

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

① 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)

③ 多変量ガウス分布

ガウス分布 (1 次元)

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.

- $p(x \mid \mu, \sigma^2)$ を, $\mathcal{N}(x \mid \mu, \sigma^2)$ 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:

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

- 大文字 X を確率変数, 小文字 x を実現値 (実際の観測値) として, 区別することが多い (このスライドでは, あまり正確に区別していない).

ガウス分布 (1 次元)

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 (1 次元)

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:

$$\int_{-\infty}^{\infty} \mathcal{N}(x \mid \mu, \sigma^2) \mathrm{d}x = 1$$

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

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

ガウス積分の公式

ガウス積分の公式 ($-\infty$ から ∞)

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

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

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

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

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

ガウス分布 (1 次元)

- 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,

$$x = \sqrt{2\sigma^2}y + \mu \quad \longrightarrow \quad \frac{dx}{dy} = \sqrt{2\sigma^2}$$

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

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

ガウス分布 (1 次元)

- よって,

$$\begin{aligned} & \int_{-\infty}^{\infty} \mathcal{N}(x \mid \mu, \sigma^2) dx \\ &= \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{1}{2\sigma^2} (x - \mu)^2\right) dx \\ &= \frac{1}{\sqrt{2\pi\sigma^2}} \underbrace{\int_{-\infty}^{\infty} \exp\left(-\frac{1}{2\sigma^2} (x - \mu)^2\right) dx}_{=\sqrt{2\pi\sigma^2}} \\ &= \frac{1}{\sqrt{2\pi\sigma^2}} \sqrt{2\pi\sigma^2} = 1 \end{aligned}$$

- 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:

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

- ただし,

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

- 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:

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

- 第 1 項は, $y = \frac{x - \mu}{\sqrt{2\sigma^2}}$ とすると, $x = \sqrt{2\sigma^2}y + \mu$, $\frac{dx}{dy} = \sqrt{2\sigma^2}$ だから,

$$\sqrt{2\sigma^2} \int_{-\infty}^{\infty} y \exp(-y^2) \frac{dx}{dy} dy = 2\sigma^2 \underbrace{\int_{-\infty}^{\infty} y \exp(-y^2) dy}_{=0} = 0$$

- $y \exp(-y^2)$ 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,

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

- ガウス分布の平均は, $\mathbb{E}[x] = \mu$.

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:

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

- ただし,

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

- 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,

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

- ガウス分布の分散は, $\mathbb{E}[(x - \mu)^2] = \text{Var}[x] = \sigma^2$.

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,

$$\begin{aligned}\mathbb{E}[x] &= \mu \\ \mathbb{E}[(x - \mu)^2] &= \text{Var}[x] = \sigma^2\end{aligned}$$

- ガウス積分を使えば, $\mathbb{E}[(x - \mu)^3]$ や $\mathbb{E}[(x - \mu)^4]$ 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.

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, $M_X(t) = \mathbb{E}[\exp(tX)]$ を, モーメント母関数という.

$$M_X(t) = \mathbb{E}[\exp(tX)] = \mathbb{E}\left[1 + tX + \frac{1}{2!}t^2X^2 + \frac{1}{3!}t^3X^3 + \dots\right]$$

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

$$\mathbb{E}[X^n] = M_X^{(n)}(t)\Big|_{t=0} = \frac{\partial^n}{\partial t^n} \mathbb{E}[\exp(tX)]\Big|_{t=0}$$

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,

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

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

- モーメント母関数:

$$\begin{aligned} M_X(t) &= \mathbb{E}[\exp(tX)] = \mathbb{E}\left[1 + tX + \frac{1}{2!}t^2X^2 + \frac{1}{3!}t^3X^3 + \dots\right] \\ &= 1 + \mathbb{E}[X]t + \frac{1}{2!}\mathbb{E}[X^2]t^2 + \frac{1}{3!}\mathbb{E}[X^3]t^3 + \dots \end{aligned}$$

- 例えば, 上式を t で $n = 3$ 回微分すれば,

$$M_X^{(3)}(t) = \frac{\partial^3}{\partial t^3} \mathbb{E}[\exp(tX)] = \mathbb{E}[X^3] + \mathbb{E}[X^4]t + \frac{1}{2!}\mathbb{E}[X^5]t^2 + \dots$$

- $t = 0$ を代入すれば,

$$M_X^{(3)}(t)\Big|_{t=0} = \frac{\partial^3}{\partial t^3} \mathbb{E}[\exp(tX)]\Big|_{t=0} = \mathbb{E}[X^3]$$

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 $M_X(t)$ を求めてみよう.

