今日のはなし

DAY-4 7/7 (7)(8) 多変量正規分布: 多次元の正規分布と線形代数(ゼロから理解する正規分布)

[午前]

- 回帰分析の流れを具体的にみる
- 決定係数
- 相関係数
- 残差プロット

[午後]

- 単変量の正規分布の復習
- 多変量の正規分布

慣れてください！(この形にギョっとしない)

$$Z := \frac{X - \mu}{\sqrt{\sigma^2}}$$

平均を引いて
標準偏差で割っただけ

$$X \xrightarrow{\text{平均引いて標準偏差で割る}} Z := \frac{X - \mu}{\sigma}$$

標準化

平均μ,分散σ²
の確率変数X

平均0,分散1
の確率変数Z

$$\mathbb{E}\{Z\} = \frac{\mathbb{E}\{X\} - \mu}{\sigma} = 0$$

$$\mathbb{E}\{(Z - \mathbb{E}\{Z\})^2\} = \mathbb{E}\{Z^2\}$$
$$= \frac{\mathbb{E}\{(X - \mu)^2\}}{\sigma^2} = 1$$

正規分布まとめ

❶

$$X \sim N(\mu, \sigma^2) \Rightarrow \boxed{\frac{X-\mu}{\sqrt{\sigma^2}}} \sim N(0,1)$$

❷

$$X_1, \ldots, X_n \overset{\text{iid}}{\sim} N(\mu, \sigma^2) \Rightarrow \bar{X} = \frac{X_1 + \cdots + X_n}{n} \sim N\left(\mu, \frac{\sigma^2}{n}\right)$$

❸

$$Z = \boxed{\frac{\bar{X}-\mu}{\sqrt{\frac{\sigma^2}{n}}}} \sim N(0,1)$$

← ❶と❷から言える❸が大事

❸

$$Z = \frac{\bar{X} - \mu}{\sqrt{\frac{\sigma^2}{n}}} \sim N(0,1)$$

← ❶と❷から言える❸が大事

中心極限定理により、独立な多数の因子の和として表される確率変数は正規分布に従う。

$$\bar{X}_n := \frac{1}{n}\sum_{i=1}^{n} X_i$$

平均μ、分散σ^2でiid

$$P\left(\frac{\bar{X}_n - \mu}{\sqrt{\frac{\sigma^2}{n}}} \leqslant \alpha\right) \to \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\alpha} \exp\left(-\frac{x^2}{2}\right) dx$$

$$\Leftrightarrow \frac{\bar{X}_n - \mu}{\sqrt{\frac{\sigma^2}{n}}} \sim N(0,1)$$

❸

$$Z = \frac{\bar{X} - \mu}{\sqrt{\frac{\sigma^2}{n}}} \sim N(0,1)$$

$$\Leftrightarrow \frac{1}{\sqrt{n}} \sum_{i=1}^{n} \left(\frac{X_i - \mu}{\sigma} \right) \sim N(0,1)$$

(最初から標準化しておくなら)

変形したりするといきなり
わかりづらくなるけどめげない

$$\mathbb{E}\{\bar{X}\} = \mu, \mathrm{var}\{\bar{X}\} = \frac{\sigma^2}{n}$$

の事実からきているだけ

$$Z = \frac{\bar{X} - \mu}{\sqrt{\frac{\sigma^2}{n}}} = \frac{\left(\frac{1}{n}\sum_{i=1}^{n} X_i - \mu\right)}{\sqrt{\frac{\sigma^2}{n}}} = \frac{\left(\sum_{i=1}^{n} X_i - n\mu\right)}{n \cdot \sqrt{\frac{\sigma^2}{n}}}$$

$$= \frac{\left(\sum_{i=1}^{n} X_i - n\mu\right)}{\sqrt{n} \cdot \sqrt{\sigma^2}} = \frac{1}{\sqrt{n}} \sum_{i=1}^{n} \left(\frac{X_i - \mu}{\sigma} \right)$$

統計学で「正規分布」が特別な地位を占める一つの背景

中心極限定理

$X_1, X_2, \ldots, X_n$ はiid確率変数列で

$\mathbb{E}\{X_i\} = \mu, \mathrm{var}\{X\} = \sigma^2 < +\infty$

$$\lim_{n\to\infty} P\left(\frac{1}{\sqrt{n}}\sum_{i=1}^{n}\left(\frac{X_i-\mu}{\sigma}\right) \leqslant \alpha\right) = \int_{-\infty}^{\alpha} \frac{1}{\sqrt{2\pi}} e^{-\frac{t^2}{2}} dt$$

$$\Leftrightarrow \lim_{n\to\infty} P\left(\frac{\sum_{i=1}^{n} X_i - n\mu}{\sqrt{n}\sigma} \leqslant \alpha\right) = \int_{-\infty}^{\alpha} \frac{1}{\sqrt{2\pi}} e^{-\frac{t^2}{2}} dt$$

ポイント：上のn個の変数はiidであれば**任意の分布**で良い。

どんな分布に従う変数でも**iid列の標本平均の分布**は正規分布に

後者の意味：任意のiid列の総和は $N(n\mu, n\sigma^2)$ に従う(n→∞)

「正規分布」のexp($-x^2/2$)はどこから来たの？

「正規分布」はあんなに見た目複雑な密度関数をもつのに
統計学において非常に特別な地位の分布である。でもなぜ？

- 《その１》手元のデータがすべて同一の何らかの確率分布に従うと仮定する。中心極限定理により、n→∞のとき、総和の標本分布は正規分布に漸近する。
- 《その２》確率pで起こる事象をn回試行したとき、生起数は二項分布に従う。n→∞のとき、二項分布は正規分布に漸近する。
- 《その３》平均値がμ、分散がσ^2となる確率密度関数 $p(x)$ のなかで、そのエントロピー「$-\int_{-\infty}^{\infty} p(x)\log p(x)dx$」が最大となるのは正規分布である。

正規分布はどこから？エントロピー最大の確率密度関数は、正規分布関数である。

http://www7a.biglobe.ne.jp/~watmas/masaru-rep/info-normal.html

正規線形モデルと回帰

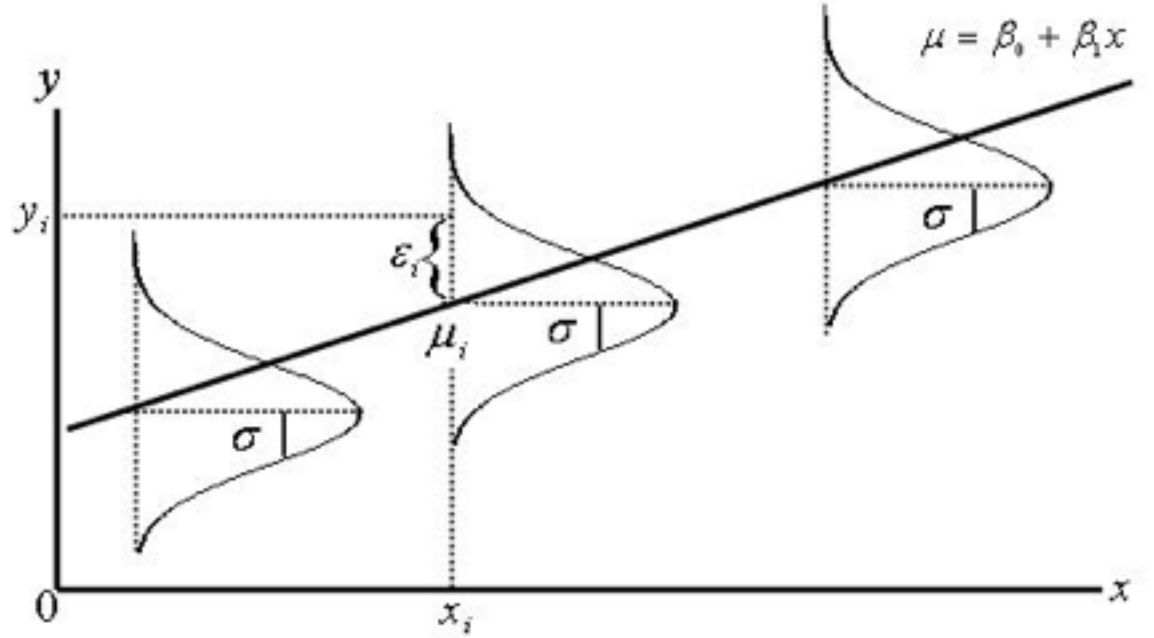
Given: $x_1, x_2, \ldots, x_n$

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

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？？



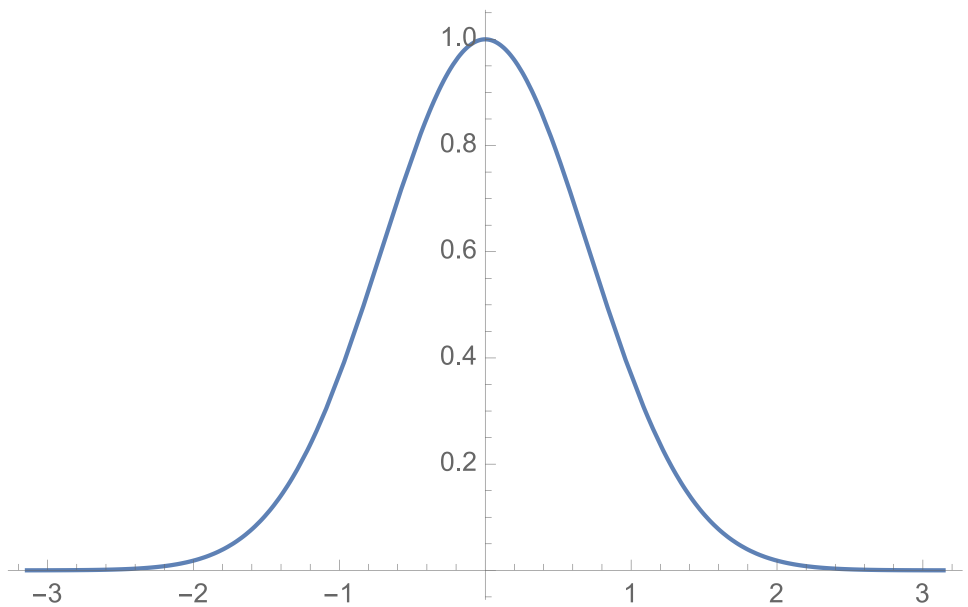
未知なる真の直線y=ax+bがあり
あるxでの観測値yは真の値まわりでN(0,σ²)に従ってずれて得られたと考える。分布はxに依らず常に同じもの。

正規分布　(ガウス分布)

normal distribution　Gaussian distribution

基本：この関数の形に比例する確率密度

$$f(x) := \exp(-x^2) = e^{-x^2}$$



確率密度なので積分したら1にするため、全面積を計算しておく

$$\boxed{\int_{-\infty}^{\infty} \exp(-x^2)dx = \sqrt{\pi}}$$

従って、この値で割って正規化した以下の関数形は密度関数！

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

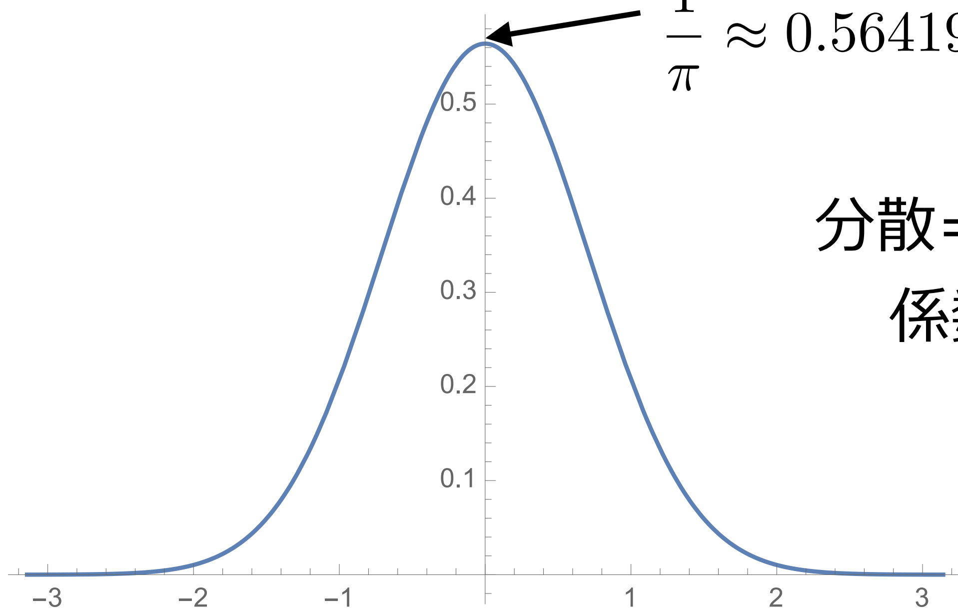
“釣鐘型”とか”ベル曲線(ベルカーブ)”とか言われる。

(…が、どっちかっつうと山型？？)

ただしこのままだと分散が1/2で不格好!?

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

$$\frac{1}{\pi} \approx 0.56419$$



平均(期待値)

$$\mathbb{E}\{x\} = \int_{-\infty}^{\infty} xp(x)dx = 0$$

分散=1/2

分散

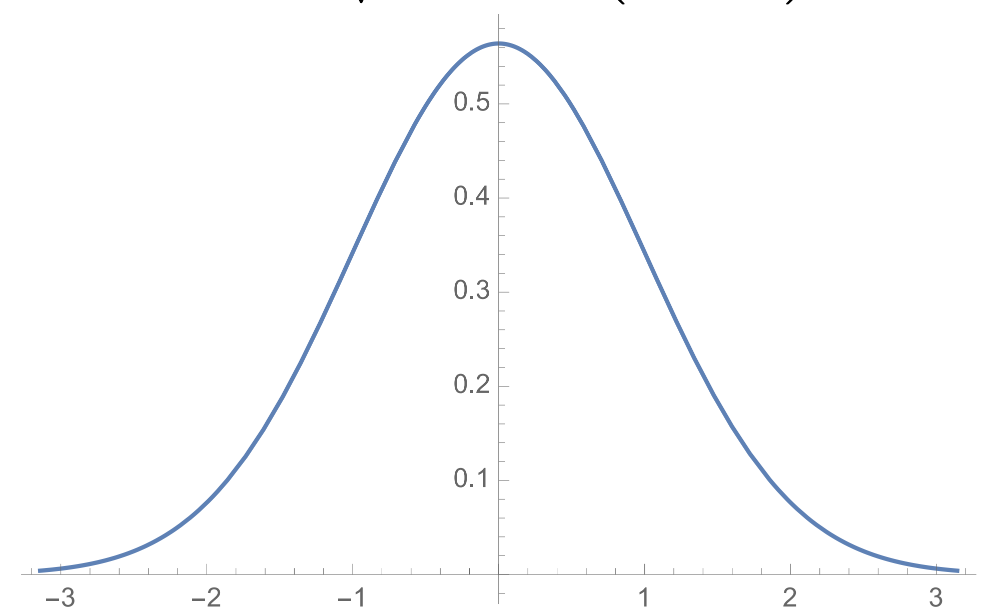
$$\mathbb{E}\{(x - \mathbb{E}\{x\})^2\} = \int_{-\infty}^{\infty} x^2 p(x)dx = \frac{1}{2}$$

$$\int_{-\infty}^{\infty} \exp\left(-\frac{1}{2}x^2\right) = \sqrt{2\pi}$$ なので

分散=1になるよう
係数をいじる

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

これが一般の単変量の正規分布を作るにも、多変量の正規分布を作るにも基本となる！！

標準正規分布 → 平均μ 分散σ²の正規分布

一般に平均μ,分散σ^2の確率変数Xを
平均0,分散1の確率変数Zに変換するには
以下の「標準化」をかます

確率変数の標準化

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

その逆、平均0,分散1の確率変数ZをXで
書き直すには単に上を代入すればよろし！

積分値=1にするための正規化定数は

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

正規分布の密度関数

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

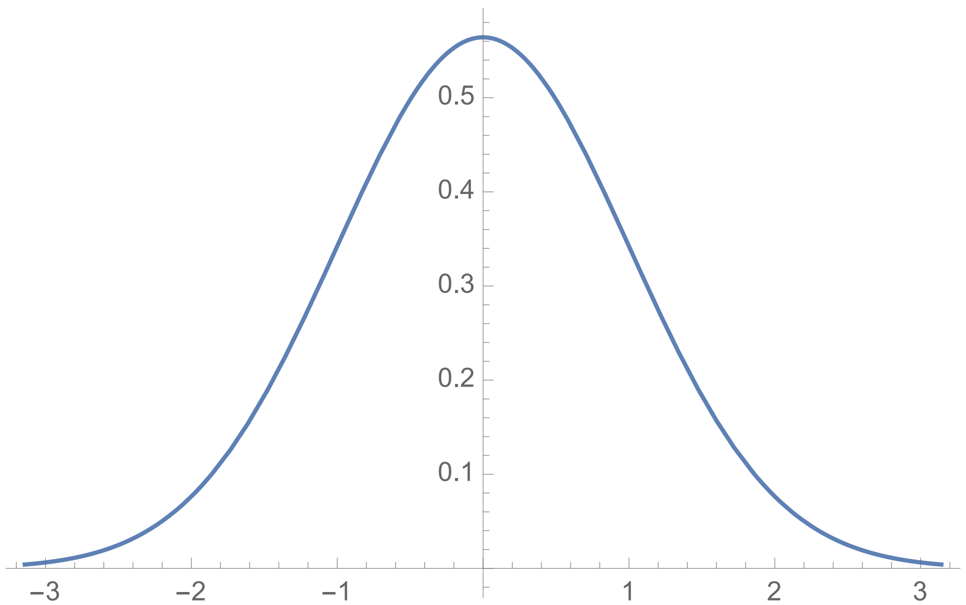
基本：この関数の形に比例する確率密度

$$f(x) := \exp(-x^2) = e^{-x^2}$$

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

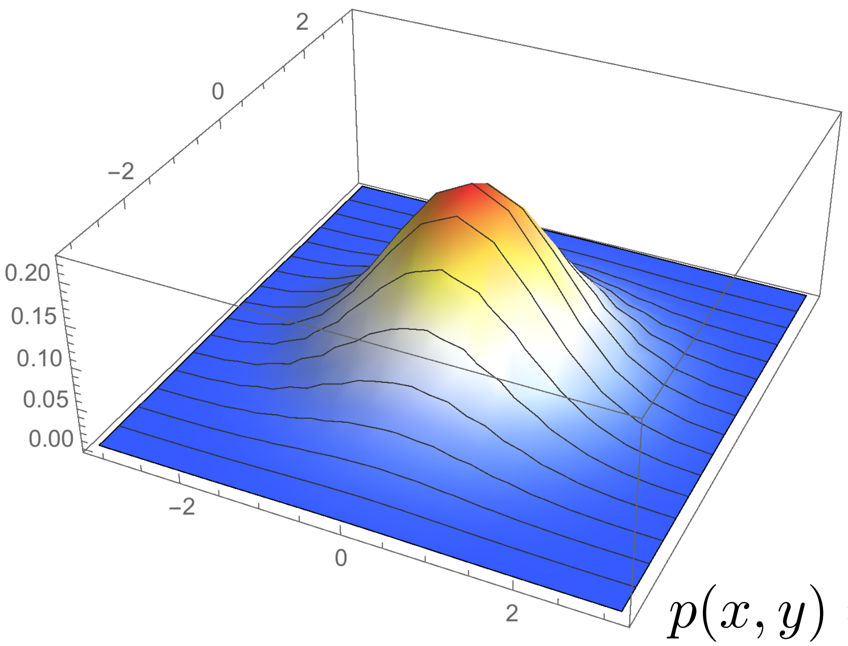
$$\boldsymbol{x} := (x_1, x_2, \ldots, x_n)'$$

