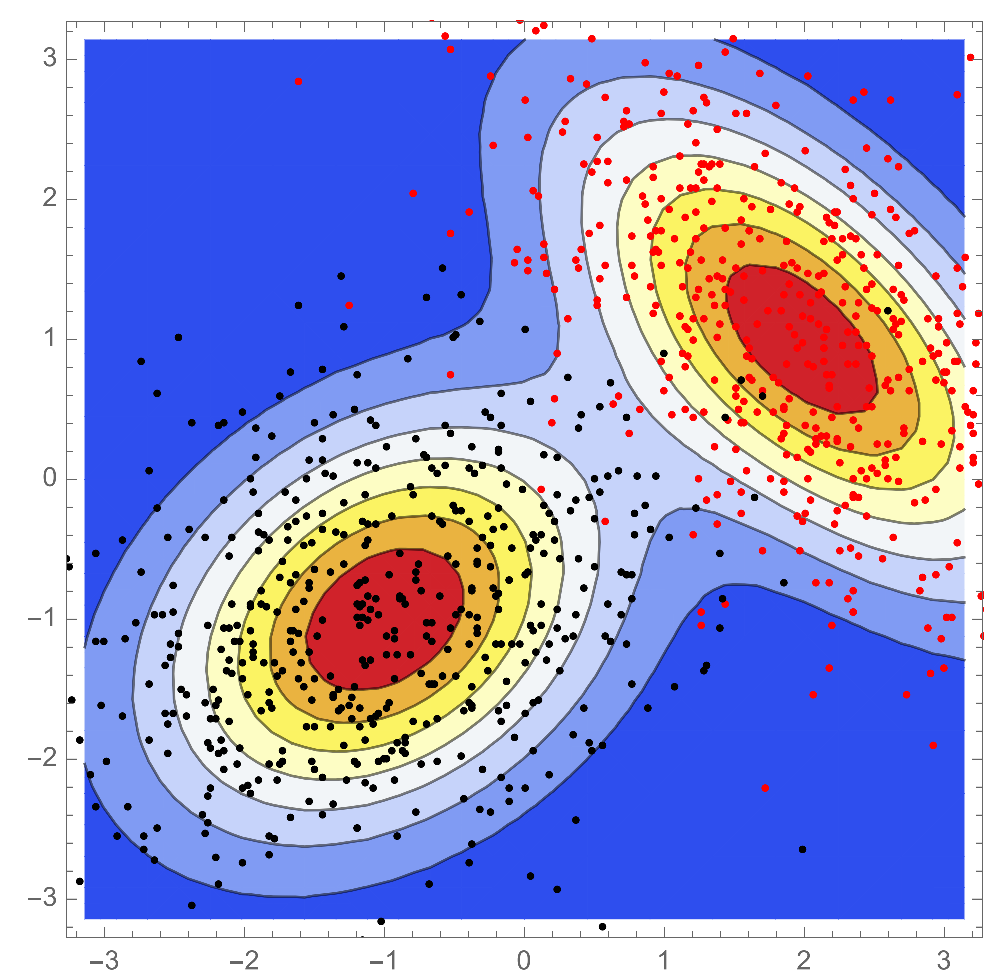
図 2.8　特徴空間での線形変換は任意の正規分布を異なる正規分布へ変換する。$\boldsymbol{A}$ なる変換は，もとの分布を $N(\boldsymbol{A}^t\boldsymbol{\mu}, \boldsymbol{A}^t\boldsymbol{\Sigma}\boldsymbol{A})$ へ変える。別の線形変換── $\boldsymbol{a}$ で定義される直線への投影 $\boldsymbol{P}$ ──は，その線に沿って $N(\mu, \sigma^2)$ となる分布を導く。変換は別の空間の分布を生むが，ここでは元の x_1–x_2 空間に重ねて表示している。白色化変換 $\boldsymbol{A}_{\mathrm{w}}$ により生じる円状の対称ガウス分布も図中にずらして示す。

判別分析

discriminant analysis

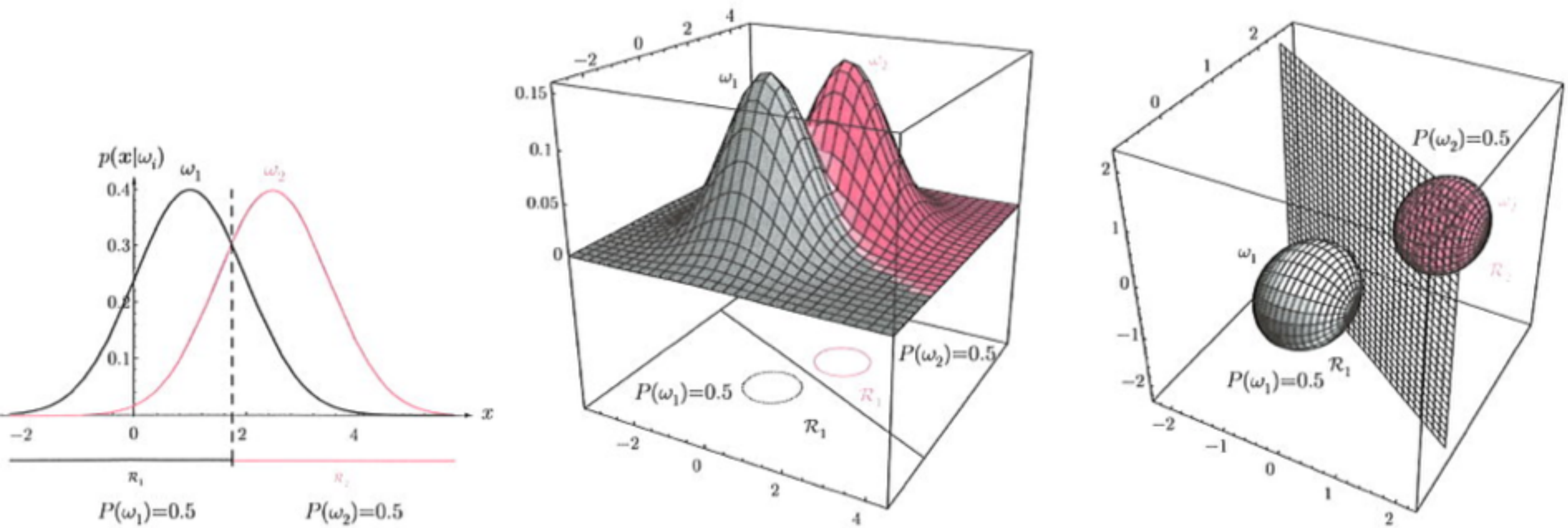


事前に与えられた二群のデータについて、各々平均と分散共分散を推定し所属が未知な検査点xがどちらの群に由来するかの判別を行う

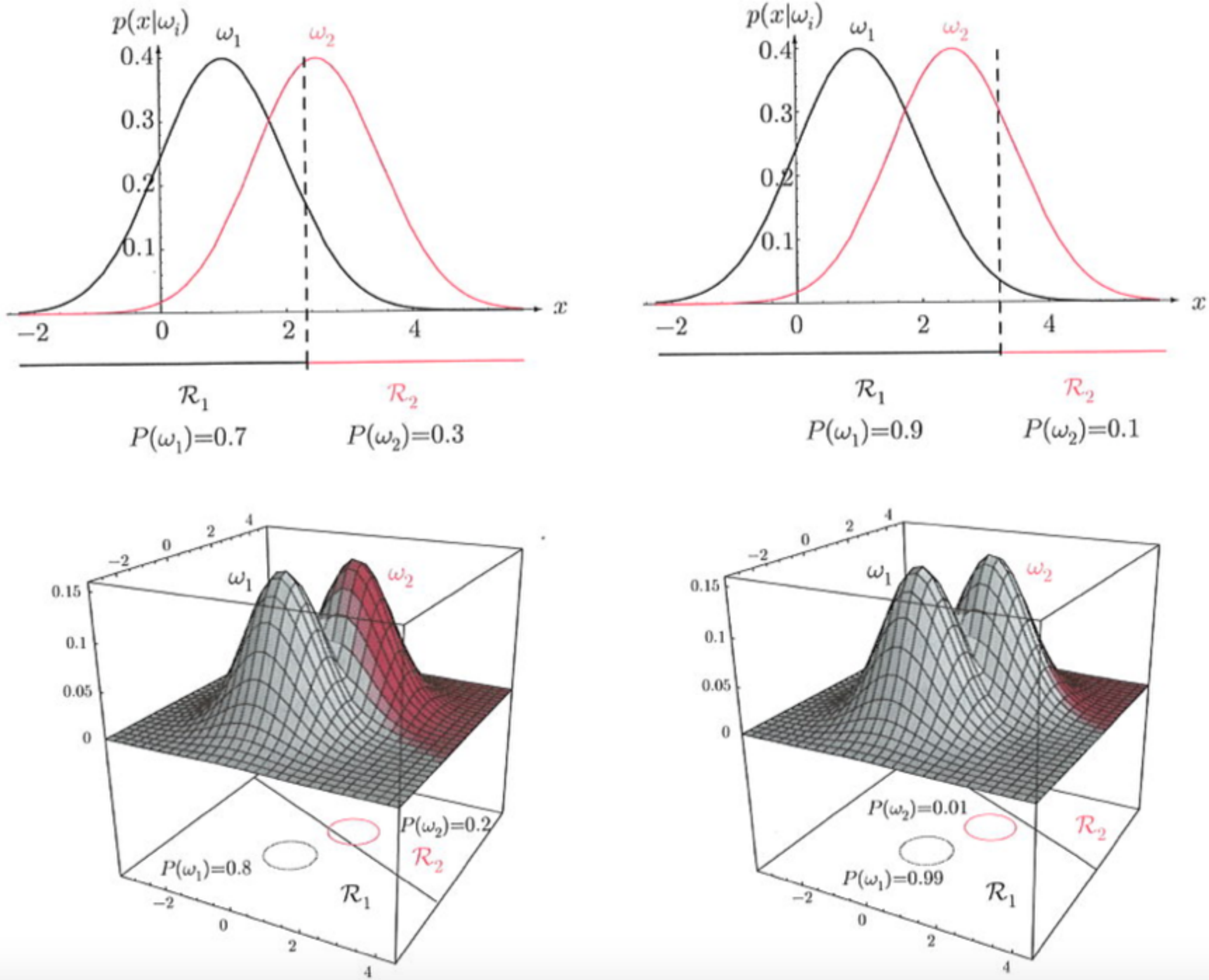


検査点xと各々の群の平均ベクトルとのマハラノビス距離を計算し、最も距離が近いものを結論とする

(パターン認識で、クラスの事前確率が等しいケースでのpluginベイズの二次判別)



2 つの分布の共分散行列が等しく，単位行列に比例していれば，それらの分布は d 次元空間で球状であり，境界は中心を結ぶ直線に垂直な $d-1$ 次元の一般化超平面となる。ここでは $P(\omega_1) = P(\omega_2)$ の場合の $p(\boldsymbol{x}|\omega_i)$ と境界を，1，2，3 次元の例で示す。3 次元の例では格子状に表した面が $\mathcal{R}_1$ と $\mathcal{R}_2$ を分割する。



$p(x|\omega_i)$
ω_1
ω_2
x
$\mathcal{R}_1$
$\mathcal{R}_2$
$P(\omega_1)$=0.7
$P(\omega_2)$=0.3
$P(\omega_1)$=0.9
$P(\omega_2)$=0.1
$P(\omega_1)$=0.8
$P(\omega_2)$=0.2
$P(\omega_1)$=0.99
$P(\omega_2)$=0.01