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

- $\exp(\cdot)$ の中身を, 平方完成する:

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

ガウス分布に対するモーメント母関数 (1 次元)

- よって,

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

- $y = \frac{x - (\mu + \sigma^2 t)}{\sqrt{2\sigma^2}}$ とおくと, $x = \sqrt{2\sigma^2}y + \mu + \sigma^2 t$ であり,
 $\frac{dx}{dy} = \sqrt{2\sigma^2}$ となるから,

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

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、

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

- 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,

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

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

$$\text{Var}[x] = \mathbb{E}[x^2] - \mathbb{E}[x]^2 = \sigma^2$$

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:

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

$$\mathbb{E}[x^2] = \mu^2 + \sigma^2$$

$$\mathbb{E}[x^3] = \mu^3 + 3\mu\sigma^2$$

$$\mathbb{E}[x^4] = \mu^4 + 6\mu^2\sigma^2 + 3\sigma^4$$

$$\mathbb{E}[x^5] = \mu^5 + 10\mu^3\sigma^2 + 15\mu\sigma^4$$

$$\mathbb{E}[x^6] = \mu^6 + 15\mu^4\sigma^2 + 45\mu^2\sigma^4 + 15\sigma^6$$

$$\mathbb{E}[x^7] = \mu^7 + 21\mu^5\sigma^2 + 105\mu^3\sigma^4 + 105\mu\sigma^6$$

$$\mathbb{E}[x^8] = \mu^8 + 28\mu^6\sigma^2 + 210\mu^4\sigma^4 + 420\mu^2\sigma^6 + 105\sigma^8$$

$$\mathbb{E}[x^9] = \mu^9 + 36\mu^7\sigma^2 + 378\mu^5\sigma^4 + 1260\mu^3\sigma^6 + 945\mu\sigma^8$$

$$\mathbb{E}[x^{10}] = \mu^{10} + 45\mu^8\sigma^2 + 630\mu^6\sigma^4 + 3150\mu^4\sigma^6 + 4725\mu^2\sigma^8 + 945\sigma^{10}$$

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:

$$\begin{aligned}\mathbb{E}[x - \mu] &= 0, & \mathbb{E}[(x - \mu)^2] &= \sigma^2 \\ \mathbb{E}[(x - \mu)^3] &= 0, & \mathbb{E}[(x - \mu)^4] &= 3\sigma^4 \\ \mathbb{E}[(x - \mu)^5] &= 0, & \mathbb{E}[(x - \mu)^6] &= 15\sigma^6 \\ \mathbb{E}[(x - \mu)^7] &= 0, & \mathbb{E}[(x - \mu)^8] &= 105\sigma^8 \\ \mathbb{E}[(x - \mu)^9] &= 0, & \mathbb{E}[(x - \mu)^{10}] &= 945\sigma^{10}\end{aligned}$$

- $y = x - \mu$ 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)

- 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,

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

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

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

高次の平均まわりのモーメントは,

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

- 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:

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

$\boldsymbol{\mu}$ は平均, $\boldsymbol{\Sigma}$ は共分散 (分散共分散行列) とよぶ.

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

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

多変量ガウス分布

多変量ガウス分布 (多変量正規分布)

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

- 定数項の分母には、いくつかの書き方がある:

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

- $\boldsymbol{\Sigma}$ は D 次行列だから, $\det(c\boldsymbol{\Sigma}) = c^D \det \boldsymbol{\Sigma}$ である.
- $\det \mathbf{A}$ を, $|\mathbf{A}|$ 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 Σ は, 次のように表される ($D \times D$ 行列):

$$\Sigma = \begin{pmatrix} \text{Var}[x_1] & \text{Cov}(x_1, x_2) & \cdots & \text{Cov}(x_1, x_D) \\ \text{Cov}(x_2, x_1) & \text{Var}[x_2] & \cdots & \text{Cov}(x_2, x_D) \\ \vdots & \vdots & \ddots & \vdots \\ \text{Cov}(x_D, x_1) & \text{Cov}(x_D, x_2) & \cdots & \text{Var}[x_D] \end{pmatrix}$$