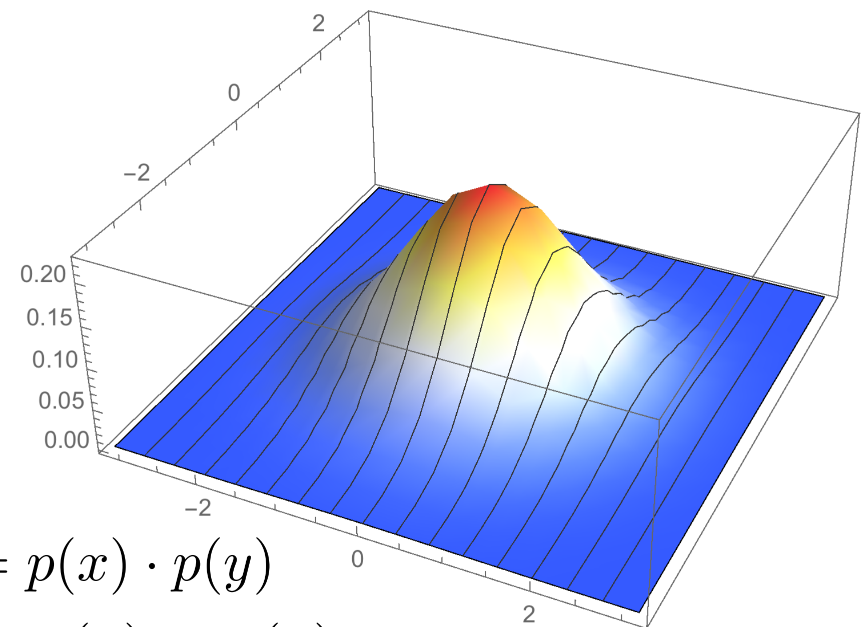
各々の変数が独立としてかけるだけ。

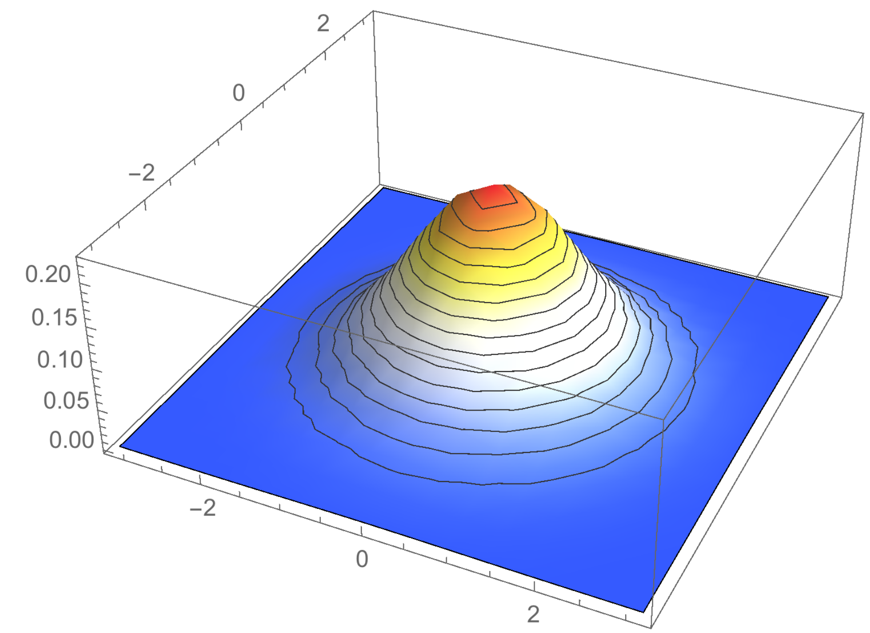
$$\begin{aligned} p(\boldsymbol{x}) &= p(x_1)p(x_2)\cdots p(x_n) \\ &= \frac{1}{(\sqrt{2\pi})^n} \exp\left(-\frac{1}{2}\sum_{i=1}^{n} x_i^2\right) \\ &= \frac{1}{(\sqrt{2\pi})^n} \exp\left(-\frac{1}{2}\boldsymbol{x}'\boldsymbol{x}\right) \end{aligned}$$

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

２変数の標準正規分布の例

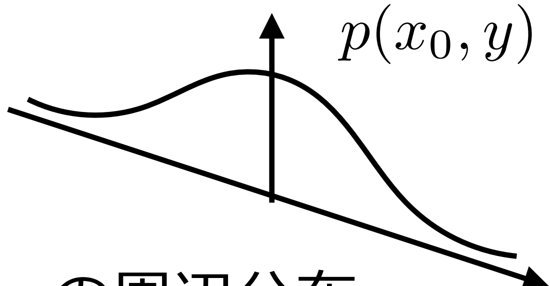




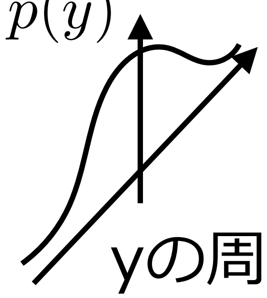


$$p(x, y) = p(x) \cdot p(y)$$

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

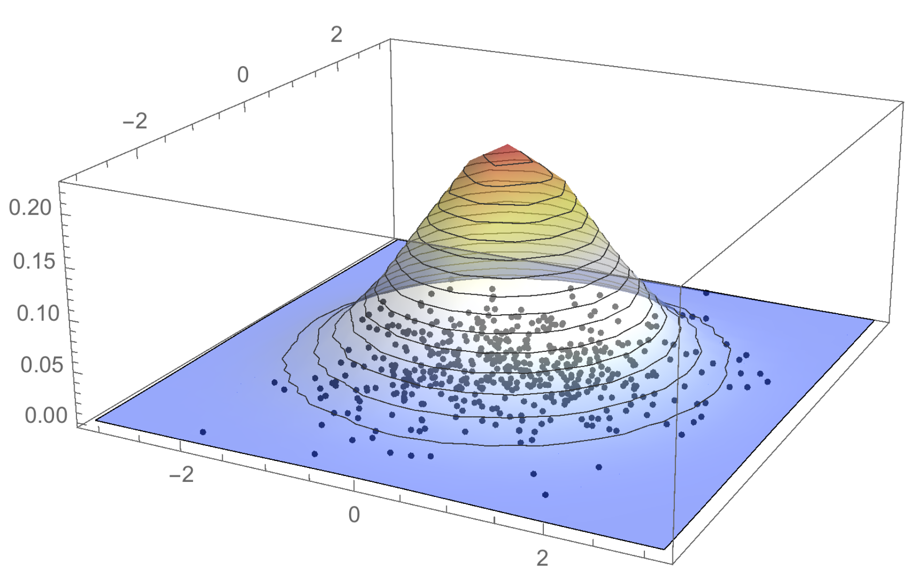


xの周辺分布



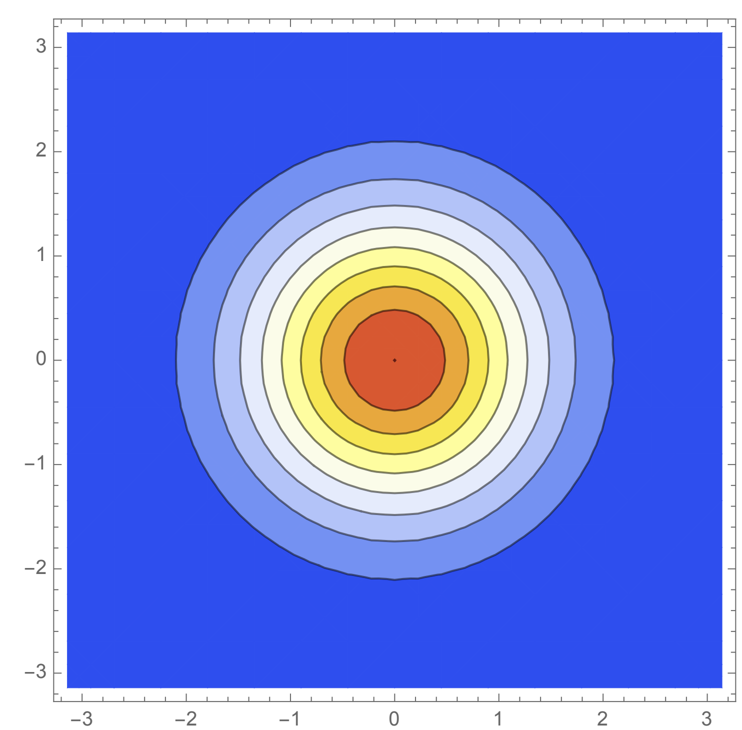
yの周辺分布

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

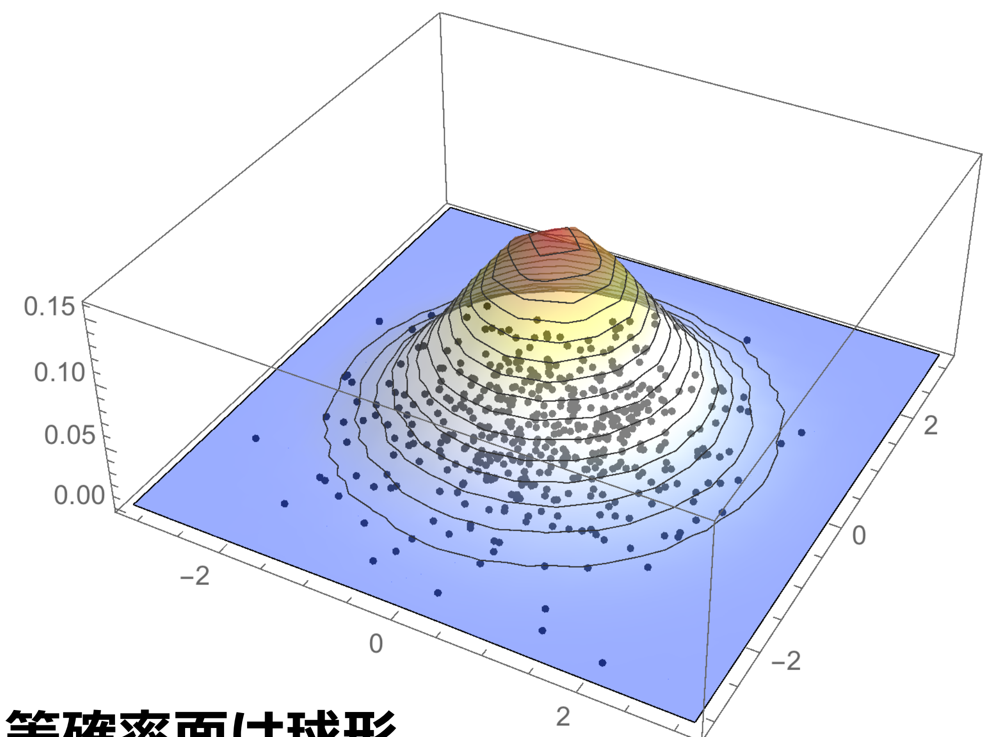
(2Dでは真円形→)



標準正規分布から一般の多変量正規分布へ

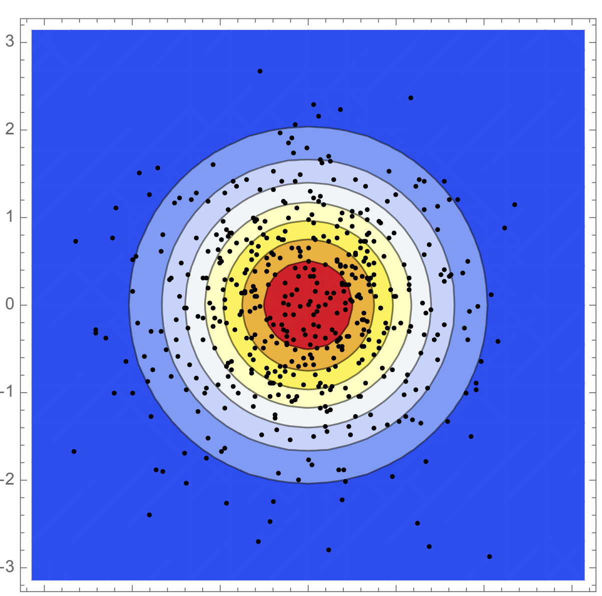
$$N(\mathbf{0}, \boldsymbol{I})$$

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



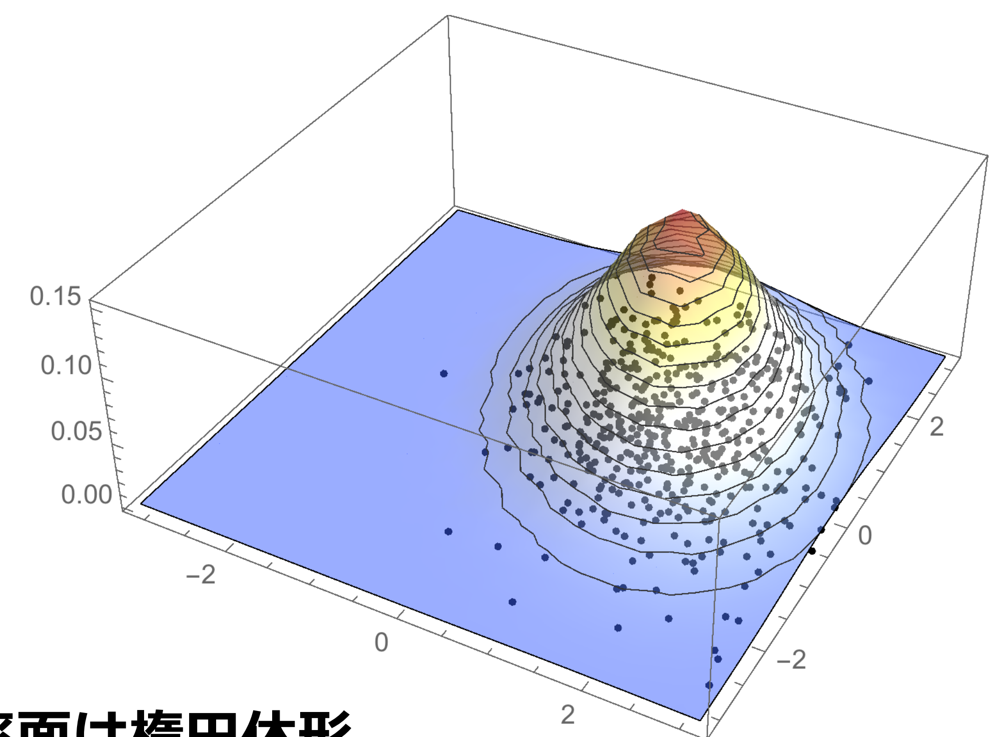
等確率面は球形

2D:等高線が真円形



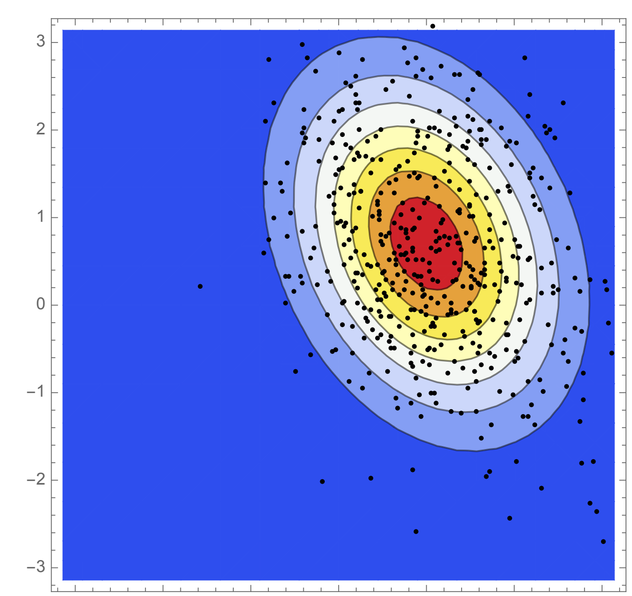
$$N(\boldsymbol{\mu}, \boldsymbol{\Sigma})$$

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



等確率面は楕円体形

2D:等高線が楕円形



基本的には「基底変換(しかも基底は直交)」の問題

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

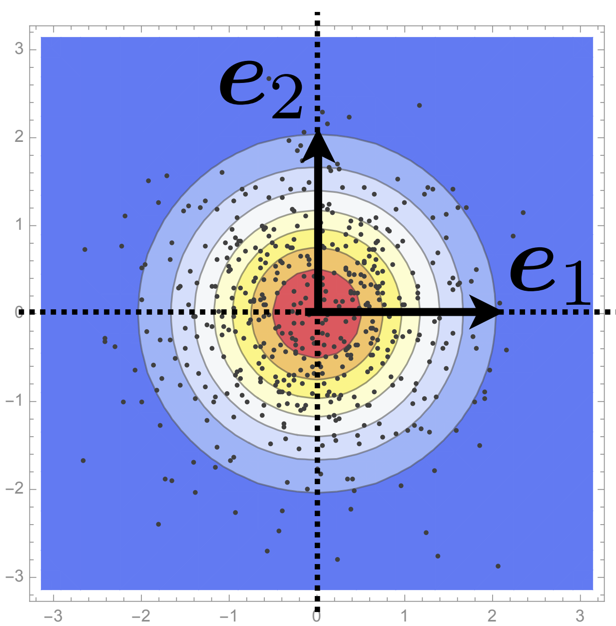
$$\frac{1}{(\sqrt{2\pi})^n} \exp\left(-\frac{1}{2}\boldsymbol{x}'\boldsymbol{x}\right)$$

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

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

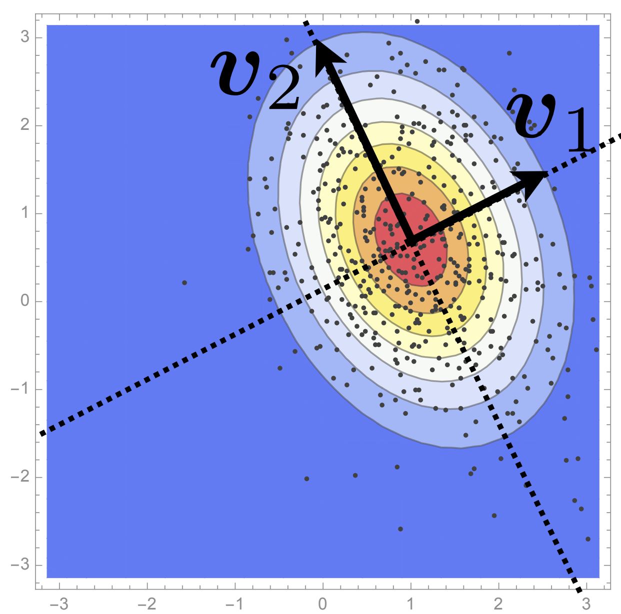
等確率面は球形

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



基底変換(しかも基底は直交)

この基底変換で
密度関数はどう変わる?

