行列輪講: 第7回 ガウス分布1

杉浦 圭祐

慶應義塾大学理工学部情報工学科 松谷研究室

September 30, 2023

目次

🕕 概要

- ② ガウス分布 (1 次元)
 - ガウス分布のモーメント母関数 (1 次元)

- ③ 多変量ガウス分布
 - 多次元ガウス分布のモーメント母関数

目次

- 1 概要
- ② ガウス分布 (1 次元)
- ③ 多変量ガウス分布

このスライドの概要

- ガウス分布について確認する
 - 基本的な事項
 - モーメント (平均, 分散)
- 以下の資料を参考に作成しました:
 - パターン認識と機械学習 (上巻)
 - State Estimation For Robotics
- 重要な分布なので、考えることがたくさんある

目次

- 1 概要
- ② ガウス分布 (1 次元)
- 多変量ガウス分布

ガウス分布 (正規分布) (1 次元)

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

 μ は平均 (Mean), σ^2 は分散 (Variance) とよぶ.

- $p(x \mid \mu, \sigma^2)$ &, $\mathcal{N}(x \mid \mu, \sigma^2)$ とかくこともある.
- 確率変数 x が、平均 μ 、分散 σ^2 のガウス分布に従うとき、次のようにかく:

$$x \sim \mathcal{N}(\mu, \sigma^2)$$

標準正規分布 (1 次元)

平均 $\mu=0$, 分散 $\sigma^2=1$ としたとき, 標準正規分布とよぶ.

$$\mathcal{N}(x \mid 0, 1) = \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{1}{2}x^2\right)$$

ガウス分布 $\mathcal{N}(x\mid\mu,\sigma^2)$ が次を満たすことを, ガウス積分を使って確かめよう:

$$\int_{-\infty}^{\infty} \mathcal{N}(x \mid \mu, \sigma^2) \, \mathrm{d}x = 1$$
$$\mathbb{E}[x] = \mu$$
$$\mathbb{E}[(x - \mu)^2] = \operatorname{Var}[x] = \sigma^2$$

ガウス積分の公式

ガウス積分の公式 $(-\infty$ から $\infty)$

$$\int_{-\infty}^{\infty} \exp\left(-ax^2\right) dx = \sqrt{\frac{\pi}{a}} \quad (a > 0)$$

$$\int_{-\infty}^{\infty} x \exp\left(-ax^2\right) dx = 0 \quad (a > 0)$$

$$\int_{-\infty}^{\infty} x^2 \exp\left(-ax^2\right) dx = \frac{1}{2} \sqrt{\frac{\pi}{a^3}} \quad (a > 0)$$

$$\int_{-\infty}^{\infty} x^3 \exp\left(-ax^2\right) dx = 0 \quad (a > 0)$$

$$\int_{-\infty}^{\infty} x^4 \exp\left(-ax^2\right) dx = \frac{3}{4} \sqrt{\frac{\pi}{a^5}} \quad (a > 0)$$

以下の積分を確認する:

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

• $y = \frac{x - \mu}{\sqrt{2\sigma^2}}$ とすると,

$$x = \sqrt{2\sigma^2}y + \mu \longrightarrow \frac{\mathrm{d}x}{\mathrm{d}y} = \sqrt{2\sigma^2}$$

• 積分を変数変換すると, $x\in (-\infty,\infty)$ のとき $y\in (-\infty,\infty)$ だから,

$$\int_{-\infty}^{\infty} \exp(-y^2) \frac{dx}{dy} dy = \sqrt{2\sigma^2} \underbrace{\int_{-\infty}^{\infty} \exp(-y^2) dy}_{=\sqrt{\pi}}$$
$$= \sqrt{2\sigma^2} \sqrt{\pi} = \sqrt{2\pi\sigma^2}$$

よって、

$$\int_{-\infty}^{\infty} \mathcal{N}(x \mid \mu, \sigma^2) dx$$

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

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

$$= \frac{1}{\sqrt{2\pi\sigma^2}} \sqrt{2\pi\sigma^2} = 1$$

• 積分すると 1 になるから、ガウス分布 $\mathcal{N}(x \mid \mu, \sigma^2)$ は確率分布である.

ガウス分布の平均(1次元)

● 続いて, 平均が, 以下を満たすことを確かめよう:

$$\mathbb{E}\left[x\right] = \mu$$

• ただし,

$$\mathbb{E}[x] = \int_{-\infty}^{\infty} x \mathcal{N}(x \mid \mu, \sigma^2) dx$$
$$= \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi\sigma^2}} x \exp\left(-\frac{1}{2\sigma^2} (x - \mu)^2\right) dx$$

次の積分に着目する:

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



ガウス分布の平均 (1 次元)

• 次のように分解する:

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

• 第1項は、
$$y=rac{x-\mu}{\sqrt{2\sigma^2}}$$
 とすると、 $x=\sqrt{2\sigma^2}y+\mu, rac{\mathrm{d}x}{\mathrm{d}y}=\sqrt{2\sigma^2}$ だから、

$$\sqrt{2\sigma^2} \int_{-\infty}^{\infty} y \exp\left(-y^2\right) \frac{\mathrm{d}x}{\mathrm{d}y} \, \mathrm{d}y = 2\sigma^2 \underbrace{\int_{-\infty}^{\infty} y \exp\left(-y^2\right) \mathrm{d}y}_{=0} = 0$$

- $ullet y \exp\left(-y^2
 ight)$ は奇関数だから, $-\infty$ から ∞ まで積分すると 0.
- 第 2 項は, $\mu\sqrt{2\pi\sigma^2}$.

ガウス分布の平均 (1次元)

よって、

$$\mathbb{E}[x] = \int_{-\infty}^{\infty} x \mathcal{N}(x \mid \mu, \sigma^2) \, \mathrm{d}x$$

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

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

$$= \frac{1}{\sqrt{2\pi\sigma^2}} \mu \sqrt{2\pi\sigma^2}$$

$$= \frac{1}{\sqrt{2\pi\sigma^2}} \mu \sqrt{2\pi\sigma^2} = \mu$$

• ガウス分布の平均は, $\mathbb{E}\left[x\right]=\mu$.

ガウス分布の分散 (1次元)

続いて、分散が、以下を満たすことを確かめよう:

$$\mathbb{E}\left[(x-\mu)^2\right] = \operatorname{Var}\left[x\right] = \sigma^2$$

ただし、

$$\mathbb{E}\left[(x-\mu)^2\right] = \int_{-\infty}^{\infty} (x-\mu)^2 \mathcal{N}(x \mid \mu, \sigma^2) \, \mathrm{d}x$$
$$= \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi\sigma^2}} (x-\mu)^2 \exp\left(-\frac{1}{2\sigma^2} (x-\mu)^2\right) \, \mathrm{d}x$$

次の積分に着目する:

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

ガウス分布の分散 (1 次元)

次の積分に着目する:

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

• 練習問題.

ガウス分布の分散 (1 次元)

よって、

$$\mathbb{E}\left[(x-\mu)^2\right] = \int_{-\infty}^{\infty} (x-\mu)^2 \mathcal{N}(x \mid \mu, \sigma^2) \, \mathrm{d}x$$

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

$$= \frac{1}{\sqrt{2\pi\sigma^2}} \underbrace{\int_{-\infty}^{\infty} (x-\mu)^2 \exp\left(-\frac{1}{2\sigma^2} (x-\mu)^2\right) \, \mathrm{d}x}_{\sigma^2 \sqrt{2\pi\sigma^2}}$$

$$= \frac{1}{\sqrt{2\pi\sigma^2}} \sigma^2 \sqrt{2\pi\sigma^2} = \sigma^2$$

• ガウス分布の分散は、 $\mathbb{E}\left[(x-\mu)^2\right] = \operatorname{Var}\left[x\right] = \sigma^2$.

ガウス分布の平均と分散 (1 次元)

ガウス分布の平均と分散 (1 次元)

$$\mathcal{N}(x \mid \mu, \sigma^2) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{1}{2\sigma^2} (x - \mu)^2\right)$$

このガウス分布の平均と分散は、

$$\mathbb{E}\left[x\right] = \mu$$

$$\mathbb{E}\left[(x - \mu)^2\right] = \operatorname{Var}\left[x\right] = \sigma^2$$

- ガウス積分を使えば、 $\mathbb{E}\left[(x-\mu)^3\right]$ や $\mathbb{E}\left[(x-\mu)^4\right]$ も同様に求まる.
- モーメント母関数 (積率母関数) を使えば、もっと楽に求められる.

目次

- ② ガウス分布 (1 次元)
 - ガウス分布のモーメント母関数 (1 次元)

モーメント母関数 (Moment-generating Function) (1 次元)

モーメント母関数

x の確率分布 p(x) があるとき, $\mathbb{E}\left[\exp(tx)\right]$ を, モーメント母関数という.

$$\mathbb{E}\left[\exp(tx)\right] = \mathbb{E}\left[1 + tx + \frac{1}{2!}t^2x^2 + \frac{1}{3!}t^3x^3 + \cdots\right]$$

t で n 回微分して, t=0 を代入すると, $\mathbb{E}\left[x^n\right]$ が得られる.

$$\mathbb{E}[x^n] = \left. \frac{\partial^n}{\partial t^n} \mathbb{E}[\exp(tx)] \right|_{t=0}$$

ガウス分布 $\mathcal{N}(x \mid \mu, \sigma^2)$ に対するモーメント母関数は,

$$\mathbb{E}\left[\exp(tx)\right] = \exp\left(\mu t + \frac{\sigma^2 t^2}{2}\right)$$

モーメント母関数 (Moment-generating Function) (1 次元)

● モーメント母関数:

$$\mathbb{E}\left[\exp(tx)\right] = \mathbb{E}\left[1 + tx + \frac{1}{2!}t^2x^2 + \frac{1}{3!}t^3x^3 + \cdots\right]$$
$$= 1 + \mathbb{E}\left[x\right]t + \frac{1}{2!}\mathbb{E}\left[x^2\right]t^2 + \frac{1}{3!}\mathbb{E}\left[x^3\right]t^3 + \cdots$$

• 例えば、上式を t で n=3 回微分すれば、

$$\frac{\partial^3}{\partial t^3} \mathbb{E}\left[\exp(tx)\right] = \mathbb{E}\left[x^3\right] + \mathbb{E}\left[x^4\right]t + \frac{1}{2!} \mathbb{E}\left[x^5\right]t^2 + \frac{1}{3!} \mathbb{E}\left[x^6\right]t^3 + \cdots$$

t = 0 を代入すれば,

$$\left. \frac{\partial^3}{\partial t^3} \, \mathbb{E} \left[\exp(tx) \right] \right|_{t=0} = \mathbb{E} \left[x^3 \right]$$

◆ロト ◆問 ト ◆ 恵 ト ◆ 恵 ・ 釣 へ ②

ガウス分布に対するモーメント母関数 (1 次元)

• ガウス分布 $\mathcal{N}(x\mid\mu,\sigma^2)$ に対する, モーメント母関数 $\mathbb{E}\left[\exp(tx)\right]$ を求めてみよう.