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

標準正規分布 (多次元)

標準正規分布 (多次元)

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

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

\mathbf{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:

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

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)

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

- 共分散 $\boldsymbol{\Sigma}$ は, **正定値対称**行列.

$$\boldsymbol{\Sigma}^\top = \boldsymbol{\Sigma}$$

$$\forall \mathbf{x} \neq \mathbf{0}, \mathbf{x}^\top \boldsymbol{\Sigma} \mathbf{x} \geq 0 \quad (\boldsymbol{\Sigma} \geq 0 \text{ ともかく})$$

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

$$(\boldsymbol{\Sigma}^{-1})^\top = \boldsymbol{\Sigma}^{-1}$$

$$\forall \mathbf{x} \neq \mathbf{0}, \mathbf{x}^\top \boldsymbol{\Sigma}^{-1} \mathbf{x} \geq 0 \quad (\boldsymbol{\Sigma}^{-1} \geq 0 \text{ ともかく})$$

多変量ガウス分布の共分散

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

$$(\Sigma^{-1})^{\top} = (\Sigma^{\top})^{-1} = \Sigma^{-1}$$

- 正定値性:

Σ が正定値 $\rightarrow \Sigma$ の固有値 $\{\lambda_i\}$ は全て正

$\rightarrow \Sigma^{-1}$ の固有値は $\{\lambda_i^{-1}\}$ であるが, それらも全て正

$\rightarrow \Sigma^{-1}$ は正定値

- Σ, Σ^{-1} は正定値行列なので, **平方根** $\Sigma^{\frac{1}{2}}, \Sigma^{-\frac{1}{2}}$ 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.

多変量ガウス分布の共分散

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

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

- Λ^{-1} は, Σ の固有値の逆数 λ_i^{-1} を対角成分にもった, 対角行列:

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

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

多変量ガウス分布の共分散

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

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

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

$$\Lambda^{\frac{1}{2}} = \text{diag}(\sqrt{\lambda_1}, \dots, \sqrt{\lambda_D})$$

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

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

多変量ガウス分布の共分散

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

$$\begin{aligned}\Sigma &= U \Lambda U^{\top}, & \Sigma^{-1} &= U \Lambda^{-1} U^{\top} \\ \Sigma^{\frac{1}{2}} &= U \Lambda^{\frac{1}{2}} U^{\top}, & \Sigma^{-\frac{1}{2}} &= U \Lambda^{-\frac{1}{2}} U^{\top}\end{aligned}$$

- Λ は, A の固有値 λ_i を対角成分にもった, 対角行列:

$$\begin{aligned}\Lambda &= \text{diag}(\lambda_1, \dots, \lambda_D), & \Lambda^{-1} &= \text{diag}\left(\frac{1}{\lambda_1}, \dots, \frac{1}{\lambda_D}\right) \\ \Lambda^{\frac{1}{2}} &= \text{diag}\left(\sqrt{\lambda_1}, \dots, \sqrt{\lambda_D}\right), & \Lambda^{-\frac{1}{2}} &= \text{diag}\left(\frac{1}{\sqrt{\lambda_1}}, \dots, \frac{1}{\sqrt{\lambda_D}}\right)\end{aligned}$$

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

多変量ガウス分布の共分散

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

Σ が正定値 $\rightarrow \Sigma$ の固有値 $\{\lambda_i\}$ は全て正

\rightarrow 行列式は, 全ての固有値の積なので正

多変量ガウス分布の無相関化

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

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

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

$$\det \mathbf{U}\mathbf{U}^\top = \det \mathbf{I} = 1$$

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

- 対角行列の行列式は, 対角成分の積であるから,

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

多変量ガウス分布の無相関化

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

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

- ここで, $\mathbf{y} = \mathbf{U}^\top (\mathbf{x} - \boldsymbol{\mu})$ とおいた. $\mathbf{\Lambda}^{-1} = \text{diag}(\lambda_1^{-1}, \dots, \lambda_D^{-1})$ だから,