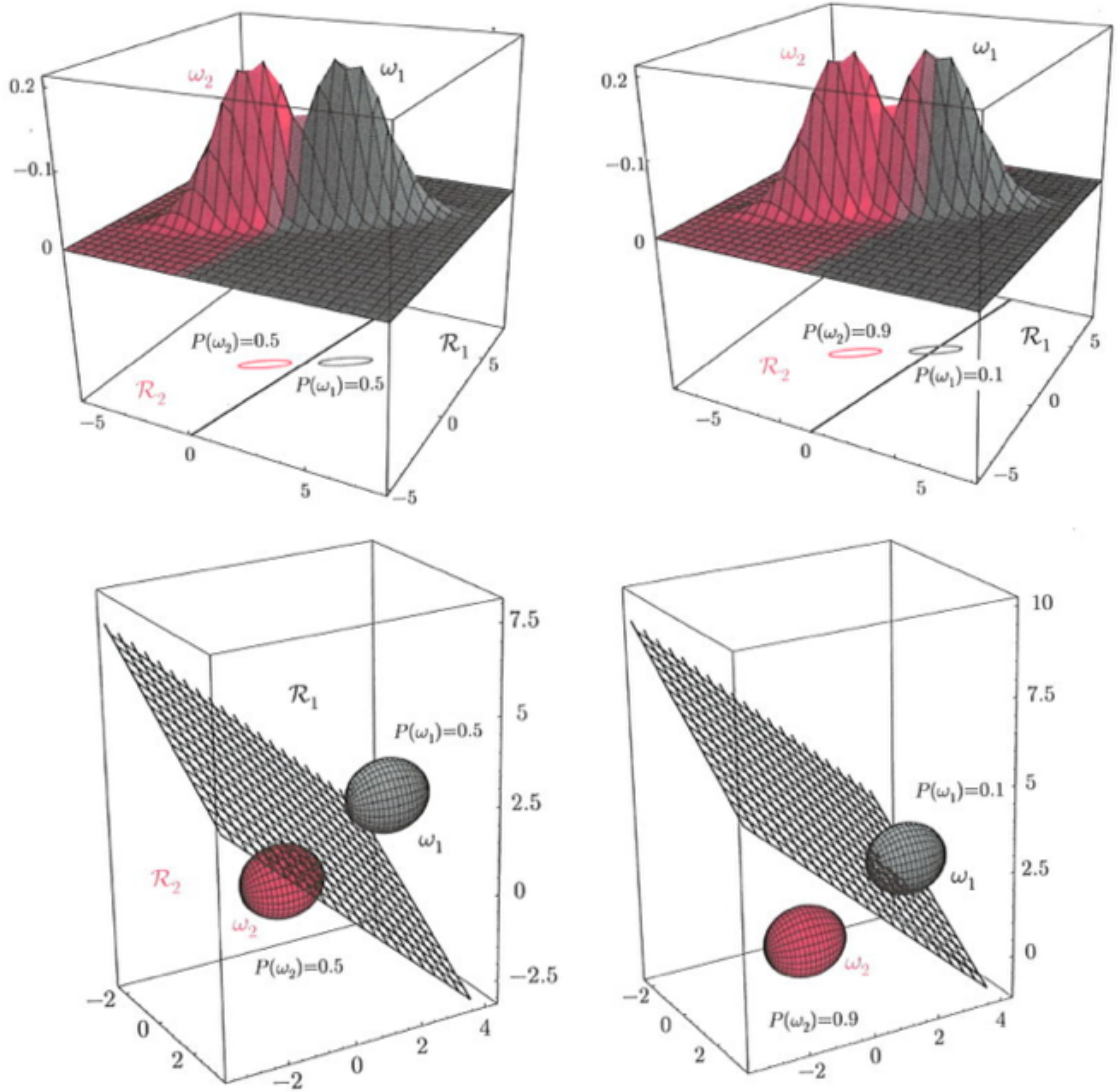
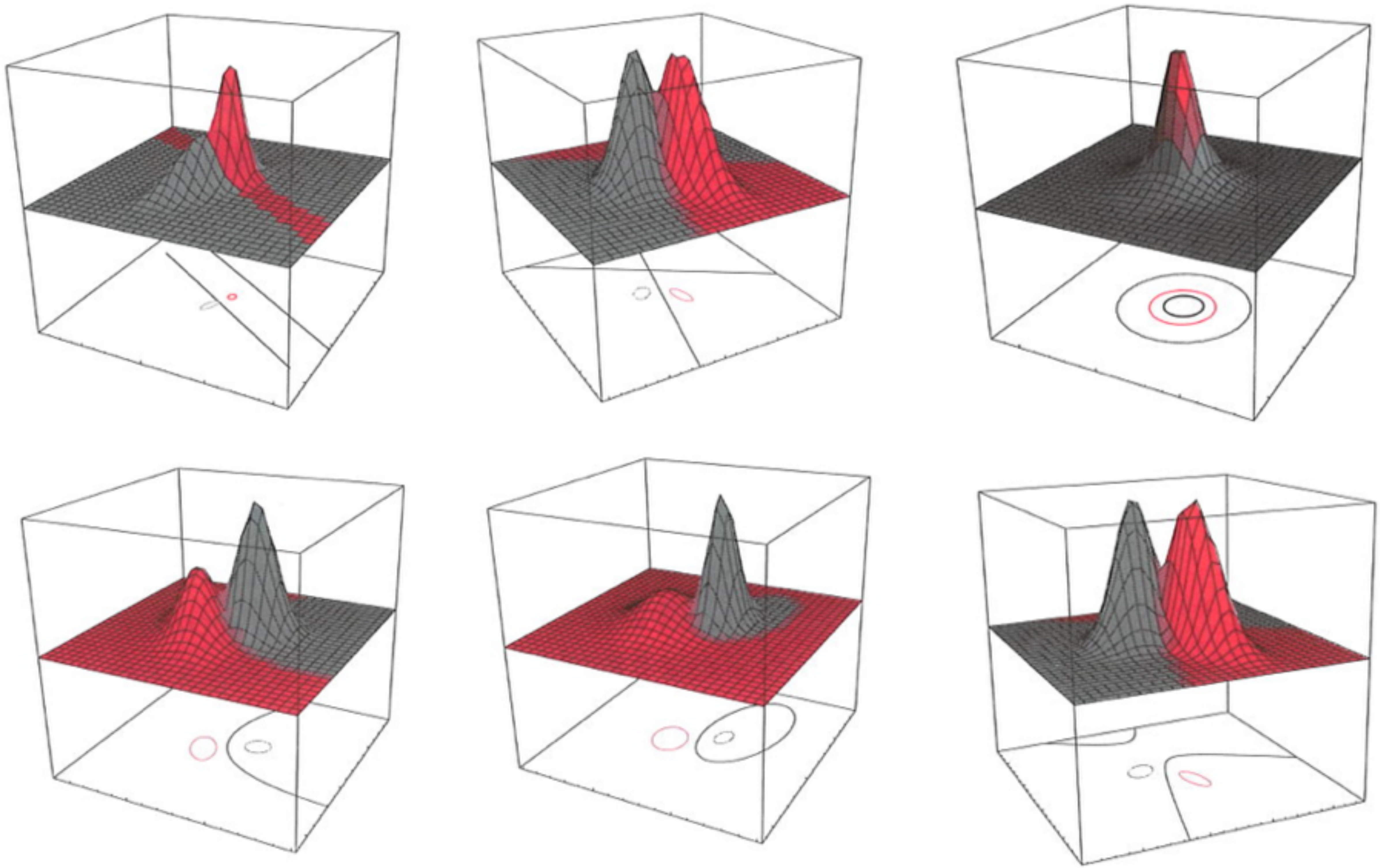


図 2.12　同一で非対称なガウス分布の確率密度（2 次元では曲面で，3 次元では楕円体面で示す）と決定領域。決定超平面は平均値を結ぶ直線と直交するとは限らない。



要確認：対角行列をかけると…

右からかける → 各列ベクトルを定数倍

$$\begin{bmatrix} a & b & c \\ d & e & f \\ g & h & i \end{bmatrix} \begin{bmatrix} \alpha & 0 & 0 \\ 0 & \beta & 0 \\ 0 & 0 & \gamma \end{bmatrix} = \begin{bmatrix} \alpha \cdot a & \beta \cdot b & \gamma \cdot c \\ \alpha \cdot d & \beta \cdot e & \gamma \cdot f \\ \alpha \cdot g & \beta \cdot h & \gamma \cdot i \end{bmatrix}$$

左からかける → 各行ベクトルを定数倍

$$\begin{bmatrix} \alpha & 0 & 0 \\ 0 & \beta & 0 \\ 0 & 0 & \gamma \end{bmatrix} \begin{bmatrix} a & b & c \\ d & e & f \\ g & h & i \end{bmatrix} = \begin{bmatrix} \alpha \cdot a & \alpha \cdot b & \alpha \cdot c \\ \beta \cdot d & \beta \cdot e & \beta \cdot f \\ \gamma \cdot g & \gamma \cdot h & \gamma \cdot i \end{bmatrix}$$

要確認：基底を変える行列の作り方

$$\begin{bmatrix}1\\0\end{bmatrix} \xrightarrow{\boldsymbol{A}} \begin{bmatrix}v_1\\v_2\end{bmatrix} \qquad \begin{bmatrix}0\\1\end{bmatrix} \xrightarrow{\boldsymbol{A}} \begin{bmatrix}u_1\\u_2\end{bmatrix}$$ となる行列 $\boldsymbol{A}$ は？

$$\begin{bmatrix}v_1\\v_2\end{bmatrix} = \boldsymbol{A}\begin{bmatrix}1\\0\end{bmatrix}$$

$$\begin{bmatrix}u_1\\u_2\end{bmatrix} = \boldsymbol{A}\begin{bmatrix}0\\1\end{bmatrix}$$

$$\begin{bmatrix}v_1 & u_1\\v_2 & u_2\end{bmatrix} = \boldsymbol{A}\begin{bmatrix}1 & 0\\0 & 1\end{bmatrix}$$

$$\boldsymbol{A} = \begin{bmatrix}a_{11} & a_{12}\\a_{21} & a_{22}\end{bmatrix}$$

$$\boldsymbol{A} = \begin{bmatrix}a_{11} & a_{12}\\a_{21} & a_{22}\end{bmatrix} = \begin{bmatrix}v_1 & u_1\\v_2 & u_2\end{bmatrix}$$

ミニレポートで確認してみよう！

1）次の関係をみたす行列 A はどんな形？

$$\begin{bmatrix} v_1 \\ v_2 \end{bmatrix} = \boldsymbol{A} \begin{bmatrix} 1 \\ 0 \end{bmatrix} \quad \text{かつ} \quad \begin{bmatrix} u_1 \\ u_2 \end{bmatrix} = \boldsymbol{A} \begin{bmatrix} 0 \\ 1 \end{bmatrix}$$

2）上の行列 A の列ベクトルが互いに直交する場合、行列 A'A および A^{-1} はどんな形？

3）上のように行列 A は直交行列とすると、次の 2 つのベクトルはどんな形？

$$\boldsymbol{A}' \begin{bmatrix} v_1 \\ v_2 \end{bmatrix} \qquad \boldsymbol{A}' \begin{bmatrix} u_1 \\ u_2 \end{bmatrix}$$

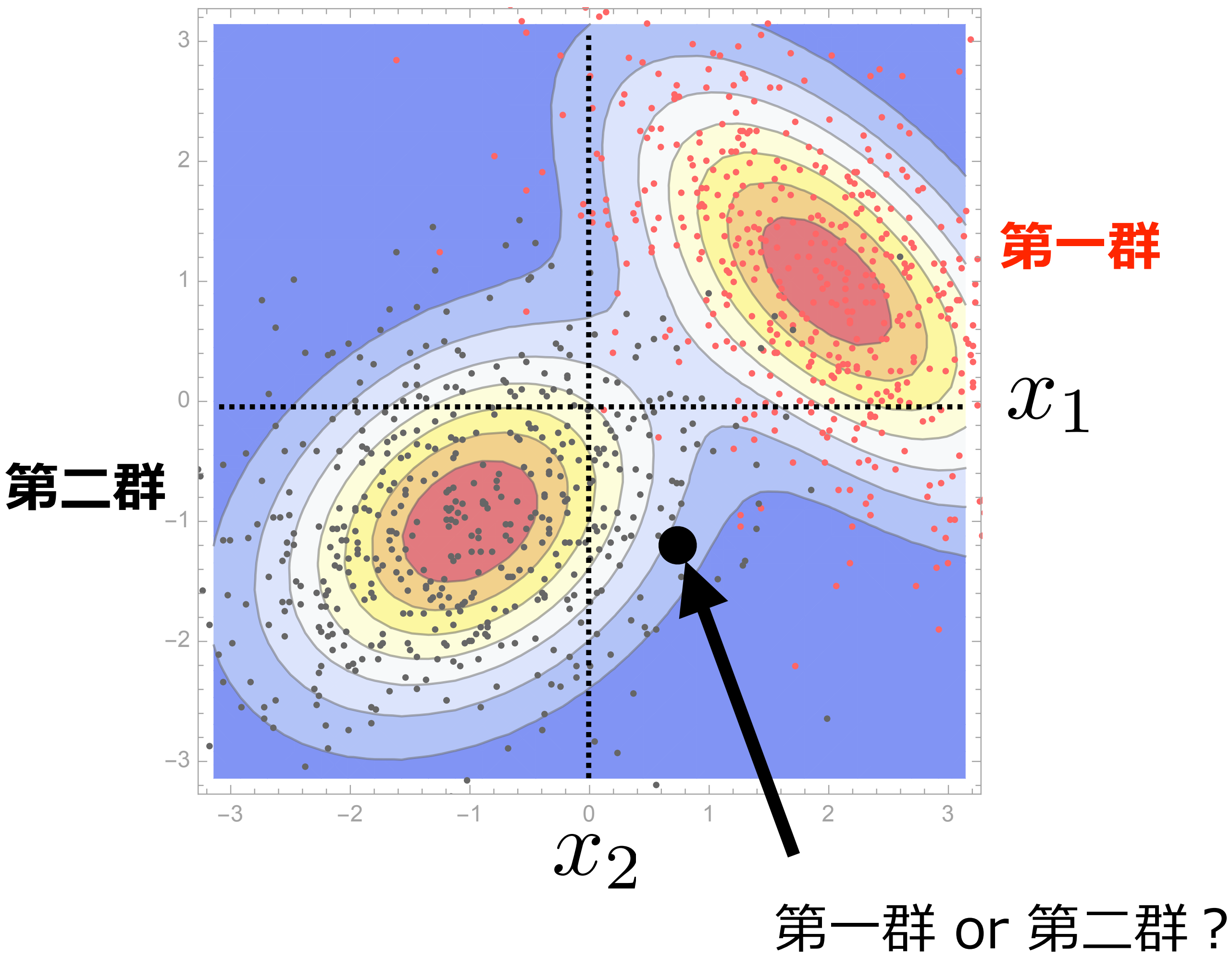
4）次の行列はどんな形？

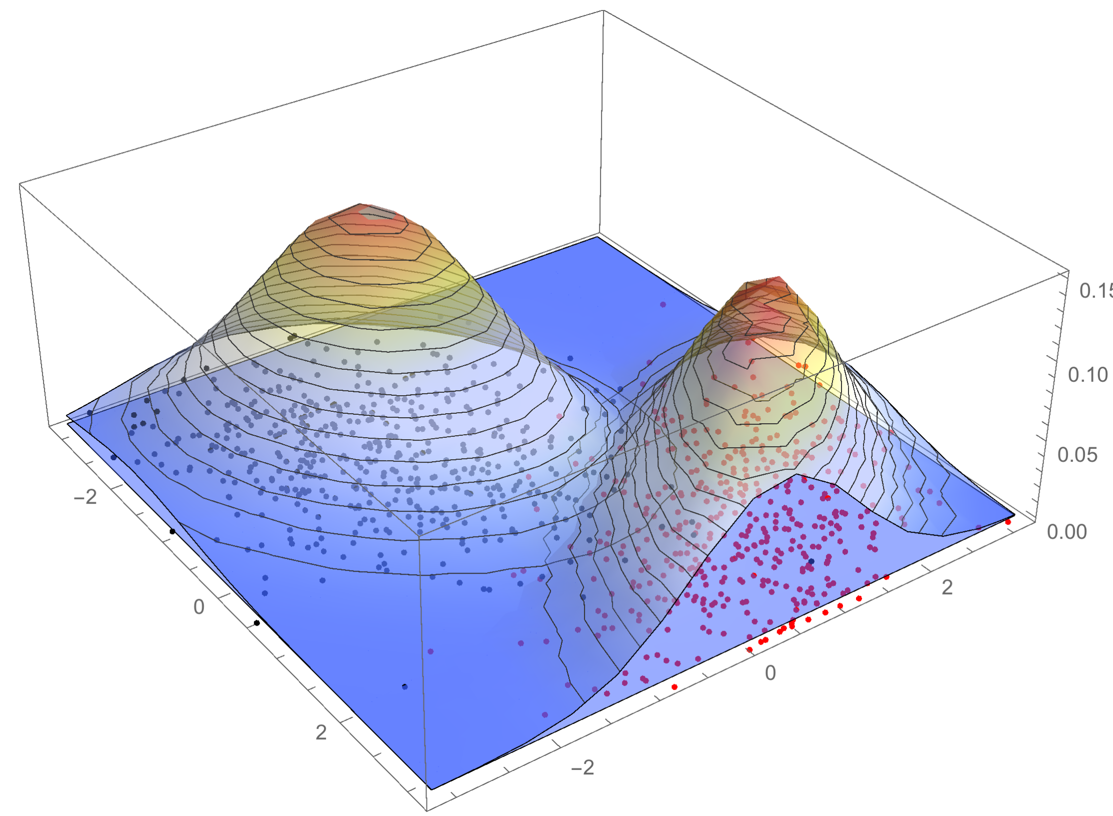
$$\begin{bmatrix} a_1 & b_1 \\ a_2 & b_2 \end{bmatrix} \begin{bmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{bmatrix} \text{および} \begin{bmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{bmatrix} \begin{bmatrix} a_1 & b_1 \\ a_2 & b_2 \end{bmatrix}'$$