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

• exp(·) の中身を, 平方完成する:

$$tx - \frac{1}{2\sigma^2} (x - \mu)^2 = (練習問題)$$
$$= \left(\mu t + \frac{\sigma^2 t^2}{2}\right) - \frac{1}{2\sigma^2} \left(x - (\mu + \sigma^2 t)\right)^2$$

ガウス分布に対するモーメント母関数 (1 次元)

よって、

$$\mathbb{E}\left[\exp(tx)\right] = \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi\sigma^2}} \exp(tx) \exp\left(-\frac{1}{2\sigma^2} (x-\mu)^2\right) dx$$
$$= \exp\left(\mu t + \frac{\sigma^2 t^2}{2}\right) \frac{1}{\sqrt{2\pi\sigma^2}} \int_{-\infty}^{\infty} \exp\left(-\frac{1}{2\sigma^2} (x - (\mu + \sigma^2 t))^2\right) dx$$

•
$$y=rac{x-\left(\mu+\sigma^2t
ight)}{\sqrt{2\sigma^2}}$$
 とおくと, $x=\sqrt{2\sigma^2}y+\mu+\sigma^2t$ であり, $rac{\mathrm{d}x}{\mathrm{d}y}=\sqrt{2\sigma^2}$ となるから,

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

4□ > 4₫ > 4½ > ½
 9<0

ガウス分布に対するモーメント母関数 (1次元)

• よって, ガウス分布に対するモーメント母関数は,

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

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

$$= \exp\left(\mu t + \frac{\sigma^2 t^2}{2}\right) \frac{1}{\sqrt{2\pi\sigma^2}} \sqrt{2\pi\sigma^2}$$

$$= \exp\left(\mu t + \frac{\sigma^2 t^2}{2}\right)$$

$$= \exp\left(\mu t + \frac{\sigma^2 t^2}{2}\right)$$

このモーメント母関数を使って、モーメントを求めてみよう。

ガウス分布に対するモーメント母関数 (1 次元)

● モーメント母関数を使うと、ガウス分布の平均と分散は、

$$\mathbb{E}\left[x\right] = \frac{\mathrm{d}}{\mathrm{d}t} \exp\left(\mu t + \frac{\sigma^2 t^2}{2}\right) \bigg|_{t=0} = (練習問題) = \mu$$

$$\mathbb{E}\left[x^2\right] = \frac{\mathrm{d}^2}{\mathrm{d}t^2} \exp\left(\mu t + \frac{\sigma^2 t^2}{2}\right) \bigg|_{t=0} = (練習問題) = \sigma^2 + \mu^2$$

$$\operatorname{Var}\left[x\right] = \mathbb{E}\left[x^2\right] - \mathbb{E}\left[x\right]^2 = \sigma^2$$

ガウス分布の高次のモーメント (1次元)

• ガウス分布 $\mathcal{N}(x \mid \mu, \sigma^2)$ のモーメント:

$$\begin{split} \mathbb{E}\left[x\right] &= \mu \\ \mathbb{E}\left[x^2\right] &= \mu^2 + \sigma^2 \\ \mathbb{E}\left[x^3\right] &= \mu^3 + 3\mu\sigma^2 \\ \mathbb{E}\left[x^4\right] &= \mu^4 + 6\mu^2\sigma^2 + 3\sigma^4 \\ \mathbb{E}\left[x^5\right] &= \mu^5 + 10\mu^3\sigma^2 + 15\mu\sigma^4 \\ \mathbb{E}\left[x^6\right] &= \mu^6 + 15\mu^4\sigma^2 + 45\mu^2\sigma^4 + 15\sigma^6 \\ \mathbb{E}\left[x^7\right] &= \mu^7 + 21\mu^5\sigma^2 + 105\mu^3\sigma^4 + 105\mu\sigma^6 \\ \mathbb{E}\left[x^8\right] &= \mu^8 + 28\mu^6\sigma^2 + 210\mu^4\sigma^4 + 420\mu^2\sigma^6 + 105\sigma^8 \\ \mathbb{E}\left[x^9\right] &= \mu^9 + 36\mu^7\sigma^2 + 378\mu^5\sigma^4 + 1260\mu^3\sigma^6 + 945\mu\sigma^8 \\ \mathbb{E}\left[x^{10}\right] &= \mu^{10} + 45\mu^8\sigma^2 + 630\mu^6\sigma^4 + 3150\mu^4\sigma^6 + 4725\mu^2\sigma^8 + 945\sigma^{10} \end{split}$$

ガウス分布の高次のモーメント (平均まわり; 1次元)

• ガウス分布 $\mathcal{N}(x \mid \mu, \sigma^2)$ の, 平均まわりのモーメント:

$$\mathbb{E}\left[x-\mu\right] = 0, \quad \mathbb{E}\left[\left(x-\mu\right)^2\right] = \sigma^2$$

$$\mathbb{E}\left[\left(x-\mu\right)^3\right] = 0, \quad \mathbb{E}\left[\left(x-\mu\right)^4\right] = 3\sigma^4$$

$$\mathbb{E}\left[\left(x-\mu\right)^5\right] = 0, \quad \mathbb{E}\left[\left(x-\mu\right)^6\right] = 15\sigma^6$$

$$\mathbb{E}\left[\left(x-\mu\right)^7\right] = 0, \quad \mathbb{E}\left[\left(x-\mu\right)^8\right] = 105\sigma^8$$

$$\mathbb{E}\left[\left(x-\mu\right)^9\right] = 0, \quad \mathbb{E}\left[\left(x-\mu\right)^{10}\right] = 945\sigma^{10}$$

- ullet $y=x-\mu$ としたとき, p(y) は平均 0, 分散 σ^2 のガウス分布 (後述).
- よって, $\mathbb{E}[y^n] = \mathbb{E}[(x-\mu)^n]$ は, 先程求めた $\mathcal{N}(x \mid \mu, \sigma^2)$ のモーメント $\mathbb{E}[x^n]$ に, $\mu = 0$ を代入したものとなる.

ガウス分布の高次のモーメント (平均まわり; 1次元)

ullet 平均 μ , 分散 σ^2 のガウス分布を考える:

$$\mathcal{N}(x \mid \mu, \sigma^2) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{1}{2\sigma^2} (x - \mu)^2\right)$$

• $y = x - \mu$ としたとき, p(y) は平均 0, 分散 σ^2 のガウス分布:

$$p(y) = \mathcal{N}(y \mid 0, \sigma^2) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{1}{2\sigma^2}y^2\right)$$

- ガウス分布の線形変換, 周辺分布, 条件付き分布なども, ガウス分布になる (後述).
- 多変量の確率分布についても, 同様の計算ができる (後述).

ガウス分布のモーメント (1 次元)

ガウス分布のモーメント (1次元)

ガウス分布 $\mathcal{N}(x \mid \mu, \sigma^2)$ について,

$$\mathbb{E}\left[x\right] = \mu$$
 (平均), $\mathbb{E}\left[\left(x - \mu\right)^2\right] = \operatorname{Var}\left[x\right] = \sigma^2$ (分散)

$$\mathbb{E}\left[\left(\frac{x-\mu}{\sigma}\right)^3\right] = 0 \quad \textbf{(歪度)}, \quad \mathbb{E}\left[\left(\frac{x-\mu}{\sigma}\right)^4\right] = 3 \quad \textbf{(尖度)}$$

$$\mathbb{E}\left[x^2\right] = \mu^2 + \sigma^2, \quad \mathbb{E}\left[\left(x - \mu\right)^3\right] = 0, \quad \mathbb{E}\left[\left(x - \mu\right)^4\right] = 3\sigma^4$$

高次の平均まわりのモーメントは,

$$\mathbb{E}\left[(x-\mu)^n\right] = \begin{cases} 0 & (n \text{ が奇数})\\ ((n-1)(n-3)(n-5)\cdots 3\cdot 1)\sigma^n & (n \text{ が偶数}) \end{cases}$$

休憩

- ここまで、1次元のガウス分布を確認した.
 - 定義
 - 積分
 - モーメント母関数
 - モーメント (平均, 分散, 歪度, 尖度)
- 既にお腹いっぱい.
- 続いて,多変量ガウス分布を確認する.

目次

- 1 概要
- ② ガウス分布 (1 次元)
- ③ 多変量ガウス分布

多変量ガウス分布

多変量ガウス分布 (多変量正規分布)

D 次元のとき, 次のように定義される:

$$p(\mathbf{x} \mid \boldsymbol{\mu}, \boldsymbol{\Sigma}) = \frac{1}{(2\pi)^{\frac{D}{2}} \sqrt{\det \boldsymbol{\Sigma}}} \exp \left(-\frac{1}{2} (\mathbf{x} - \boldsymbol{\mu})^{\top} \boldsymbol{\Sigma}^{-1} (\mathbf{x} - \boldsymbol{\mu})\right)$$

 μ は平均, Σ は共分散 (分散共分散行列) とよぶ.

- 平均 μ は D 次ベクトル, 共分散 Σ は D × D 行列.
- exp(·) の中身は、スカラー。
- $p(\mathbf{x} \mid \boldsymbol{\mu}, \boldsymbol{\Sigma})$ &, $\mathcal{N}(\mathbf{x} \mid \boldsymbol{\mu}, \boldsymbol{\Sigma})$ とかくこともある.
- 確率変数 x が, 平均 μ , 共分散 Σ の多変量ガウス分布に従うとき, 次のようにかく:

$$\mathbf{x} \sim \mathcal{N}(oldsymbol{\mu}, oldsymbol{\Sigma})$$



多変量ガウス分布

多変量ガウス分布 (多変量正規分布)

$$\mathcal{N}(\mathbf{x} \mid \boldsymbol{\mu}, \boldsymbol{\Sigma}) = \frac{1}{(2\pi)^{\frac{D}{2}} \sqrt{\det \boldsymbol{\Sigma}}} \exp \left(-\frac{1}{2} \left(\mathbf{x} - \boldsymbol{\mu} \right)^{\top} \boldsymbol{\Sigma}^{-1} \left(\mathbf{x} - \boldsymbol{\mu} \right) \right)$$

定数項の分母には、いくつかの書き方がある:

$$(2\pi)^{\frac{D}{2}}\sqrt{\det \Sigma} = \sqrt{(2\pi)^D \det \Sigma} = \sqrt{\det(2\pi\Sigma)}$$

- Σ は D 次行列だから, $\det(c\Sigma) = c^D \det \Sigma$ である.
- det A を, |A| と書くこともある.
- 絶対値との混同を避けるため、ここでは det と表記する。

共分散行列 (分散共分散行列)

- $\mathbf{x} = (x_1, \dots, x_D)$ と表す (D 次元).
- 共分散 ∑ は、次のように表される (D × D 行列):

$$\Sigma = \begin{pmatrix} \operatorname{Var}\left[x_{1}\right] & \operatorname{Cov}\left(x_{1}, x_{2}\right) & \cdots & \operatorname{Cov}\left(x_{1}, x_{D}\right) \\ \operatorname{Cov}\left(x_{2}, x_{1}\right) & \operatorname{Var}\left[x_{2}\right] & \cdots & \operatorname{Cov}\left(x_{2}, x_{D}\right) \\ \vdots & \vdots & \ddots & \vdots \\ \operatorname{Cov}\left(x_{D}, x_{1}\right) & \operatorname{Cov}\left(x_{D}, x_{2}\right) & \cdots & \operatorname{Var}\left[x_{D}\right] \end{pmatrix}$$

- Σ の (i,j) 成分は、 $\operatorname{Cov}(x_i,x_j) = \mathbb{E}[(x_i \mathbb{E}[x_i])(x_j \mathbb{E}[x_j])]$.
- $Cov(x_i, x_j) = Cov(x_j, x_i)$ だから, Σ は対称.
- $\operatorname{Var}[x_i] = \operatorname{Cov}(x_i, x_i)$ に注意.
- ∑の対角成分は分散、それ以外の成分は共分散となる。
- 重要: ${f x}$ の各成分が互いに独立 (無相関) ならば, ${f \Sigma}$ は対角行列となる $(i
 eq j
 ightarrow {
 m Cov}\,(x_i,x_j)=0).$

標準正規分布 (多次元)

標準正規分布 (多次元)

平均 $\mu=0$, 共分散 $\Sigma=\mathrm{I}$ としたとき, 標準正規分布とよぶ.