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

多変量ガウス分布の無相関化

- 係数部分:

$$\begin{aligned}\det \boldsymbol{\Sigma} = \det \boldsymbol{\Lambda} = \prod_{i=1}^D \lambda_i &\longrightarrow \frac{1}{\sqrt{\det \boldsymbol{\Sigma}}} = \prod_{i=1}^D \frac{1}{\sqrt{\lambda_i}} \\ &\longrightarrow \frac{1}{(2\pi)^{\frac{D}{2}} \sqrt{\det \boldsymbol{\Sigma}}} = \prod_{i=1}^D \frac{1}{\sqrt{2\pi \lambda_i}}\end{aligned}$$

- $\exp(\cdot)$ の中身:

$$\begin{aligned}(\mathbf{x} - \boldsymbol{\mu})^\top \boldsymbol{\Sigma}^{-1} (\mathbf{x} - \boldsymbol{\mu}) &= \mathbf{y}^\top \boldsymbol{\Lambda}^{-1} \mathbf{y} = \sum_{i=1}^D \frac{y_i^2}{\lambda_i} \\ \longrightarrow \exp\left(-\frac{1}{2} (\mathbf{x} - \boldsymbol{\mu})^\top \boldsymbol{\Sigma}^{-1} (\mathbf{x} - \boldsymbol{\mu})\right) &= \prod_{i=1}^D \exp\left(-\frac{y_i^2}{2\lambda_i}\right)\end{aligned}$$

多変量ガウス分布の無相関化

- これらの結果を基に、ガウス分布を書き換えると、

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

- $\mathbf{y} = \mathbf{U}^\top (\mathbf{x} - \boldsymbol{\mu})$ 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。
- \mathbf{y} の各成分は、互いに**独立**だから、**無相関**でもある。

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

多変量ガウス分布の無相関化

- 次のように、行列の形でも書き換えておく：

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

- $\mathbf{y} = \mathbf{U}^\top (\mathbf{x} - \boldsymbol{\mu})$ 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 。
- 共分散 $\boldsymbol{\Lambda}$ の (i, j) 成分は、 $\text{Cov}(y_i, y_j)$ である (対角成分は $\text{Var}[y_i]$)。
- $i \neq j$ であるとき、 $\text{Cov}(y_i, y_j) = 0$ 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)

- \mathbf{x} についてのガウス分布:

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

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

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

- \mathbf{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:

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

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

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

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

多変量ガウス分布

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

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

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

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

- \mathbf{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,

$$\begin{aligned}\int_{-\infty}^{\infty} \mathcal{N}(\mathbf{x} \mid \boldsymbol{\mu}, \boldsymbol{\Sigma}) d\mathbf{x} &= \int_{-\infty}^{\infty} \mathcal{N}(\mathbf{y} \mid \mathbf{0}, \mathbf{\Lambda}) |\det \mathbf{J}| d\mathbf{y} \\ &= \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} \prod_{i=1}^D \mathcal{N}(y_i \mid 0, \lambda_i) dy_1 \cdots dy_D \\ &= \int_{-\infty}^{\infty} \mathcal{N}(y_1 \mid 0, \lambda_1) dy_1 \cdots \int_{-\infty}^{\infty} \mathcal{N}(y_D \mid 0, \lambda_1) dy_D = 1\end{aligned}$$

- 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.
- 積分すると1になるから、ガウス分布 $\mathcal{N}(\mathbf{x} \mid \boldsymbol{\mu}, \boldsymbol{\Sigma})$ 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:

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

ただし,

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

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

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

多変量ガウス分布の平均

ヤコビ行列 \mathbf{J} は, $\mathbf{x} = \mathbf{z} + \boldsymbol{\mu}$ だから, $\frac{\partial \mathbf{x}}{\partial \mathbf{z}} = \mathbf{I}$ 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,

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

$\mathbf{f}(\mathbf{z}) = \exp\left(-\frac{1}{2} \mathbf{z}^{\top} \boldsymbol{\Sigma}^{-1} \mathbf{z}\right) \mathbf{z}$ とおくと, $\mathbf{f}(-\mathbf{z}) = -\mathbf{f}(\mathbf{z})$ より奇関数なので,