5）左辺を展開し右辺と等しいことを確認せよ。

$$\begin{bmatrix} a_1 & b_1 \\ a_2 & b_2 \end{bmatrix} \begin{bmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{bmatrix} \begin{bmatrix} a_1 & b_1 \\ a_2 & b_2 \end{bmatrix}' = \lambda_1 \begin{bmatrix} a_1 \\ a_2 \end{bmatrix} \begin{bmatrix} a_1 & a_2 \end{bmatrix} + \lambda_2 \begin{bmatrix} b_1 \\ b_2 \end{bmatrix} \begin{bmatrix} b_1 & b_2 \end{bmatrix}$$

判別分析: Discriminant Analysis

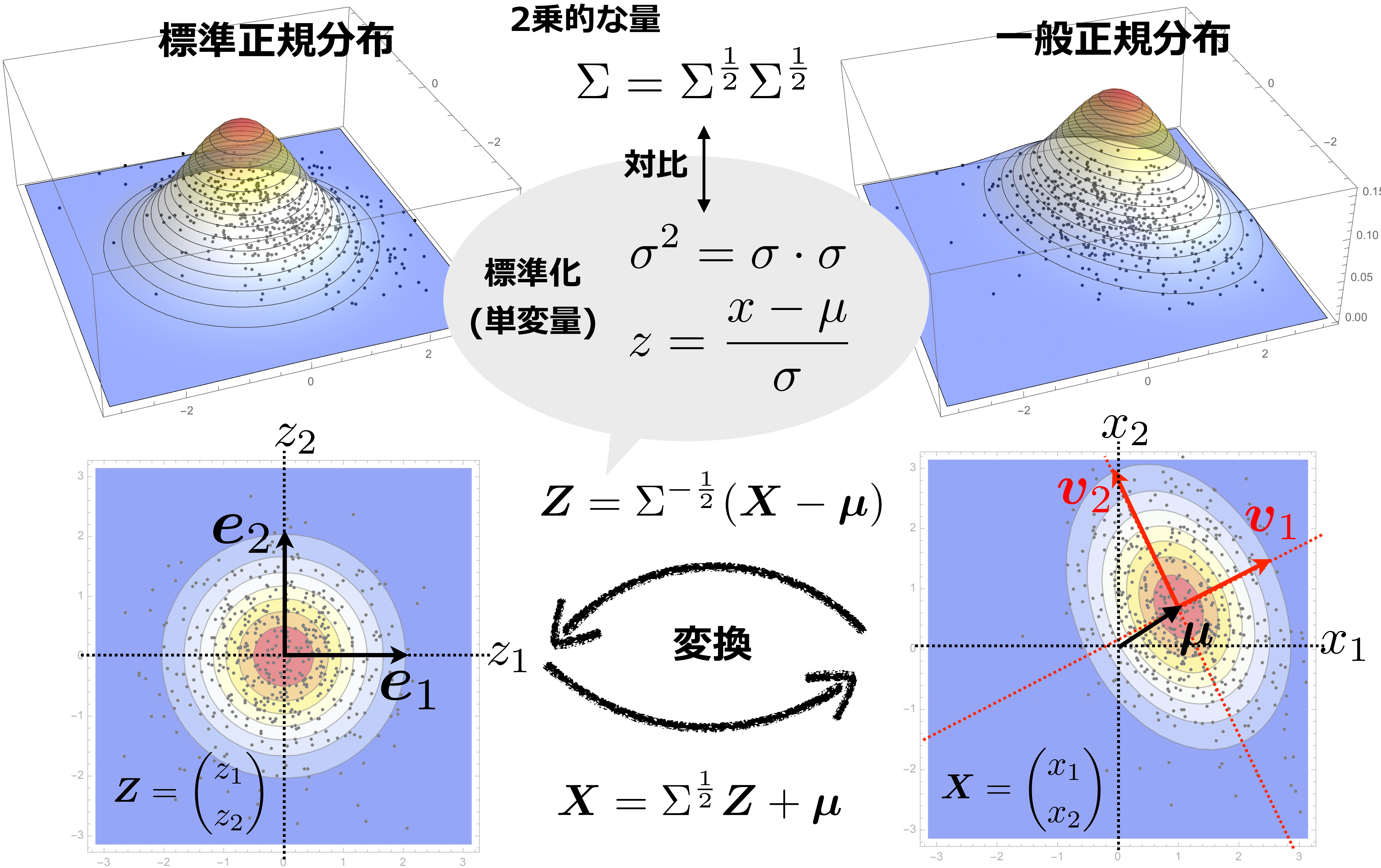




知りたいこと
(教科書 p.4)

- 検査値 (x_1, x_2) から第１群か第２群か判別できるか？
- 判別できるとすればその精度はどれくらいか？
- 例えば $x_1 = 0.7, x_2 = -1.1$ ならどう判別されるか？

基底の変換～分散共分散行列～正規分布

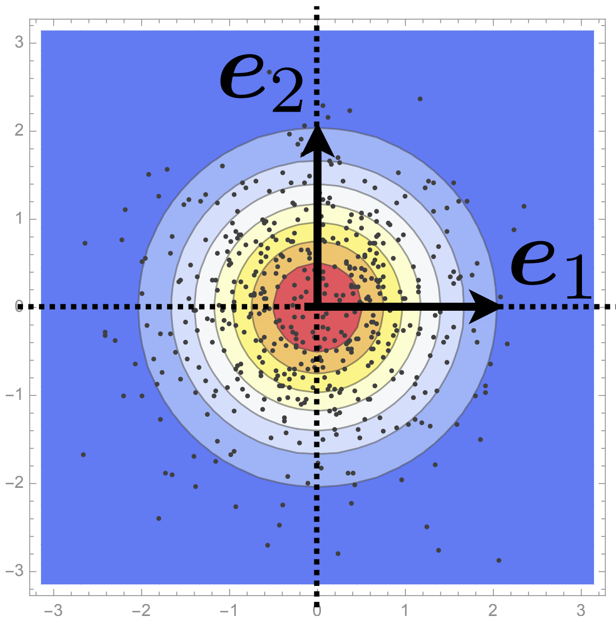


前回資料

標準化変換 (基底変換+平行移動)

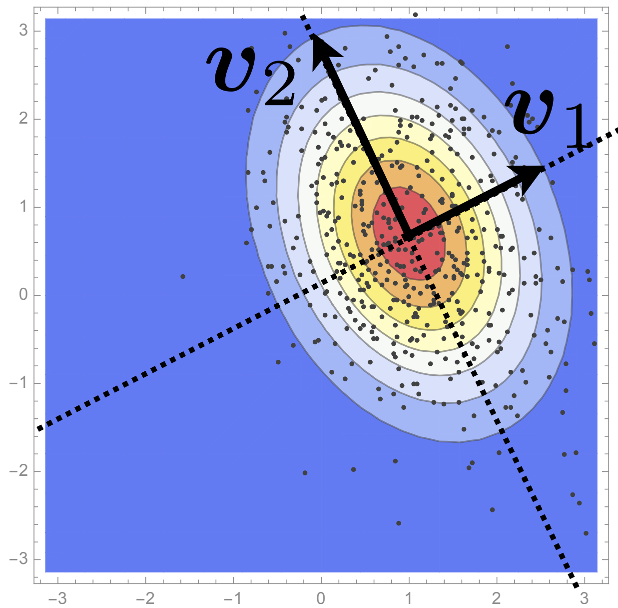
等確率面は球形

2D:等高線が真円形



等確率面は楕円体形

2D:等高線が楕円形



参考：確率変数の標準化 (単変量)

$$\sigma^2 = \sigma \cdot \sigma$$

$$Z := \frac{X - \mu}{\sigma}$$

Zの平均=0
Zの分散=1

確率変数の標準化 (多変量)

$$\boldsymbol{Z} := \boldsymbol{\Sigma}^{-\frac{1}{2}}(\boldsymbol{X} - \boldsymbol{\mu})$$

$$\boldsymbol{\Sigma} = \boldsymbol{\Sigma}^{\frac{1}{2}}\boldsymbol{\Sigma}^{\frac{1}{2}}$$

Zの平均ベクトル=0
Zの分散共分散行列=単位行列

分散共分散行列

多変量の分布の"散らばり"具合 (2次の統計量) を司る

$$\Sigma = \begin{bmatrix} \sigma_{x_1,x_1} & \sigma_{x_2,x_1} & \cdots & \sigma_{x_p,x_1} \\ \sigma_{x_1,x_2} & \sigma_{x_2,x_2} & \cdots & \sigma_{x_p,x_2} \\ \vdots & & & \\ \sigma_{x_1,x_p} & \sigma_{x_2,x_p} & \cdots & \sigma_{x_p,x_p} \end{bmatrix}$$

i, j 要素は x_i と x_j の共分散　σ_{x_i,x_j}

$$\sigma_{x_i,x_j} := \mathbb{E}\{(x_i - \mu_i)(x_j - \mu_j)\} \qquad \mu_i := \mathbb{E}\{x_i\}$$

→ 対角要素は分散　$\sigma_{x_i,x_i} = \mathbb{E}\{(x_i - \mu_i)^2\} = \sigma_{x_i}^2$

分散共分散行列：ベクトル・行列表記

$$\boldsymbol{X} := \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_p \end{bmatrix} \quad \boldsymbol{\mu} := \mathbb{E}\{\boldsymbol{X}\} = \begin{bmatrix} \mathbb{E}\{x_1\} \\ \mathbb{E}\{x_2\} \\ \vdots \\ \mathbb{E}\{x_p\} \end{bmatrix} = \begin{bmatrix} \mu_1 \\ \mu_2 \\ \vdots \\ \mu_p \end{bmatrix}$$ とベクトル表記すれば

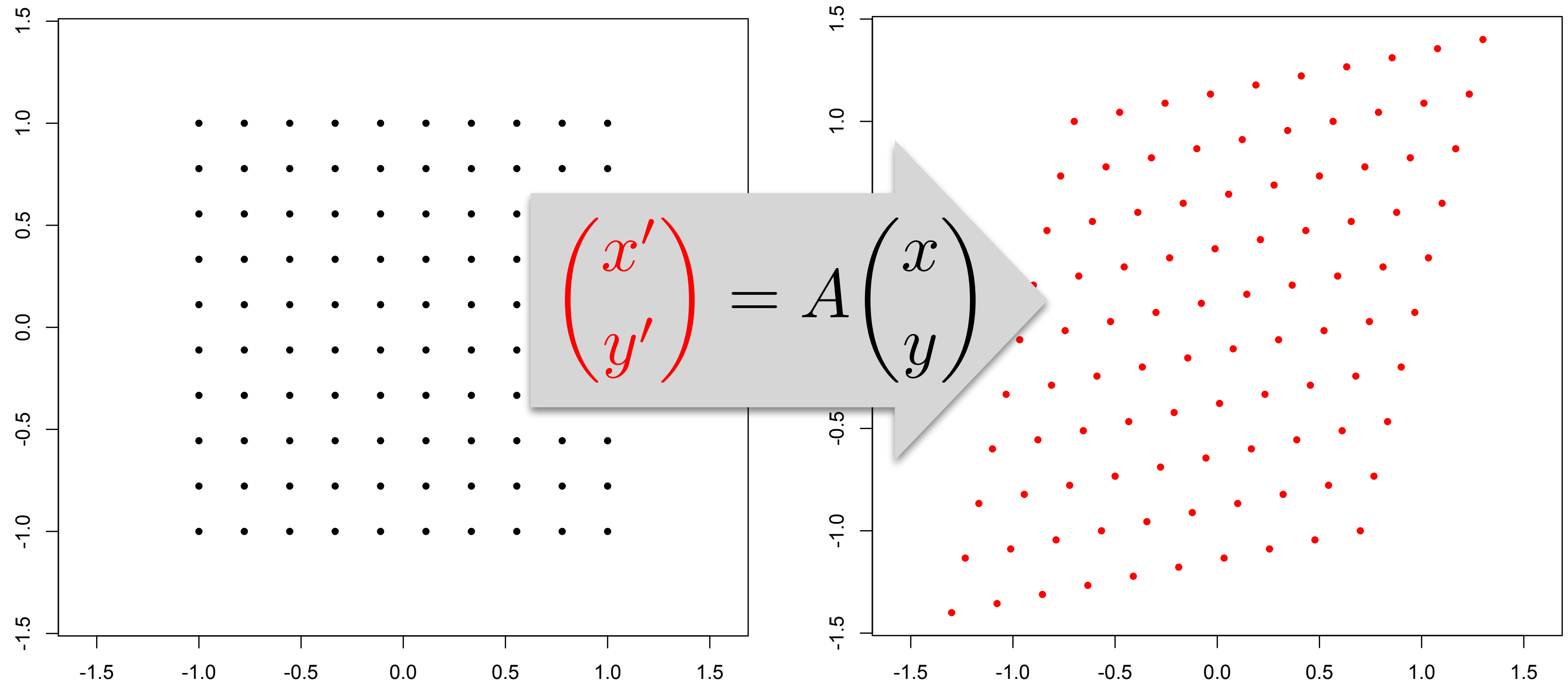
$$\Sigma = \mathbb{E}\{(\boldsymbol{X}-\boldsymbol{\mu})(\boldsymbol{X}-\boldsymbol{\mu})'\} = \mathbb{E}\left\{ \begin{bmatrix} x_1-\mu_1 \\ x_2-\mu_2 \\ \vdots \\ x_p-\mu_p \end{bmatrix} \begin{bmatrix} x_1-\mu_1 & x_2-\mu_2 & \cdots & x_p-\mu_p \end{bmatrix} \right\}$$

$$= \mathbb{E}\left\{ \begin{bmatrix} (x_1-\mu_1)(x_1-\mu_1) & (x_1-\mu_1)(x_2-\mu_2) & \cdots & (x_1-\mu_1)(x_p-\mu_p) \\ (x_2-\mu_2)(x_1-\mu_1) & (x_2-\mu_2)(x_2-\mu_2) & \cdots & (x_2-\mu_2)(x_p-\mu_p) \\ \vdots & & & \\ (x_p-\mu_p)(x_1-\mu_1) & (x_p-\mu_p)(x_2-\mu_2) & \cdots & (x_p-\mu_p)(x_p-\mu_p) \end{bmatrix} \right\}$$

固有値と固有ベクトル

$$A = \begin{bmatrix} 1.0 & 0.3 \\ 0.2 & 1.2 \end{bmatrix}$$

点がAによる線形写像で
どこへ移るか観察



$$\begin{pmatrix} x' \\ y' \end{pmatrix} = A \begin{pmatrix} x \\ y \end{pmatrix}$$