$$\mathcal{N}(\mathbf{x} \mid \mathbf{0}, \mathbf{I}) = \frac{1}{(2\pi)^{\frac{D}{2}}} \exp\left(-\frac{1}{2}\mathbf{x}^{\top}\mathbf{x}\right)$$
$$= \prod_{i=1}^{D} \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{1}{2}x_i^2\right) = \prod_{i=1}^{D} \mathcal{N}(x_i \mid 0, 1)$$

x の各成分についてのガウス分布に分解される. 従って、各成分は互いに独立で、無相関:

$$\forall i \neq j \ \mathbb{E}[x_i x_j] = \mathbb{E}[x_i] \mathbb{E}[x_j], \operatorname{Cov}(x_i, x_j) = 0$$

各成分は、平均0、分散1のガウス分布に従う: $x_i \sim \mathcal{N}(0,1)$

多変量ガウス分布の共分散

多変量ガウス分布 (多変量正規分布)

$$\mathcal{N}(\mathbf{x} \mid \boldsymbol{\mu}, \boldsymbol{\Sigma}) = \frac{1}{(2\pi)^{\frac{D}{2}} \sqrt{\det \boldsymbol{\Sigma}}} \exp \left(-\frac{1}{2} \left(\mathbf{x} - \boldsymbol{\mu} \right)^{\top} \boldsymbol{\Sigma}^{-1} \left(\mathbf{x} - \boldsymbol{\mu} \right) \right)$$

共分散 ∑ は, 正定値対称行列.

$$oldsymbol{\Sigma}^{ op} = oldsymbol{\Sigma}$$
 $orall \mathbf{x}
eq \mathbf{0}, \ \mathbf{x}^{ op} oldsymbol{\Sigma} \mathbf{x} \geq \mathbf{0} \quad (oldsymbol{\Sigma} \geq 0 \ extbf{ともかく})$

ullet 逆行列 $oldsymbol{\Sigma}^{-1}$ も, 正定値対称行列になる.

$$\left(\mathbf{\Sigma}^{-1}\right)^{\top} = \mathbf{\Sigma}^{-1}$$
 $\forall \mathbf{x} \neq \mathbf{0}, \ \mathbf{x}^{\top}\mathbf{\Sigma}^{-1}\mathbf{x} \geq \mathbf{0} \ \left(\mathbf{\Sigma}^{-1} \geq 0 \ \mathsf{ともかく}\right)$

- 共分散 ∑ は, 正定値対称行列.
- ullet このとき, 逆行列 Σ^{-1} も, 正定値対称行列になる.
- 対称性:

$$\left(\mathbf{\Sigma}^{-1}
ight)^{ op} = \left(\mathbf{\Sigma}^{ op}
ight)^{-1} = \mathbf{\Sigma}^{-1}$$

• 正定値性:

$$m{\Sigma}$$
 が正定値 o $m{\Sigma}$ の固有値 $\{\lambda_i\}$ は全て正 $m{ o}$ $m{\Sigma}^{-1}$ の固有値は $m{\lambda}_i^{-1}$ であるが, それらも全て正 $m{ o}$ $m{\Sigma}^{-1}$ は正定値

- ullet Σ, Σ^{-1} は正定値行列なので、 $rac{f v}{2}, \Sigma^{-rac{1}{2}}$ が存在する (後述).
- 注意: 正定値性は, 対称行列について定義される.
- 注意: 対称行列の固有値は,全て実数となる.

- ullet Σ の固有値が $\{\lambda_1,\ldots,\lambda_D\}$ であるとき, Σ^{-1} の固有値は $\{\lambda_1^{-1},\ldots,\lambda_D^{-1}\}$ となることを確認しよう.
- ullet 実対称行列 Σ は,直交行列 U で, $oldsymbol{\Lambda} = U^{ op} \Sigma U$ と対角化できる.
- \bullet Λ は, Σ の固有値 λ_i を対角成分にもった, 対角行列:

$$\mathbf{\Lambda} = \operatorname{diag}(\lambda_1, \dots, \lambda_D)$$

ullet $oldsymbol{\Lambda}^{-1}$ は, $oldsymbol{\Sigma}$ の固有値の逆数 λ_i^{-1} を対角成分にもった, 対角行列:

$$\boldsymbol{\Lambda}^{-1} = \operatorname{diag}(\lambda_1^{-1}, \dots, \lambda_D^{-1})$$

- $\mathbf{\Lambda}^{-1} = \mathbf{U}^{\top} \mathbf{\Sigma}^{-1} \mathbf{U}$ だから,実対称行列 $\mathbf{\Sigma}^{-1}$ も,直交行列 \mathbf{U} で対角化される $(\mathbf{U}^{-1} = \mathbf{U}$ に注意).
- ullet $oldsymbol{\Lambda}^{-1}$ の対角成分は, $oldsymbol{\Sigma}^{-1}$ の固有値 ($=oldsymbol{\Sigma}$ の固有値の逆数)



- ullet Σ, Σ^{-1} は正定値行列なので、 $rac{f v}{2}, \Sigma^{-1}$ が存在する.
- A を半正定値対称行列とする.
- ullet 対称行列なので、直交行列 $oldsymbol{\mathrm{U}}$ で、 $oldsymbol{\Lambda} = oldsymbol{\mathrm{U}}^{ op} oldsymbol{\mathrm{A}} oldsymbol{\mathrm{U}}$ と対角化できる.
- ullet Λ は、 Λ の固有値 λ_i を対角成分にもった、対角行列:

$$\mathbf{\Lambda} = \operatorname{diag}(\lambda_1, \dots, \lambda_D)$$

ullet A は半正定値だから,固有値は全て ≥ 0 . $oldsymbol{\Lambda}$ には平方根 $oldsymbol{\Lambda}^{rac{1}{2}}$ が存在し,

$$\mathbf{\Lambda}^{\frac{1}{2}} = \operatorname{diag}\left(\sqrt{\lambda_1}, \dots, \sqrt{\lambda_D}\right)$$

 $oldsymbol{A}^{rac{1}{2}} = oldsymbol{U}oldsymbol{\Lambda}^{rac{1}{2}}oldsymbol{U}^ op$ は $oldsymbol{A}$ の平方根となる $(oldsymbol{U}^{-1} = oldsymbol{U}^ op, oldsymbol{U}^ op oldsymbol{U} = oldsymbol{I})$:

$$\mathbf{A}^{\frac{1}{2}}\mathbf{A}^{\frac{1}{2}} = \mathbf{U}\mathbf{\Lambda}^{\frac{1}{2}}\mathbf{U}^{\top}\mathbf{U}\mathbf{\Lambda}^{\frac{1}{2}}\mathbf{U}^{\top} = \mathbf{U}\mathbf{\Lambda}^{\frac{1}{2}}\mathbf{\Lambda}^{\frac{1}{2}}\mathbf{U}^{\top} = \mathbf{U}\mathbf{\Lambda}\mathbf{U}^{\top} = \mathbf{A}$$

- ullet Σ, Σ^{-1} は正定値行列なので、平方根 $\Sigma^{rac{1}{2}}, \Sigma^{-rac{1}{2}}$ が存在する.
- ullet 対角化 $oldsymbol{\Lambda} = oldsymbol{\mathrm{U}}^ op oldsymbol{\Sigma} oldsymbol{\mathrm{U}}$ の結果を使えば,

$$\begin{split} \boldsymbol{\Sigma} &= \mathbf{U}\boldsymbol{\Lambda}\mathbf{U}^{\top}, \quad \boldsymbol{\Sigma}^{-1} &= \mathbf{U}\boldsymbol{\Lambda}^{-1}\mathbf{U}^{\top} \\ \boldsymbol{\Sigma}^{\frac{1}{2}} &= \mathbf{U}\boldsymbol{\Lambda}^{\frac{1}{2}}\mathbf{U}^{\top}, \quad \boldsymbol{\Sigma}^{-\frac{1}{2}} &= \mathbf{U}\boldsymbol{\Lambda}^{-\frac{1}{2}}\mathbf{U}^{\top} \end{split}$$

Λ は, A の固有値 λ_i を対角成分にもった, 対角行列:

$$\mathbf{\Lambda} = \operatorname{diag}(\lambda_1, \dots, \lambda_D), \quad \mathbf{\Lambda}^{-1} = \operatorname{diag}\left(\frac{1}{\lambda_1}, \dots, \frac{1}{\lambda_D}\right)$$
$$\mathbf{\Lambda}^{\frac{1}{2}} = \operatorname{diag}\left(\sqrt{\lambda_1}, \dots, \sqrt{\lambda_D}\right), \quad \mathbf{\Lambda}^{-\frac{1}{2}} = \operatorname{diag}\left(\frac{1}{\sqrt{\lambda_1}}, \dots, \frac{1}{\sqrt{\lambda_D}}\right)$$

• \mathbf{U} は直交行列 $(\mathbf{U}^{-1} = \mathbf{U}^{\top}, \mathbf{U}\mathbf{U}^{\top} = \mathbf{U}^{\top}\mathbf{U} = \mathbf{I}).$

• Σ は正定値行列なので、行列式は常に正: $\det \Sigma > 0$.

 $oldsymbol{\Sigma}$ が正定値 ightarrow $oldsymbol{\Sigma}$ の固有値 $\{\lambda_i\}$ は全て正

→ 行列式は、全ての固有値の積なので正

ullet 対角化の結果 $oldsymbol(\Sigma = \mathbf{U} oldsymbol{\Lambda} \mathbf{U}^{ op})$ を使って, ガウス分布を書き換え:

$$\mathcal{N}(\mathbf{x} \mid \boldsymbol{\mu}, \boldsymbol{\Sigma}) = \frac{1}{(2\pi)^{\frac{D}{2}} \sqrt{\det \boldsymbol{\Sigma}}} \exp \left(-\frac{1}{2} \left(\mathbf{x} - \boldsymbol{\mu} \right)^{\top} \boldsymbol{\Sigma}^{-1} \left(\mathbf{x} - \boldsymbol{\mu} \right) \right)$$

• $\mathbf{U}\mathbf{U}^{\top} = \mathbf{I}$ と、行列式の性質 $(\det \mathbf{A}^{\top} = \det \mathbf{A}, \det \mathbf{A}\mathbf{B} = \det \mathbf{A} \det \mathbf{B})$ から、

$$\det \mathbf{U}\mathbf{U}^{\top} = \det \mathbf{I} = 1$$
$$\det \mathbf{U}\mathbf{U}^{\top} = \det \mathbf{U} \det \mathbf{U}^{\top} = (\det \mathbf{U})^2 = 1 \longrightarrow \det \mathbf{U} = \pm 1$$

• 対角行列の行列式は、対角成分の積であるから、

$$\det \mathbf{\Sigma} = \det \left(\mathbf{U} \mathbf{\Lambda} \mathbf{U}^{\top} \right) = \det \mathbf{U} \det \mathbf{\Lambda} \det \mathbf{U}^{\top} = \det \mathbf{\Lambda} = \prod_{i=1}^{D} \lambda_{i}$$

• $\exp(\cdot)$ の中身 $(\Sigma^{-1} = \mathbf{U} \mathbf{\Lambda}^{-1} \mathbf{U}^{\top})$:

$$(\mathbf{x} - \boldsymbol{\mu})^{\top} \, \boldsymbol{\Sigma}^{-1} \, (\mathbf{x} - \boldsymbol{\mu}) = (\mathbf{x} - \boldsymbol{\mu})^{\top} \, \mathbf{U} \boldsymbol{\Lambda}^{-1} \mathbf{U}^{\top} \, (\mathbf{x} - \boldsymbol{\mu}) = \mathbf{y}^{\top} \boldsymbol{\Lambda}^{-1} \mathbf{y}$$