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

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,

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

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

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

多変量ガウス分布の共分散

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

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

ただし,

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

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

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

多変量ガウス分布の共分散

先程と同様に、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,

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

$\mathbf{f}(\mathbf{z}) = \exp\left(-\frac{1}{2}\mathbf{z}^\top \Sigma^{-1} \mathbf{z}\right) \boldsymbol{\mu} \mathbf{z}^\top$, $\mathbf{g}(\mathbf{z}) = \exp\left(-\frac{1}{2}\mathbf{z}^\top \Sigma^{-1} \mathbf{z}\right) \mathbf{z} \boldsymbol{\mu}^\top$ 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,

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

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

多変量ガウス分布の共分散

また,

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

ガウス分布を積分すると 1 になるので, 容易に求まる.

多変量ガウス分布の共分散

続いて、以下の積分を考える:

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

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

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

多変量ガウス分布の共分散

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

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

ヤコビ行列は $\mathbf{J} = \frac{\partial \mathbf{z}}{\partial \mathbf{y}} = \mathbf{U}$ である. \mathbf{U} は直交行列だから, $\det \mathbf{U} = \pm 1$.

また $\mathcal{N}(\mathbf{y} \mid \mathbf{0}, \boldsymbol{\Lambda})$ は, 各成分のガウス分布の積に分解される:

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

多変量ガウス分布の共分散

これらの結果を使うと,

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

以下の積分を示そう:

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

多変量ガウス分布の共分散

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

$$\begin{aligned} & \left(\int_{-\infty}^{\infty} \prod_{k=1}^D \mathcal{N}(y_k \mid 0, \lambda_k) \mathbf{y}\mathbf{y}^\top \, d\mathbf{y} \right)_{ij} \\ &= \int_{-\infty}^{\infty} \prod_{j=1}^D \mathcal{N}(y_k \mid 0, \lambda_k) y_i y_j \, d\mathbf{y} \\ &= \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} \mathcal{N}(y_1 \mid 0, \lambda_1) \cdots \mathcal{N}(y_D \mid 0, \lambda_D) y_i y_j \, dy_1 \cdots dy_D \end{aligned}$$

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、

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

多変量ガウス分布の共分散

$i \neq j$ ならば,

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

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

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

$i = j$ ならば, $\int_{-\infty}^{\infty} x^2 \exp(-ax^2) \, dx = \frac{1}{2} \sqrt{\frac{\pi}{a^3}}$ を使って,

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

多変量ガウス分布の共分散

$i = j$ のときの積分は、次のようにも求められる:

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

$\text{Var}[x] = \mathbb{E}[x^2] - \mathbb{E}[x]^2$ と, $\mathbb{E}[y_i] = 0$, $\text{Var}[y_i] = \lambda_i$ 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,

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

多変量ガウス分布の共分散

よって,

$$\begin{aligned} & \frac{1}{(2\pi)^{\frac{D}{2}} \sqrt{\det \Sigma}} \int_{-\infty}^{\infty} \exp\left(-\frac{1}{2} \mathbf{z}^{\top} \Sigma^{-1} \mathbf{z}\right) \mathbf{z} \mathbf{z}^{\top} d\mathbf{z} \\ &= \mathbf{U} \left(\int_{-\infty}^{\infty} \prod_{i=1}^D \mathcal{N}(y_i \mid 0, \lambda_i) \mathbf{y} \mathbf{y}^{\top} d\mathbf{y} \right) \mathbf{U}^{\top} = \mathbf{U} \mathbf{\Lambda} \mathbf{U}^{\top} = \Sigma \end{aligned}$$

多変量ガウス分布の共分散

以上より,

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

ただし,

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

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

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

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

多変量ガウス分布の共分散

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

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

ただし,

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

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

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

多変量ガウス分布の共分散

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,

$$\begin{aligned} & \mathbb{E} \left[(\mathbf{x} - \boldsymbol{\mu}) (\mathbf{x} - \boldsymbol{\mu})^\top \right] \\ &= \frac{1}{(2\pi)^{\frac{D}{2}} \sqrt{\det \boldsymbol{\Sigma}}} \int_{-\infty}^{\infty} \exp \left(-\frac{1}{2} \mathbf{z}^\top \boldsymbol{\Sigma}^{-1} \mathbf{z} \right) \mathbf{z} \mathbf{z}^\top |\det \mathbf{J}| d\mathbf{z} \\ &= \frac{1}{(2\pi)^{\frac{D}{2}} \sqrt{\det \boldsymbol{\Sigma}}} \int_{-\infty}^{\infty} \exp \left(-\frac{1}{2} \mathbf{z}^\top \boldsymbol{\Sigma}^{-1} \mathbf{z} \right) \mathbf{z} \mathbf{z}^\top d\mathbf{z} \\ &= \boldsymbol{\Sigma} \quad (\text{先程行った積分と同じ}) \end{aligned}$$