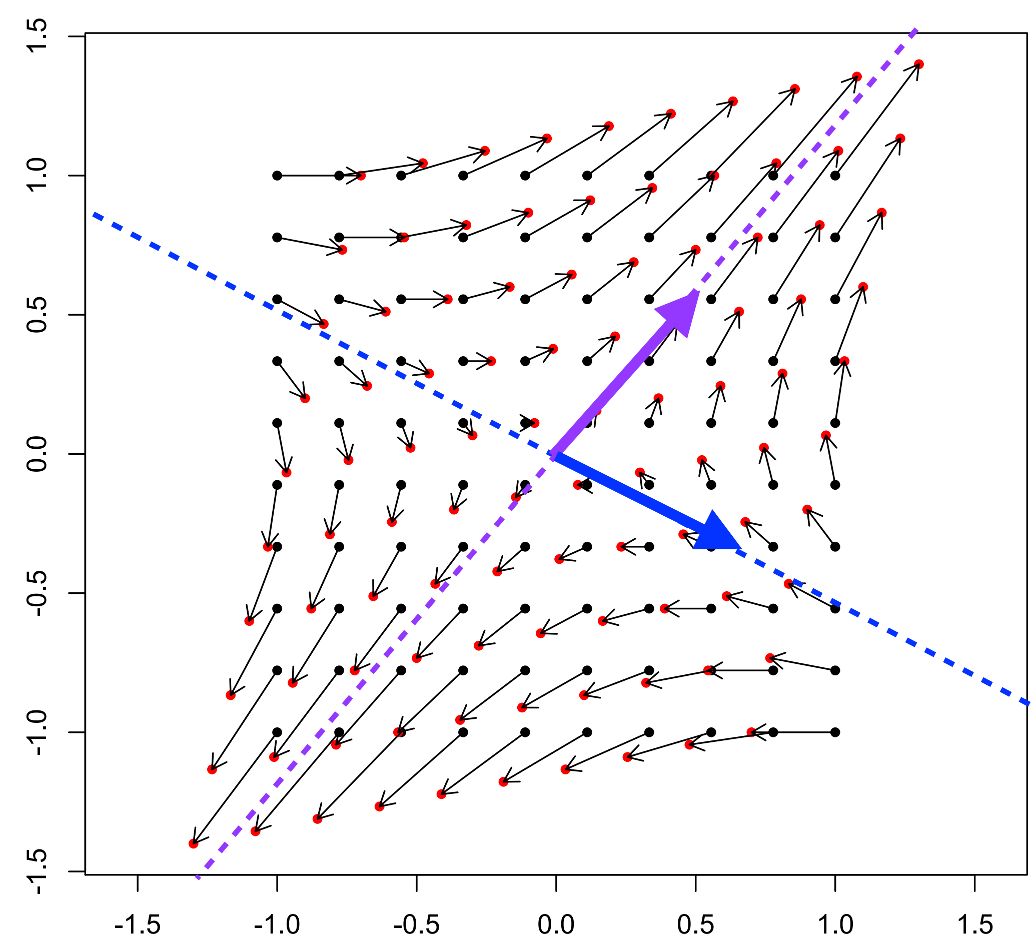
$$A \begin{pmatrix} v_1 \\ v_2 \end{pmatrix} = \lambda_1 \begin{pmatrix} v_1 \\ v_2 \end{pmatrix}$$

$$A \begin{pmatrix} u_1 \\ u_2 \end{pmatrix} = \lambda_2 \begin{pmatrix} u_1 \\ u_2 \end{pmatrix}$$

方向が変わらない
ところが存在
(倍率のみ)

n個あるとき対角化可能

$$A \begin{bmatrix} v_1 & u_1 \\ v_2 & u_2 \end{bmatrix} = \begin{bmatrix} v_1 & u_1 \\ v_2 & u_2 \end{bmatrix} \begin{bmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{bmatrix}$$



$$A\begin{bmatrix} v_1 & u_1 \\ v_2 & u_2 \end{bmatrix} = \begin{bmatrix} v_1 & u_1 \\ v_2 & u_2 \end{bmatrix}\begin{bmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{bmatrix}$$

対角化

$$\begin{bmatrix} v_1 & u_1 \\ v_2 & u_2 \end{bmatrix}^{-1} A\begin{bmatrix} v_1 & u_1 \\ v_2 & u_2 \end{bmatrix} = \begin{bmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{bmatrix}$$

固有値分解

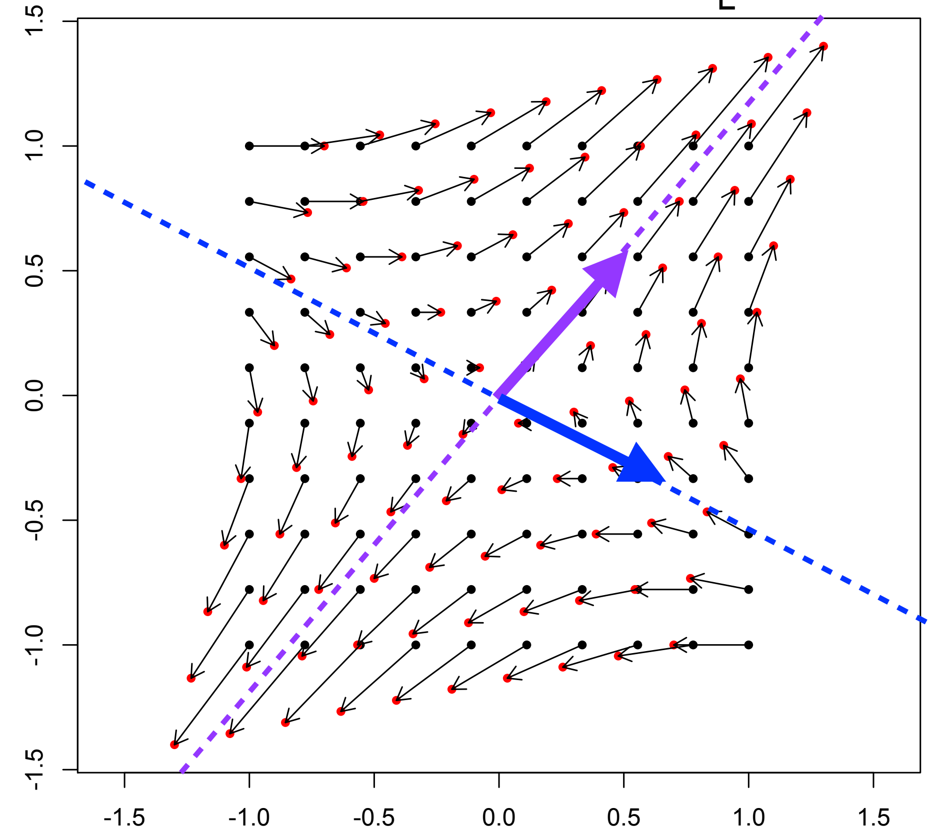
$$A = \begin{bmatrix} v_1 & u_1 \\ v_2 & u_2 \end{bmatrix}\begin{bmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{bmatrix}\begin{bmatrix} v_1 & u_1 \\ v_2 & u_2 \end{bmatrix}^{-1}$$

対称行列なら固有ベクトルは直交

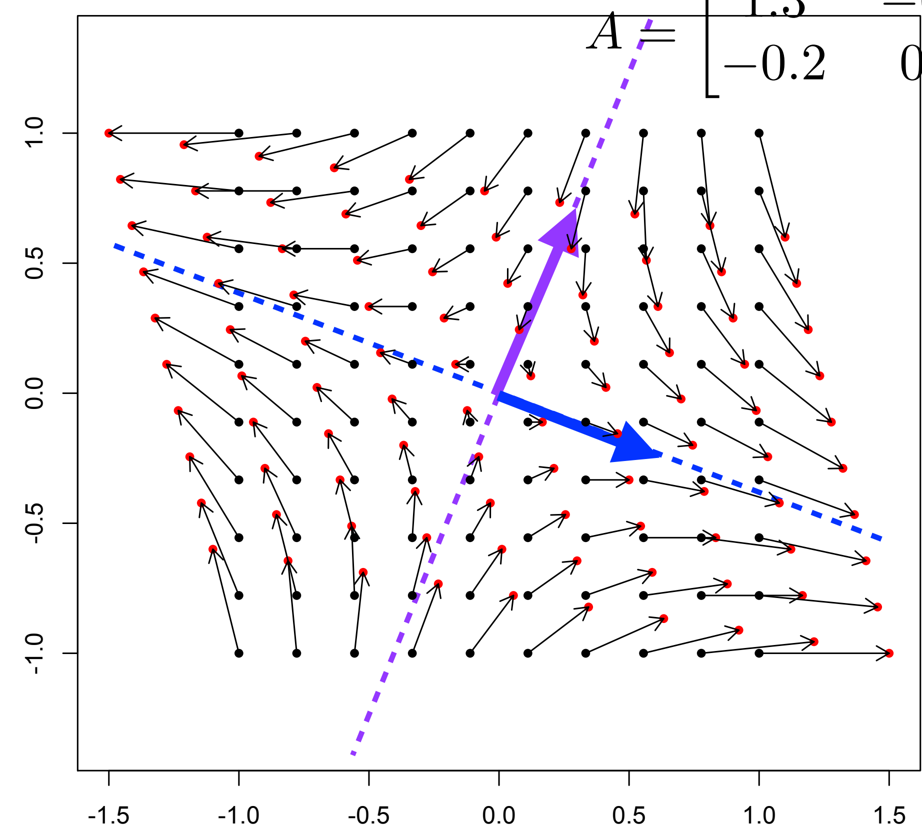
$$\begin{bmatrix} v_1 & u_1 \\ v_2 & u_2 \end{bmatrix}' \begin{bmatrix} v_1 & u_1 \\ v_2 & u_2 \end{bmatrix} = \begin{bmatrix} v_1 & v_2 \\ u_1 & u_2 \end{bmatrix}\begin{bmatrix} v_1 & u_1 \\ v_2 & u_2 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

$$A = \begin{bmatrix} v_1 & u_1 \\ v_2 & u_2 \end{bmatrix}\begin{bmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{bmatrix}\begin{bmatrix} v_1 & u_1 \\ v_2 & u_2 \end{bmatrix}'$$

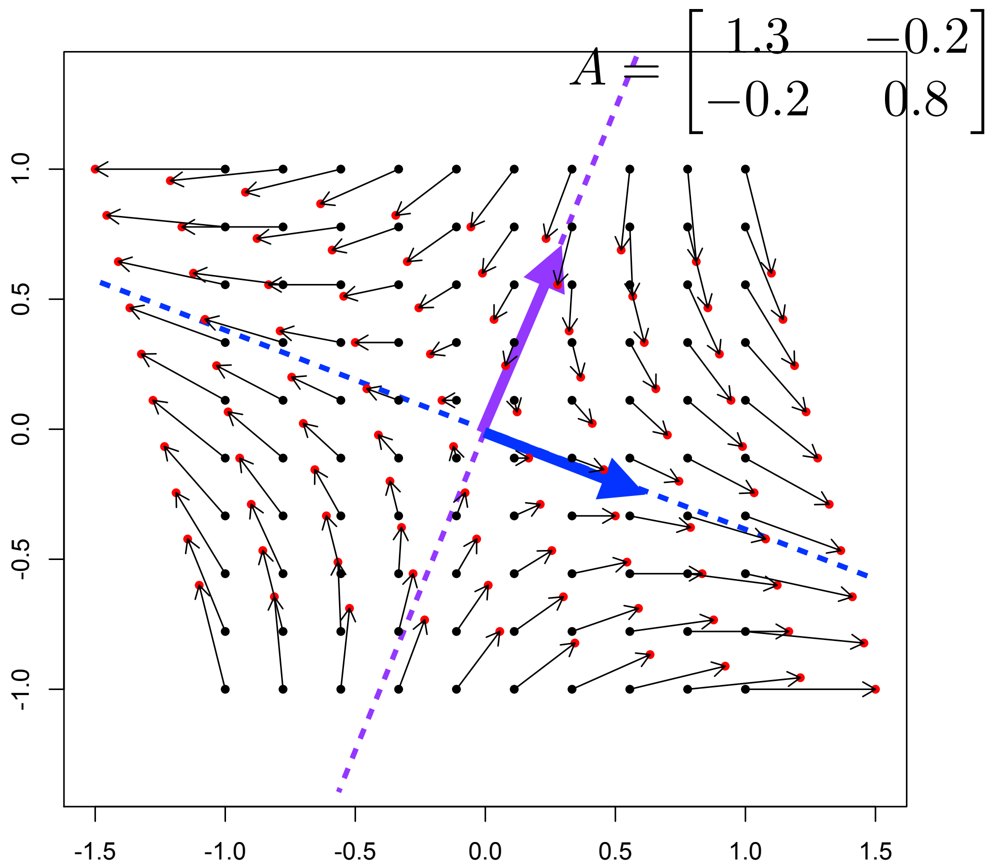
$$A = \begin{bmatrix} 1.0 & 0.3 \\ 0.2 & 1.2 \end{bmatrix}$$



$$A = \begin{bmatrix} 1.3 & -0.2 \\ -0.2 & 0.8 \end{bmatrix}$$



$$\begin{bmatrix} v_1 & u_1 \\ v_2 & u_2 \end{bmatrix} \qquad \begin{bmatrix} 1 \\ 0 \end{bmatrix} \to \begin{bmatrix} v_1 \\ v_2 \end{bmatrix} \qquad \begin{bmatrix} 0 \\ 1 \end{bmatrix} \to \begin{bmatrix} u_1 \\ u_2 \end{bmatrix}$$



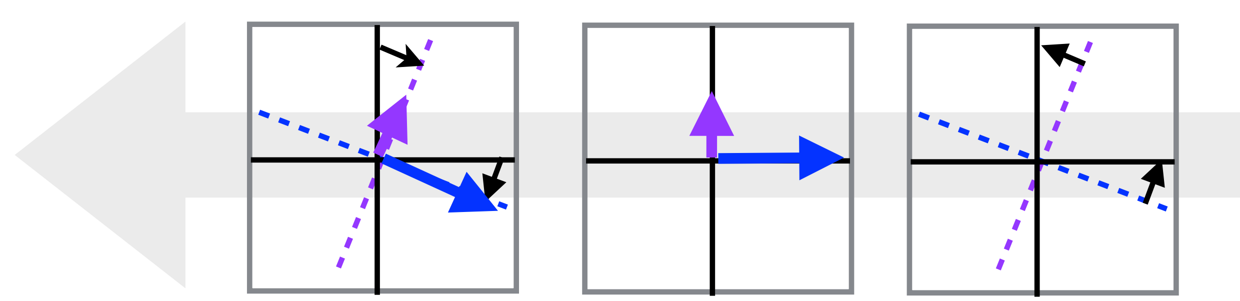
$$A = \begin{bmatrix} v_1 & u_1 \\ v_2 & u_2 \end{bmatrix} \begin{bmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{bmatrix} \underbrace{\begin{bmatrix} v_1 & u_1 \\ v_2 & u_2 \end{bmatrix}'}$$

$$= \begin{bmatrix} v_1 & u_1 \\ v_2 & u_2 \end{bmatrix}^{-1}$$

固有値分解の直感的意味

$$A \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \underbrace{\begin{bmatrix} v_1 & u_1 \\ v_2 & u_2 \end{bmatrix}}_{\text{基底変換}} \underbrace{\begin{bmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{bmatrix}}_{\text{倍率}} \underbrace{\begin{bmatrix} v_1 & u_1 \\ v_2 & u_2 \end{bmatrix}'}_{\text{逆基底変換}} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$$