単位直交基底 $\boldsymbol{e}_1, \boldsymbol{e}_2, \ldots, \boldsymbol{e}_n$

新しい直交基底 $\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_n$

ポイント① 平行移動では形は変わらない

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

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

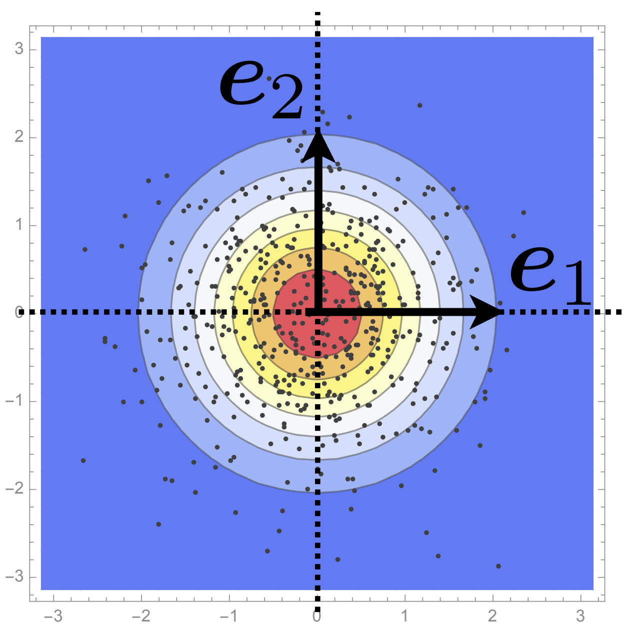
$N(\boldsymbol{\mu}, \boldsymbol{I})$

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

変わらない！

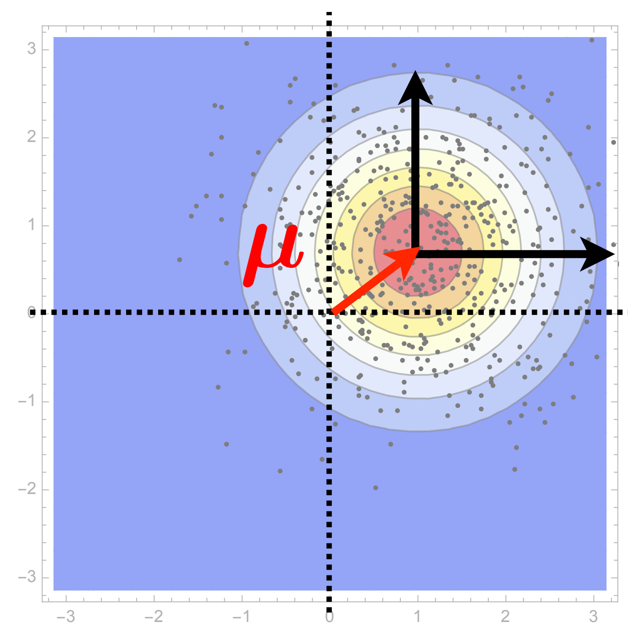
等確率面は球形

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



平行移動のみ

ズレるだけで形は変わらないので
積分値=1にする正規化定数も変わらない！

ポイント② 基底変換のところがミソ

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

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

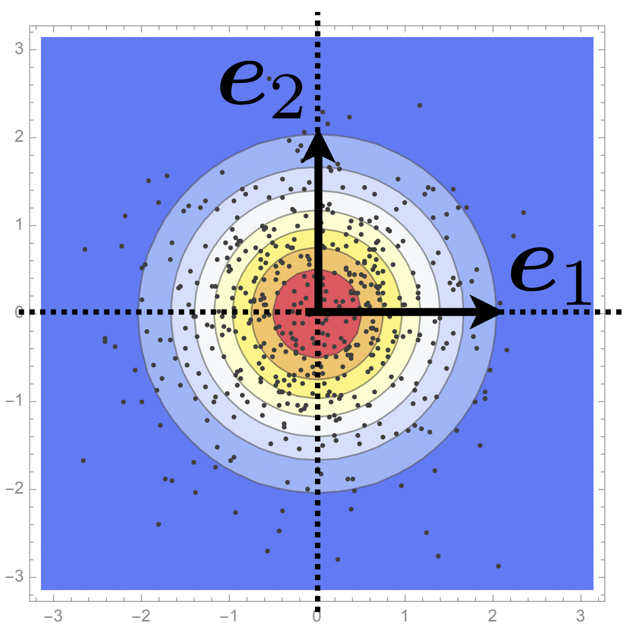
$N(\mathbf{0}, \boldsymbol{\Sigma})$

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

変わる！

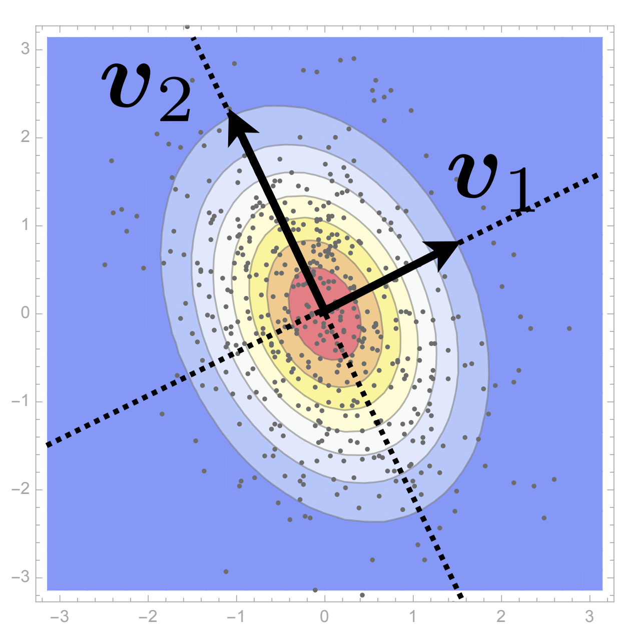
等確率面は球形

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



直交する基底への変換のみ

分散共分散行列 Σ
がこれを決める！

分布がひしゃげたり、形が変わるので
積分値=1にする正規化定数も変わる！

ポイント② 求める基底変換=多変量版の標準化

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

平均ベクトル

$$\mathbb{E}\{\boldsymbol{Z}\} = \mathbb{E}\{\boldsymbol{\Sigma}^{-\frac{1}{2}}(\boldsymbol{X} - \boldsymbol{\mu})\} = \boldsymbol{\Sigma}^{-\frac{1}{2}}(\mathbb{E}\{\boldsymbol{X}\} - \boldsymbol{\mu}) = \boldsymbol{0}$$

分散共分散行列

$$\begin{aligned}\mathbb{E}\{(\boldsymbol{Z} - \mathbb{E}\{\boldsymbol{Z}\})(\boldsymbol{Z} - \mathbb{E}\{\boldsymbol{Z}\})'\} &= \mathbb{E}\{\boldsymbol{Z}\boldsymbol{Z}'\} = \mathbb{E}\{\boldsymbol{\Sigma}^{-\frac{1}{2}}(\boldsymbol{X} - \boldsymbol{\mu})(\boldsymbol{\Sigma}^{-\frac{1}{2}}(\boldsymbol{X} - \boldsymbol{\mu}))'\} \\ &= \boldsymbol{\Sigma}^{-\frac{1}{2}}\mathbb{E}\{(\boldsymbol{X} - \boldsymbol{\mu})(\boldsymbol{X} - \boldsymbol{\mu})'\}\boldsymbol{\Sigma}^{-\frac{1}{2}} \\ &= \boldsymbol{\Sigma}^{-\frac{1}{2}}\boldsymbol{\Sigma}\boldsymbol{\Sigma}^{-\frac{1}{2}} = \boldsymbol{I}\end{aligned}$$

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

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

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

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

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

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

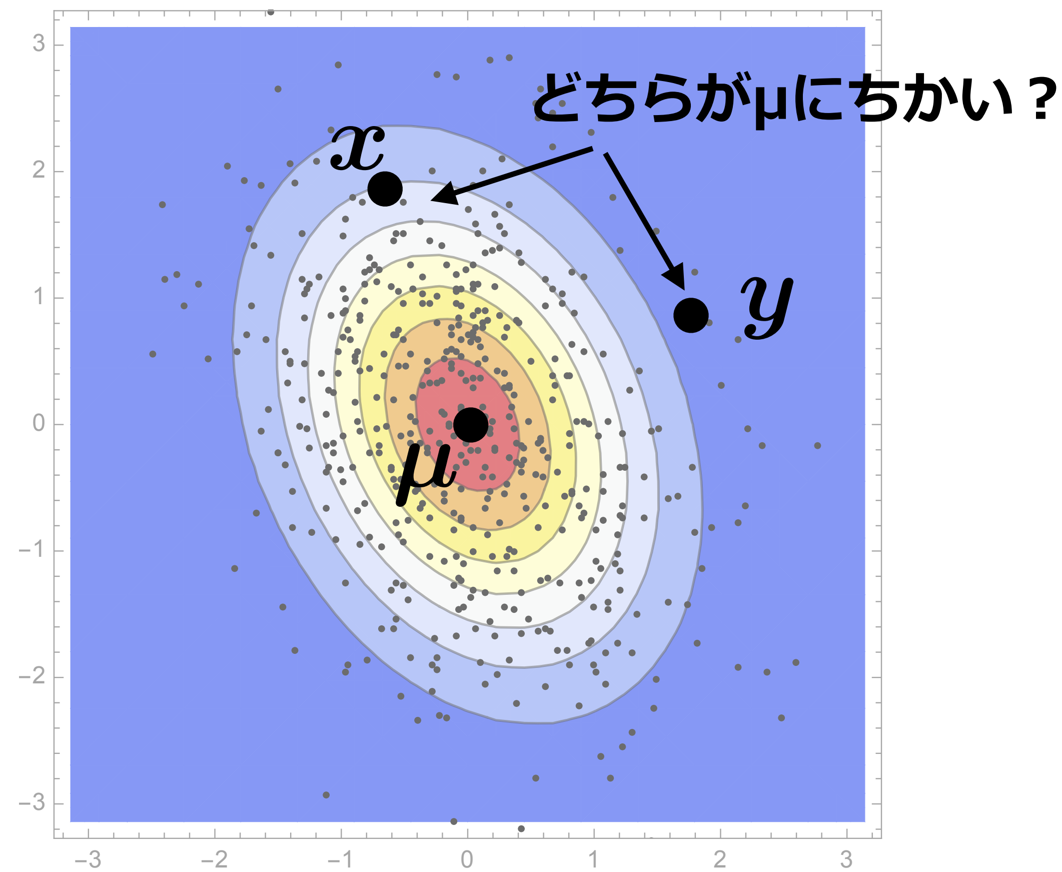
$N(\boldsymbol{\mu}, \boldsymbol{\Sigma})$

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

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**

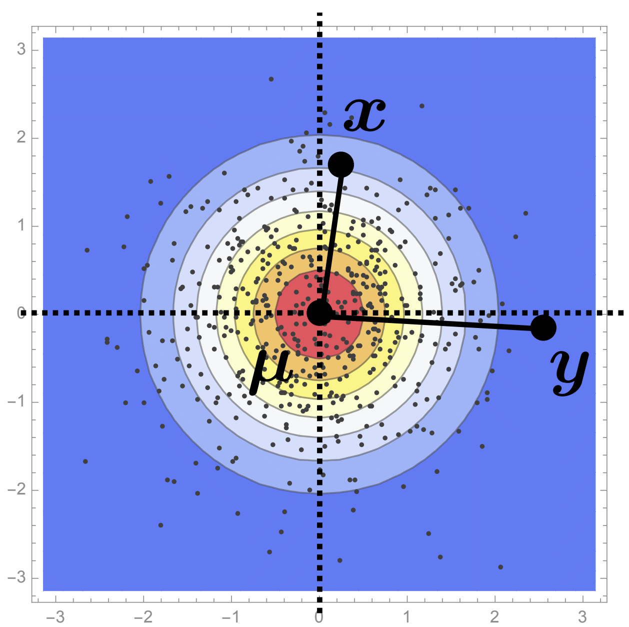
共分散を考慮して距離を測りたい。



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

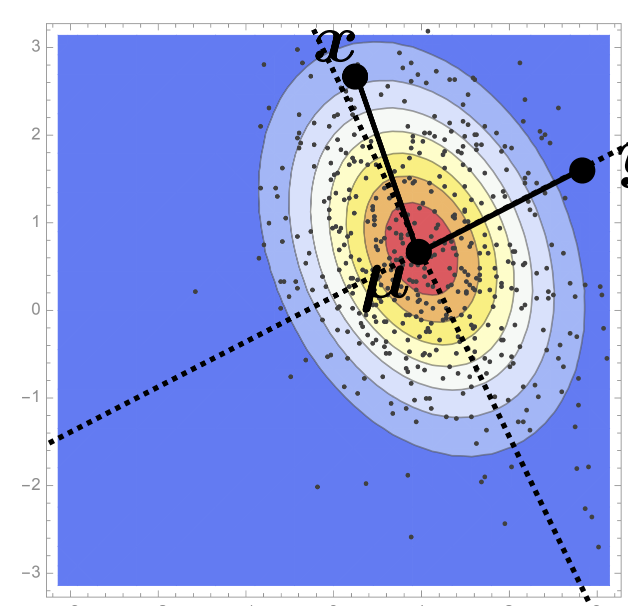


解釈：

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)

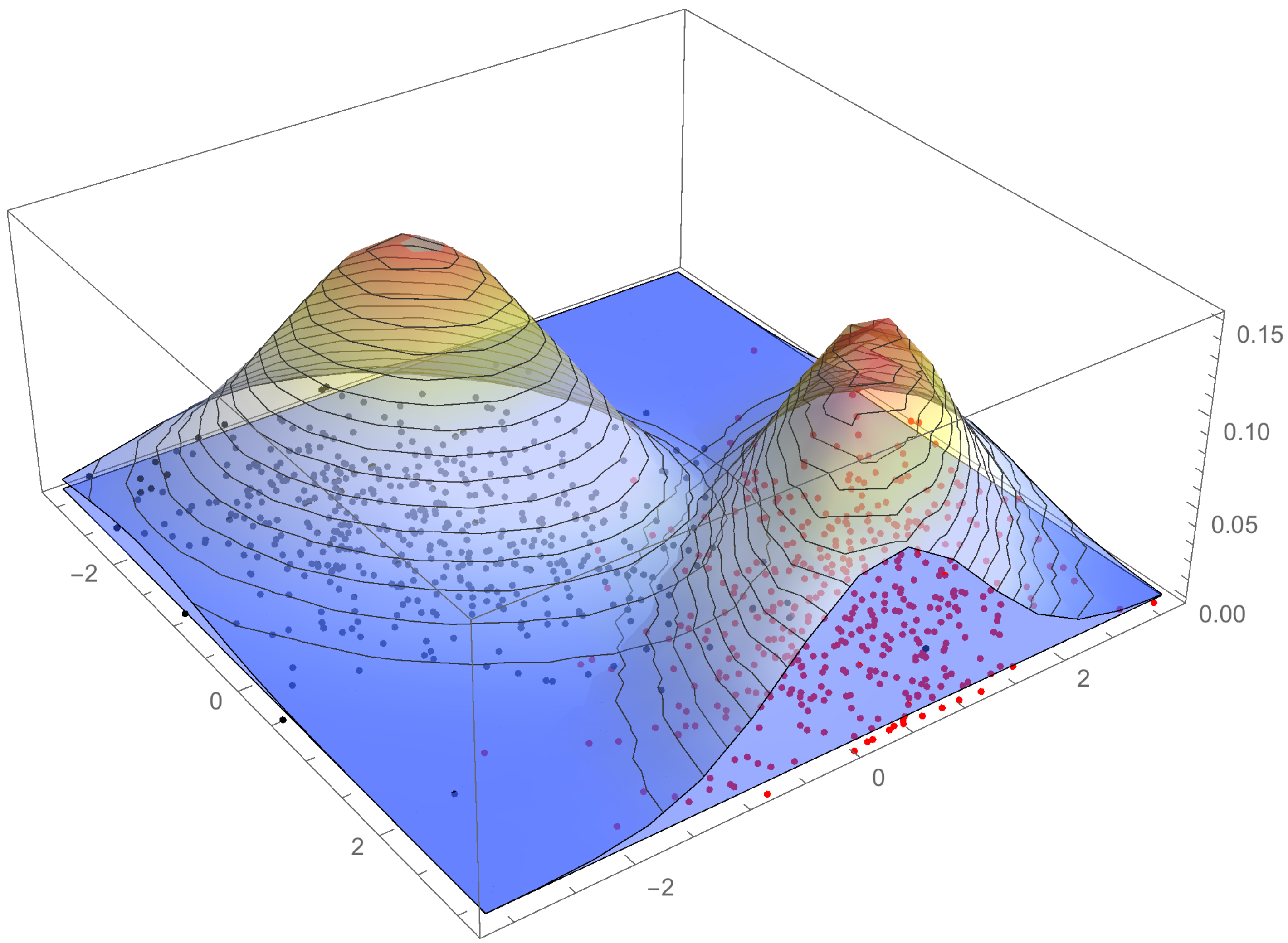
$$\boldsymbol{Z} := \boldsymbol{\Sigma}^{-\frac{1}{2}}(\boldsymbol{X}-\boldsymbol{\mu})$$

ユークリッド距離

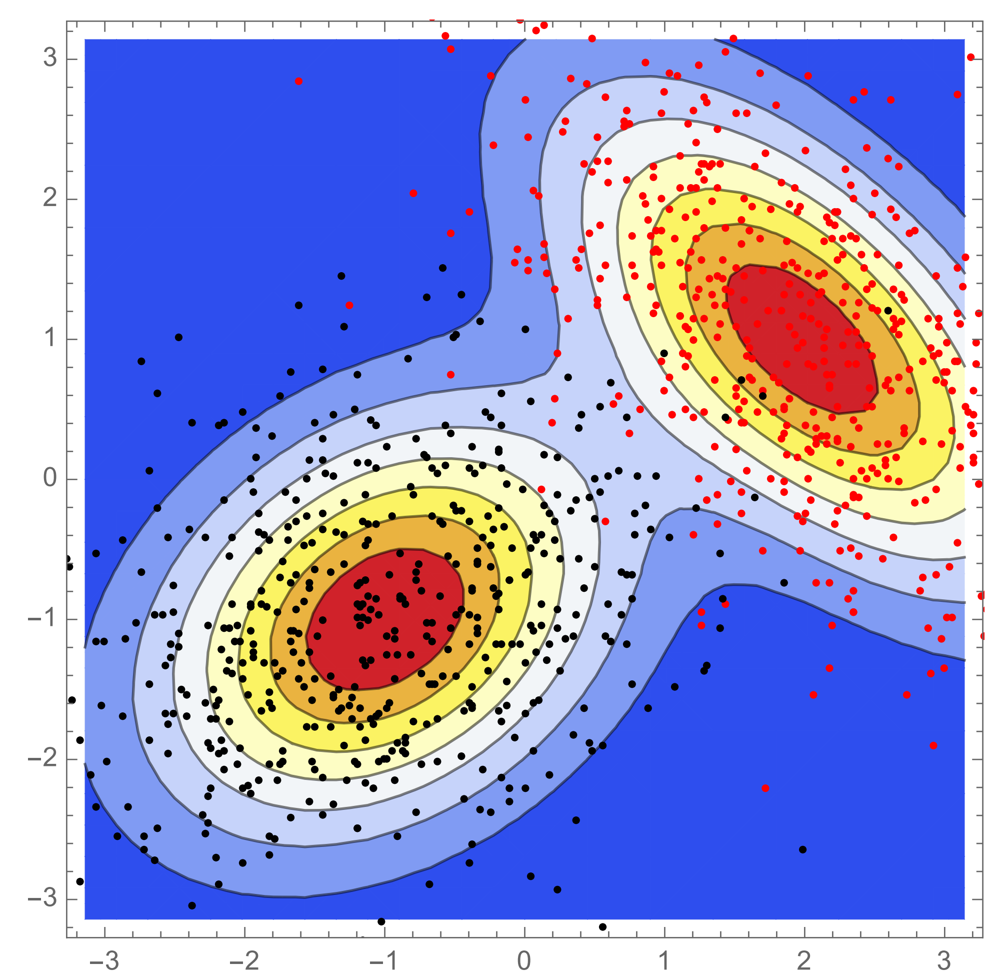


判別分析

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

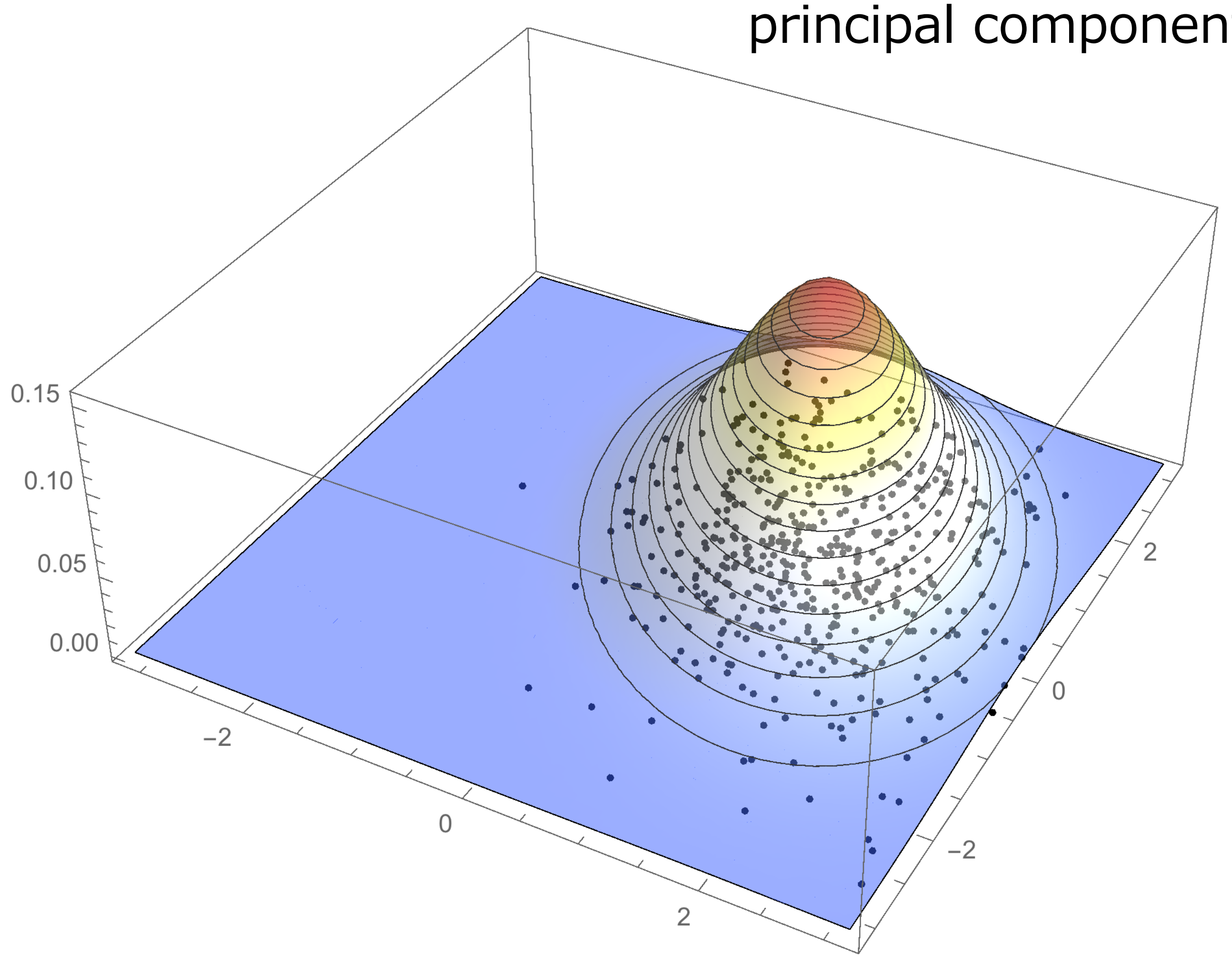


検査点xと各々の群の平均ベクトルとのマハラノビス距離を計算し、最も距離が近いものを結論とする

(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)