• ここで、 $\mathbf{y}=\mathbf{U}^{\top}\left(\mathbf{x}-\boldsymbol{\mu}\right)$ とおいた. $\boldsymbol{\Lambda}^{-1}=\mathrm{diag}(\lambda_1^{-1},\ldots,\lambda_D^{-1})$ だから、

$$\mathbf{y}^{\top} \mathbf{\Lambda}^{-1} \mathbf{y} = \begin{pmatrix} y_1 & \cdots & y_D \end{pmatrix} \begin{pmatrix} \lambda_1^{-1} & & \\ & \ddots & \\ & & \lambda_D^{-1} \end{pmatrix} \begin{pmatrix} y_1 \\ \vdots \\ y_D \end{pmatrix}$$
$$= \begin{pmatrix} y_1 & \cdots & y_D \end{pmatrix} \begin{pmatrix} \lambda_1^{-1} y_1 \\ \vdots \\ \lambda_D^{-1} y_D \end{pmatrix} = \sum_{i=1}^D \frac{y_i^2}{\lambda_i}$$

• 係数部分:

$$\det \mathbf{\Sigma} = \det \mathbf{\Lambda} = \prod_{i=1}^{D} \lambda_{i} \longrightarrow \frac{1}{\sqrt{\det \mathbf{\Sigma}}} = \prod_{i=1}^{D} \frac{1}{\sqrt{\lambda_{i}}}$$

$$\longrightarrow \frac{1}{(2\pi)^{\frac{D}{2}} \sqrt{\det \mathbf{\Sigma}}} = \prod_{i=1}^{D} \frac{1}{\sqrt{2\pi\lambda_{i}}}$$

exp(·) の中身:

$$(\mathbf{x} - \boldsymbol{\mu})^{\top} \boldsymbol{\Sigma}^{-1} (\mathbf{x} - \boldsymbol{\mu}) = \mathbf{y}^{\top} \boldsymbol{\Lambda}^{-1} \mathbf{y} = \sum_{i=1}^{D} \frac{y_i^2}{\lambda_i}$$

$$\longrightarrow \exp\left(-\frac{1}{2} (\mathbf{x} - \boldsymbol{\mu})^{\top} \boldsymbol{\Sigma}^{-1} (\mathbf{x} - \boldsymbol{\mu})\right) = \prod_{i=1}^{D} \exp\left(-\frac{y_i^2}{2\lambda_i}\right)$$

これらの結果を基に、ガウス分布を書き換えると、

$$\mathcal{N}(\mathbf{x} \mid \boldsymbol{\mu}, \boldsymbol{\Sigma}) = \frac{1}{(2\pi)^{\frac{D}{2}} \sqrt{\det \boldsymbol{\Sigma}}} \exp\left(-\frac{1}{2} (\mathbf{x} - \boldsymbol{\mu})^{\top} \boldsymbol{\Sigma}^{-1} (\mathbf{x} - \boldsymbol{\mu})\right)$$
$$= \left(\prod_{i=1}^{D} \frac{1}{\sqrt{2\pi\lambda_{i}}}\right) \left(\prod_{i=1}^{D} \exp\left(-\frac{y_{i}^{2}}{2\lambda_{i}}\right)\right)$$
$$= \prod_{i=1}^{D} \frac{1}{\sqrt{2\pi\lambda_{i}}} \exp\left(-\frac{y_{i}^{2}}{2\lambda_{i}}\right) = \prod_{i=1}^{D} \mathcal{N}(y_{i} \mid 0, \lambda_{i})$$

- $\mathbf{y} = \mathbf{U}^{\top} (\mathbf{x} \boldsymbol{\mu})$ の変数変換を施すと, \mathbf{x} のガウス分布は, \mathbf{y} の各成分 についてのガウス分布に分解された.
- y の各成分は, 互いに独立だから, 無相関でもある.

$$\forall i \neq j \ \mathbb{E}[y_i y_j] = \mathbb{E}[y_i] \mathbb{E}[y_j], \operatorname{Cov}(y_i, y_j) = 0$$

次のように、行列の形でも書き換えておく:

$$\mathcal{N}(\mathbf{x} \mid \boldsymbol{\mu}, \boldsymbol{\Sigma}) = \frac{1}{(2\pi)^{\frac{D}{2}} \sqrt{\det \boldsymbol{\Sigma}}} \exp\left(-\frac{1}{2} (\mathbf{x} - \boldsymbol{\mu})^{\top} \boldsymbol{\Sigma}^{-1} (\mathbf{x} - \boldsymbol{\mu})\right)$$
$$= \frac{1}{(2\pi)^{\frac{D}{2}} \sqrt{\det \boldsymbol{\Lambda}}} \exp\left(-\frac{1}{2} \mathbf{y}^{\top} \boldsymbol{\Lambda}^{-1} \mathbf{y}\right) = \mathcal{N}(\mathbf{y} \mid \mathbf{0}, \boldsymbol{\Lambda})$$

- $\mathbf{y} = \mathbf{U}^{\top} (\mathbf{x} \boldsymbol{\mu})$ の変数変換を施すと、平均 $\mathbf{0}$ 、共分散 $\boldsymbol{\Lambda}$ のガウス分布となる. $\boldsymbol{\Lambda}$ は対角行列だから、対角成分以外は全て 0.
- ullet 共分散 $oldsymbol{\Lambda}$ の (i,j) 成分は, $\mathrm{Cov}\,(y_i,y_j)$ である (対角成分は $\mathrm{Var}\,[y_i]$).
- ullet i
 eq j であるとき, $\mathrm{Cov}\left(y_i,y_j
 ight) = 0$ だから, y_i と y_j は無相関である.

多変量ガウス分布の無相関化 (まとめ)

x についてのガウス分布:

$$\mathcal{N}(\mathbf{x} \mid \boldsymbol{\mu}, \boldsymbol{\Sigma}) = \frac{1}{(2\pi)^{\frac{D}{2}} \sqrt{\det \boldsymbol{\Sigma}}} \exp \left(-\frac{1}{2} \left(\mathbf{x} - \boldsymbol{\mu} \right)^{\top} \boldsymbol{\Sigma}^{-1} \left(\mathbf{x} - \boldsymbol{\mu} \right) \right)$$

- ullet 共分散を,直交行列で対角化する $(oldsymbol{\Sigma} = \mathbf{U} oldsymbol{\Lambda} \mathbf{U}^{ op})$.
- Λ は, Σ の固有値を斜めに並べた, 対角行列.
- ullet $\mathbf{y} = \mathbf{U}^{ op}\left(\mathbf{x} oldsymbol{\mu}
 ight)$ の変数変換を施すと、 \mathbf{y} についてのガウス分布となる:

$$\mathcal{N}(\mathbf{y} \mid \mathbf{0}, \mathbf{\Lambda}) = \prod_{i=1}^{D} \mathcal{N}(y_i \mid 0, \lambda_i)$$

- y の各成分は, 互いに独立, 無相関となる.
- このような手続きを, 無相関化という.



多変量ガウス分布

多変量ガウス分布 $\mathcal{N}(\mathbf{x} \mid \boldsymbol{\mu}, \boldsymbol{\Sigma})$ が次を満たすことを確かめよう:

$$\int_{-\infty}^{\infty} \mathcal{N}(\mathbf{x} \mid \boldsymbol{\mu}, \boldsymbol{\Sigma}) \, d\mathbf{x} = 1$$

$$\mathbb{E} \left[\mathbf{x} \right] = \boldsymbol{\mu}$$

$$\mathbb{E} \left[\mathbf{x} \mathbf{x}^{\top} \right] = \boldsymbol{\Sigma} + \boldsymbol{\mu} \boldsymbol{\mu}^{\top}$$

$$\mathbb{E} \left[(\mathbf{x} - \boldsymbol{\mu}) (\mathbf{x} - \boldsymbol{\mu})^{\top} \right] = \operatorname{Var} \left[\mathbf{x} \right] = \boldsymbol{\Sigma}$$

多変量ガウス分布

- ullet 共分散を, 直交行列 $oldsymbol{\mathrm{U}}$ で対角化する $(oldsymbol{\Sigma} = oldsymbol{\mathrm{U}}oldsymbol{\Lambda}oldsymbol{\mathrm{U}}^{ op})$.
- ullet $\mathbf{y} = \mathbf{U}^{ op} (\mathbf{x} oldsymbol{\mu})$ の変数変換を施すと,積分の変数変換から,

$$\int_{-\infty}^{\infty} \mathcal{N}(\mathbf{x} \mid \boldsymbol{\mu}, \boldsymbol{\Sigma}) \, \mathrm{d}\mathbf{x} = \int_{-\infty}^{\infty} \mathcal{N}(\mathbf{y} \mid \mathbf{0}, \boldsymbol{\Lambda}) \left| \det \mathbf{J} \right| \mathrm{d}\mathbf{y}$$

• ヤコビ行列 J は, $\mathbf{x} = \mathbf{U}(\mathbf{y} + \boldsymbol{\mu})$ だから, $\frac{\partial \mathbf{A}\mathbf{x}}{\partial \mathbf{x}} = \mathbf{A}$ (分子レイアウト) を使うと,

$$\mathbf{J} = \frac{\partial \mathbf{x}}{\partial \mathbf{y}} = \frac{\partial}{\partial \mathbf{y}} \mathbf{U} (\mathbf{y} + \boldsymbol{\mu}) = \mathbf{U}$$

ullet U は直交行列だから, $\det \mathbf{J} = \det \mathbf{U} = \pm 1$ (先述).

多変量ガウス分布

•
$$|\det \mathbf{J}| = |\det \mathbf{U}| = 1$$
 と $\mathcal{N}(\mathbf{y} \mid \mathbf{0}, \mathbf{\Lambda}) = \prod_{i=1}^{D} \mathcal{N}(y_i \mid 0, \lambda_i)$ から、

$$\int_{-\infty}^{\infty} \mathcal{N}(\mathbf{x} \mid \boldsymbol{\mu}, \boldsymbol{\Sigma}) \, d\mathbf{x} = \int_{-\infty}^{\infty} \mathcal{N}(\mathbf{y} \mid \mathbf{0}, \boldsymbol{\Lambda}) |\det \mathbf{J}| \, d\mathbf{y}$$

$$= \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} \prod_{i=1}^{D} \mathcal{N}(y_i \mid 0, \lambda_i) \, dy_1 \cdots dy_D$$

$$= \int_{-\infty}^{\infty} \mathcal{N}(y_1 \mid 0, \lambda_1) \, dy_1 \cdots \int_{-\infty}^{\infty} \mathcal{N}(y_D \mid 0, \lambda_1) \, dy_D = 1$$

- 1次元のガウス分布について、積分すると1になることを用いた。
- ullet 積分すると 1 になるから, ガウス分布 $\mathcal{N}(\mathbf{x} \mid oldsymbol{\mu}, oldsymbol{\Sigma})$ は確率分布である.