$\to \begin{bmatrix} v_1 \\ v_2 \end{bmatrix}, \begin{bmatrix} u_1 \\ u_2 \end{bmatrix}$ $\qquad\qquad$ $\to \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \end{bmatrix}$



$$A \begin{bmatrix} v_1 & u_1 \\ v_2 & u_2 \end{bmatrix} = \begin{bmatrix} v_1 & u_1 \\ v_2 & u_2 \end{bmatrix} \begin{bmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{bmatrix}$$

対角化

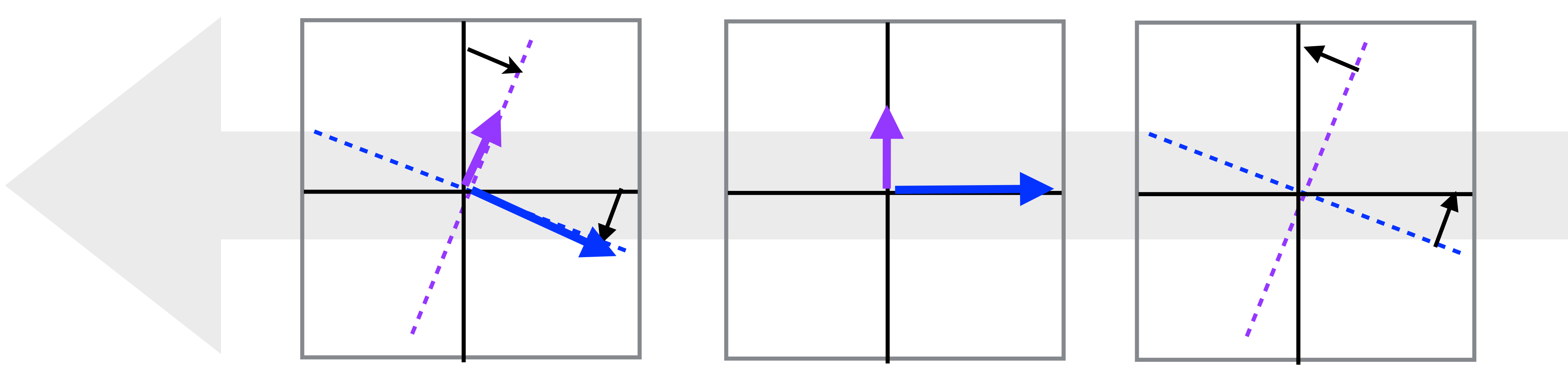
$$\begin{bmatrix} v_1 & u_1 \\ v_2 & u_2 \end{bmatrix}^{-1} A \begin{bmatrix} v_1 & u_1 \\ v_2 & u_2 \end{bmatrix} = \begin{bmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{bmatrix}$$

固有値分解

$$A = \begin{bmatrix} v_1 & u_1 \\ v_2 & u_2 \end{bmatrix} \begin{bmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{bmatrix} \begin{bmatrix} v_1 & u_1 \\ v_2 & u_2 \end{bmatrix}^{-1}$$

$$A\begin{bmatrix}x_1\\x_2\end{bmatrix}=\underbrace{\begin{bmatrix}v_1 & u_1\\v_2 & u_2\end{bmatrix}}_{\text{基底変換}}\underbrace{\begin{bmatrix}\color{red}{\lambda_1} & 0\\0 & \color{purple}{\lambda_2}\end{bmatrix}}_{\text{倍率}}\underbrace{\begin{bmatrix}v_1 & u_1\\v_2 & u_2\end{bmatrix}'}_{\text{逆基底変換}}\begin{bmatrix}x_1\\x_2\end{bmatrix}$$

$$\rightarrow\begin{bmatrix}v_1\\v_2\end{bmatrix},\begin{bmatrix}u_1\\u_2\end{bmatrix}\qquad\qquad\rightarrow\begin{bmatrix}1\\0\end{bmatrix},\begin{bmatrix}0\\1\end{bmatrix}$$



分散共分散行列の固有構造 (スペクトル分解)

$$\Sigma = P \Lambda P'$$

$$\begin{bmatrix} \sigma_{x_1,x_1} & \sigma_{x_2,x_1} & \cdots & \sigma_{x_p,x_1} \\ \sigma_{x_1,x_2} & \sigma_{x_2,x_2} & \cdots & \sigma_{x_p,x_2} \\ \vdots & & & \\ \sigma_{x_1,x_p} & \sigma_{x_2,x_p} & \cdots & \sigma_{x_p,x_p} \end{bmatrix} = \underbrace{\begin{bmatrix} p_{11} & p_{12} & \cdots & p_{1p} \\ p_{21} & p_{22} & \cdots & p_{p2} \\ \vdots & & & \\ p_{p1} & p_{p2} & \cdots & p_{pp} \end{bmatrix}}_{\text{固有ベクトル}} \underbrace{\begin{bmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \vdots & & & \\ 0 & 0 & \cdots & \lambda_p \end{bmatrix}}_{\text{固有値}} \underbrace{\begin{bmatrix} p_{11} & p_{21} & \cdots & p_{p1} \\ p_{12} & p_{22} & \cdots & p_{p2} \\ \vdots & & & \\ p_{1p} & p_{2p} & \cdots & p_{pp} \end{bmatrix}}_{\text{固有ベクトル(転置)}}$$

$$= \lambda_1 \begin{bmatrix} p_{11} \\ p_{21} \\ \vdots \\ p_{p1} \end{bmatrix} \begin{bmatrix} p_{11} & p_{21} & \cdots & p_{p1} \end{bmatrix} + \lambda_2 \begin{bmatrix} p_{12} \\ p_{22} \\ \vdots \\ p_{p2} \end{bmatrix} \begin{bmatrix} p_{12} & p_{22} & \cdots & p_{p2} \end{bmatrix} + \cdots + \lambda_p \begin{bmatrix} p_{1p} \\ p_{2p} \\ \vdots \\ p_{pp} \end{bmatrix} \begin{bmatrix} p_{1p} & p_{2p} & \cdots & p_{pp} \end{bmatrix}$$

- 対称行列なので固有ベクトルは直交する
- 正定値行列なので正定値な平方根行列が唯一存在

分散共分散行列の平方根行列

$$\Sigma^{\frac{1}{2}} = P \quad \Lambda^{\frac{1}{2}} \quad P'$$

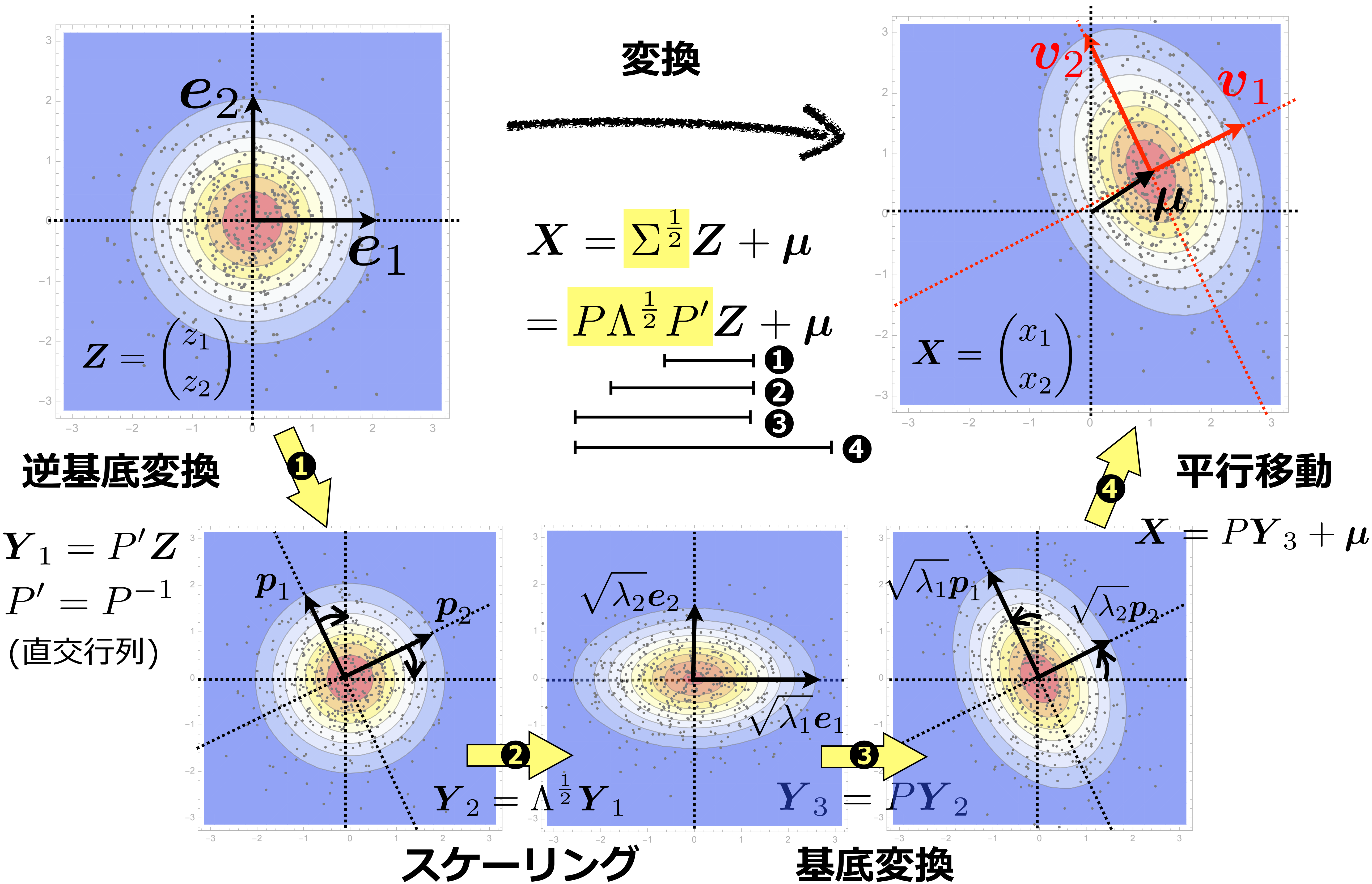
$$\begin{bmatrix} s_{11} & s_{12} & \cdots & s_{1p} \\ s_{21} & s_{22} & \cdots & s_{2p} \\ \vdots & & & \\ s_{p1} & s_{p2} & \cdots & s_{pp} \end{bmatrix} = \begin{bmatrix} p_{11} & p_{12} & \cdots & p_{1p} \\ p_{21} & p_{22} & \cdots & p_{p2} \\ \vdots & & & \\ p_{p1} & p_{p2} & \cdots & p_{pp} \end{bmatrix} \begin{bmatrix} \sqrt{\lambda_1} & 0 & \cdots & 0 \\ 0 & \sqrt{\lambda_2} & \cdots & 0 \\ \vdots & & & \\ 0 & 0 & \cdots & \sqrt{\lambda_p} \end{bmatrix} \begin{bmatrix} p_{11} & p_{21} & \cdots & p_{p1} \\ p_{12} & p_{22} & \cdots & p_{p2} \\ \vdots & & & \\ p_{1p} & p_{2p} & \cdots & p_{pp} \end{bmatrix}$$

- 二乗したらΣ

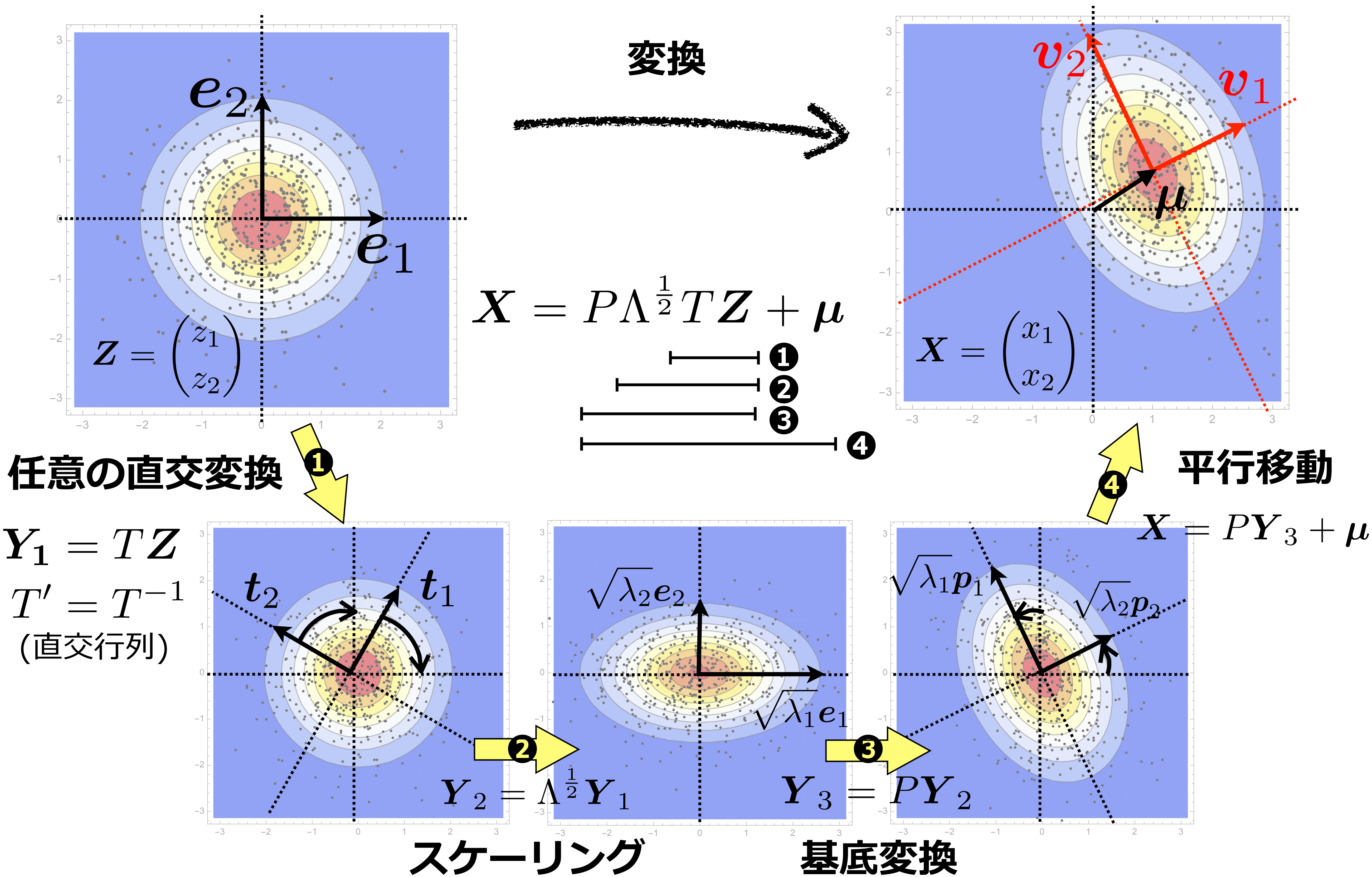
  $$\Sigma^{1/2}\Sigma^{1/2} = P\Lambda^{1/2}P'P\Lambda^{1/2}P' = P\Lambda^{1/2}\Lambda^{1/2}P' = \Sigma$$