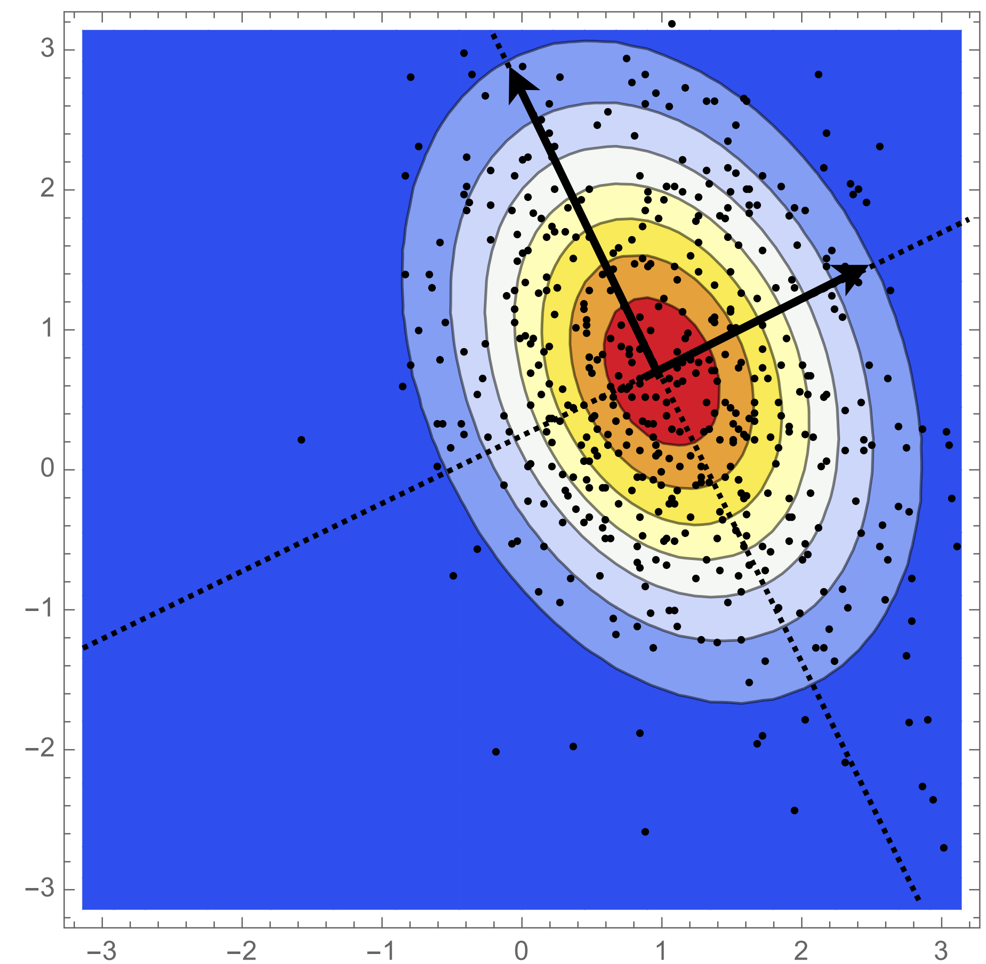
主成分分析 (主軸変換)

principal component analysis (PCA)



データが多変量正規的に分布していると仮定して、主軸を各々の楕円軸に取り直し、軸が長い順に、各軸での座標値を第一主成分、第二主成分、…として分析する。

第一主成分

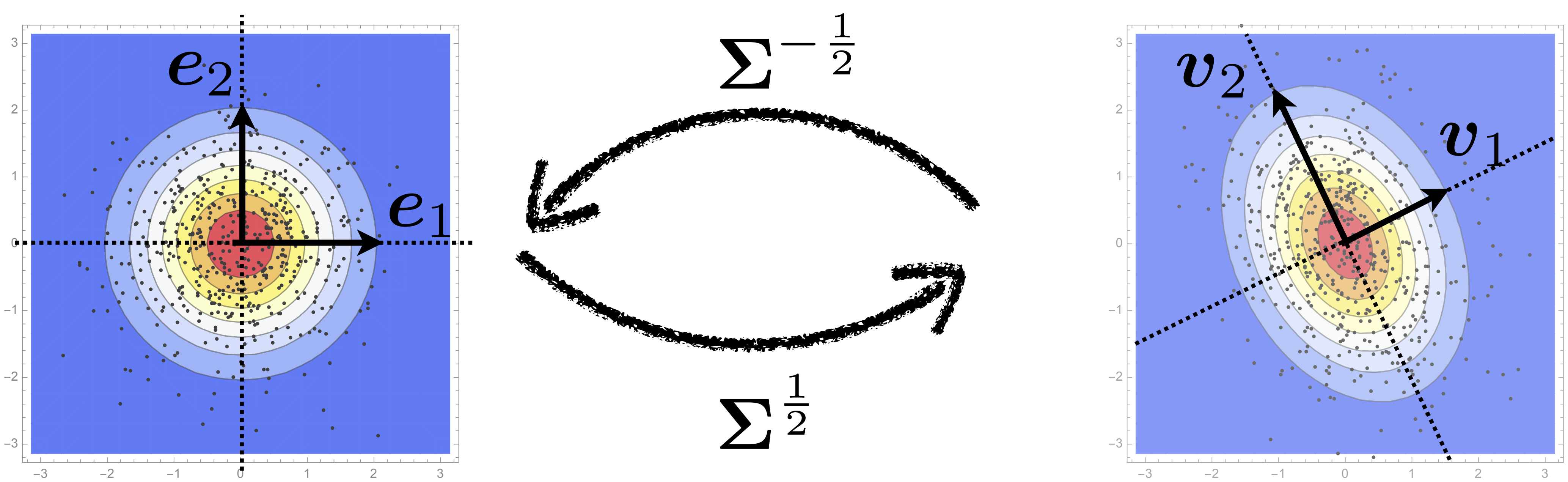


第二主成分

高次元のデータになれば、実質的なデータの分布は低次元の部分空間にのみ存在することにより、主成分表示を適当に上位kで打ち切ることにより次元削減ができる。

(パターン認識ではKarhunen-Loeve展開、KL展開による次元削減と呼ばれる)

基底変換や軸方向の計算 = 固有値・固有ベクトル！



以下の線形代数の内容を復習のこと

- 線形写像の表現としての行列
- 行列の平方根、行列の相似、基底変換、直交変換
- 固有値、固有ベクトル、対角化、三角化
- 実対称行列のスペクトル分解 (Rank-1行列分解)
- 二次形式の最大・最小と固有値・固有ベクトル

$\Sigma = P\Lambda P'$ **固有値分解**

$P'P = I$ **直交行列**

$\Lambda := \mathrm{diag}(\lambda_1, \lambda_2, \ldots, \lambda_n)$

$\Sigma^{1/2} = P\Lambda^{1/2}P'$ **平方根行列**

$\Lambda^{1/2} := \mathrm{diag}(\sqrt{\lambda_1}, \sqrt{\lambda_2}, \ldots, \sqrt{\lambda_n})$

重要　**分散共分散行列Σは、正方、実対称、半正定値な行列**

(従って、全ての固有値≧0で一意な平方根行列を持ち常に直交行列による対角化が可能。半正定値なら対称、対称なら正方だけど)

多変量データの具体例 (復習)

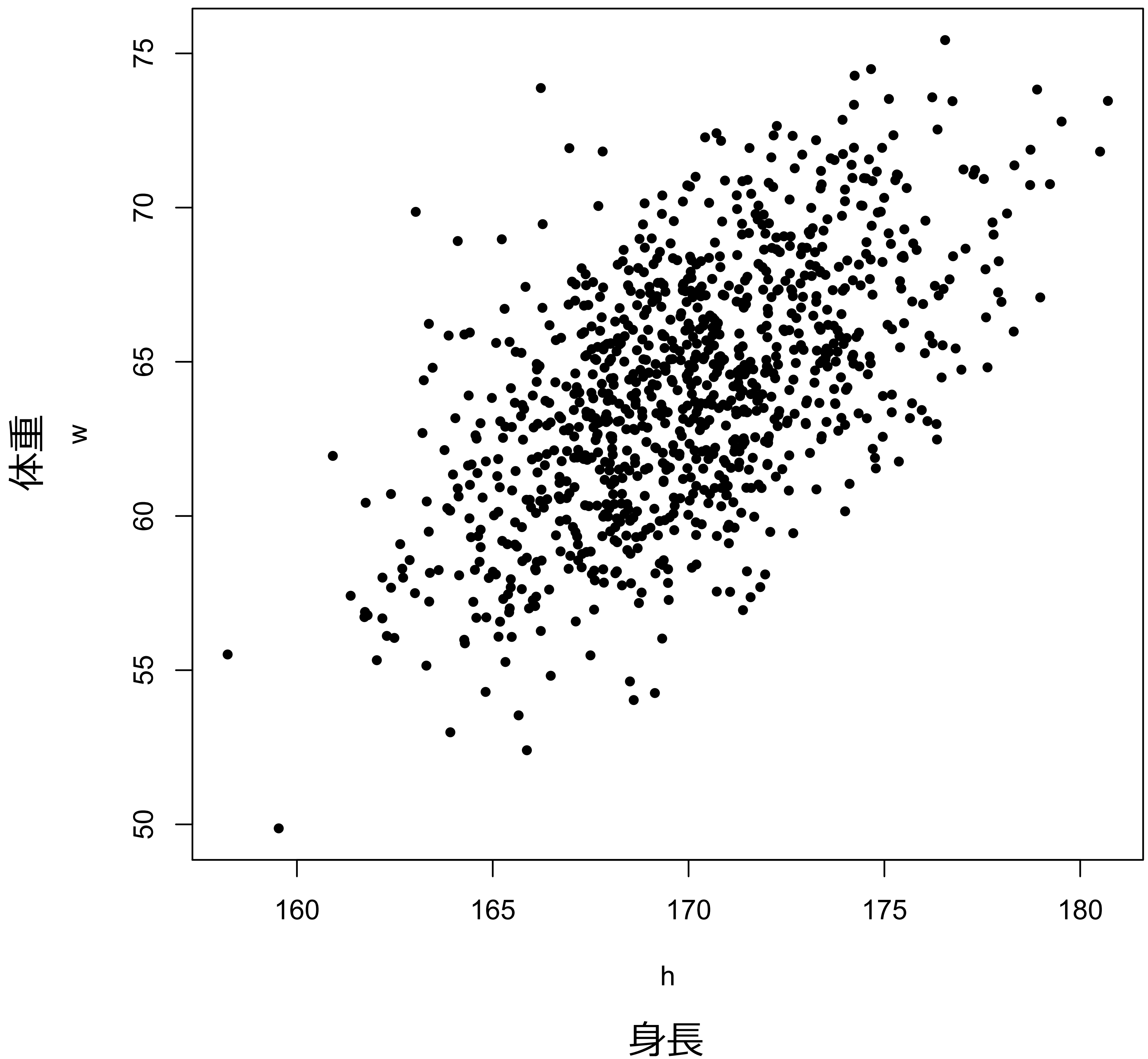
体重 w および身長 h を n 人に調査

集まったデータ $\rightarrow$ $w_1, w_2, w_3, \ldots, w_n$
$h_1, h_2, h_3, \ldots, h_n$

これを各々個人ごとに (w, h) という対データを得ると見て、2変量のベクトル値データとみなす

$$\begin{pmatrix} w_1 \\ h_1 \end{pmatrix}, \begin{pmatrix} w_2 \\ h_2 \end{pmatrix}, \begin{pmatrix} w_3 \\ h_3 \end{pmatrix}, \cdots, \begin{pmatrix} w_n \\ h_n \end{pmatrix} \qquad \in \mathbb{R}^2$$

各々の変量をxy軸にとると散布図



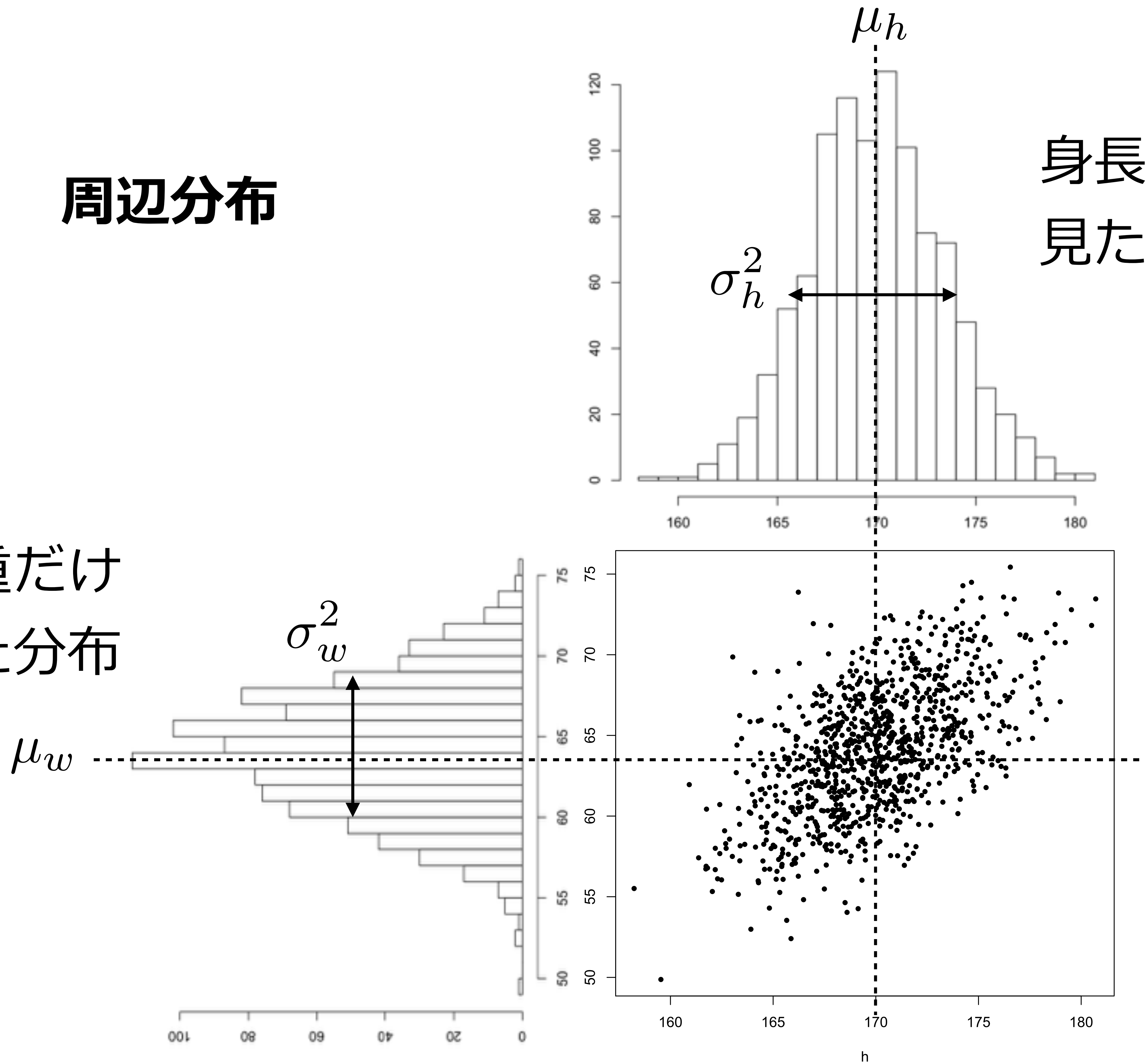
2変量データ

	身長 h	体重
1	174.0	64.1
2	169.6	65.4
3	168.4	67.4
4	171.7	63.4
5	172.1	72.3
6	167.0	63.4
7	167.0	62.5
:	:	:
999	172.7	64.9
1000	167.3	61.97

周辺分布

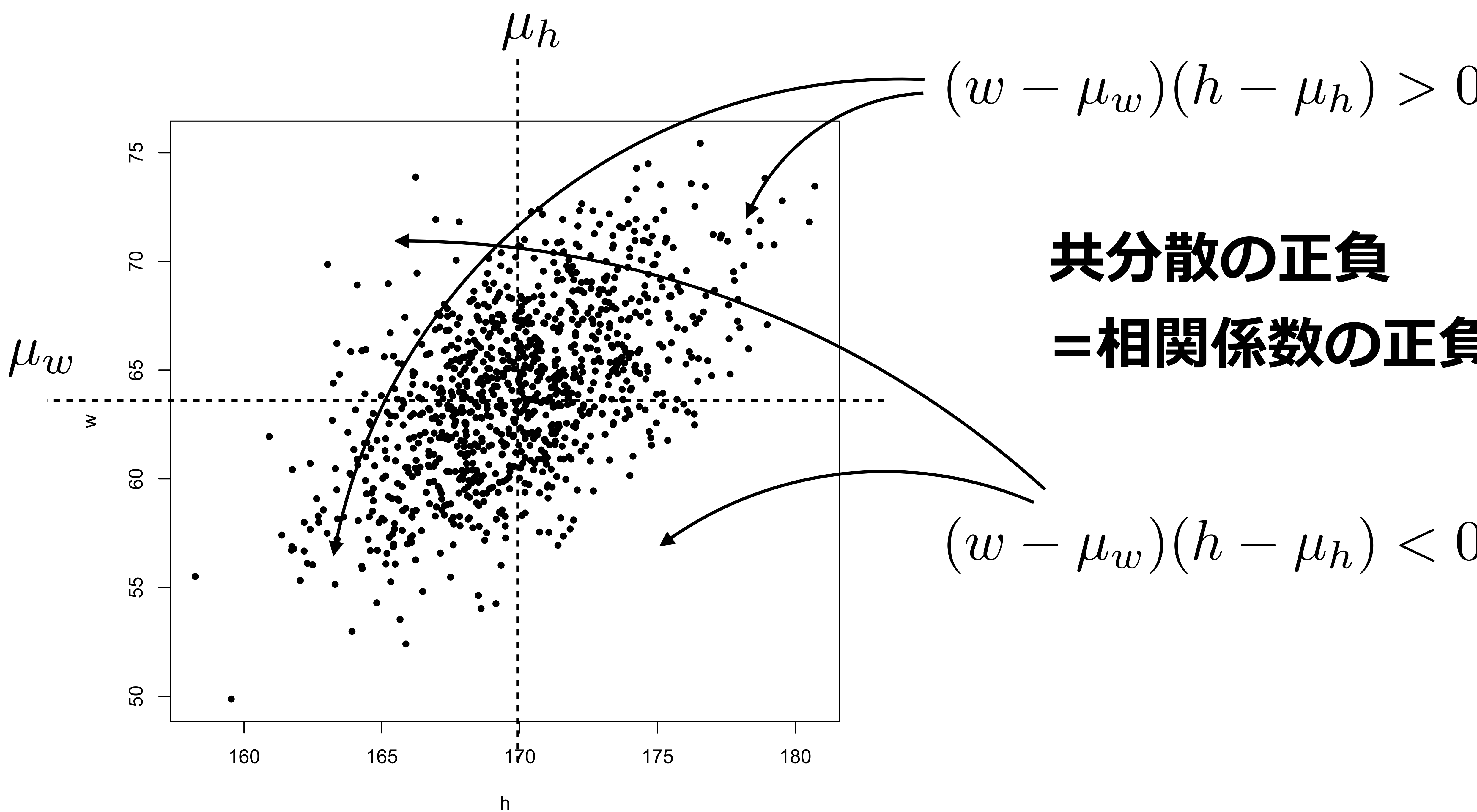
身長だけ
見た分布

体重だけ
見た分布



共分散 $\sigma_{w,h} := \mathbb{E}\{(w - \mu_w)(h - \mu_h)\}$

相関係数 $r_{w,h} := \dfrac{\sigma_{w,h}}{\sigma_w \sigma_h}$ ← 共分散を正規化



共分散の正負
＝相関係数の正負

分散共分散行列

$$\Sigma := \begin{bmatrix} \sigma_{w,w} & \sigma_{w,h} \\ \sigma_{h,w} & \sigma_{h,h} \end{bmatrix}$$