多変量ガウス分布の平均

続いて, 平均が, 以下を満たすことを確かめよう:

$$\mathbb{E}\left[\mathbf{x}
ight]=oldsymbol{\mu}$$

ただし、

$$\mathbb{E}\left[\mathbf{x}\right] = \int_{-\infty}^{\infty} \mathcal{N}(\mathbf{x} \mid \boldsymbol{\mu}, \boldsymbol{\Sigma}) \mathbf{x} \, d\mathbf{x}$$

$$= \frac{1}{(2\pi)^{\frac{D}{2}} \sqrt{\det \boldsymbol{\Sigma}}} \int_{-\infty}^{\infty} \exp\left(-\frac{1}{2} \left(\mathbf{x} - \boldsymbol{\mu}\right)^{\top} \boldsymbol{\Sigma}^{-1} \left(\mathbf{x} - \boldsymbol{\mu}\right)\right) \mathbf{x} \, d\mathbf{x}$$

 $\mathbf{z} = \mathbf{x} - \boldsymbol{\mu}$ とすると, $\mathbf{x} = \mathbf{z} + \boldsymbol{\mu}$ だから,

$$\mathbb{E}\left[\mathbf{x}\right] = \frac{1}{(2\pi)^{\frac{D}{2}} \sqrt{\det \mathbf{\Sigma}}} \int_{-\infty}^{\infty} \exp\left(-\frac{1}{2}\mathbf{z}^{\top} \mathbf{\Sigma}^{-1} \mathbf{z}\right) (\mathbf{z} + \boldsymbol{\mu}) \left|\det \mathbf{J}\right| d\mathbf{z}$$



多変量ガウス分布の平均

ヤコビ行列 ${f J}$ は、 ${f x}={f z}+{m \mu}$ だから, ${\partial {f x}\over \partial {f x}}={f I}$ を使うと,

$$\mathbf{J} = \frac{\partial \mathbf{x}}{\partial \mathbf{z}} = \frac{\partial}{\partial \mathbf{z}} (\mathbf{z} + \boldsymbol{\mu}) = \mathbf{I}$$

 $|\det \mathbf{J}| = |\det \mathbf{I}| = |1| = 1$ だから,

$$\mathbb{E}\left[\mathbf{x}\right] = \frac{1}{(2\pi)^{\frac{D}{2}} \sqrt{\det \mathbf{\Sigma}}} \int_{-\infty}^{\infty} \exp\left(-\frac{1}{2}\mathbf{z}^{\top}\mathbf{\Sigma}^{-1}\mathbf{z}\right) (\mathbf{z} + \boldsymbol{\mu}) |\det \mathbf{J}| d\mathbf{z}$$
$$= \frac{1}{(2\pi)^{\frac{D}{2}} \sqrt{\det \mathbf{\Sigma}}} \int_{-\infty}^{\infty} \exp\left(-\frac{1}{2}\mathbf{z}^{\top}\mathbf{\Sigma}^{-1}\mathbf{z}\right) (\mathbf{z} + \boldsymbol{\mu}) d\mathbf{z}$$

$$\mathbf{f}(\mathbf{z}) = \exp\left(-rac{1}{2}\mathbf{z}^{ op}\mathbf{\Sigma}^{-1}\mathbf{z}
ight)\mathbf{z}$$
 とおくと、 $\mathbf{f}(-\mathbf{z}) = -\mathbf{f}(\mathbf{z})$ より奇関数なので、

$$\frac{1}{(2\pi)^{\frac{D}{2}}\sqrt{\det \boldsymbol{\Sigma}}}\int_{-\boldsymbol{\infty}}^{\boldsymbol{\infty}} \exp\biggl(-\frac{1}{2}\mathbf{z}^{\top}\boldsymbol{\Sigma}^{-1}\mathbf{z}\biggr)\mathbf{z}\,\mathrm{d}\mathbf{z} = \mathbf{0}$$

◆ロト ◆個ト ◆意ト ◆意ト 意 めので

多変量ガウス分布の平均

また, ガウス分布の積分は1となるから,

$$\begin{split} &\frac{1}{(2\pi)^{\frac{D}{2}}\sqrt{\det \boldsymbol{\Sigma}}} \int_{-\infty}^{\infty} \exp\left(-\frac{1}{2}\mathbf{z}^{\top}\boldsymbol{\Sigma}^{-1}\mathbf{z}\right) \boldsymbol{\mu} \, \mathrm{d}\mathbf{z} \\ &= \boldsymbol{\mu} \int_{-\infty}^{\infty} \frac{1}{(2\pi)^{\frac{D}{2}}\sqrt{\det \boldsymbol{\Sigma}}} \exp\left(-\frac{1}{2}\mathbf{z}^{\top}\boldsymbol{\Sigma}^{-1}\mathbf{z}\right) \mathrm{d}\mathbf{z} \\ &= \boldsymbol{\mu} \int_{-\infty}^{\infty} \mathcal{N}(\mathbf{z} \mid \mathbf{0}, \boldsymbol{\Sigma}) \, \mathrm{d}\mathbf{z} = \boldsymbol{\mu} \end{split}$$

よって, ガウス分布の平均は $\mathbb{E}\left[\mathbf{x}
ight]=oldsymbol{\mu}$:

$$\mathbb{E}\left[\mathbf{x}\right] = \frac{1}{(2\pi)^{\frac{D}{2}} \sqrt{\det \boldsymbol{\Sigma}}} \int_{-\infty}^{\infty} \exp\left(-\frac{1}{2}\mathbf{z}^{\top}\boldsymbol{\Sigma}^{-1}\mathbf{z}\right) (\mathbf{z} + \boldsymbol{\mu}) \, \mathrm{d}\mathbf{z} = \mathbf{0} + \boldsymbol{\mu} = \boldsymbol{\mu}$$

続いて, $\mathbb{E}\left[\mathbf{x}\mathbf{x}^{\top}\right]$ が, 以下を満たすことを確かめよう:

$$\mathbb{E}\left[\mathbf{x}\mathbf{x}^{\top}\right] = \boldsymbol{\Sigma} + \boldsymbol{\mu}\boldsymbol{\mu}^{\top}$$

ただし,

$$\mathbb{E}\left[\mathbf{x}\mathbf{x}^{\top}\right] = \int_{-\infty}^{\infty} \mathcal{N}(\mathbf{x} \mid \boldsymbol{\mu}, \boldsymbol{\Sigma})\mathbf{x}\mathbf{x}^{\top} d\mathbf{x}$$
$$= \frac{1}{(2\pi)^{\frac{D}{2}} \sqrt{\det \boldsymbol{\Sigma}}} \int_{-\infty}^{\infty} \exp\left(-\frac{1}{2} \left(\mathbf{x} - \boldsymbol{\mu}\right)^{\top} \boldsymbol{\Sigma}^{-1} \left(\mathbf{x} - \boldsymbol{\mu}\right)\right) \mathbf{x}\mathbf{x}^{\top} d\mathbf{x}$$

 $\mathbf{z} = \mathbf{x} - \boldsymbol{\mu}$ とすると, $\mathbf{x} = \mathbf{z} + \boldsymbol{\mu}$ だから,

$$\frac{1}{(2\pi)^{\frac{D}{2}}\sqrt{\det \boldsymbol{\Sigma}}} \int_{-\infty}^{\infty} \exp\left(-\frac{1}{2}\mathbf{z}^{\top}\boldsymbol{\Sigma}^{-1}\mathbf{z}\right) \left(\mathbf{z} + \boldsymbol{\mu}\right) \left(\mathbf{z} + \boldsymbol{\mu}\right)^{\top} \left|\det \mathbf{J}\right| d\mathbf{z}$$

先程と同様に、ヤコビ行列は ${f J}=rac{\partial {f x}}{\partial {f z}}={f I}$ 、ゆえにヤコビアンは $|\det {f J}|=1$ だから、

$$\mathbb{E}\left[\mathbf{x}\mathbf{x}^{\top}\right] = \frac{1}{(2\pi)^{\frac{D}{2}}\sqrt{\det \boldsymbol{\Sigma}}} \int_{-\infty}^{\infty} \exp\!\left(-\frac{1}{2}\mathbf{z}^{\top}\boldsymbol{\Sigma}^{-1}\mathbf{z}\right) \left(\mathbf{z} + \boldsymbol{\mu}\right) \left(\mathbf{z} + \boldsymbol{\mu}\right)^{\top} \mathrm{d}\mathbf{z}$$

$$\mathbf{f}(\mathbf{z}) = \exp\left(-\frac{1}{2}\mathbf{z}^{\top}\mathbf{\Sigma}^{-1}\mathbf{z}\right)\boldsymbol{\mu}\mathbf{z}^{\top},\ \mathbf{g}(\mathbf{z}) = \exp\left(-\frac{1}{2}\mathbf{z}^{\top}\mathbf{\Sigma}^{-1}\mathbf{z}\right)\mathbf{z}\boldsymbol{\mu}^{\top}$$
 とおく. $\mathbf{f}(-\mathbf{z}) = -\mathbf{f}(\mathbf{z}),\ \mathbf{g}(-\mathbf{z}) = -\mathbf{g}(\mathbf{z})$ だから、どちらも奇関数. よって、

$$\frac{1}{(2\pi)^{\frac{D}{2}} \sqrt{\det \Sigma}} \int_{-\infty}^{\infty} \exp\left(-\frac{1}{2} \mathbf{z}^{\top} \mathbf{\Sigma}^{-1} \mathbf{z}\right) \boldsymbol{\mu} \mathbf{z}^{\top} d\mathbf{z} = \mathbf{0}$$
$$\frac{1}{(2\pi)^{\frac{D}{2}} \sqrt{\det \Sigma}} \int_{-\infty}^{\infty} \exp\left(-\frac{1}{2} \mathbf{z}^{\top} \mathbf{\Sigma}^{-1} \mathbf{z}\right) \mathbf{z} \boldsymbol{\mu}^{\top} d\mathbf{z} = \mathbf{0}$$

また,

$$\begin{split} &\frac{1}{(2\pi)^{\frac{D}{2}}\sqrt{\det \boldsymbol{\Sigma}}} \int_{-\infty}^{\infty} \exp\left(-\frac{1}{2}\mathbf{z}^{\top}\boldsymbol{\Sigma}^{-1}\mathbf{z}\right) \boldsymbol{\mu} \boldsymbol{\mu}^{\top} d\mathbf{z} \\ &= \boldsymbol{\mu} \boldsymbol{\mu}^{\top} \int_{-\infty}^{\infty} \frac{1}{(2\pi)^{\frac{D}{2}}\sqrt{\det \boldsymbol{\Sigma}}} \exp\left(-\frac{1}{2}\mathbf{z}^{\top}\boldsymbol{\Sigma}^{-1}\mathbf{z}\right) d\mathbf{z} \\ &= \boldsymbol{\mu} \boldsymbol{\mu}^{\top} \int_{-\infty}^{\infty} \mathcal{N}(\mathbf{z} \mid \mathbf{0}, \boldsymbol{\Sigma}) d\mathbf{z} = \boldsymbol{\mu} \boldsymbol{\mu}^{\top} \end{split}$$

ガウス分布を積分すると1になるので、容易に求まる.