- 固有ベクトルはΣと同じ
- 正定値行列

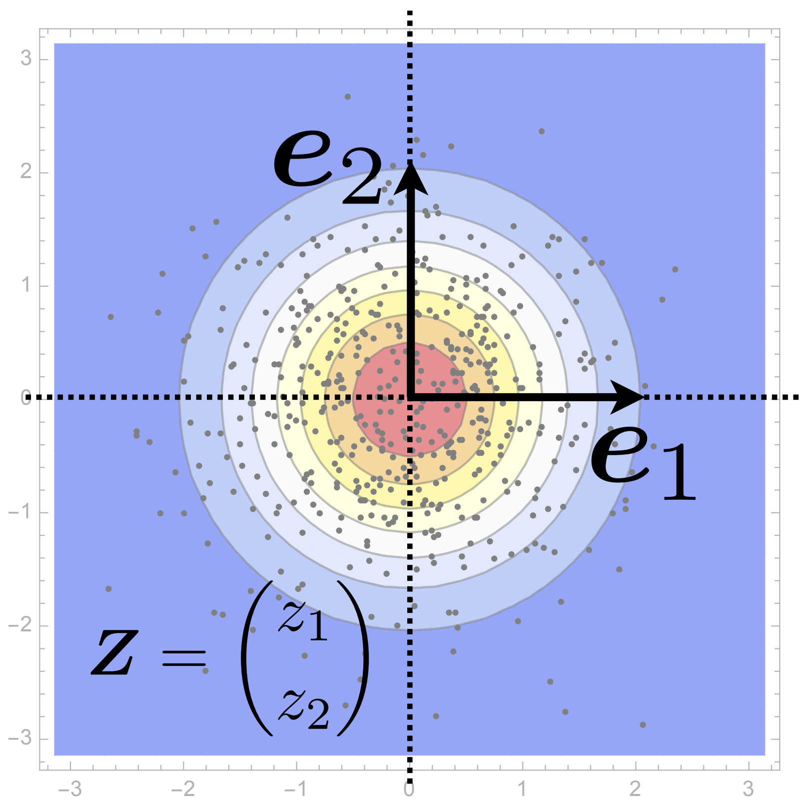
基底変換と固有値・固有ベクトル



注意：実は直交変換分の自由度がある



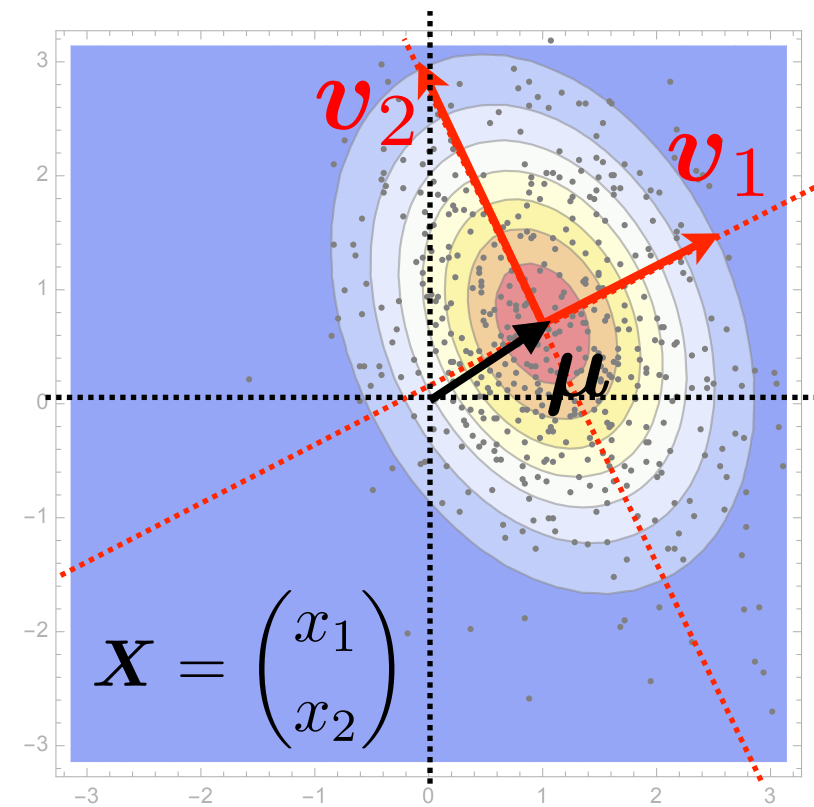
平方根行列だと対称+固有構造不変かつ逆変換も奇麗



$$Z = \Sigma^{-\frac{1}{2}}(X - \mu)$$

変換

$$X = \Sigma^{\frac{1}{2}} Z + \mu$$



いずれにしても言える事

$$v_1 = \sqrt{\lambda_1} p_1$$

$$v_2 = \sqrt{\lambda_2} p_2$$

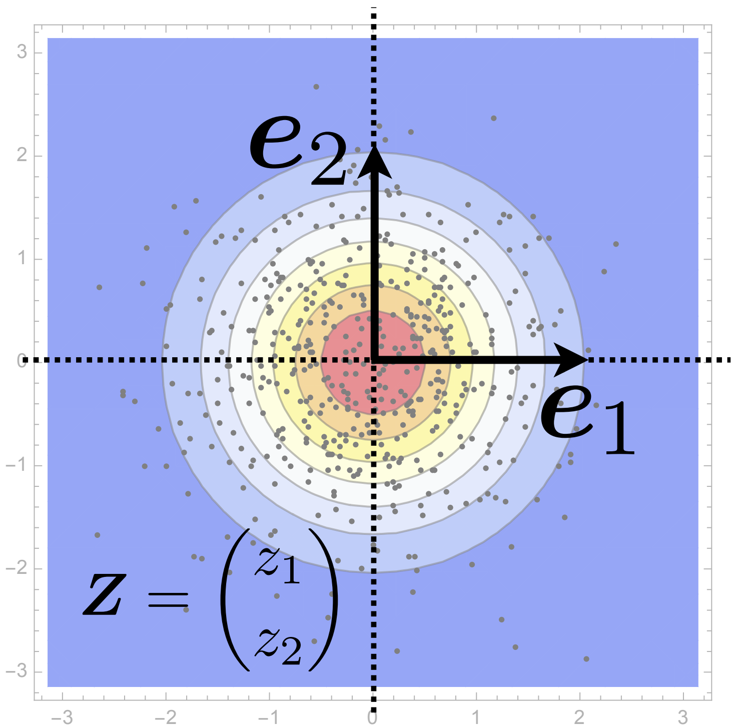
λ_i　Σ の固有値

p_i　Σ の固有ベクトル

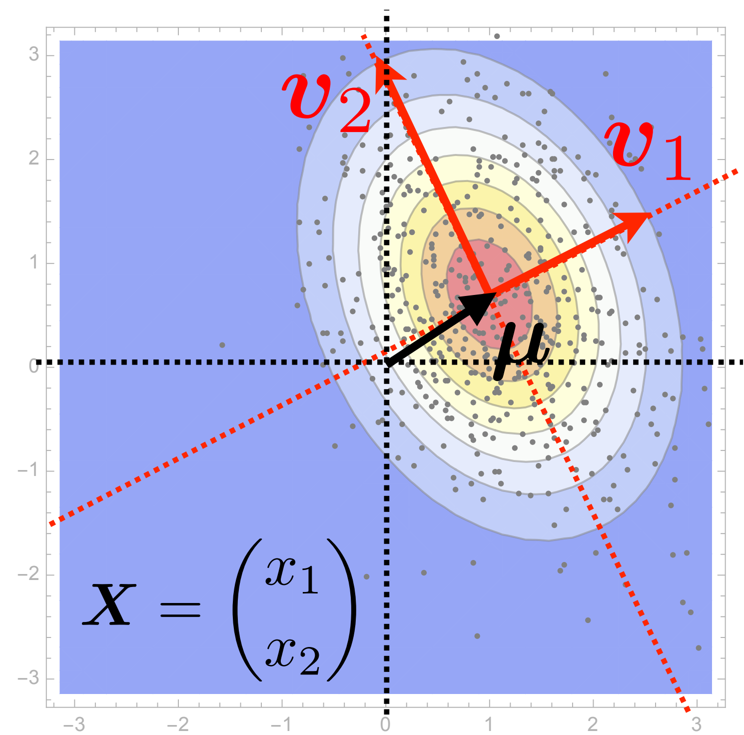
同確率面(2Dでは等高線)の楕円体の長軸・短軸は
分散共分散行列の固有値・固有ベクトルで決まる！

等確率面の楕円体の式は？

逆変換した後に原点から等距離な点の集合



$$Z = \Sigma^{-\frac{1}{2}}(\boldsymbol{X} - \boldsymbol{\mu})$$



$$\|\boldsymbol{Z}\|^2 = \text{Const.} \quad \Leftrightarrow \quad \|\Sigma^{-\frac{1}{2}}(\boldsymbol{X} - \boldsymbol{\mu})\|^2 = \text{Const.}$$

マハラノビス距離

$$\|\Sigma^{-\frac{1}{2}}(\boldsymbol{X} - \boldsymbol{\mu})\|^2 = (\boldsymbol{X} - \boldsymbol{\mu})'\Sigma^{-1}(\boldsymbol{X} - \boldsymbol{\mu})$$

正規分布の密度も一定 $\quad p(\boldsymbol{x}) \propto \exp\left(-\frac{1}{2}(\boldsymbol{X} - \boldsymbol{\mu})'\Sigma^{-1}(\boldsymbol{X} - \boldsymbol{\mu})\right)$

前回資料

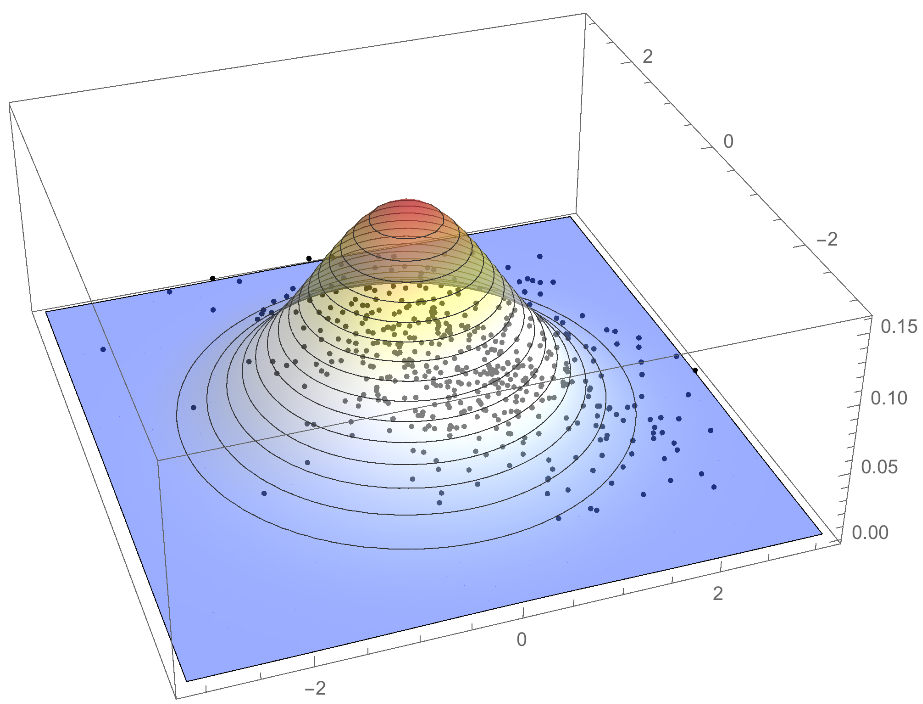
標準正規分布 vs 一般の多変量正規分布

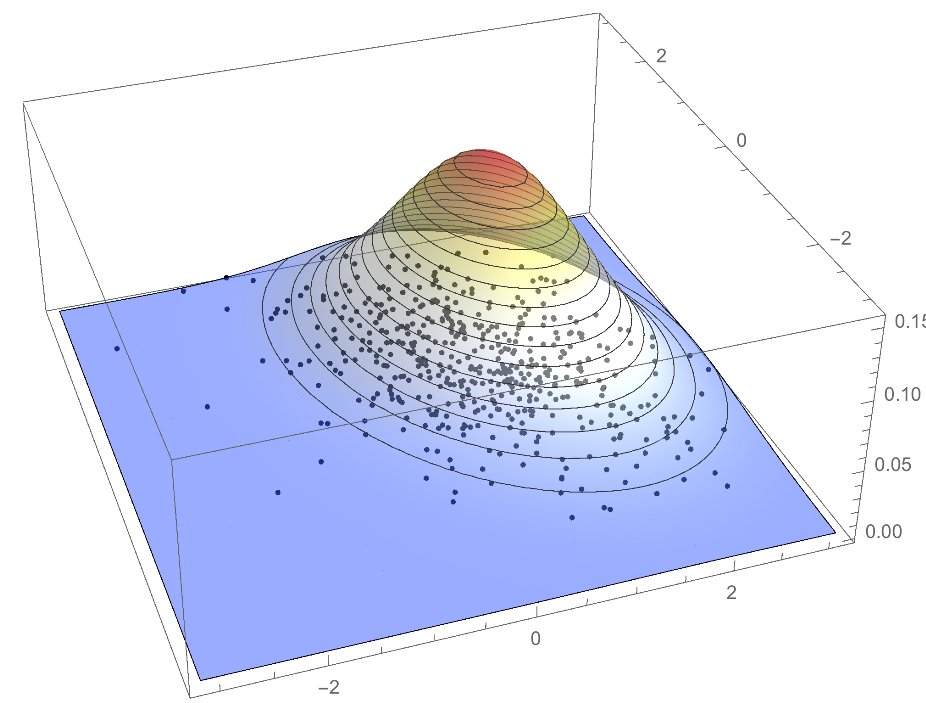
$$N(\boldsymbol{0}, \boldsymbol{I})$$

$$\frac{1}{(\sqrt{2\pi})^n} \exp\left(-\frac{1}{2}\boldsymbol{x}'\boldsymbol{x}\right)$$

$$N(\boldsymbol{\mu}, \boldsymbol{\Sigma})$$

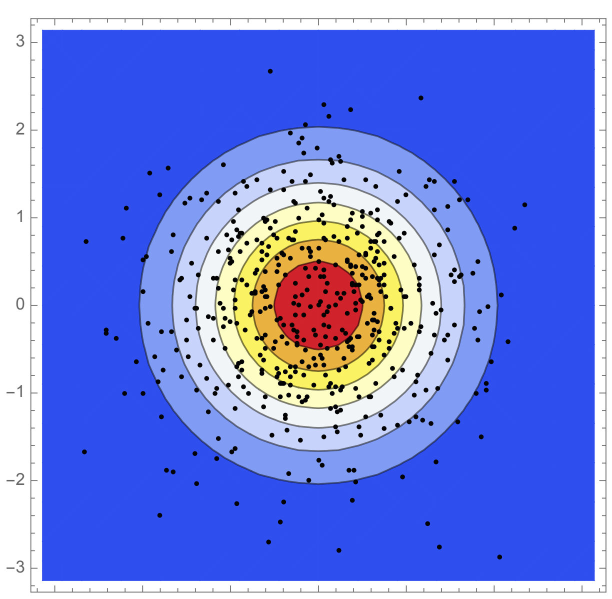
$$\frac{1}{(\sqrt{2\pi})^n \sqrt{\det(\boldsymbol{\Sigma})}} \exp\left(-\frac{1}{2}(\boldsymbol{x}-\boldsymbol{\mu})'\boldsymbol{\Sigma}^{-1}(\boldsymbol{x}-\boldsymbol{\mu})\right)$$