対角成分は分散

非対角成分は共分散

$$= \mathbb{E}\left\{ \begin{bmatrix} (w-\mu_w)(w-\mu_w) & (w-\mu_w)(h-\mu_h) \\ (h-\mu_h)(w-\mu_w) & (h-\mu_h)(h-\mu_h) \end{bmatrix} \right\}$$

$$= \begin{bmatrix} \mathbb{E}\{(w-\mu_w)(w-\mu_w)\} & \mathbb{E}\{(w-\mu_w)(h-\mu_h)\} \\ \mathbb{E}\{(h-\mu_h)(w-\mu_w)\} & \mathbb{E}\{(h-\mu_h)(h-\mu_h)\} \end{bmatrix}$$

※ 対角成分は分散

$$\sigma_{w,w} = \mathbb{E}\{(w-\mu_w)(w-\mu_w)\} = \mathbb{E}\{(w-\mu_w)^2\} = \sigma_w^2$$

$$\sigma_{h,h} = \mathbb{E}\{(h-\mu_h)(h-\mu_h)\} = \mathbb{E}\{(h-\mu_h)^2\} = \sigma_h^2$$

分散共分散行列 (一般のp変量の場合)

$$\boldsymbol{X} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_p \end{bmatrix} \quad \boldsymbol{\mu} = \begin{bmatrix} \mu_1 \\ \mu_2 \\ \vdots \\ \mu_p \end{bmatrix} \quad \Sigma = \begin{bmatrix} \sigma_{x_1,x_1} & \sigma_{x_2,x_1} & \cdots & \sigma_{x_p,x_1} \\ \sigma_{x_1,x_2} & \sigma_{x_2,x_2} & \cdots & \sigma_{x_p,x_2} \\ \vdots & & & \\ \sigma_{x_1,x_p} & \sigma_{x_2,x_p} & \cdots & \sigma_{x_p,x_p} \end{bmatrix}$$

$$= \mathbb{E}\{(\boldsymbol{X} - \boldsymbol{\mu})(\boldsymbol{X} - \boldsymbol{\mu})'\}$$

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

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

分散共分散行列

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

相関行列

$$R = \begin{bmatrix} r_{x_1,x_1} & r_{x_2,x_1} & \cdots & r_{x_p,x_1} \\ r_{x_1,x_2} & r_{x_2,x_2} & \cdots & r_{x_p,x_2} \\ \vdots & & & \\ r_{x_1,x_p} & r_{x_2,x_p} & \cdots & r_{x_p,x_p} \end{bmatrix}$$

相関係数 $r_{x_i,x_j} = \dfrac{\sigma_{x_i,x_j}}{\sigma_{x_i}\sigma_{x_j}}$ より、Rの対角成分は全て1

また、全ての分散が $\sigma_{x_i} = 1$ のとき、R=Σ

あるいは、全ての変数が標準化されていれば、R=Σ

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

$$\Sigma = DRD$$

$$D = \mathrm{diag}(\sigma_{x_1}, \sigma_{x_2}, \ldots, \sigma_{x_p})$$

$$= \begin{bmatrix} \sigma_{x_1} & 0 & \cdots & 0 \\ 0 & \sigma_{x_2} & \cdots & 0 \\ \vdots & & & \\ 0 & 0 & \cdots & \sigma_{x_p} \end{bmatrix} \begin{bmatrix} \frac{\sigma_{x_1,x_1}}{\sigma_{x_1}\sigma_{x_1}} & \frac{\sigma_{x_2,x_1}}{\sigma_{x_2}\sigma_{x_1}} & \cdots & \frac{\sigma_{x_p,x_1}}{\sigma_{x_p}\sigma_{x_1}} \\ \frac{\sigma_{x_1,x_2}}{\sigma_{x_1}\sigma_{x_2}} & \frac{\sigma_{x_2,x_2}}{\sigma_{x_2}\sigma_{x_2}} & \cdots & \frac{\sigma_{x_p,x_2}}{\sigma_{x_p}\sigma_{x_2}} \\ \vdots & & & \\ \frac{\sigma_{x_1,x_p}}{\sigma_{x_1}\sigma_{x_p}} & \frac{\sigma_{x_2,x_p}}{\sigma_{x_2}\sigma_{x_p}} & \cdots & \frac{\sigma_{x_p,x_p}}{\sigma_{x_p}\sigma_{x_p}} \end{bmatrix} \begin{bmatrix} \sigma_{x_1} & 0 & \cdots & 0 \\ 0 & \sigma_{x_2} & \cdots & 0 \\ \vdots & & & \\ 0 & 0 & \cdots & \sigma_{x_p} \end{bmatrix}$$

$$= \begin{bmatrix} \sigma_{x_1} & 0 & \cdots & 0 \\ 0 & \sigma_{x_2} & \cdots & 0 \\ \vdots & & & \\ 0 & 0 & \cdots & \sigma_{x_p} \end{bmatrix} \begin{bmatrix} r_{x_1,x_1} & r_{x_2,x_1} & \cdots & r_{x_p,x_1} \\ r_{x_1,x_2} & r_{x_2,x_2} & \cdots & r_{x_p,x_2} \\ \vdots & & & \\ r_{x_1,x_p} & r_{x_2,x_p} & \cdots & r_{x_p,x_p} \end{bmatrix} \begin{bmatrix} \sigma_{x_1} & 0 & \cdots & 0 \\ 0 & \sigma_{x_2} & \cdots & 0 \\ \vdots & & & \\ 0 & 0 & \cdots & \sigma_{x_p} \end{bmatrix}$$

$$R = \begin{bmatrix} r_{x_1,x_1} & r_{x_2,x_1} & \cdots & r_{x_p,x_1} \\ r_{x_1,x_2} & r_{x_2,x_2} & \cdots & r_{x_p,x_2} \\ \vdots & & & \\ r_{x_1,x_p} & r_{x_2,x_p} & \cdots & r_{x_p,x_p} \end{bmatrix} = \begin{bmatrix} 1 & r_{x_2,x_1} & \cdots & r_{x_p,x_1} \\ r_{x_1,x_2} & 1 & \cdots & r_{x_p,x_2} \\ \vdots & & & \\ r_{x_1,x_p} & r_{x_2,x_p} & \cdots & 1 \end{bmatrix}$$

分散共分散行列の性質（１）

$$\Sigma = \mathbb{E}\{(\boldsymbol{X} - \boldsymbol{\mu})(\boldsymbol{X} - \boldsymbol{\mu})'\} = \mathbb{E}\{\boldsymbol{X}\boldsymbol{X}'\} - \boldsymbol{\mu}\boldsymbol{\mu}'$$

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

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

$$= \begin{bmatrix} \mathbb{E}\{x_1x_1\} - \mu_1\mu_1 & \mathbb{E}\{x_2x_1\} - \mu_2\mu_1 & \cdots & \mathbb{E}\{x_px_1\} - \mu_p\mu_1 \\ \mathbb{E}\{x_2x_1\} - \mu_2\mu_1 & \mathbb{E}\{x_2x_2\} - \mu_2\mu_2 & \cdots & \mathbb{E}\{x_px_2\} - \mu_p\mu_2 \\ \vdots & & & \\ \mathbb{E}\{x_px_1\} - \mu_p\mu_1 & \mathbb{E}\{x_2x_p\} - \mu_2\mu_p & \cdots & \mathbb{E}\{x_px_p\} - \mu_p\mu_p \end{bmatrix}$$

分散共分散行列の性質（2）

実対称行列　$\Sigma = \Sigma'$

定義からも明らかであるが…

$$(\mathbb{E}\{(\boldsymbol{X}-\boldsymbol{\mu})(\boldsymbol{X}-\boldsymbol{\mu})'\})' = \mathbb{E}\{((\boldsymbol{X}-\boldsymbol{\mu})(\boldsymbol{X}-\boldsymbol{\mu})')'\}$$
$$= \mathbb{E}\{((\boldsymbol{X}-\boldsymbol{\mu})')'(\boldsymbol{X}-\boldsymbol{\mu})'\} = \mathbb{E}\{(\boldsymbol{X}-\boldsymbol{\mu})(\boldsymbol{X}-\boldsymbol{\mu})'\}$$

実対称行列の性質

- 固有値は必ず実数
- 異なる固有値に属する固有ベクトルは直交する
- 直交行列による対角化が可能

分散共分散行列の性質（3）

二次形式が常に非負

→**半正定値行列(または非不定値行列)という**

※二次形式はベクトル量や行列量ではなく、**単なる実数値**であることに注意

$$
\begin{aligned}
\boldsymbol{z}'\Sigma\boldsymbol{z} &= \boldsymbol{z}'\mathbb{E}\{(\boldsymbol{X}-\boldsymbol{\mu})(\boldsymbol{X}-\boldsymbol{\mu})'\}\boldsymbol{z} \\
&= \mathbb{E}\{\boldsymbol{z}'(\boldsymbol{X}-\boldsymbol{\mu})(\boldsymbol{X}-\boldsymbol{\mu})'\boldsymbol{z}\} = \mathbb{E}\{(\boldsymbol{z}'(\boldsymbol{X}-\boldsymbol{\mu}))^2\} \geqslant 0
\end{aligned}
$$

半正定値行列Aの性質

- すべての固有値が非負
- すべての主小行列式(Principal Minors)が非負
- A = X′ X なる行列Xが存在
- A = X X なる半正定値の平方根行列Xが1つだけ存在

標本分散共分散行列

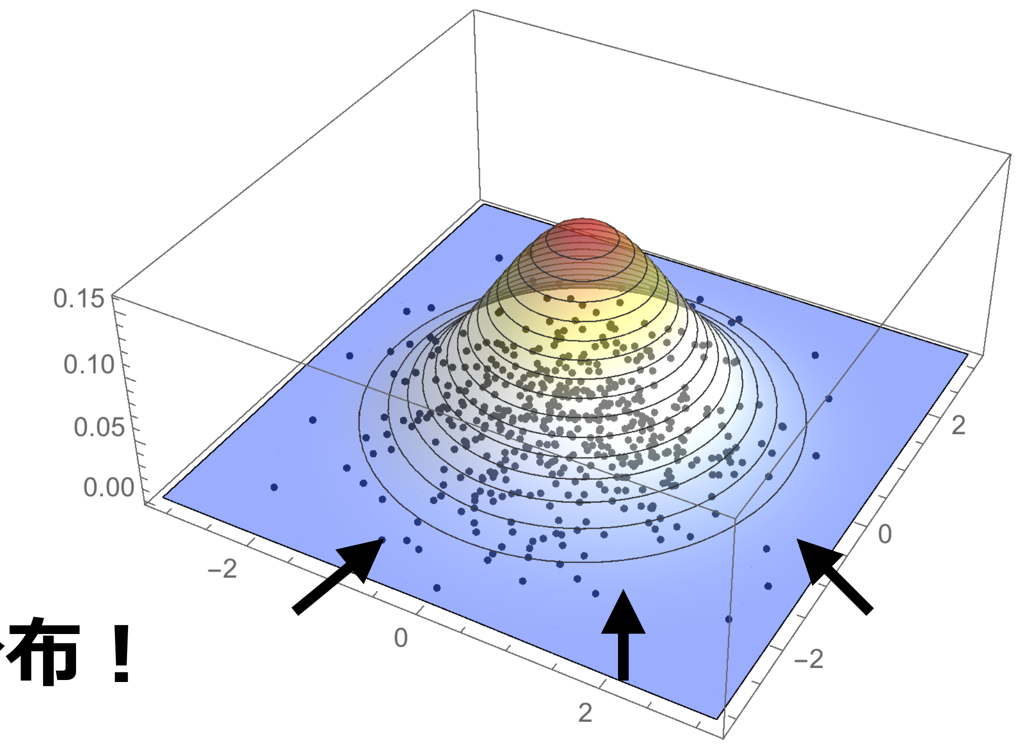
$$\hat{\Sigma} = \frac{1}{n-1}\sum_{i=1}^{n}(\boldsymbol{x}_i - \bar{\boldsymbol{x}})(\boldsymbol{x}_i - \bar{\boldsymbol{x}})' \quad \text{ただし} \ \bar{\boldsymbol{x}} := \frac{1}{n}\sum_{i=1}^{n}\boldsymbol{x}_i$$

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

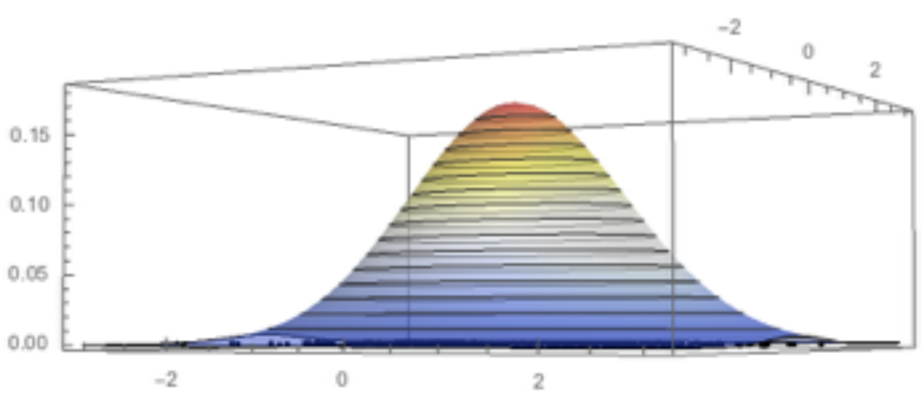
$$\begin{aligned}
\mathbb{E}\{\hat{\Sigma}\} &= \frac{1}{n-1}\mathbb{E}\left\{\sum_{i=1}^{n}(\boldsymbol{x}_i - \boldsymbol{\mu})(\boldsymbol{x}_i - \boldsymbol{\mu})' - 2n(\bar{\boldsymbol{x}} - \boldsymbol{\mu})(\bar{\boldsymbol{x}} - \boldsymbol{\mu})' + n(\bar{\boldsymbol{x}} - \boldsymbol{\mu})(\bar{\boldsymbol{x}}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

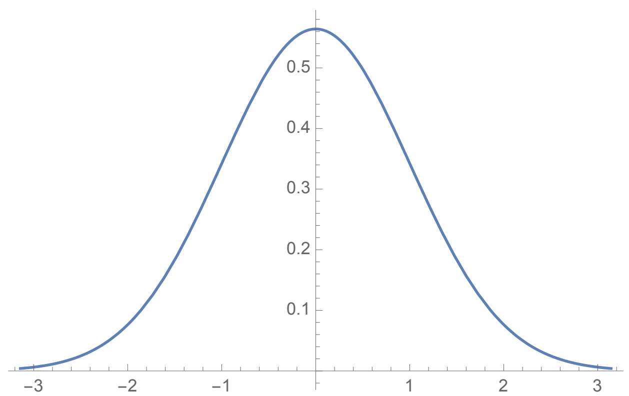
標準正規分布

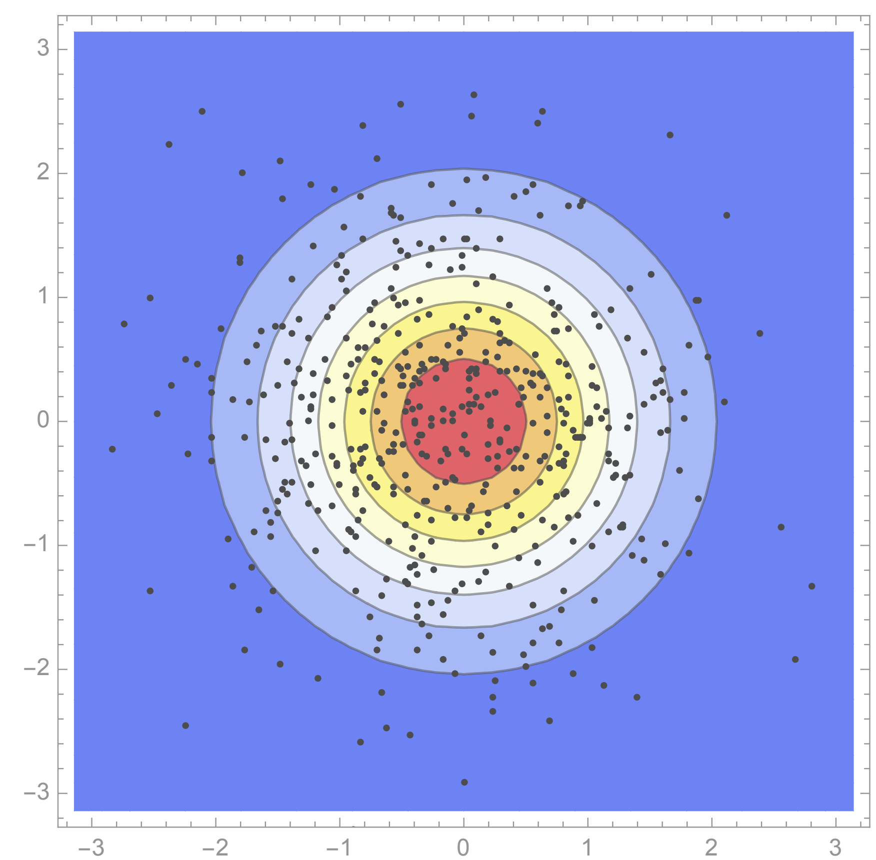
$$\Sigma = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$



どの角度から見ても単変量の正規分布！
(分布は直交変換で不変)

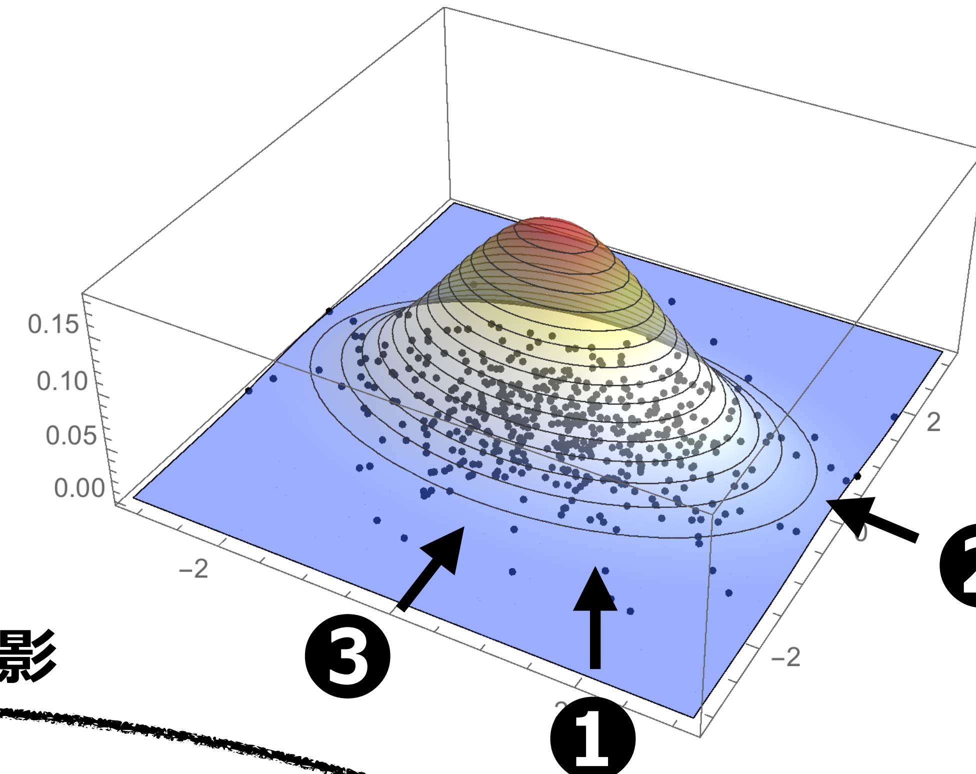


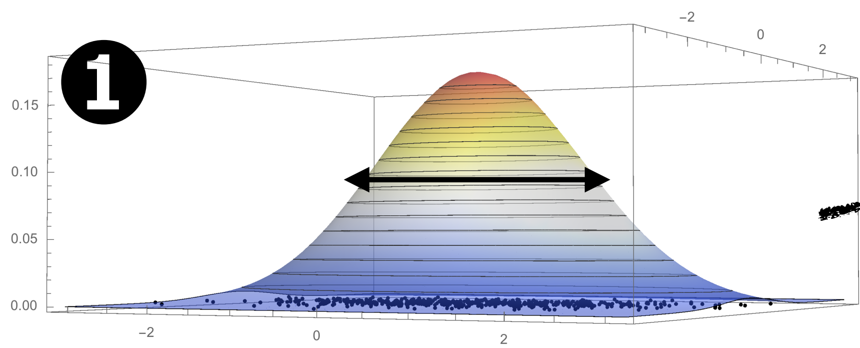




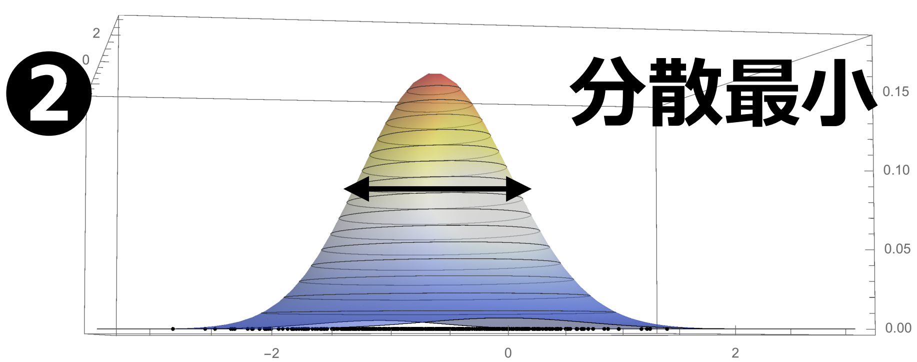
共分散がすべてゼロ = 分散共分散行列が対角行列

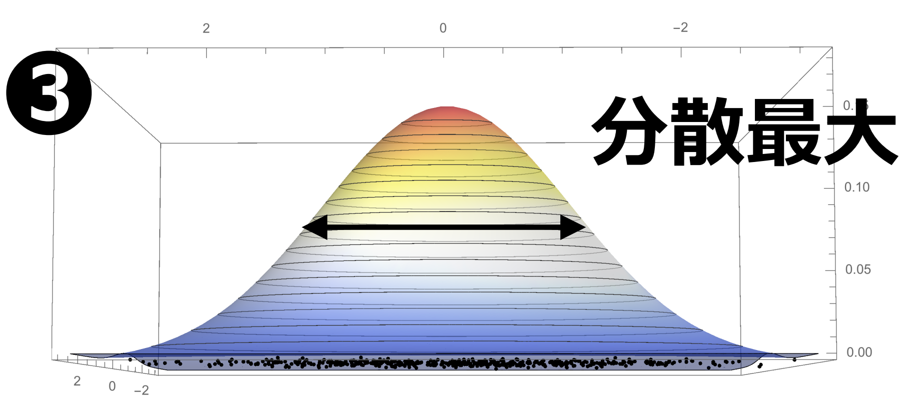
$$\Sigma = \begin{bmatrix} 1.5 & 0 \\ 0 & 0.5 \end{bmatrix}$$