続いて,以下の積分を考える:

$$\frac{1}{(2\pi)^{\frac{D}{2}}\sqrt{\det\boldsymbol{\Sigma}}}\int_{-\infty}^{\infty}\exp\biggl(-\frac{1}{2}\mathbf{z}^{\top}\boldsymbol{\Sigma}^{-1}\mathbf{z}\biggr)\mathbf{z}\mathbf{z}^{\top}\,\mathrm{d}\mathbf{z}$$

直交行列 \mathbf{U} による対角化 $\left(\mathbf{\Sigma} = \mathbf{U}\mathbf{\Lambda}\mathbf{U}^{ op}
ight)$ と, $\mathbf{y} = \mathbf{U}^{ op}\left(\mathbf{x} - \boldsymbol{\mu}\right)$ を用いる. このとき, $\mathcal{N}(\mathbf{x} \mid \boldsymbol{\mu}, \boldsymbol{\Sigma}) = \mathcal{N}(\mathbf{y} \mid \mathbf{0}, \boldsymbol{\Lambda})$ であったから,

$$\begin{split} &\frac{1}{(2\pi)^{\frac{D}{2}}\sqrt{\det \boldsymbol{\Sigma}}}\exp\left(-\frac{1}{2}\mathbf{z}^{\top}\boldsymbol{\Sigma}^{-1}\mathbf{z}\right) \quad (\mathbf{z}=\mathbf{x}-\boldsymbol{\mu} \ \boldsymbol{\varepsilon}$$
使って、元に戻す)
$$&=\frac{1}{(2\pi)^{\frac{D}{2}}\sqrt{\det \boldsymbol{\Sigma}}}\exp\left(-\frac{1}{2}\left(\mathbf{x}-\boldsymbol{\mu}\right)^{\top}\boldsymbol{\Sigma}^{-1}\left(\mathbf{x}-\boldsymbol{\mu}\right)\right)=\mathcal{N}(\mathbf{x}\mid\boldsymbol{\mu},\boldsymbol{\Sigma}) \\ &=\frac{1}{(2\pi)^{\frac{D}{2}}\sqrt{\det \boldsymbol{\Lambda}}}\exp\left(-\frac{1}{2}\mathbf{y}^{\top}\boldsymbol{\Lambda}^{-1}\mathbf{y}\right)=\mathcal{N}(\mathbf{y}\mid\boldsymbol{0},\boldsymbol{\Lambda}) \end{split}$$

 $\mathbf{z} = \mathbf{x} - \boldsymbol{\mu}$, $\mathbf{y} = \mathbf{U}^{\top} (\mathbf{x} - \boldsymbol{\mu})$ より, $\mathbf{y} = \mathbf{U}^{\top} \mathbf{z}$. $\mathbf{U}^{-1} = \mathbf{U}^{\top}$ に注意すれば, $\mathbf{z} = \mathbf{U} \mathbf{y}$. 積分を変数変換すると,

$$\frac{1}{(2\pi)^{\frac{D}{2}}} \sqrt{\det \mathbf{\Sigma}} \int_{-\infty}^{\infty} \exp\left(-\frac{1}{2} \mathbf{z}^{\top} \mathbf{\Sigma}^{-1} \mathbf{z}\right) \mathbf{z} \mathbf{z}^{\top} d\mathbf{z}$$
$$= \int_{-\infty}^{\infty} \mathcal{N}(\mathbf{y} \mid \mathbf{0}, \mathbf{\Lambda}) \mathbf{U} \mathbf{y} \mathbf{y}^{\top} \mathbf{U}^{\top} |\det \mathbf{J}| d\mathbf{y}$$

ヤコビ行列は $\mathbf{J}=rac{\partial \mathbf{z}}{\partial \mathbf{y}}=\mathbf{U}$ である. \mathbf{U} は直交行列だから, $\det \mathbf{U}=\pm 1$. また $\mathcal{N}(\mathbf{y}\mid \mathbf{0}, \mathbf{\Lambda})$ は, 各成分のガウス分布の積に分解される:

$$\frac{1}{(2\pi)^{\frac{D}{2}}\sqrt{\det \boldsymbol{\Sigma}}}\exp\biggl(-\frac{1}{2}\mathbf{z}^{\top}\boldsymbol{\Sigma}^{-1}\mathbf{z}\biggr) = \mathcal{N}(\mathbf{y}\mid \boldsymbol{0},\boldsymbol{\Lambda}) = \prod_{i=1}^{D}\mathcal{N}(y_i\mid 0,\lambda_i)$$

これらの結果を使うと,

$$\frac{1}{(2\pi)^{\frac{D}{2}} \sqrt{\det \mathbf{\Sigma}}} \int_{-\infty}^{\infty} \exp\left(-\frac{1}{2} \mathbf{z}^{\top} \mathbf{\Sigma}^{-1} \mathbf{z}\right) \mathbf{z} \mathbf{z}^{\top} d\mathbf{z}$$

$$= \int_{-\infty}^{\infty} \mathcal{N}(\mathbf{y} \mid \mathbf{0}, \mathbf{\Lambda}) \mathbf{U} \mathbf{y} \mathbf{y}^{\top} \mathbf{U}^{\top} |\det \mathbf{J}| d\mathbf{y}$$

$$= \int_{-\infty}^{\infty} \prod_{i=1}^{D} \mathcal{N}(y_i \mid 0, \lambda_i) \mathbf{U} \mathbf{y} \mathbf{y}^{\top} \mathbf{U}^{\top} d\mathbf{y}$$

$$= \mathbf{U} \left(\int_{-\infty}^{\infty} \prod_{i=1}^{D} \mathcal{N}(y_i \mid 0, \lambda_i) \mathbf{y} \mathbf{y}^{\top} d\mathbf{y} \right) \mathbf{U}^{\top}$$

以下の積分を示そう:

$$\int_{-\infty}^{\infty} \prod_{i=1}^{D} \mathcal{N}(y_i \mid 0, \lambda_i) \mathbf{y} \mathbf{y}^{\top} \, \mathrm{d} \mathbf{y} = \mathbf{\Lambda}$$

 $\mathbf{y}\mathbf{y}^{ op}$ の (i,j) 成分は y_iy_j だから, 積分の (i,j) 成分を考えると,

$$\left(\int_{-\infty}^{\infty} \prod_{k=1}^{D} \mathcal{N}(y_k \mid 0, \lambda_k) \mathbf{y} \mathbf{y}^{\top} d\mathbf{y}\right)_{ij}$$

$$= \int_{-\infty}^{\infty} \prod_{j=1}^{D} \mathcal{N}(y_k \mid 0, \lambda_k) y_i y_j d\mathbf{y}$$

$$= \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} \mathcal{N}(y_1 \mid 0, \lambda_1) \cdots \mathcal{N}(y_D \mid 0, \lambda_D) y_i y_j dy_1 \cdots dy_D$$

全ての $k \neq i, j$ について, $\mathcal{N}(y_k \mid 0, \lambda_k)$ は積分により 1 となるから,

$$\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \mathcal{N}(y_i \mid 0, \lambda_i) \mathcal{N}(y_j \mid 0, \lambda_j) y_i y_j \, dy_i \, dy_j = \mathbb{E}[y_i y_j]$$

 $i \neq j$ ならば,

$$\int_{-\infty}^{\infty} \mathcal{N}(y_i \mid 0, \lambda_i) y_i \, dy_i \int_{-\infty}^{\infty} \mathcal{N}(y_j \mid 0, \lambda_j) y_j \, dy_j = 0$$

 $\mathcal{N}(y_i \mid 0, \lambda_i)y_i$ は奇関数なので、積分すると 0 になることを用いた:

$$\mathcal{N}(y_i \mid 0, \lambda_i)y_i = \frac{1}{\sqrt{2\pi\lambda_i}} \exp\left(-\frac{1}{2\lambda_i}y_i^2\right) y_i$$
 (奇関数)

$$i=j$$
 ならば、 $\int_{-\infty}^{\infty}x^2\exp\left(-ax^2
ight)\mathrm{d}x=rac{1}{2}\sqrt{rac{\pi}{a^3}}$ を使って、

$$\int_{-\infty}^{\infty} \mathcal{N}(y_i \mid 0, \lambda_i) y_i^2 \, dy_i = \frac{1}{\sqrt{2\pi\lambda_i}} \int_{-\infty}^{\infty} \exp\left(-\frac{1}{2\lambda_i} y_i^2\right) y_i^2 \, dy_i$$
$$= \frac{1}{\sqrt{2\pi\lambda_i}} \cdot \frac{1}{2} \sqrt{8\lambda_i^3 \pi} = \lambda_i$$

i = j のときの積分は、次のようにも求められる:

$$\int_{-\infty}^{\infty} \mathcal{N}(y_i \mid 0, \lambda_i) y_i^2 \, dy_i = \mathbb{E}\left[y_i^2\right] = \operatorname{Var}\left[y_i\right] + \mathbb{E}\left[y_i\right]^2 = \lambda_i + 0^2 = \lambda_i$$

 $\operatorname{Var}\left[x
ight]=\mathbb{E}\left[x^2
ight]-\mathbb{E}\left[x
ight]^2$ と, $\mathbb{E}\left[y_i
ight]=0$, $\operatorname{Var}\left[y_i
ight]=\lambda_i$ を用いた. 以上より, 次の積分は, 対角要素が λ_i , それ以外の成分が 0 だから,

$$\int_{-\infty}^{\infty} \prod_{k=1}^{D} \mathcal{N}(y_k \mid 0, \lambda_k) \mathbf{y} \mathbf{y}^{\top} d\mathbf{y} = \operatorname{diag}(\lambda_1, \dots, \lambda_D) = \mathbf{\Lambda}$$

よって,

$$\frac{1}{(2\pi)^{\frac{D}{2}} \sqrt{\det \mathbf{\Sigma}}} \int_{-\infty}^{\infty} \exp\left(-\frac{1}{2} \mathbf{z}^{\top} \mathbf{\Sigma}^{-1} \mathbf{z}\right) \mathbf{z} \mathbf{z}^{\top} d\mathbf{z}$$

$$= \mathbf{U} \left(\int_{-\infty}^{\infty} \prod_{i=1}^{D} \mathcal{N}(y_i \mid 0, \lambda_i) \mathbf{y} \mathbf{y}^{\top} d\mathbf{y} \right) \mathbf{U}^{\top} = \mathbf{U} \mathbf{\Lambda} \mathbf{U}^{\top} = \mathbf{\Sigma}$$

以上より,

$$\mathbb{E}\left[\mathbf{x}\mathbf{x}^{\top}\right] = \frac{1}{(2\pi)^{\frac{D}{2}}\sqrt{\det \Sigma}} \int_{-\infty}^{\infty} \exp\left(-\frac{1}{2}\mathbf{z}^{\top}\Sigma^{-1}\mathbf{z}\right) (\mathbf{z} + \boldsymbol{\mu}) (\mathbf{z} + \boldsymbol{\mu})^{\top} d\mathbf{z}$$
$$= \Sigma + \boldsymbol{\mu}\boldsymbol{\mu}^{\top}$$

ただし,

$$\frac{1}{(2\pi)^{\frac{D}{2}}\sqrt{\det \Sigma}} \int_{-\infty}^{\infty} \exp\left(-\frac{1}{2}\mathbf{z}^{\top}\boldsymbol{\Sigma}^{-1}\mathbf{z}\right) \boldsymbol{\mu} \mathbf{z}^{\top} d\mathbf{z} = \mathbf{0}$$

$$\frac{1}{(2\pi)^{\frac{D}{2}}\sqrt{\det \Sigma}} \int_{-\infty}^{\infty} \exp\left(-\frac{1}{2}\mathbf{z}^{\top}\boldsymbol{\Sigma}^{-1}\mathbf{z}\right) \mathbf{z} \boldsymbol{\mu}^{\top} d\mathbf{z} = \mathbf{0}$$

$$\frac{1}{(2\pi)^{\frac{D}{2}}\sqrt{\det \Sigma}} \int_{-\infty}^{\infty} \exp\left(-\frac{1}{2}\mathbf{z}^{\top}\boldsymbol{\Sigma}^{-1}\mathbf{z}\right) \boldsymbol{\mu} \boldsymbol{\mu}^{\top} d\mathbf{z} = \boldsymbol{\mu} \boldsymbol{\mu}^{\top}$$

$$\frac{1}{(2\pi)^{\frac{D}{2}}\sqrt{\det \Sigma}} \int_{-\infty}^{\infty} \exp\left(-\frac{1}{2}\mathbf{z}^{\top}\boldsymbol{\Sigma}^{-1}\mathbf{z}\right) \mathbf{z} \mathbf{z}^{\top} d\mathbf{z} = \boldsymbol{\Sigma}$$

最後に, $\mathbb{E}\left[\left(\mathbf{x}-oldsymbol{\mu}\right)\left(\mathbf{x}-oldsymbol{\mu}\right)^{ op}
ight]$ が, 以下を満たすことを確かめよう:

$$\mathbb{E}\left[\left(\mathbf{x} - \boldsymbol{\mu}\right)\left(\mathbf{x} - \boldsymbol{\mu}\right)^{\top}\right] = \boldsymbol{\Sigma}$$