次のようにも確かめられる:

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

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

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

このガウス分布の平均と共分散は,

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

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

- 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.

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)

モーメント母関数 (多次元)

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

$$M_{\mathbf{X}}(\mathbf{t}) = \mathbb{E} \left[\exp(\mathbf{t}^\top \mathbf{X}) \right] = \mathbb{E} \left[1 + \mathbf{t}^\top \mathbf{X} + \frac{1}{2!} (\mathbf{t}^\top \mathbf{X})^2 + \dots \right]$$

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

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

$$\mathbb{E}[\mathbf{X}] = \left. \frac{\partial}{\partial \mathbf{t}} M_{\mathbf{X}}(\mathbf{t}) \right|_{\mathbf{t}=0}$$
$$\mathbb{E}[\mathbf{X}\mathbf{X}^\top] = \left. \frac{\partial^2}{\partial \mathbf{t}^\top \partial \mathbf{t}} M_{\mathbf{X}}(\mathbf{t}) \right|_{\mathbf{t}=0}$$

モーメント母関数 (多次元)

- 分子レイアウトでは, 次が成り立つ (スカラーのベクトルによる微分):

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

- 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,

$$\begin{aligned}\frac{\partial (\mathbf{a}^\top \mathbf{x})^n}{\partial x_i} &= \frac{\partial (\mathbf{x}^\top \mathbf{a})^n}{\partial x_i} = n (\mathbf{x}^\top \mathbf{a})^{n-1} \frac{\partial \mathbf{x}^\top \mathbf{a}}{\partial x_i} \\ &= n (\mathbf{x}^\top \mathbf{a})^{n-1} \left(\frac{\partial \mathbf{x}^\top \mathbf{a}}{\partial \mathbf{x}} \right)_i = n (\mathbf{x}^\top \mathbf{a})^{n-1} (\mathbf{a})_i\end{aligned}$$

モーメント母関数 (多次元)

- スカラの \mathbf{x}^\top による微分は, \mathbf{x} による微分を転置したものと考える:

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

モーメント母関数 (多次元)

- モーメント母関数:

$$\begin{aligned}M_{\mathbf{X}}(\mathbf{t}) &= \mathbb{E} \left[\exp(\mathbf{t}^\top \mathbf{X}) \right] = \mathbb{E} \left[1 + \mathbf{t}^\top \mathbf{X} + \frac{1}{2!} (\mathbf{t}^\top \mathbf{X})^2 + \frac{1}{3!} (\mathbf{t}^\top \mathbf{X})^3 + \dots \right] \\&= 1 + \mathbb{E} [\mathbf{t}^\top \mathbf{X}] + \frac{1}{2!} \mathbb{E} \left[(\mathbf{t}^\top \mathbf{X})^2 \right] + \frac{1}{3!} \mathbb{E} \left[(\mathbf{t}^\top \mathbf{X})^3 \right] + \dots \\&= 1 + \mathbf{t}^\top \mathbb{E} [\mathbf{X}] + \frac{1}{2!} \mathbf{t}^\top \mathbb{E} [\mathbf{X} \mathbf{X}^\top] \mathbf{t} + \frac{1}{3!} \mathbf{t}^\top \mathbb{E} [\mathbf{X} \mathbf{t}^\top \mathbf{X} \mathbf{X}^\top] \mathbf{t} + \dots\end{aligned}$$

- \mathbf{t} で 1 回微分して $\mathbf{t} = \mathbf{0}$ を代入すれば, $\mathbb{E} [\mathbf{X}]^\top$ を得る.