この軸へ直交射影

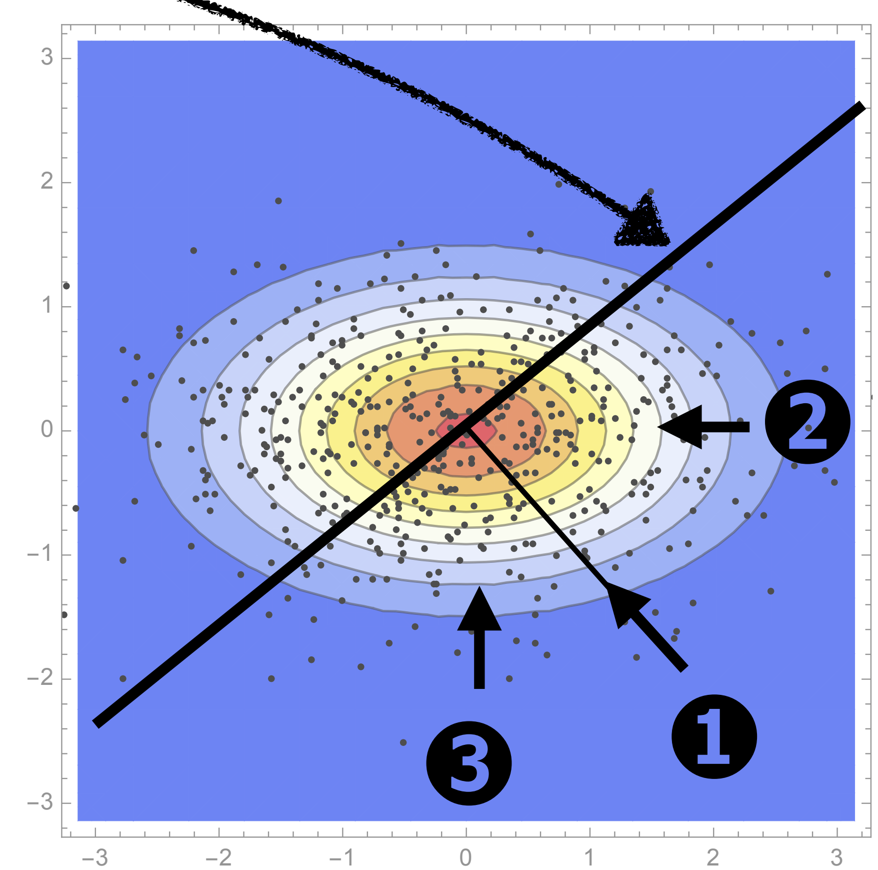




等確率面の
楕円が**軸平行**

$\times 1.5$

$\times 0.5$



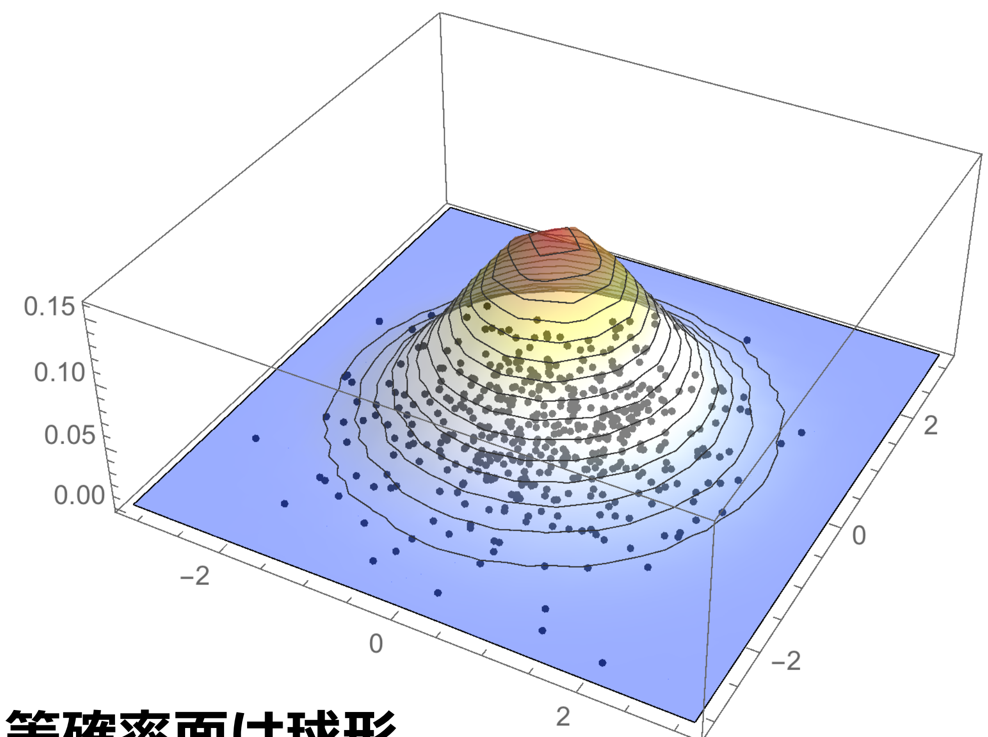
標準正規分布から一般の多変量正規分布へ

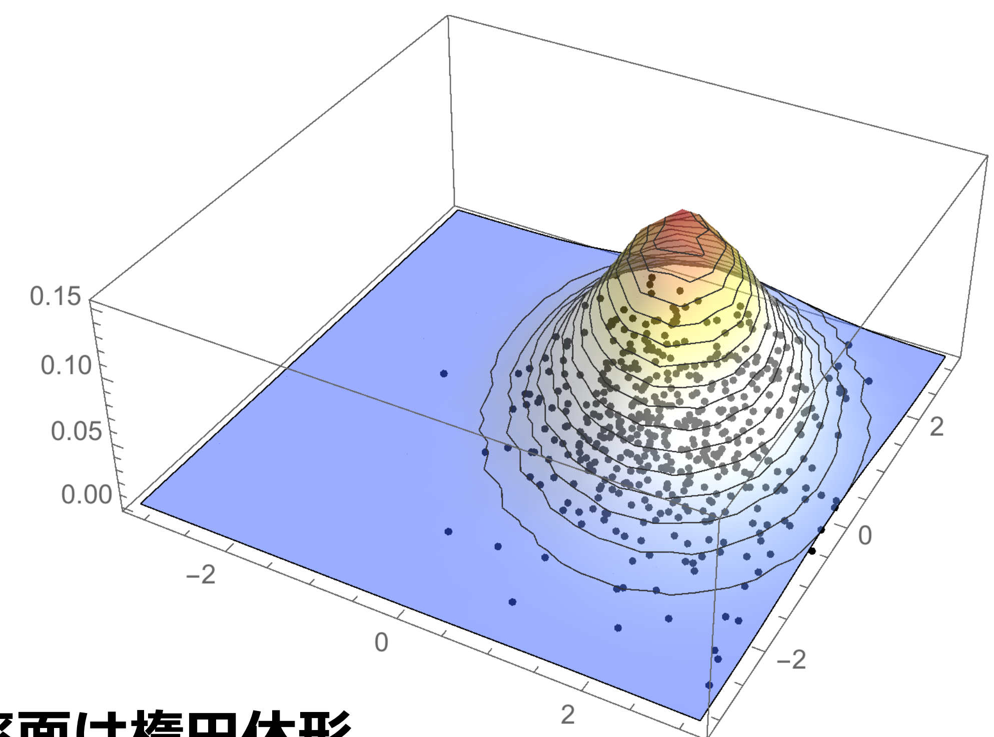
$N(\mathbf{0}, \boldsymbol{I})$

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

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

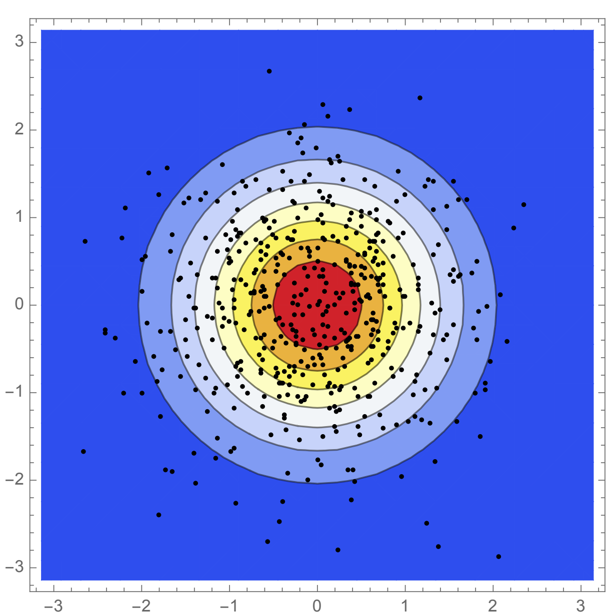
$$\frac{1}{(\sqrt{2\pi})^n\sqrt{\det(\boldsymbol{\Sigma})}}\exp\left(-\frac{1}{2}(\boldsymbol{x}-\boldsymbol{\mu})'\boldsymbol{\Sigma}^{-1}(\boldsymbol{x}-\boldsymbol{\mu})\right)$$





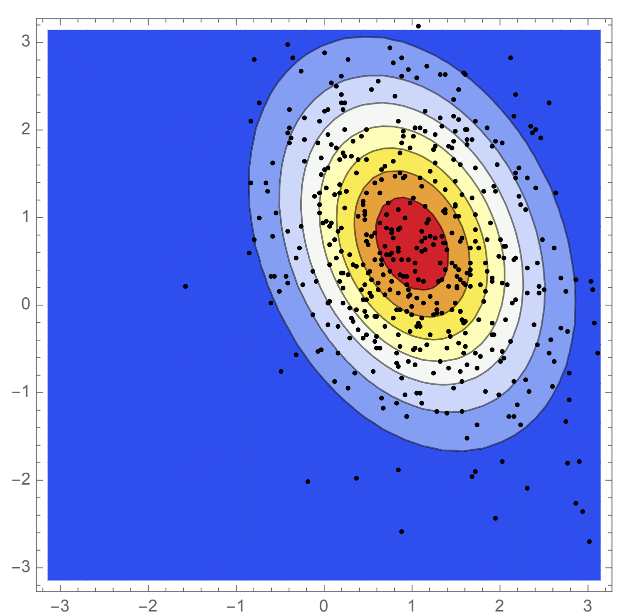
等確率面は球形

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



基本的には「基底変換(しかも基底は直交)」の問題

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

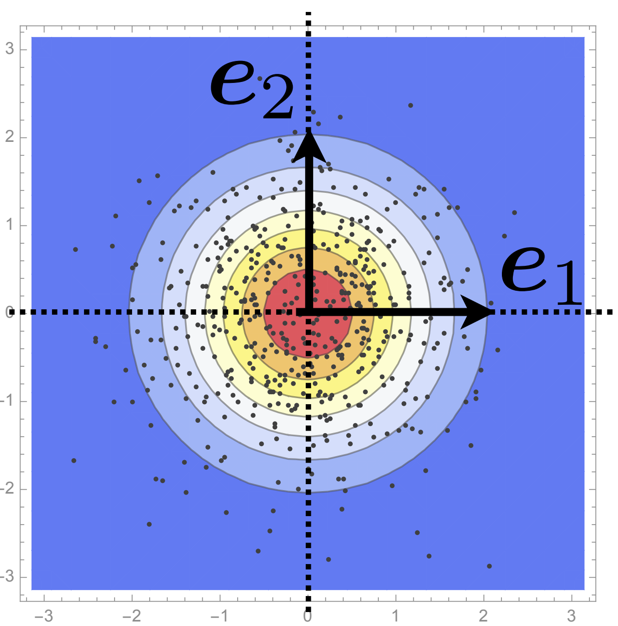
$$\frac{1}{(\sqrt{2\pi})^n} \exp\left(-\frac{1}{2}\boldsymbol{x}'\boldsymbol{x}\right)$$

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

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

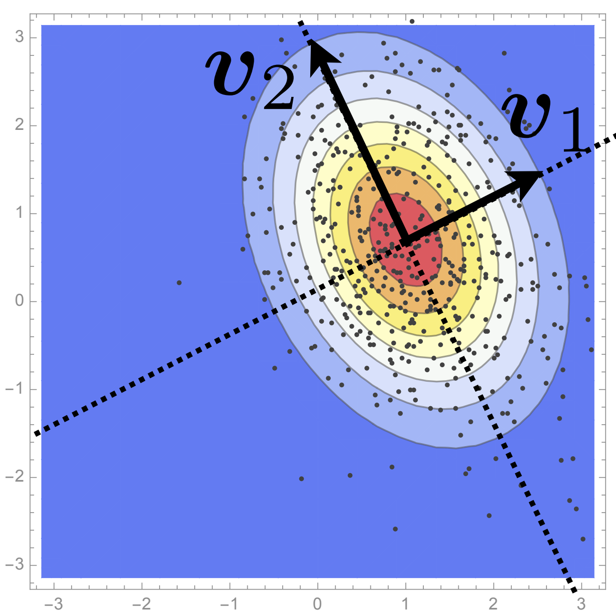
等確率面は球形

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



基底変換(しかも基底は直交)

この基底変換で
密度関数はどう変わる?