ただし、

$$\mathbb{E}\left[\left(\mathbf{x} - \boldsymbol{\mu}\right)\left(\mathbf{x} - \boldsymbol{\mu}\right)^{\top}\right] = \int_{-\infty}^{\infty} \mathcal{N}(\mathbf{x} \mid \boldsymbol{\mu}, \boldsymbol{\Sigma}) \left(\mathbf{x} - \boldsymbol{\mu}\right) \left(\mathbf{x} - \boldsymbol{\mu}\right)^{\top} d\mathbf{x}$$

$$= \frac{1}{(2\pi)^{\frac{D}{2}} \sqrt{\det \boldsymbol{\Sigma}}} \int_{-\infty}^{\infty} \exp\left(-\frac{1}{2} \left(\mathbf{x} - \boldsymbol{\mu}\right)^{\top} \boldsymbol{\Sigma}^{-1} \left(\mathbf{x} - \boldsymbol{\mu}\right)\right) \left(\mathbf{x} - \boldsymbol{\mu}\right) \left(\mathbf{x} - \boldsymbol{\mu}\right)^{\top} d\mathbf{x}$$

 $\mathbf{z} = \mathbf{x} - \boldsymbol{\mu}$ とすると, $\mathbf{x} = \mathbf{z} + \boldsymbol{\mu}$ だから,

$$\frac{1}{(2\pi)^{\frac{D}{2}}\sqrt{\det \boldsymbol{\Sigma}}} \int_{-\infty}^{\infty} \exp\left(-\frac{1}{2}\mathbf{z}^{\top}\boldsymbol{\Sigma}^{-1}\mathbf{z}\right) \mathbf{z}\mathbf{z}^{\top} |\det \mathbf{J}| \, d\mathbf{z}$$

◆ロト ◆母 ト ◆ 恵 ト ◆ 恵 ・ 釣 へ ○

ヤコビ行列は
$$\mathbf{J} = \frac{\partial \mathbf{x}}{\partial \mathbf{z}} = \mathbf{I}$$
, ゆえにヤコビアンは $|\det \mathbf{J}| = 1$ だから,
$$\mathbb{E}\left[(\mathbf{x} - \boldsymbol{\mu}) (\mathbf{x} - \boldsymbol{\mu})^{\top} \right]$$
$$= \frac{1}{(2\pi)^{\frac{D}{2}} \sqrt{\det \boldsymbol{\Sigma}}} \int_{-\infty}^{\infty} \exp\left(-\frac{1}{2} \mathbf{z}^{\top} \boldsymbol{\Sigma}^{-1} \mathbf{z} \right) \mathbf{z} \mathbf{z}^{\top} |\det \mathbf{J}| d\mathbf{z}$$
$$= \frac{1}{(2\pi)^{\frac{D}{2}} \sqrt{\det \boldsymbol{\Sigma}}} \int_{-\infty}^{\infty} \exp\left(-\frac{1}{2} \mathbf{z}^{\top} \boldsymbol{\Sigma}^{-1} \mathbf{z} \right) \mathbf{z} \mathbf{z}^{\top} d\mathbf{z}$$
$$= \boldsymbol{\Sigma} \quad (先程行った積分と同じ)$$

次のようにも確かめられる:

$$\mathbb{E}\left[\left(\mathbf{x} - \boldsymbol{\mu}\right)\left(\mathbf{x} - \boldsymbol{\mu}\right)^{\top}\right] = \mathbb{E}\left[\left(\mathbf{x} - \mathbb{E}\left[\mathbf{x}\right]\right)\left(\mathbf{x} - \mathbb{E}\left[\mathbf{x}\right]\right)^{\top}\right] = \operatorname{Var}\left[\mathbf{x}\right]$$
$$= \mathbb{E}\left[\mathbf{x}\mathbf{x}^{\top}\right] - \mathbb{E}\left[\mathbf{x}\right]\mathbb{E}\left[\mathbf{x}\right]^{\top} = \boldsymbol{\Sigma} + \boldsymbol{\mu}\boldsymbol{\mu}^{\top} - \boldsymbol{\mu}\boldsymbol{\mu}^{\top} = \boldsymbol{\Sigma}$$

66 / 82

多変量ガウス分布の平均と共分散

多変量ガウス分布の平均と共分散

$$\mathcal{N}(\mathbf{x} \mid \boldsymbol{\mu}, \boldsymbol{\Sigma}) = \frac{1}{(2\pi)^{\frac{D}{2}} \sqrt{\det \boldsymbol{\Sigma}}} \exp \left(-\frac{1}{2} \left(\mathbf{x} - \boldsymbol{\mu} \right)^{\top} \boldsymbol{\Sigma}^{-1} \left(\mathbf{x} - \boldsymbol{\mu} \right) \right)$$

このガウス分布の平均と共分散は、

$$\mathbb{E}\left[\mathbf{x}\right] = \boldsymbol{\mu}$$

$$\mathbb{E}\left[\left(\mathbf{x} - \boldsymbol{\mu}\right)\left(\mathbf{x} - \boldsymbol{\mu}\right)^{\top}\right] = \operatorname{Var}\left[\mathbf{x}\right] = \boldsymbol{\Sigma}$$

スカラの場合と同じように、モーメント母関数 (積率母関数) を使うこともできる。

目次

- ③ 多変量ガウス分布
 - 多次元ガウス分布のモーメント母関数

モーメント母関数 (多次元)

確率分布 $p(\mathbf{x})$ があるとき, $\mathbb{E}\left[\exp\left(\mathbf{t}^{\top}\mathbf{x}\right)\right]$ を, モーメント母関数という.

$$\mathbb{E}\left[\exp\left(\mathbf{t}^{\top}\mathbf{x}\right)\right] = \mathbb{E}\left[1 + \mathbf{t}^{\top}\mathbf{x} + \frac{1}{2!}\left(\mathbf{t}^{\top}\mathbf{x}\right)^{2} + \frac{1}{3!}\left(\mathbf{t}^{\top}\mathbf{x}\right)^{3} + \cdots\right]$$

 \mathbf{t} で 1 回微分してから $\mathbf{t}=\mathbf{0}$ を代入すると, $\mathbb{E}\left[\mathbf{x}\right]$ が得られる.

 \mathbf{t} で 2 回微分してから $\mathbf{t} = \mathbf{0}$ を代入すると, $\mathbb{E}\left[\mathbf{x}\mathbf{x}^{ op}
ight]$ が得られる.

$$\mathbb{E}\left[\mathbf{x}\right] = \frac{\partial}{\partial \mathbf{t}} \, \mathbb{E}\left[\exp\left(\mathbf{t}^{\top}\mathbf{x}\right)\right] \bigg|_{\mathbf{t} = \mathbf{0}}$$

$$\mathbb{E}\left[\mathbf{x}\mathbf{x}^{\top}\right] = \frac{\partial^{2}}{\partial \mathbf{t}^{\top}\partial \mathbf{t}} \, \mathbb{E}\left[\exp\left(\mathbf{t}^{\top}\mathbf{x}\right)\right] \bigg|_{\mathbf{t} = \mathbf{0}}$$

分子レイアウトでは、次が成り立つ (スカラのベクトルによる微分):

$$\begin{split} \frac{\partial \mathbf{a}^{\top} \mathbf{x}}{\partial \mathbf{x}} &= \frac{\partial \mathbf{x}^{\top} \mathbf{a}}{\partial \mathbf{x}} = \mathbf{a}^{\top}, \\ \frac{\partial \left(\mathbf{a}^{\top} \mathbf{x}\right)^{2}}{\partial \mathbf{x}} &= \frac{\partial \left(\mathbf{x}^{\top} \mathbf{a}\right)^{2}}{\partial \mathbf{x}} = 2 \left(\mathbf{x}^{\top} \mathbf{a}\right) \mathbf{a}^{\top} \\ \frac{\partial \left(\mathbf{a}^{\top} \mathbf{x}\right)^{n}}{\partial \mathbf{x}} &= \frac{\partial \left(\mathbf{x}^{\top} \mathbf{a}\right)^{n}}{\partial \mathbf{x}} = n \left(\mathbf{x}^{\top} \mathbf{a}\right)^{n-1} \mathbf{a}^{\top} \end{split}$$

ullet 各要素 x_i に対する微分を考えると, 合成関数の微分から,

$$\frac{\partial \left(\mathbf{a}^{\top}\mathbf{x}\right)^{n}}{\partial x_{i}} = \frac{\partial \left(\mathbf{x}^{\top}\mathbf{a}\right)^{n}}{\partial x_{i}} = n\left(\mathbf{x}^{\top}\mathbf{a}\right)^{n-1} \frac{\partial \mathbf{x}^{\top}\mathbf{a}}{\partial x_{i}}$$
$$= n\left(\mathbf{x}^{\top}\mathbf{a}\right)^{n-1} \left(\frac{\partial \mathbf{x}^{\top}\mathbf{a}}{\partial \mathbf{x}}\right)_{i} = n\left(\mathbf{x}^{\top}\mathbf{a}\right)^{n-1} \left(\mathbf{a}\right)_{i}$$

ullet スカラの $\mathbf{x}^{ op}$ による微分は、 \mathbf{x} による微分を転置したものと考える:

$$\begin{split} \frac{\partial \mathbf{a}^{\top} \mathbf{x}}{\partial \mathbf{x}^{\top}} &= \frac{\partial \mathbf{x}^{\top} \mathbf{a}}{\partial \mathbf{x}^{\top}} = \mathbf{a}, \\ \frac{\partial \left(\mathbf{a}^{\top} \mathbf{x}\right)^{2}}{\partial \mathbf{x}^{\top}} &= \frac{\partial \left(\mathbf{x}^{\top} \mathbf{a}\right)^{2}}{\partial \mathbf{x}^{\top}} = 2 \left(\mathbf{x}^{\top} \mathbf{a}\right) \mathbf{a} \\ \frac{\partial \left(\mathbf{a}^{\top} \mathbf{x}\right)^{n}}{\partial \mathbf{x}^{\top}} &= \frac{\partial \left(\mathbf{x}^{\top} \mathbf{a}\right)^{n}}{\partial \mathbf{x}^{\top}} = n \left(\mathbf{x}^{\top} \mathbf{a}\right)^{n-1} \mathbf{a} \end{split}$$

• モーメント母関数:

$$\mathbb{E}\left[\exp\left(\mathbf{t}^{\top}\mathbf{x}\right)\right] = \mathbb{E}\left[1 + \mathbf{t}^{\top}\mathbf{x} + \frac{1}{2!}\left(\mathbf{t}^{\top}\mathbf{x}\right)^{2} + \frac{1}{3!}\left(\mathbf{t}^{\top}\mathbf{x}\right)^{3} + \cdots\right]$$

$$= 1 + \mathbb{E}\left[\mathbf{t}^{\top}\mathbf{x}\right] + \frac{1}{2!}\mathbb{E}\left[\left(\mathbf{t}^{\top}\mathbf{x}\right)^{2}\right] + \frac{1}{3!}\mathbb{E}\left[\left(\mathbf{t}^{\top}\mathbf{x}\right)^{3}\right] + \cdots$$

$$= 1 + \mathbf{t}^{\top}\mathbb{E}\left[\mathbf{x}\right] + \frac{1}{2!}\mathbf{t}^{\top}\mathbb{E}\left[\mathbf{x}\mathbf{x}^{\top}\right]\mathbf{t} + \frac{1}{3!}\mathbf{t}^{\top}\mathbb{E}\left[\mathbf{x}\mathbf{t}^{\top}\mathbf{x}\mathbf{x}^{\top}\right]\mathbf{t} + \cdots$$

ullet t で 1 回微分して $\mathbf{t} = \mathbf{0}$ を代入すれば, $\mathbb{E}\left[\mathbf{x}\right]^{ op}$ を得る.

$$\frac{\partial}{\partial \mathbf{t}} \mathbb{E} \left[\exp \left(\mathbf{t}^{\top} \mathbf{x} \right) \right] = \mathbb{E} \left[\mathbf{x} \right]^{\top} + \mathbf{t}^{\top} \mathbb{E} \left[\mathbf{x} \mathbf{x}^{\top} \right] + \cdots$$

tで1回微分すると,

$$\frac{\partial}{\partial \mathbf{t}} \mathbb{E} \left[\exp \left(\mathbf{t}^{\top} \mathbf{x} \right) \right] = \mathbb{E} \left[\mathbf{x} \right]^{\top} + \mathbf{t}^{\top} \mathbb{E} \left[\mathbf{x} \mathbf{x}^{\top} \right] + \cdots$$

 \bullet \mathbf{t}^{\top} で, さらにもう 1 回微分すると,

$$\frac{\partial^2}{\partial \mathbf{t}^\top \partial \mathbf{t}} \mathbb{E} \left[\exp \left(\mathbf{t}^\top \mathbf{x} \right) \right] = \mathbb{E} \left[\mathbf{x} \mathbf{x}^\top \right] + \cdots$$

ullet $\mathbf{t} = \mathbf{0}$ を代入すれば、 $\mathbb{E}\left[\mathbf{x}\mathbf{x}^{ op}
ight]$ を得る.