$$\frac{\partial}{\partial \mathbf{t}} M_{\mathbf{X}}(\mathbf{t}) = \mathbb{E} [\mathbf{X}]^\top + \mathbf{t}^\top \mathbb{E} [\mathbf{X} \mathbf{X}^\top] + \dots$$

モーメント母関数 (多次元)

- \mathbf{t} で 1 回微分すると,

$$\frac{\partial}{\partial \mathbf{t}} M_{\mathbf{X}}(\mathbf{t}) = \mathbb{E}[\mathbf{X}]^{\top} + \mathbf{t}^{\top} \mathbb{E}[\mathbf{X}\mathbf{X}^{\top}] + \dots$$

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

$$\frac{\partial^2}{\partial \mathbf{t}^{\top} \partial \mathbf{t}} M_{\mathbf{X}}(\mathbf{t}) = \mathbb{E}[\mathbf{X}\mathbf{X}^{\top}] + \dots$$

- $\mathbf{t} = \mathbf{0}$ を代入すれば, $\mathbb{E}[\mathbf{X}\mathbf{X}^{\top}]$ を得る.

多次元ガウス分布に対するモーメント母関数

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,

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

ただし,

$$\begin{aligned} M_{\mathbf{X}}(\mathbf{t}) &= \int_{-\infty}^{\infty} \mathcal{N}(\mathbf{x} \mid \boldsymbol{\mu}, \boldsymbol{\Sigma}) \exp(\mathbf{t}^\top \mathbf{x}) \, d\mathbf{x} \\ &= \frac{1}{(2\pi)^{\frac{D}{2}} \sqrt{\det \boldsymbol{\Sigma}}} \int_{-\infty}^{\infty} \exp \left(-\frac{1}{2} (\mathbf{x} - \boldsymbol{\mu})^\top \boldsymbol{\Sigma}^{-1} (\mathbf{x} - \boldsymbol{\mu}) \right) \\ &\quad \exp(\mathbf{t}^\top \mathbf{x}) \, d\mathbf{x} \end{aligned}$$

多次元ガウス分布に対するモーメント母関数

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

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

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

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

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

多次元ガウス分布に対するモーメント母関数

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

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

多次元ガウ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,

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

多次元ガウス分布に対するモーメント母関数

以上より、モーメント母関数は、

$$\begin{aligned} M_{\mathbf{X}}(\mathbf{t}) &= \frac{1}{(2\pi)^{\frac{D}{2}} \sqrt{\det \Sigma}} \int_{-\infty}^{\infty} \exp\left(-\frac{1}{2} (\mathbf{x} - \boldsymbol{\mu})^{\top} \Sigma^{-1} (\mathbf{x} - \boldsymbol{\mu})\right) \\ &\quad \exp(\mathbf{t}^{\top} \mathbf{x}) d\mathbf{x} \\ &= \frac{1}{(2\pi)^{\frac{D}{2}} \sqrt{\det \Sigma}} \exp\left(\boldsymbol{\mu}^{\top} \mathbf{t} + \frac{1}{2} \mathbf{t}^{\top} \Sigma \mathbf{t}\right) \\ &\quad \int_{-\infty}^{\infty} \exp\left(-\frac{1}{2} (\mathbf{x} - (\boldsymbol{\mu} + \Sigma \mathbf{t}))^{\top} \Sigma^{-1} (\mathbf{x} - (\boldsymbol{\mu} + \Sigma \mathbf{t}))\right) d\mathbf{x} \\ &= \exp\left(\boldsymbol{\mu}^{\top} \mathbf{t} + \frac{1}{2} \mathbf{t}^{\top} \Sigma \mathbf{t}\right) \int_{-\infty}^{\infty} \mathcal{N}(\mathbf{x} \mid \boldsymbol{\mu} + \Sigma \mathbf{t}, \Sigma) d\mathbf{x} \\ &= \exp\left(\boldsymbol{\mu}^{\top} \mathbf{t} + \frac{1}{2} \mathbf{t}^{\top} \Sigma \mathbf{t}\right) \quad (\because \text{ガウス分布の積分は } 1) \end{aligned}$$

多次元ガウス分布に対するモーメント母関数

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

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

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

多次元ガウス分布に対するモーメント母関数

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,

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

\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.

多次元ガウス分布に対するモーメント母関数

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

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

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