単位直交基底 $\boldsymbol{e}_1, \boldsymbol{e}_2, \ldots, \boldsymbol{e}_n$

新しい直交基底 $\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_n$

基底変換のところがミソ

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

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

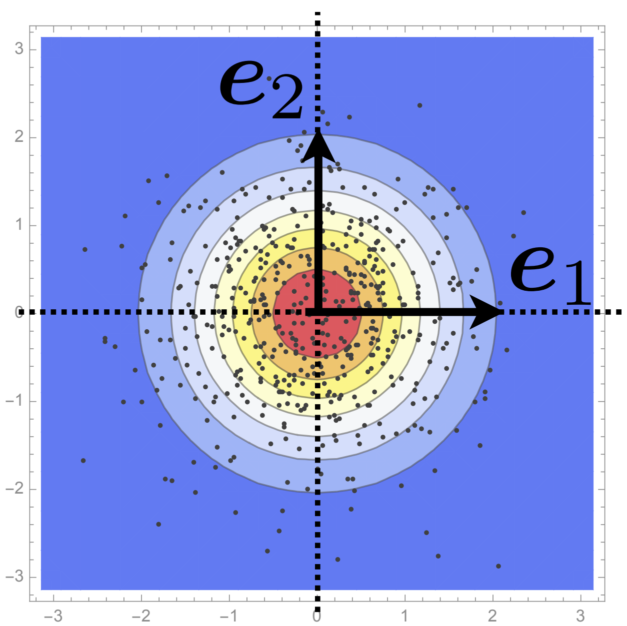
$N(\mathbf{0}, \boldsymbol{\Sigma})$

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

変わる！

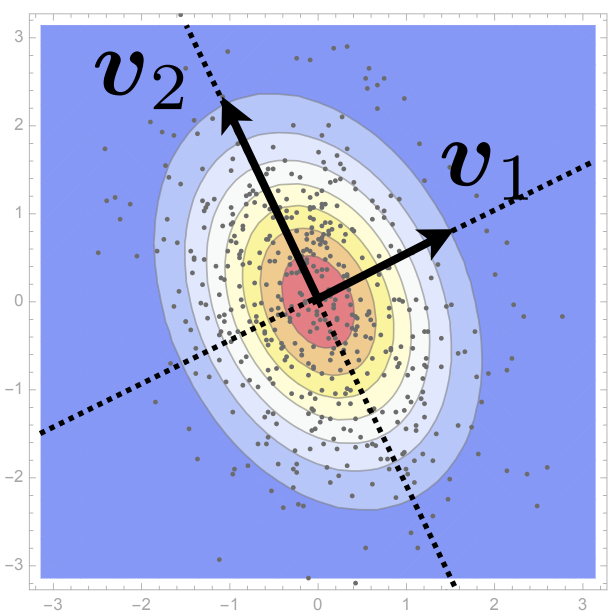
等確率面は球形

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

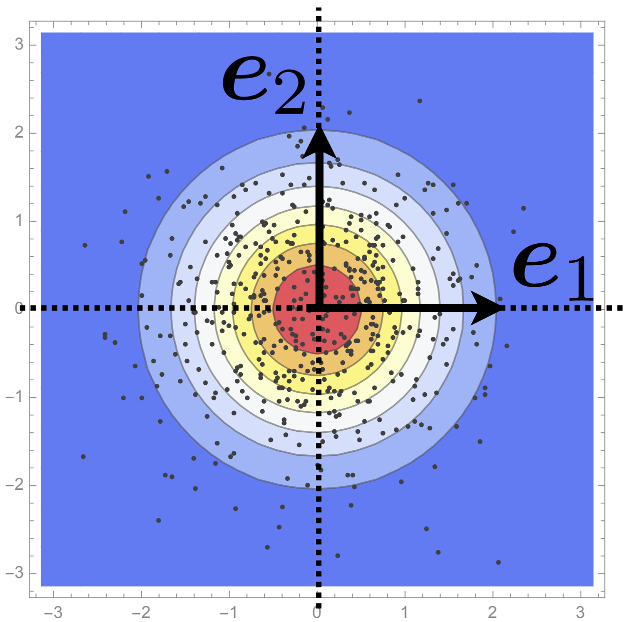


直交する基底への変換のみ

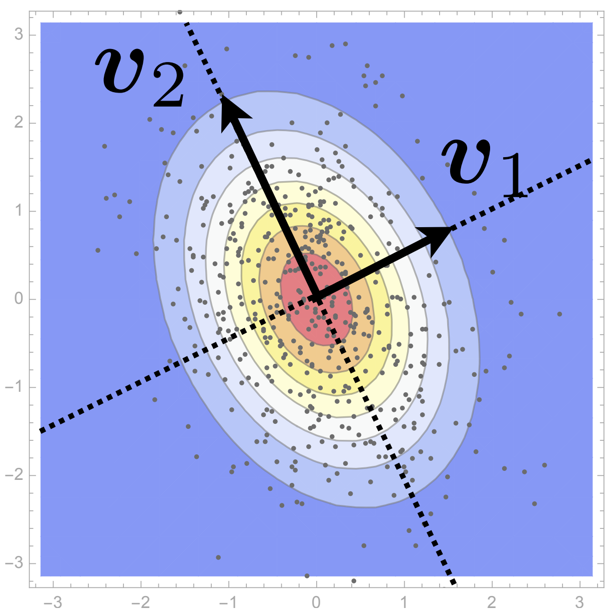
分散共分散行列 $\boldsymbol{\Sigma}$
がこれを決める！

分布がひしゃげたり、形が変わるので
積分値=1にする正規化定数も変わる！

標準化変換



S



$$\boldsymbol{Z} = S(\boldsymbol{X} - \boldsymbol{\mu})$$

平均0,分散I

$$\mathbb{E}\{\boldsymbol{Z}\} = \boldsymbol{0}$$

$$\mathbb{E}\{\boldsymbol{Z}\boldsymbol{Z}'\} = S\Sigma S' = I$$

$$\Leftrightarrow \Sigma = (S^{-1})(S^{-1})'$$

$\boldsymbol{Z}$ **標準正規分布**

$T\boldsymbol{Z}$ **標準正規分布**

直交変換Tをかけても
分布は標準正規分布のまま！

$\Sigma = P\Lambda P'$ **固有値分解**

$\downarrow$

一意ではない→

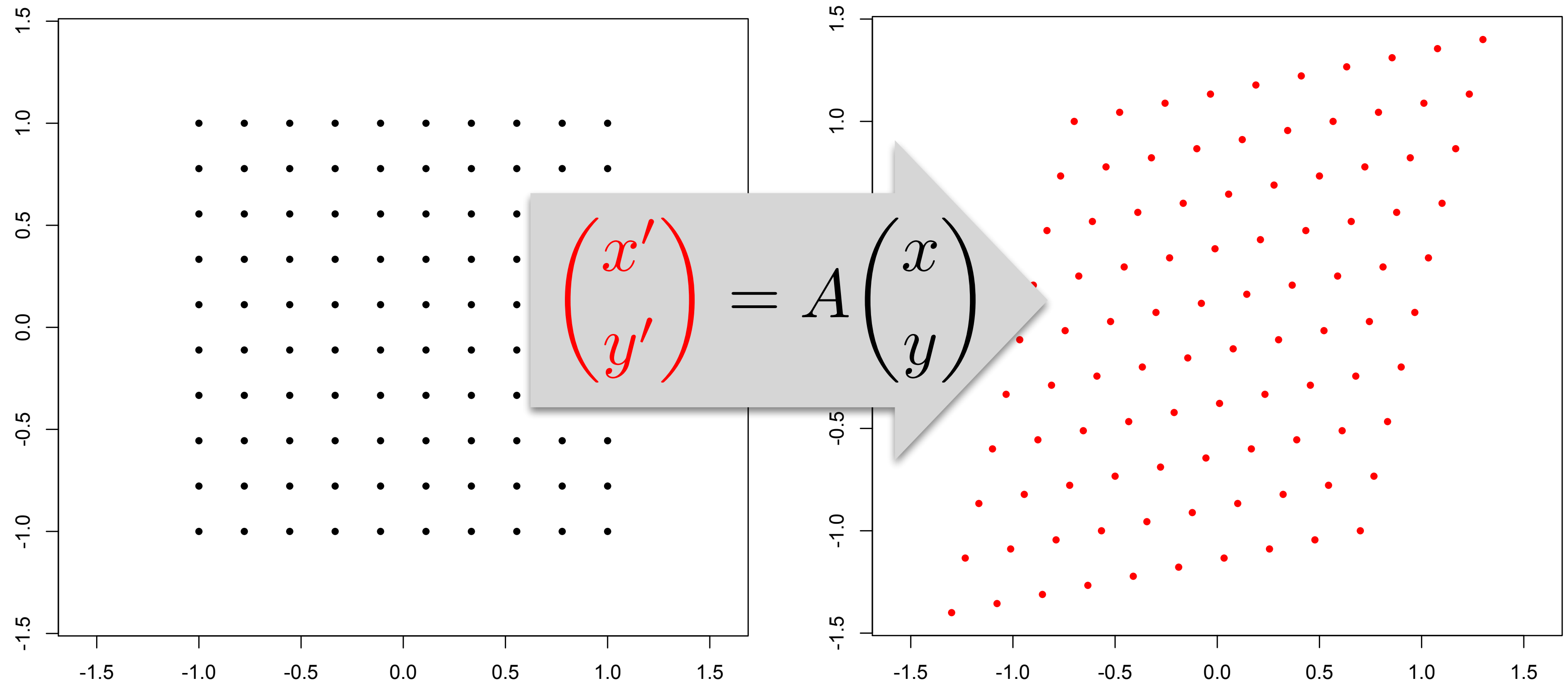
$S^{-1} = P\Lambda^{1/2}$ **非対称**

$S^{-1} = P\Lambda^{1/2}P'$ **平方根行列(対称)**

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



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

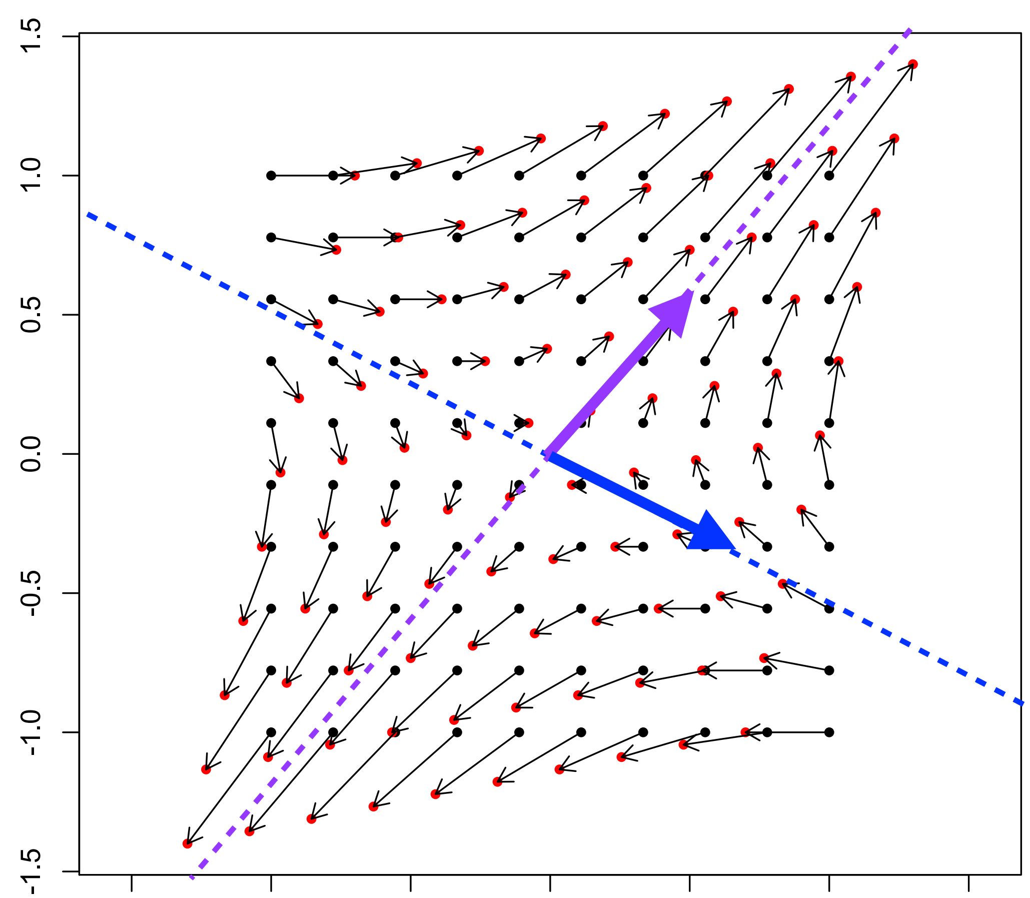
$$A \begin{pmatrix} v_1 \\ v_2 \end{pmatrix} = \lambda_1 \begin{pmatrix} v_1 \\ v_2 \end{pmatrix}$$

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

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

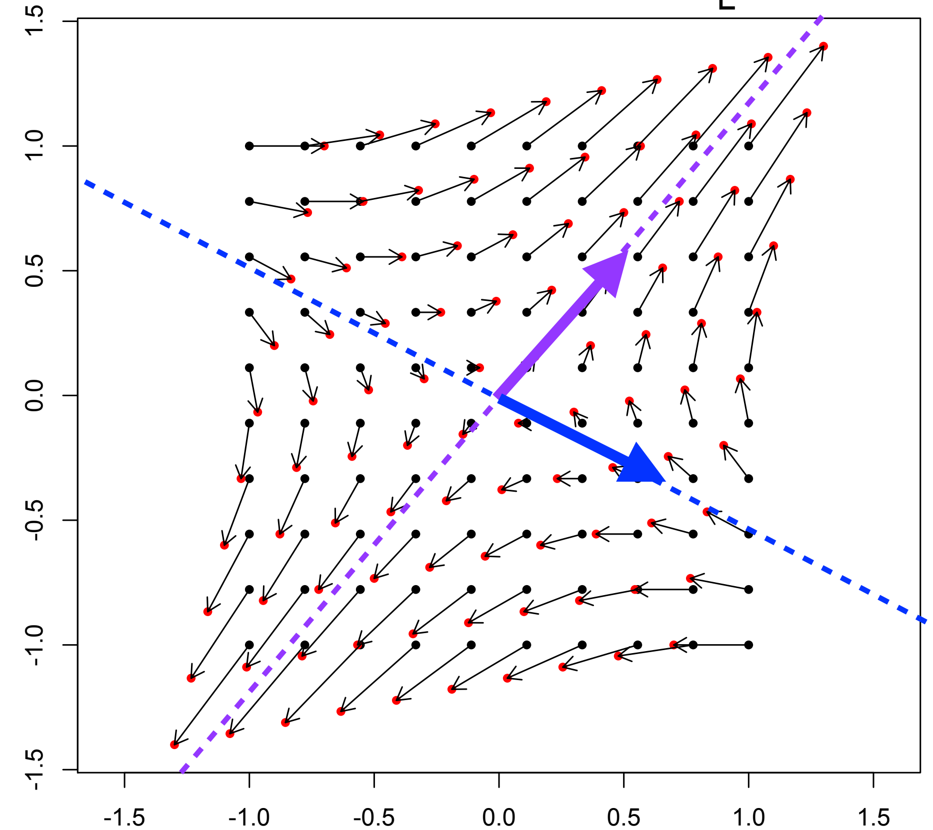
$$A = \begin{bmatrix} v_1 & u_1 \\ v_2 & u_2 \end{bmatrix}\begin{bmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{bmatrix}\begin{bmatrix} v_1 & u_1 \\ v_2 & u_2 \end{bmatrix}^{-1}$$

対称行列なら固有ベクトルは直交

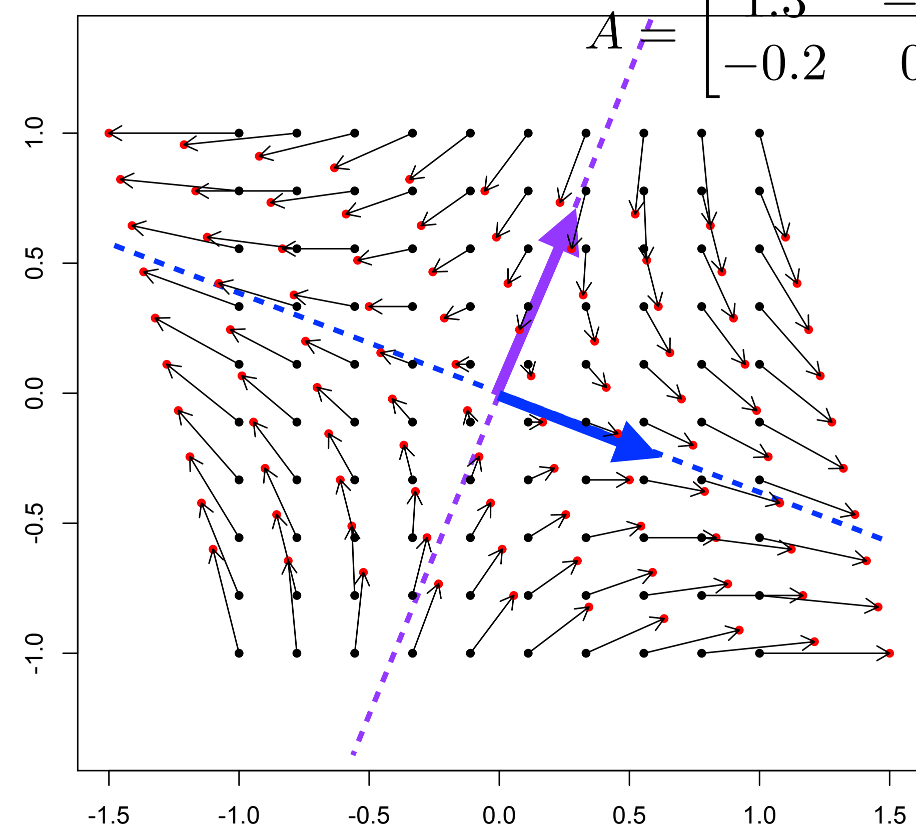
$$\begin{bmatrix} v_1 & u_1 \\ v_2 & u_2 \end{bmatrix}'\begin{bmatrix} v_1 & u_1 \\ v_2 & u_2 \end{bmatrix} = \begin{bmatrix} v_1 & v_2 \\ u_1 & u_2 \end{bmatrix}\begin{bmatrix} v_1 & u_1 \\ v_2 & u_2 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

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

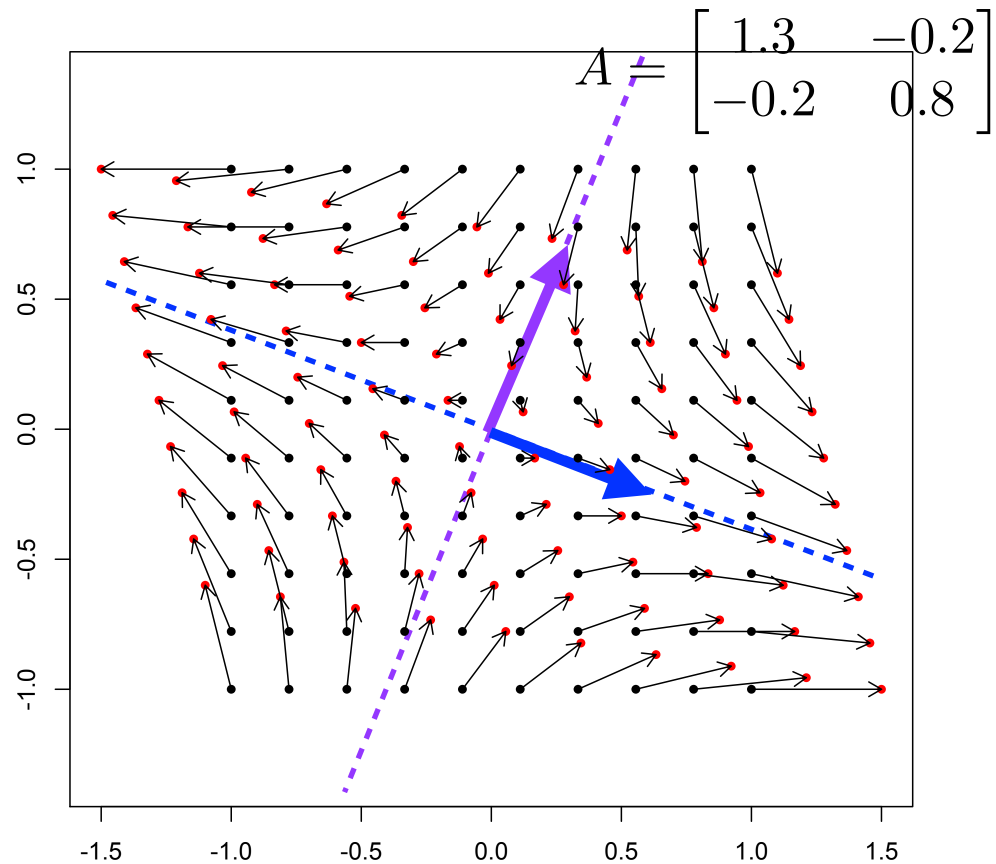
$$A = \begin{bmatrix} 1.0 & 0.3 \\ 0.2 & 1.2 \end{bmatrix}$$



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



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



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

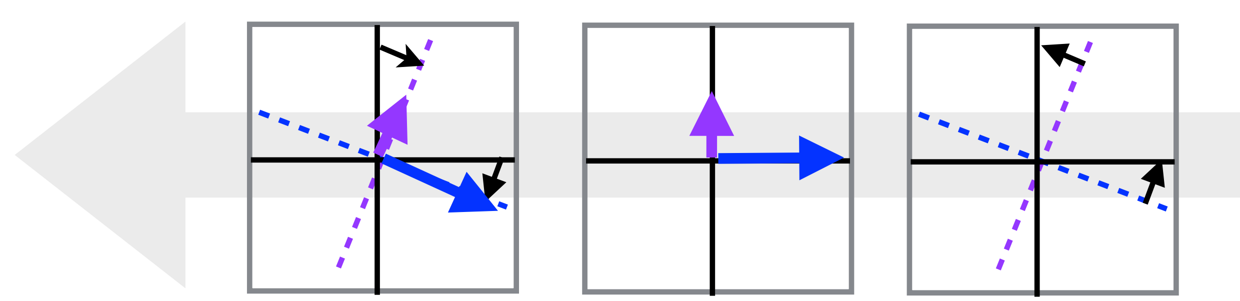
固有値分解の直感的意味

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

基底変換 $\to \begin{bmatrix} v_1 \\ v_2 \end{bmatrix}, \begin{bmatrix} u_1 \\ u_2 \end{bmatrix}$

倍率

逆基底変換 $\to \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \end{bmatrix}$



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

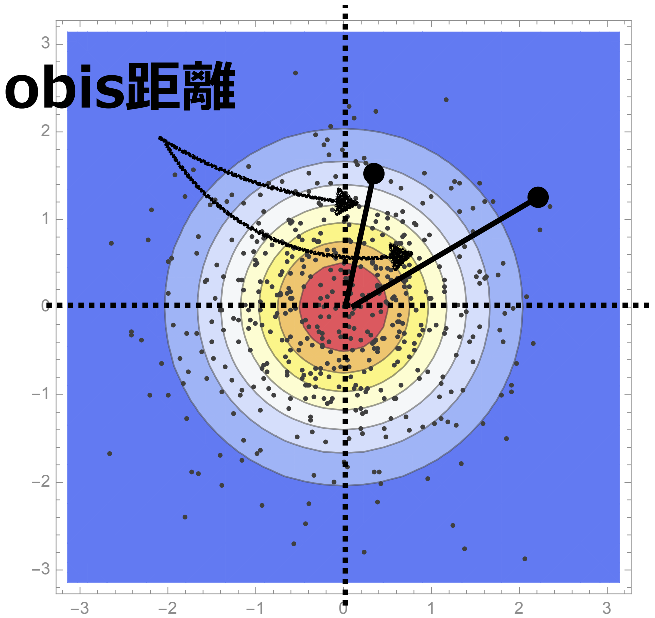
Zの分散共分散行列=単位行列

マハラノビス距離：分散共分散行列の二次形式

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

標準化した後のベクトルの大きさ

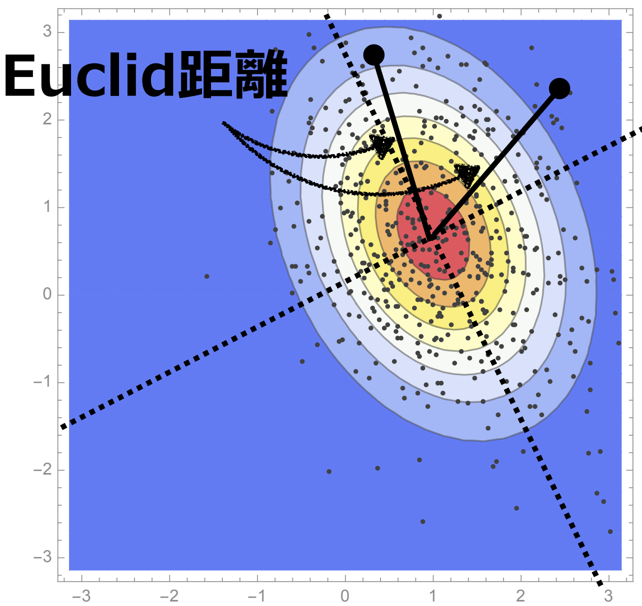
Mahalanobis距離



変数変換 (標準化)

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

Euclid距離

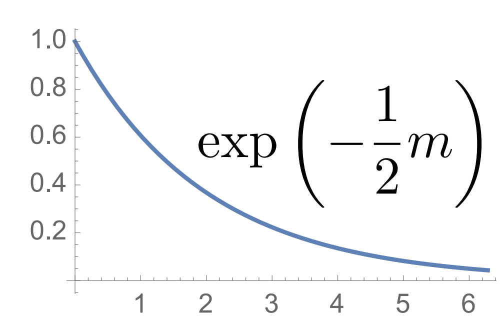


多変量正規分布の密度関数 $N(\boldsymbol{\mu}, \boldsymbol{\Sigma})$

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

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

分散共分散行列の平方根行列

$\Sigma = P\Lambda P' \quad P'P = I$　**直交行列による固有値分解**

$$\Sigma^{1/2} = P\Lambda^{1/2}P' \iff \Sigma^{-1/2} = P\Lambda^{-1/2}P'$$
$$\Lambda^{1/2} = \mathrm{diag}(\sqrt{\lambda_1}, \sqrt{\lambda_2}, \cdots, \sqrt{\lambda_n})$$