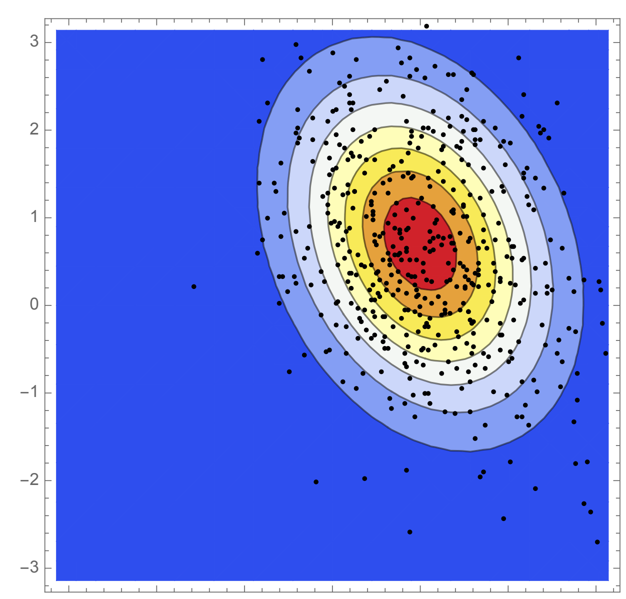
等確率面は球形

2D:等高線が真円形



等確率面は楕円体形

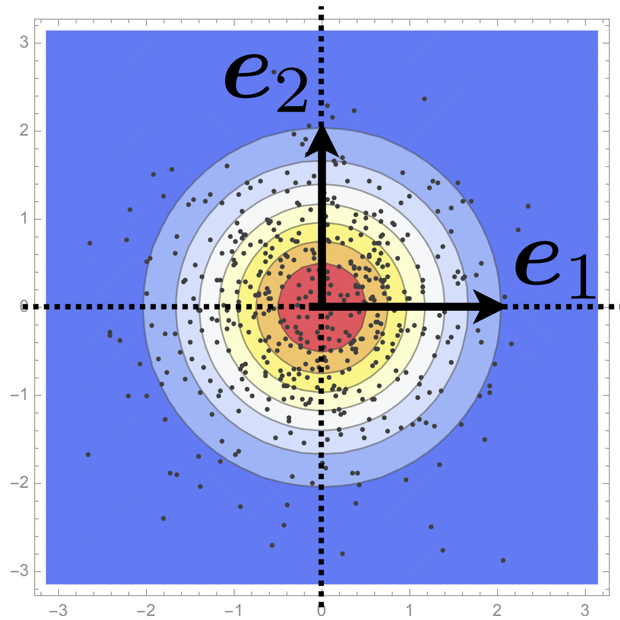
2D:等高線が楕円形



標準化変換 (基底変換+平行移動)

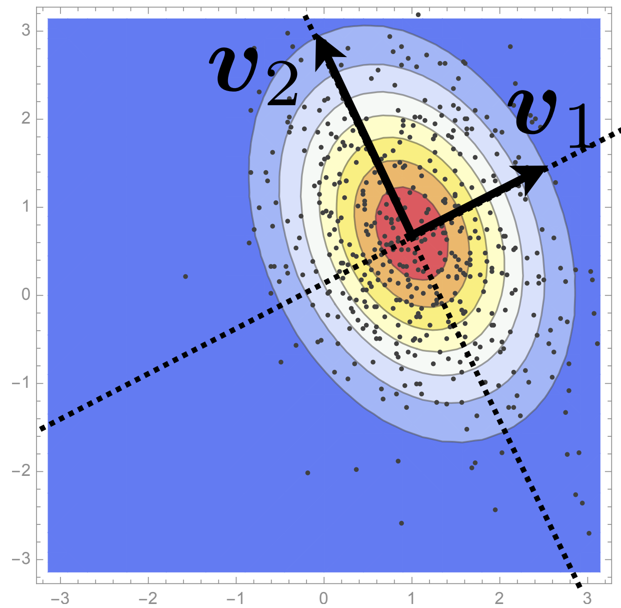
等確率面は球形

2D:等高線が真円形



等確率面は楕円体形

2D:等高線が楕円形



Whitening

"白色化"

参考：確率変数の標準化 (単変量)

$$\sigma^2 = \sigma \cdot \sigma$$

$$Z := \frac{X - \mu}{\sigma}$$

Zの平均=0

Zの分散=1

確率変数の標準化 (多変量)

$$\boldsymbol{Z} := \boldsymbol{\Sigma}^{-\frac{1}{2}}(\boldsymbol{X} - \boldsymbol{\mu})$$

$$\boldsymbol{\Sigma} = \boldsymbol{\Sigma}^{\frac{1}{2}}\boldsymbol{\Sigma}^{\frac{1}{2}}$$

Zの平均ベクトル=0

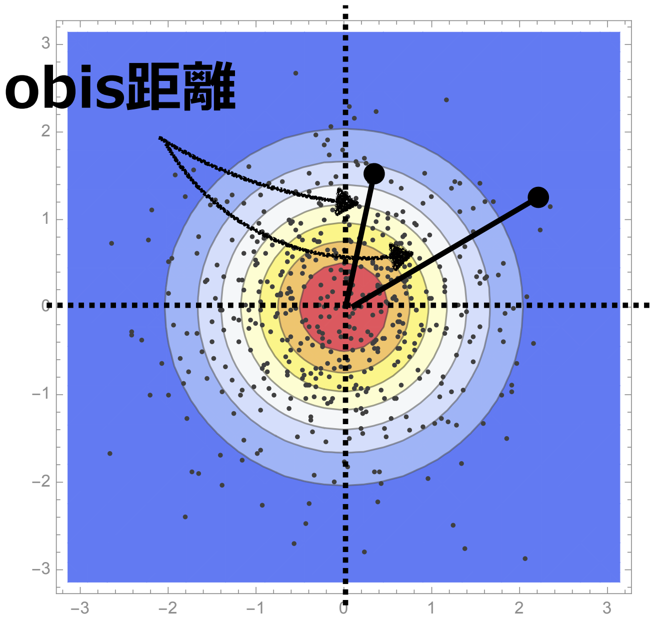
Zの分散共分散行列=単位行列

マハラノビス距離：分散共分散行列の二次形式

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

標準化した後のベクトルの大きさ

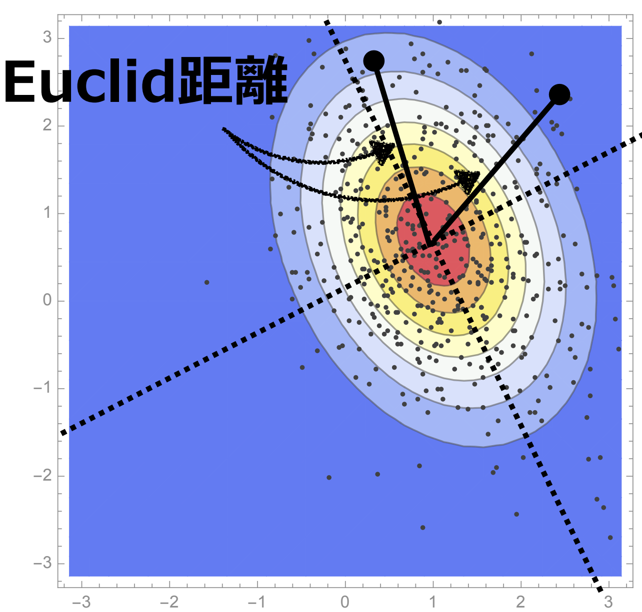
Mahalanobis距離



変数変換 (標準化)

$$\Sigma^{-1/2}(\boldsymbol{x}-\boldsymbol{\mu})$$

Euclid距離

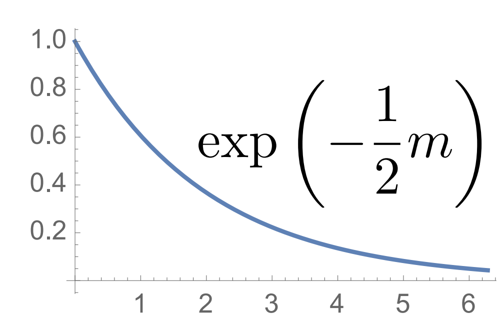


多変量正規分布の密度関数 $N(\boldsymbol{\mu}, \boldsymbol{\Sigma})$

$$\sqrt{\det(\Sigma)} \qquad \frac{1}{(\sqrt{2\pi})^n\sqrt{\det(\boldsymbol{\Sigma})}} \exp\left(-\frac{1}{2}(\boldsymbol{x}-\boldsymbol{\mu})'\boldsymbol{\Sigma}^{-1}(\boldsymbol{x}-\boldsymbol{\mu})\right)$$

変数変換の
Jacobian

積分=1にするための
正規化項(定数)



マハラノビス距離 m

平均からのマハラノビス距離

確率変数の標準化 (多変量)

$$\boldsymbol{\Sigma} = \boldsymbol{\Sigma}^{\frac{1}{2}}\boldsymbol{\Sigma}^{\frac{1}{2}}$$

$$\boldsymbol{Z} := \boldsymbol{\Sigma}^{-\frac{1}{2}}(\boldsymbol{X} - \boldsymbol{\mu})$$

Zの平均ベクトル=0
Zの分散共分散行列=単位行列

アナロジー

確率変数の標準化 (単変量)

$$\sigma^2 = \sigma \cdot \sigma$$

$$Z := \frac{X - \mu}{\sigma}$$

Zの平均=0
Zの分散=1

内積

$$\begin{aligned} \boldsymbol{z}'\boldsymbol{z} &= (\boldsymbol{\Sigma}^{-\frac{1}{2}}(\boldsymbol{x} - \boldsymbol{\mu}))'\boldsymbol{\Sigma}^{-\frac{1}{2}}(\boldsymbol{x} - \boldsymbol{\mu}) \\ &= (\boldsymbol{x} - \boldsymbol{\mu})'\underbrace{(\boldsymbol{\Sigma}^{-\frac{1}{2}})'}_{\text{対称}}\boldsymbol{\Sigma}^{-\frac{1}{2}}(\boldsymbol{x} - \boldsymbol{\mu}) \\ &= (\boldsymbol{x} - \boldsymbol{\mu})'\boldsymbol{\Sigma}^{-1}(\boldsymbol{x} - \boldsymbol{\mu}) \end{aligned}$$

$$N(\mu, \sigma^2)$$

$$\frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{1}{2}\left(\frac{x - \mu}{\sigma}\right)^2\right)$$

$$N(\boldsymbol{0}, \boldsymbol{I})$$

$$\frac{1}{(\sqrt{2\pi})^n} \exp\left(-\frac{1}{2}\boldsymbol{z}'\boldsymbol{z}\right)$$

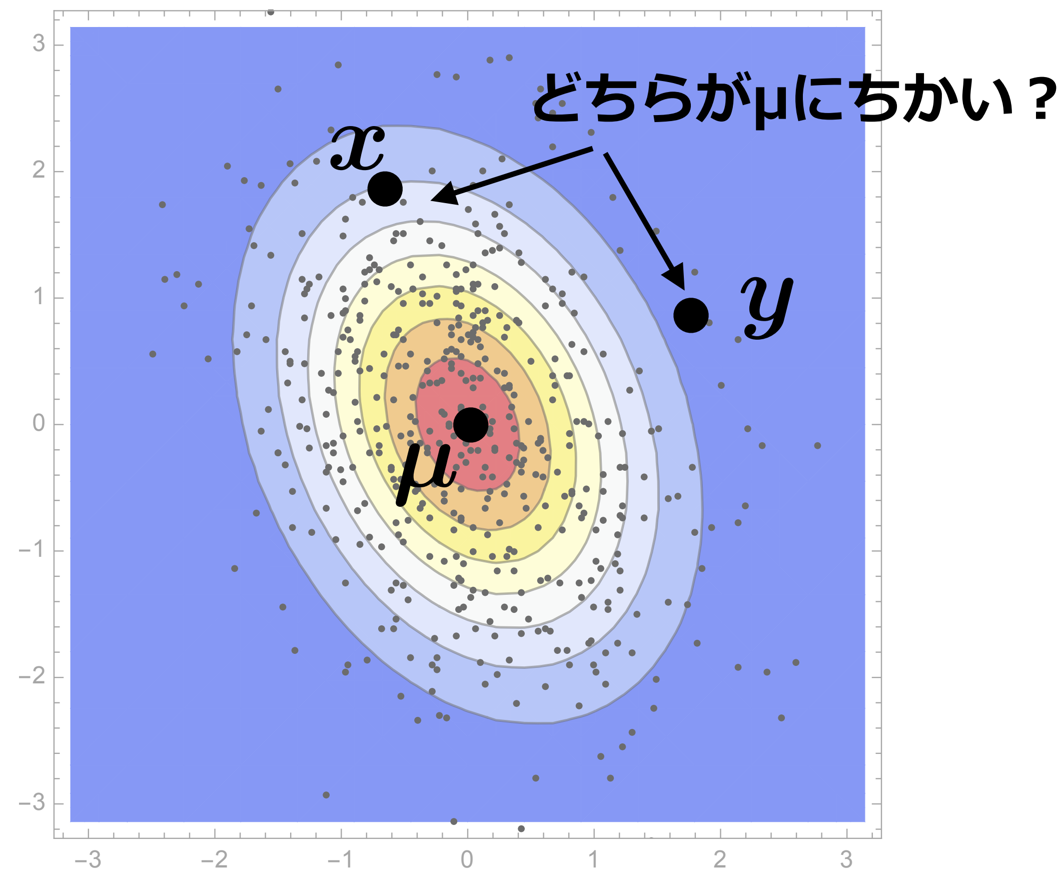
$$N(\boldsymbol{\mu}, \boldsymbol{\Sigma})$$

$$\frac{1}{(\sqrt{2\pi})^n\sqrt{\det(\boldsymbol{\Sigma})}} \exp\left(-\frac{1}{2}(\boldsymbol{x} - \boldsymbol{\mu})'\boldsymbol{\Sigma}^{-1}(\boldsymbol{x} - \boldsymbol{\mu})\right)$$