ガウス分布のモーメント母関数 (多次元)

ガウス分布 $\mathcal{N}(\mathbf{x} \mid \boldsymbol{\mu}, \boldsymbol{\Sigma})$ に対するモーメント母関数は,

$$\mathbb{E}\left[\exp\!\left(\mathbf{t}^{\top}\mathbf{x}\right)\right] = \exp\!\left(\boldsymbol{\mu}^{\top}\mathbf{t} + \frac{1}{2}\mathbf{t}^{\top}\boldsymbol{\Sigma}\mathbf{t}\right)$$

ただし,

$$\mathbb{E}\left[\exp\left(\mathbf{t}^{\top}\mathbf{x}\right)\right] = \int_{-\infty}^{\infty} \mathcal{N}(\mathbf{x} \mid \boldsymbol{\mu}, \boldsymbol{\Sigma}) \exp\left(\mathbf{t}^{\top}\mathbf{x}\right) d\mathbf{x}$$
$$= \frac{1}{(2\pi)^{\frac{D}{2}} \sqrt{\det \boldsymbol{\Sigma}}} \int_{-\infty}^{\infty} \exp\left(-\frac{1}{2} \left(\mathbf{x} - \boldsymbol{\mu}\right)^{\top} \boldsymbol{\Sigma}^{-1} \left(\mathbf{x} - \boldsymbol{\mu}\right)\right)$$
$$\exp\left(\mathbf{t}^{\top}\mathbf{x}\right) d\mathbf{x}$$

 $\exp(\cdot)$ の中身を平方完成させる $(\mathbf{t}^{\top}\mathbf{x} = \mathbf{x}^{\top}\mathbf{t})$:

$$\begin{aligned} &-\frac{1}{2} \left(\mathbf{x} - \boldsymbol{\mu} \right)^{\top} \boldsymbol{\Sigma}^{-1} \left(\mathbf{x} - \boldsymbol{\mu} \right) + \mathbf{t}^{\top} \mathbf{x} \\ &= -\frac{1}{2} \left(\mathbf{x}^{\top} \boldsymbol{\Sigma}^{-1} \mathbf{x} - \boldsymbol{\mu}^{\top} \boldsymbol{\Sigma}^{-1} \mathbf{x} - \mathbf{x}^{\top} \boldsymbol{\Sigma}^{-1} \boldsymbol{\mu} + \boldsymbol{\mu}^{\top} \boldsymbol{\Sigma}^{-1} \boldsymbol{\mu} - 2 \mathbf{t}^{\top} \mathbf{x} \right) \\ &= -\frac{1}{2} \left(\mathbf{x}^{\top} \boldsymbol{\Sigma}^{-1} \mathbf{x} - \boldsymbol{\mu}^{\top} \boldsymbol{\Sigma}^{-1} \mathbf{x} - \mathbf{t}^{\top} \mathbf{x} - \mathbf{x}^{\top} \boldsymbol{\Sigma}^{-1} \boldsymbol{\mu} - \mathbf{x}^{\top} \mathbf{t} + \boldsymbol{\mu}^{\top} \boldsymbol{\Sigma}^{-1} \boldsymbol{\mu} \right) \end{aligned}$$

 Σ, Σ^{-1} には平方根 $\Sigma^{rac{1}{2}}, \Sigma^{-rac{1}{2}}$ が存在する. 対角化 $\Lambda=\mathbf{U}^{ op}\Sigma\mathbf{U}$ の結果を使えば (\mathbf{U} は直交行列, Λ は Σ の固有値を並べた対角行列),

$$\mathbf{\Sigma}^{rac{1}{2}} = \mathbf{U} \mathbf{\Lambda}^{rac{1}{2}} \mathbf{U}^{ op}, \quad \mathbf{\Sigma}^{-rac{1}{2}} = \mathbf{U} \mathbf{\Lambda}^{-rac{1}{2}} \mathbf{U}^{ op}$$

 $oldsymbol{\Lambda}^{rac{1}{2}}, oldsymbol{\Lambda}^{-rac{1}{2}}$ は対角行列だから対称. よって, $oldsymbol{\Sigma}^{rac{1}{2}}, oldsymbol{\Sigma}^{-rac{1}{2}}$ も対称.

平方根 $oldsymbol{\Sigma}^{rac{1}{2}}, oldsymbol{\Sigma}^{-rac{1}{2}}$ を使い (対称性に注意), さらに平方完成すると,

$$\begin{split} &-\frac{1}{2}\left(\mathbf{x}^{\top}\boldsymbol{\Sigma}^{-1}\mathbf{x}-\boldsymbol{\mu}^{\top}\boldsymbol{\Sigma}^{-1}\mathbf{x}-\mathbf{t}^{\top}\mathbf{x}-\mathbf{x}^{\top}\boldsymbol{\Sigma}^{-1}\boldsymbol{\mu}-\mathbf{x}^{\top}\mathbf{t}+\boldsymbol{\mu}^{\top}\boldsymbol{\Sigma}^{-1}\boldsymbol{\mu}\right)\\ &=\left(練習問題\right)\\ &=-\frac{1}{2}\bigg(\left(\boldsymbol{\Sigma}^{-\frac{1}{2}}\mathbf{x}-\left(\boldsymbol{\Sigma}^{-\frac{1}{2}}\boldsymbol{\mu}+\boldsymbol{\Sigma}^{\frac{1}{2}}\mathbf{t}\right)\right)^{\top}\left(\boldsymbol{\Sigma}^{-\frac{1}{2}}\mathbf{x}-\left(\boldsymbol{\Sigma}^{-\frac{1}{2}}\boldsymbol{\mu}+\boldsymbol{\Sigma}^{\frac{1}{2}}\mathbf{t}\right)\right)\\ &-\boldsymbol{\mu}^{\top}\mathbf{t}-\mathbf{t}^{\top}\boldsymbol{\mu}-\mathbf{t}^{\top}\boldsymbol{\Sigma}\mathbf{t}\bigg) \end{split}$$

<u>多次元ガウ</u>ス分布に対するモーメント母関数

以上より, $\exp(\cdot)$ の中身は,

$$\begin{split} &-\frac{1}{2}(\mathbf{x}-\boldsymbol{\mu})^{\top} \, \boldsymbol{\Sigma}^{-1} \, (\mathbf{x}-\boldsymbol{\mu}) + \mathbf{t}^{\top} \mathbf{x} \\ &= -\frac{1}{2} \bigg(\left(\boldsymbol{\Sigma}^{-\frac{1}{2}} \mathbf{x} - \left(\boldsymbol{\Sigma}^{-\frac{1}{2}} \boldsymbol{\mu} + \boldsymbol{\Sigma}^{\frac{1}{2}} \mathbf{t} \right) \right)^{\top} \left(\boldsymbol{\Sigma}^{-\frac{1}{2}} \mathbf{x} - \left(\boldsymbol{\Sigma}^{-\frac{1}{2}} \boldsymbol{\mu} + \boldsymbol{\Sigma}^{\frac{1}{2}} \mathbf{t} \right) \right) \\ &- \boldsymbol{\mu}^{\top} \mathbf{t} - \mathbf{t}^{\top} \boldsymbol{\mu} - \mathbf{t}^{\top} \boldsymbol{\Sigma} \mathbf{t} \bigg) \\ &= (\!\! \boldsymbol{k} \!\! \boldsymbol{\Xi} \!\!$$

以上より、モーメント母関数は、

- ◆ロト ◆昼 ト ◆ 恵 ト - 恵 - かへで

ガウス分布 $\mathcal{N}(\mathbf{x} \mid \boldsymbol{\mu}, \boldsymbol{\Sigma})$ の平均は,

$$\mathbb{E}\left[\mathbf{x}\right]^{\top} = \frac{\mathrm{d}}{\mathrm{d}\mathbf{t}} \exp\left(\boldsymbol{\mu}^{\top}\mathbf{t} + \frac{1}{2}\mathbf{t}^{\top}\boldsymbol{\Sigma}\mathbf{t}\right) \bigg|_{\mathbf{t}=\mathbf{0}}$$
$$= (練習問題) = \boldsymbol{\mu}^{\top}$$

であるから, $\mathbb{E}\left[\mathbf{x}
ight]=oldsymbol{\mu}$.

ガウス分布 $\mathcal{N}(\mathbf{x} \mid \boldsymbol{\mu}, \boldsymbol{\Sigma})$ の 2 次モーメントは、合成関数の微分より、

$$\mathbb{E}\left[\mathbf{x}\mathbf{x}^{\top}\right] = \frac{\mathrm{d}^{2}}{\mathrm{d}\mathbf{t}^{\top}\mathrm{d}\mathbf{t}} \exp\left(\boldsymbol{\mu}^{\top}\mathbf{t} + \frac{1}{2}\mathbf{t}^{\top}\boldsymbol{\Sigma}\mathbf{t}\right)\Big|_{\mathbf{t}=\mathbf{0}}$$
$$= (練習問題)$$
$$= \boldsymbol{\mu}\boldsymbol{\mu}^{\top} + \boldsymbol{\Sigma}$$

 \mathbf{t}^{\top} による微分は, \mathbf{t} による微分を転置させたものとなる.

ガウス分布 $\mathcal{N}(\mathbf{x} \mid \boldsymbol{\mu}, \boldsymbol{\Sigma})$ の分散は, $\mathbb{E}\left[\mathbf{x}\mathbf{x}^{\top}\right]$, $\mathbb{E}\left[\mathbf{x}\right]$ を使えば簡単に求められる:

$$\begin{aligned} \operatorname{Var}\left[\mathbf{x}\right] &= \mathbb{E}\left[\left(\mathbf{x} - \mathbb{E}\left[\mathbf{x}\right]\right)\left(\mathbf{x} - \mathbb{E}\left[\mathbf{x}\right]\right)^{\top}\right] \\ &= \mathbb{E}\left[\mathbf{x}\mathbf{x}^{\top}\right] - \mathbb{E}\left[\mathbf{x}\right]\mathbb{E}\left[\mathbf{x}\right]^{\top} \\ &= \boldsymbol{\mu}\boldsymbol{\mu}^{\top} + \boldsymbol{\Sigma} - \boldsymbol{\mu}\boldsymbol{\mu}^{\top} = \boldsymbol{\Sigma} \end{aligned}$$

モーメント母関数を使うことで、積分を使った最初の方法よりも楽に、平均と共分散を計算できた.

まとめ

- ガウス分布 (1 次元)
 - 平均,分散,高次のモーメント,モーメント母関数
- 多変量ガウス分布
 - 平均, 共分散, 共分散の性質, 無相関化, モーメント母関数
- まだ, モーメントしか確認できていない!
- 次回以降, ガウス分布の他の性質について確認していく:
 - 条件付き分布, 周辺分布, 線形変換, 非線形変換