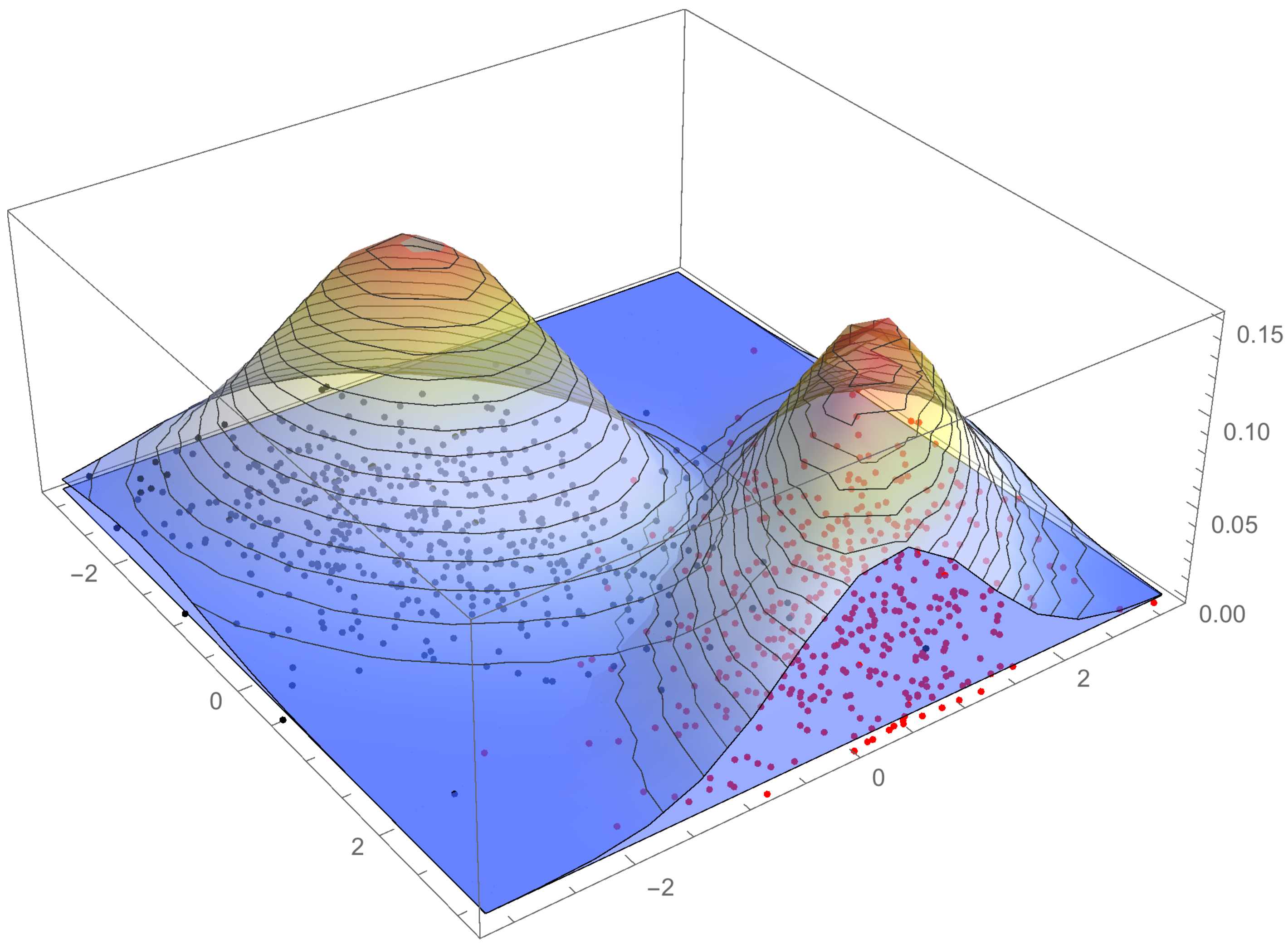
実は $A = (P\Lambda^{1/2})^{-1} = \Lambda^{-1/2}P'$ でも標準化可能

$$\mathbb{E}\{A(\boldsymbol{X}-\boldsymbol{\mu})(\boldsymbol{X}-\boldsymbol{\mu})A'\}$$
$$= A\Sigma A' = (\Lambda^{-1/2}P')(P\Lambda P')P\Lambda^{-1/2} = I$$

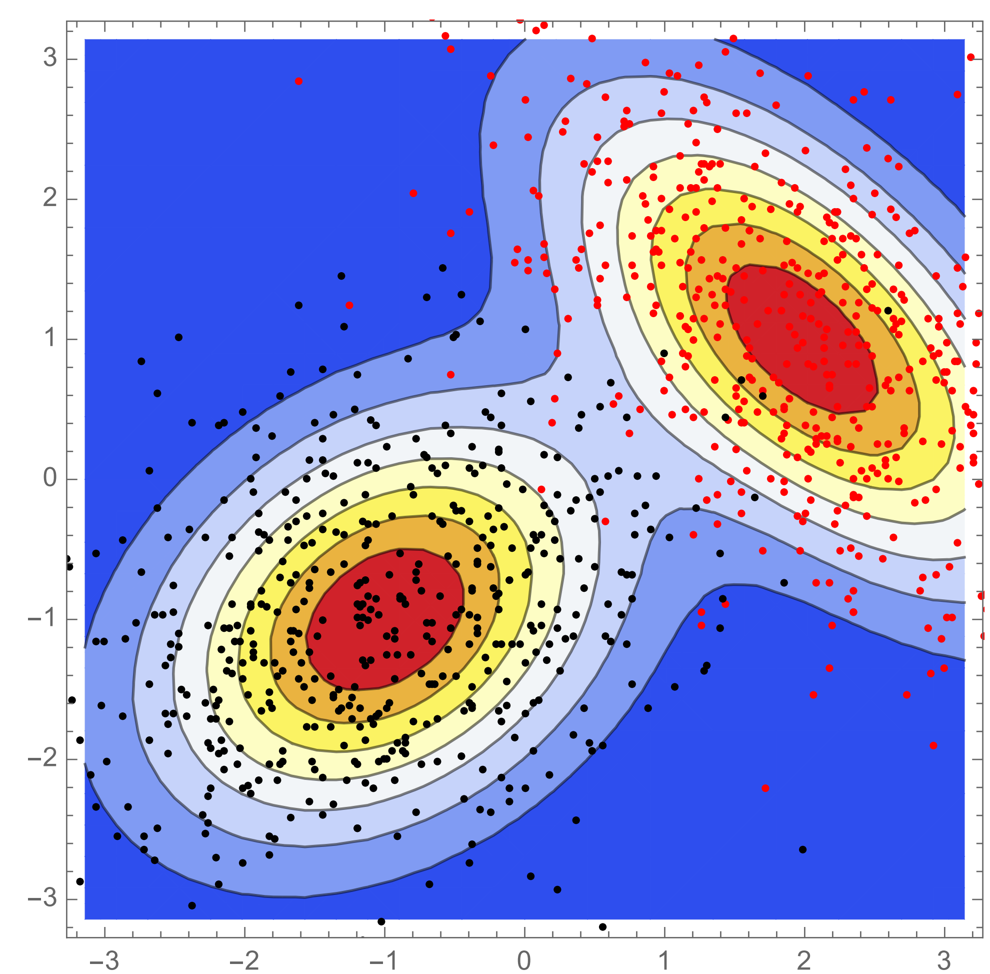
しかしこの行列は非対称であり、また意味合いとして直交変換1つ分足りない。標準正規分布に従う確率変数は直交変換をかけても標準正規分布のままなことに注意

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

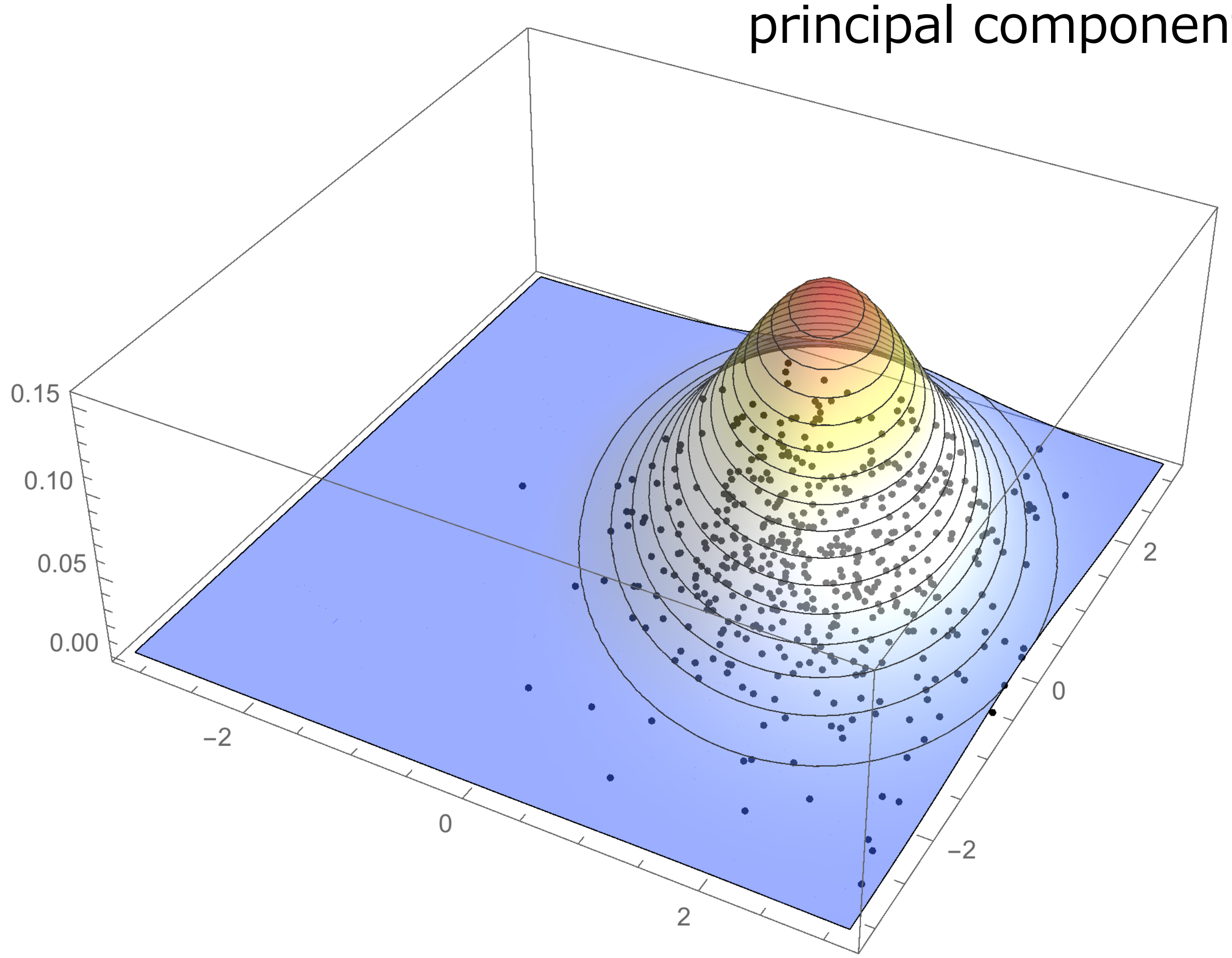


検査点xと各々の群の平均ベクトルとのマハラノビス距離を計算し、最も距離が近いものを結論とする

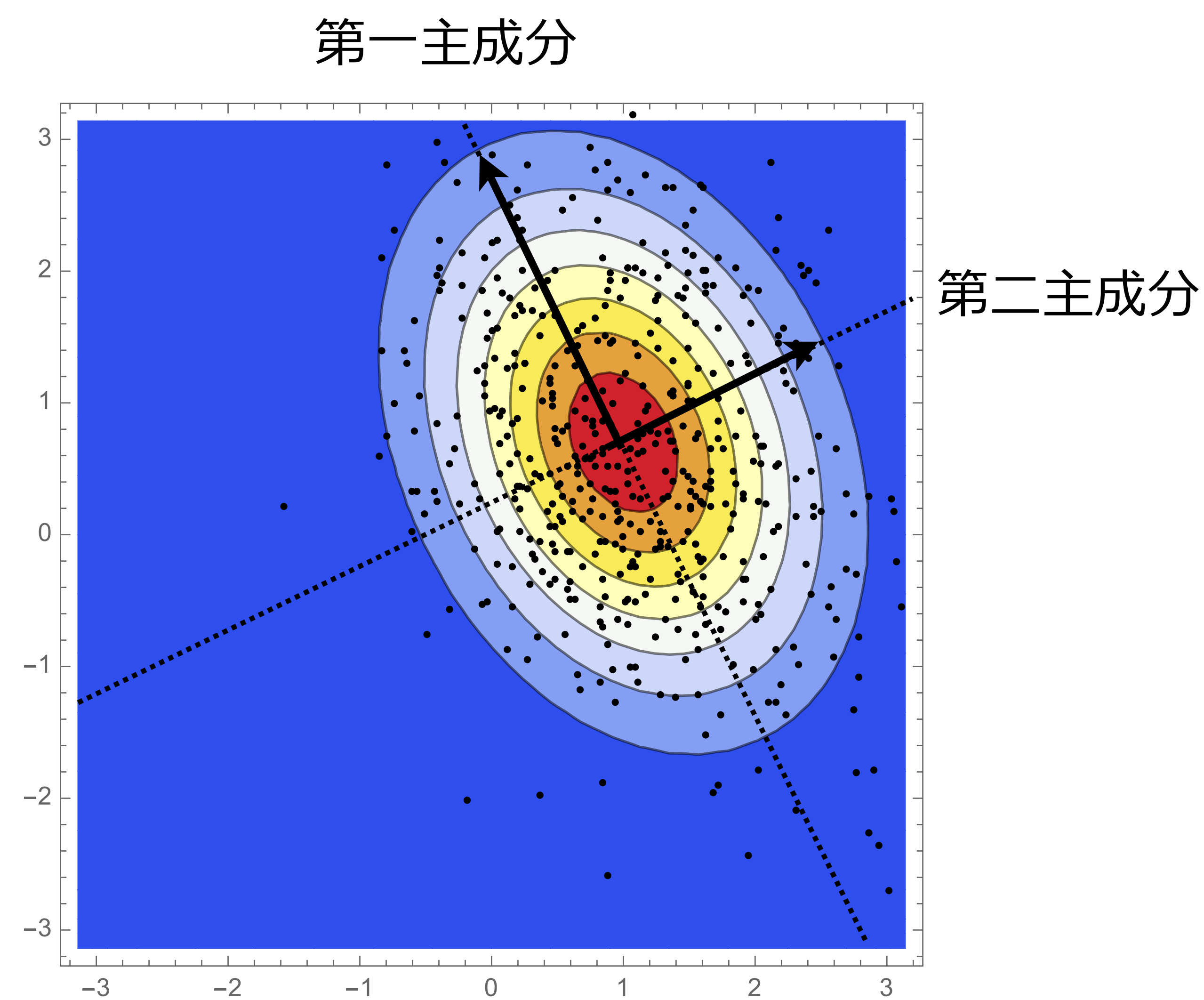
(パターン認識で、クラスの事前確率が等しいケースでのpluginベイズの二次判別)

主成分分析 (主軸変換)

principal component analysis (PCA)



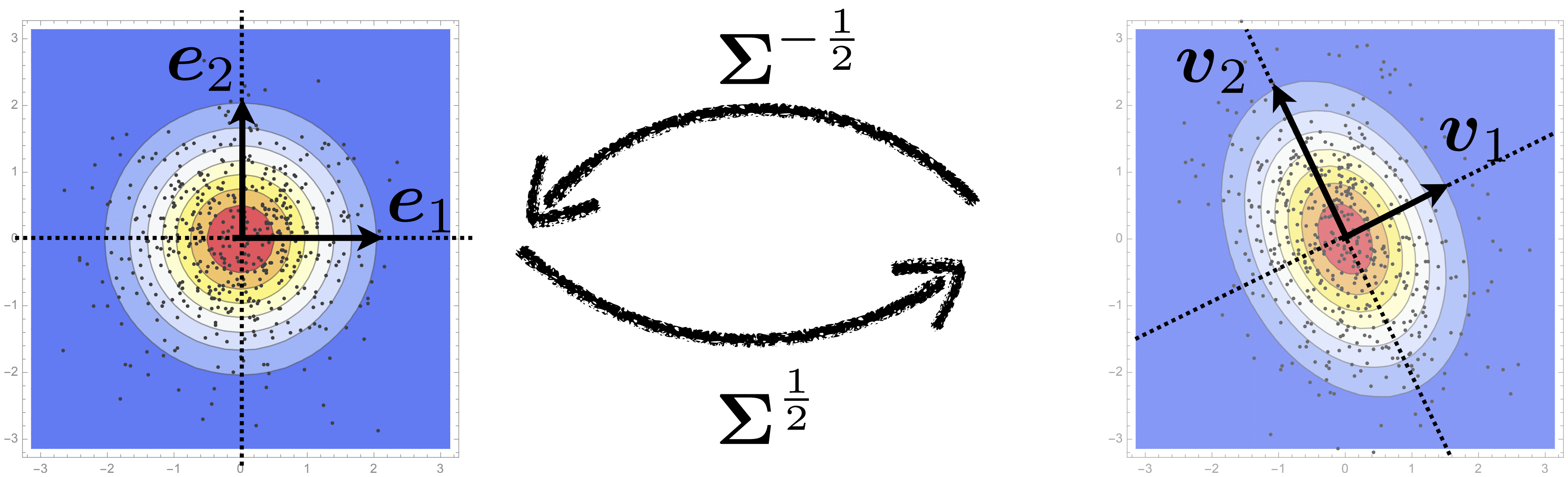
データが多変量正規的に分布していると仮定して、主軸を各々の楕円軸に取り直し、軸が長い順に、各軸での座標値を第一主成分、第二主成分、…として分析する。



高次元のデータになれば、実質的なデータの分布は低次元の部分空間にのみ存在することにより、主成分表示を適当に上位kで打ち切ることにより次元削減ができる。

(パターン認識ではKarhunen-Loeve展開、KL展開による次元削減と呼ばれる)

基底変換や軸方向の計算 = 固有値・固有ベクトル！



以下の線形代数の内容を復習のこと

- 線形写像の表現としての行列
- 行列の平方根、行列の相似、基底変換、直交変換
- 固有値、固有ベクトル、対角化、三角化
- 実対称行列のスペクトル分解 (Rank-1行列分解)
- 二次形式の最大・最小と固有値・固有ベクトル

$\boldsymbol{\Sigma} = P\Lambda P'$ **固有値分解**

$P'P = I$ **直交行列**

$\Lambda := \mathrm{diag}(\lambda_1, \lambda_2, \ldots, \lambda_n)$

$\boldsymbol{\Sigma}^{1/2} = P\Lambda^{1/2}P'$ **平方根行列**

$\Lambda^{1/2} := \mathrm{diag}(\sqrt{\lambda_1}, \sqrt{\lambda_2}, \ldots, \sqrt{\lambda_n})$

重要　**分散共分散行列Σは、正方、実対称、半正定値な行列**

(従って、全ての固有値≧0で一意な平方根行列を持ち常に直交行列による対角化が可能。半正定値なら対称、対称なら正方だけど)

授業計画

Prologue: データを読み解くとは何なのか
(多変量解析とデータサイエンスと統計学とパターン認識と機械学習とデータマイニング)

DAY-1 6/16 (01)(02) **単回帰: 点群への直線当てはめを"真剣に"考える**
(見えない世界へようこそ)
DAY-2 6/23 (03)(04) **重回帰と線形代数: 回帰の行列計算とその意味**
(データの計算とデータの解釈)
DAY-3 6/30 (05)(06) **重回帰と確率統計: なぜ回帰に確率が必要?**
(推測統計入門:データの向こう側について語るための代償)
DAY-4 7/07 (07)(08) **多変量正規分布: 多次元の正規分布と線形代数**
(ゼロから理解する正規分布)
DAY-5 7/14 (09)(10) **マハラノビス距離と判別分析: 線形代数を使う1**
(最適な判別とは)
DAY-6 7/28 (11)(12) **固有値分解と主成分分析: 線形代数を使う2**
(高次元データがかかえる大問題)
DAY-7 8/04 (13)(14) **特異値分解と数量化: 線形代数を使う3**
(数値じゃない対象に統計を効かすには)

Epilogue: 基礎の上に在る世界(話したことと話さなかったこと)