平均からのマハラノビス距離

右図で $(\boldsymbol{x}-\boldsymbol{\mu})^2 = (\boldsymbol{y}-\boldsymbol{\mu})^2$ **のとき**

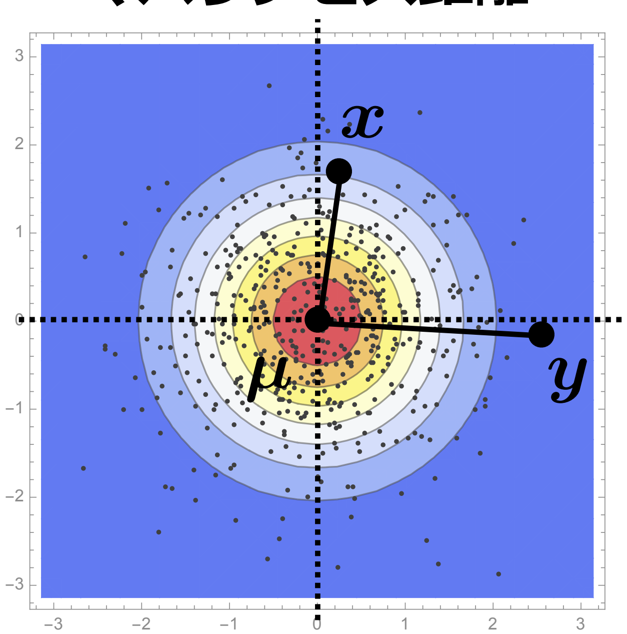
共分散を考慮して距離を測りたい。



右図の等確率の等高線がしめすとおり、正規分布の密度関数の値はexp(-マハラノビス距離)に比例するので、p(x) > p(y)

$$(\boldsymbol{x}-\boldsymbol{\mu})'\boldsymbol{\Sigma}^{-1}(\boldsymbol{x}-\boldsymbol{\mu}) < (\boldsymbol{y}-\boldsymbol{\mu})'\boldsymbol{\Sigma}^{-1}(\boldsymbol{y}-\boldsymbol{\mu})$$

マハラノビス距離

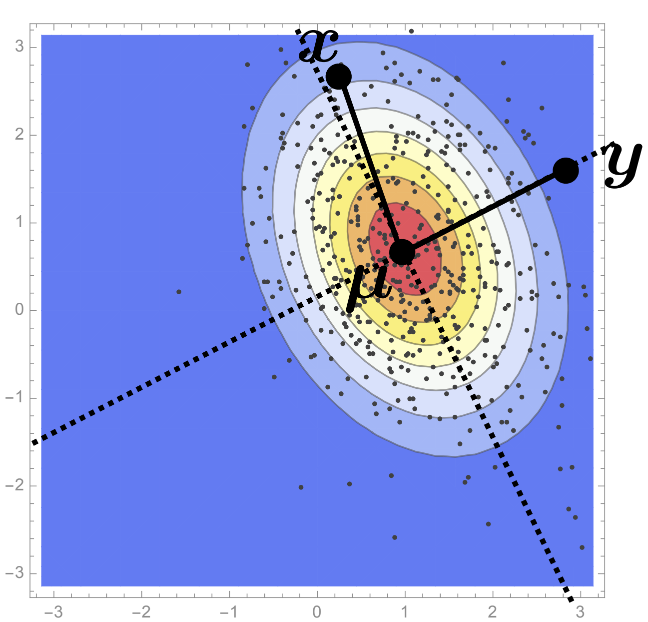


解釈：

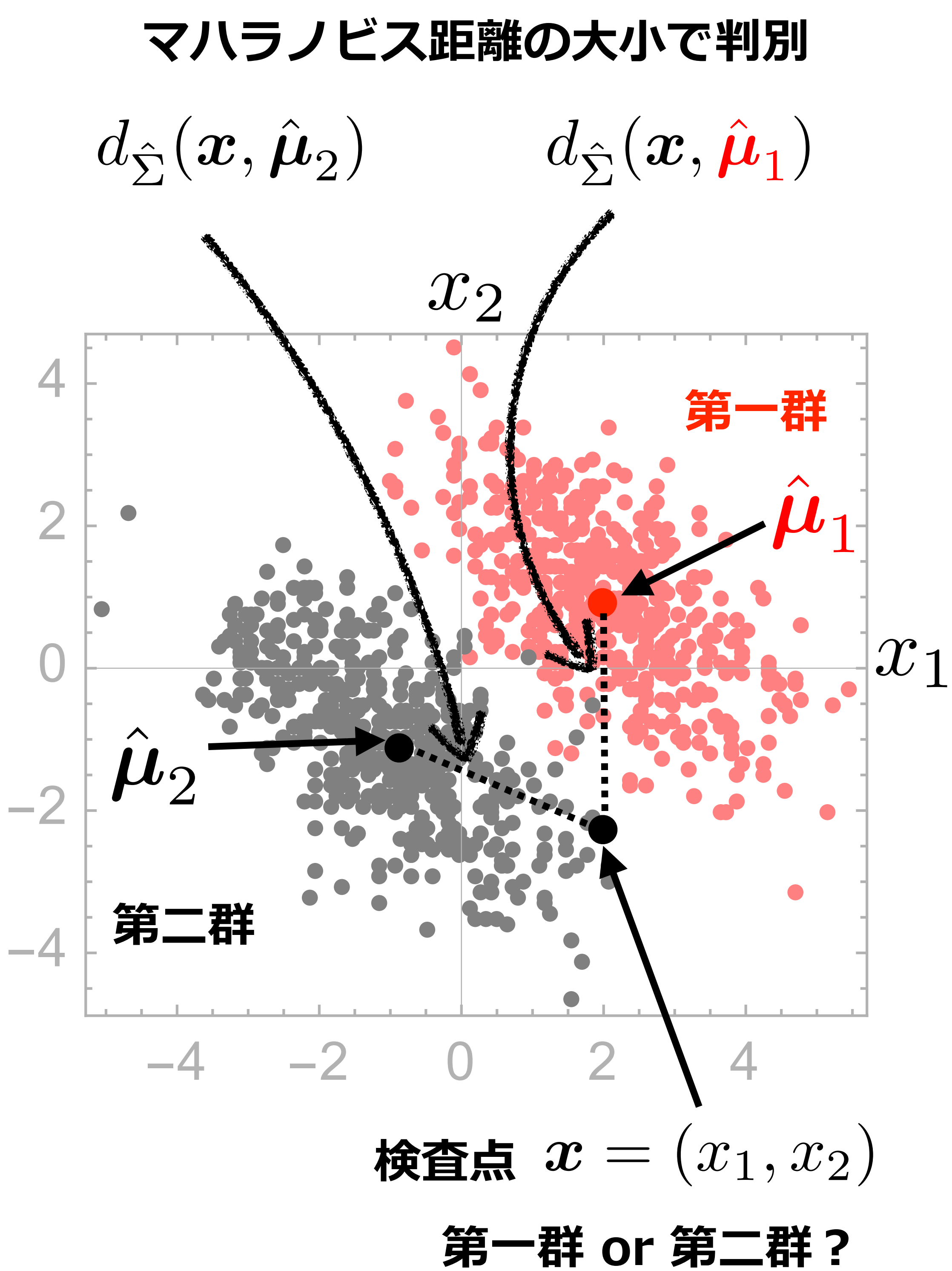
標準化で戻したあとの座標での原点からの二乗距離(ベクトルの長さの二乗)

$$\boldsymbol{Z} := \boldsymbol{\Sigma}^{-\frac{1}{2}}(\boldsymbol{X}-\boldsymbol{\mu})$$

ユークリッド距離



判別分析 (Fisherの線形判別分析)



1) 各々の群で平均ベクトルを標本平均で推定

$$\hat{\boldsymbol{\mu}}_1 = \frac{1}{n_1}\sum_{i=1}^{n_1} \boldsymbol{x}_i \qquad \hat{\boldsymbol{\mu}}_2 = \frac{1}{n_2}\sum_{i=1}^{n_2} \boldsymbol{x}_i$$

2) 分散共分散行列は第一群も第二群も同じと仮定して、
標本分散共分散行列で推定

$$\hat{\Sigma} = \frac{1}{(n_1-1)+(n_2-1)}\left\{\sum_{i=1}^{n_1}(\boldsymbol{x}_i - \hat{\boldsymbol{\mu}}_1)(\boldsymbol{x}_i - \hat{\boldsymbol{\mu}}_1)' + \sum_{i=1}^{n_2}(\boldsymbol{x}_i - \hat{\boldsymbol{\mu}}_2)(\boldsymbol{x}_i - \hat{\boldsymbol{\mu}}_2)'\right\}$$

3) 検査点 $\boldsymbol{x} = (x_1, x_2)$ の群はマハラノビス距離で判別

$$d_{\hat{\Sigma}}(\boldsymbol{x}, \hat{\boldsymbol{\mu}}_i) := (\boldsymbol{x} - \hat{\boldsymbol{\mu}}_i)'\hat{\Sigma}^{-1}(\boldsymbol{x} - \hat{\boldsymbol{\mu}}_i)$$

$$d_{\hat{\Sigma}}(\boldsymbol{x}, \hat{\boldsymbol{\mu}}_1) < d_{\hat{\Sigma}}(\boldsymbol{x}, \hat{\boldsymbol{\mu}}_2) \Rightarrow \text{第一群 else 第二群}$$

分散共分散行列の不偏推定

2) 分散共分散行列は第一群も第二群も同じと仮定して、標本分散共分散行列で推定 ⇒ 偏差を計算し群を区別せずに平均する！

$$\hat{\Sigma} = \frac{1}{(n_1 - 1) + (n_2 - 1)} \left\{ \sum_{i=1}^{n_1} (\boldsymbol{x}_i - \hat{\boldsymbol{\mu}}_1)(\boldsymbol{x}_i - \hat{\boldsymbol{\mu}}_1)' + \sum_{i=1}^{n_2} (\boldsymbol{x}_i - \hat{\boldsymbol{\mu}}_2)(\boldsymbol{x}_i - \hat{\boldsymbol{\mu}}_2)' \right\}$$

(不偏推定量にするのに)平均ベクトルに真値の代わりに標本からの推定量を代入していることで各々で自由度が一つ下がる

※ 正確な理屈は以前の配布資料の以下の2つの例を参照

母分散の推定量としての標本分散

$$\mathbb{E}\left\{ \frac{1}{n} \sum_{i=1}^{n} (X_i - \hat{X})^2 \right\} = \frac{n-1}{n} \sigma^2$$

p変量のときの回帰誤差の不偏推定量

$$\hat{\sigma}^2 := \frac{1}{n-p-1} \|\boldsymbol{y} - \boldsymbol{X}\hat{\boldsymbol{\beta}}\|^2$$

線形判別式：マハラノビス距離の差

3) 検査点 $x = (x_1, x_2)$ **の群はマハラノビス距離で判別**

$$d_{\hat{\Sigma}}(\boldsymbol{x}, \hat{\boldsymbol{\mu}}_i) := (\boldsymbol{x} - \hat{\boldsymbol{\mu}}_i)' \hat{\Sigma}^{-1} (\boldsymbol{x} - \hat{\boldsymbol{\mu}}_i)$$

$d_{\hat{\Sigma}}(\boldsymbol{x}, \hat{\boldsymbol{\mu}}_1) < d_{\hat{\Sigma}}(\boldsymbol{x}, \hat{\boldsymbol{\mu}}_2) \longrightarrow$ **検査点** x **は第一群と判別**

$d_{\hat{\Sigma}}(\boldsymbol{x}, \hat{\boldsymbol{\mu}}_1) \geqslant d_{\hat{\Sigma}}(\boldsymbol{x}, \hat{\boldsymbol{\mu}}_2) \longrightarrow$ **検査点** x **は第二群と判別**

判別式 $d_{\hat{\Sigma}}(\boldsymbol{x}, \hat{\boldsymbol{\mu}}_1) - d_{\hat{\Sigma}}(\boldsymbol{x}, \hat{\boldsymbol{\mu}}_2)$ **(この正負で群を判別)**

$$= (\boldsymbol{x} - \hat{\boldsymbol{\mu}}_1)' \Sigma^{-1} (\boldsymbol{x} - \hat{\boldsymbol{\mu}}_1) - (\boldsymbol{x} - \hat{\boldsymbol{\mu}}_2)' \Sigma^{-1} (\boldsymbol{x} - \hat{\boldsymbol{\mu}}_2)$$

$$= 2(\hat{\boldsymbol{\mu}}_1 - \hat{\boldsymbol{\mu}}_2)' \Sigma^{-1} \left(\boldsymbol{x} - \frac{\hat{\boldsymbol{\mu}}_1 + \hat{\boldsymbol{\mu}}_2}{2} \right)$$

誤判別率の計算

検査点 x の判別式

$$d(\boldsymbol{x}) = (\hat{\boldsymbol{\mu}}_1 - \hat{\boldsymbol{\mu}}_2)'\Sigma^{-1}\left(\boldsymbol{x} - \frac{\hat{\boldsymbol{\mu}}_1 + \hat{\boldsymbol{\mu}}_2}{2}\right)$$

誤判別率：本当は第一群の標本なのに第二群だと判別されてしまう確率

検査点が第一群の標本 $\boldsymbol{x} \sim N(\hat{\boldsymbol{\mu}}_1, \Sigma)$ のとき
統計量 $d(\boldsymbol{x})$ の分布を求め $d(\boldsymbol{x}) < 0$ となる確率を計算

$\boldsymbol{\delta} := \hat{\boldsymbol{\mu}}_1 - \hat{\boldsymbol{\mu}}_2$ および $\bar{\boldsymbol{\mu}} = (\hat{\boldsymbol{\mu}}_1 + \hat{\boldsymbol{\mu}}_2)/2$ と置くと

$$\mu_d = \mathbb{E}\{d(\boldsymbol{x})\} = \mathbb{E}\left\{\boldsymbol{\delta}'\Sigma^{-1}(\boldsymbol{x} - \bar{\boldsymbol{\mu}})\right\} = \boldsymbol{\delta}'\Sigma^{-1}\boldsymbol{\delta}/2$$

$$\sigma_d^2 = \mathbb{E}\left\{(d(\boldsymbol{x}) - \mu_d)^2\right\} = \boldsymbol{\delta}'\Sigma^{-1}\mathbb{E}\left\{(\boldsymbol{x} - \hat{\boldsymbol{\mu}}_1)(\boldsymbol{x} - \hat{\boldsymbol{\mu}}_1)'\right\}\Sigma^{-1}\boldsymbol{\delta} = \boldsymbol{\delta}'\Sigma^{-1}\boldsymbol{\delta}$$

で、正規分布の線形変換は正規分布なので $N(\mu_d, \sigma_d^2)$ で負値を取る確率