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

目次

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

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:

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

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：

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

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

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

ガウス積分の公式

ガウス積分の公式 ($-\infty$ から ∞)

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

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

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

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

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

ガウス分布 (1 次元)

- 以下の積分を確認する:

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

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

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

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

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

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

ガウス分布 (1 次元)

- よって,

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

- 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,

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

- 次の積分に着目する:

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

ガウス分布の平均 (1 次元)

- 次のように分解する:

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

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

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

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

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

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

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:

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

- ただし,

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

- 次の積分に着目する:

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

ガウス分布の分散 (1 次元)

- 次の積分に着目する:

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

- $y = \dfrac{x-\mu}{\sqrt{2\sigma^2}}$ とすると, $x = \sqrt{2\sigma^2}y + \mu$, $\dfrac{\mathrm{d}x}{\mathrm{d}y} = \sqrt{2\sigma^2}$ だから,

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

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

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

ガウス分布の分散 (1 次元)

- よって,

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

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

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,

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

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

- 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

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．

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

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:

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

- 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,

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

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．

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

- $\exp(\cdot)$ の中身を，平方完成する:

$$\begin{aligned} tx - \frac{1}{2\sigma^2}(x-\mu)^2 &= -\frac{1}{2\sigma^2}\left(x^2 - 2\mu x - 2\sigma^2 tx + \mu^2\right) \\ &= -\frac{1}{2\sigma^2}\left(\left(x - (\mu + \sigma^2 t)\right)^2 - 2\mu\sigma^2 t - \sigma^4 t^2\right) \\ &= \left(\mu t + \frac{\sigma^2 t^2}{2}\right) - \frac{1}{2\sigma^2}\left(x - (\mu + \sigma^2 t)\right)^2 \end{aligned}$$

ガウス分布に対するモーメント母関数 (1 次元)

- よって，

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

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

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

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，

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

- 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,

$$
\begin{aligned}
\mathbb{E}[x] &= \frac{\mathrm{d}}{\mathrm{d}t} \exp\left(\mu t + \frac{\sigma^2 t^2}{2}\right)\bigg|_{t=0} \\
&= \exp\left(\mu t + \frac{\sigma^2 t^2}{2}\right)\left(\mu + \sigma^2 t\right)\bigg|_{t=0} = \mu \\
\mathbb{E}\left[x^2\right] &= \frac{\mathrm{d}^2}{\mathrm{d}t^2} \exp\left(\mu t + \frac{\sigma^2 t^2}{2}\right)\bigg|_{t=0} \\
&= \exp\left(\mu t + \frac{\sigma^2 t^2}{2}\right)\sigma^2 + \exp\left(\mu t + \frac{\sigma^2 t^2}{2}\right)\left(\mu + \sigma^2 t\right)^2\bigg|_{t=0} \\
&= \sigma^2 + \mu^2 \\
\mathrm{Var}[x] &= \mathbb{E}\left[x^2\right] - \mathbb{E}[x]^2 = \sigma^2
\end{aligned}
$$

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:

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

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:

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

- $y = x - \mu$ 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, $\mathbb{E}\left[y^n\right] = \mathbb{E}\left[(x-\mu)^n\right]$ は, 先程求めた $\mathcal{N}(x \mid \mu, \sigma^2)$ のモーメント $\mathbb{E}\left[x^n\right]$ 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,

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

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

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

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

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)

- $\mathbf{x} = (x_1, \ldots, x_D)$ と表す (D 次元).
- 共分散 $\boldsymbol{\Sigma}$ は, 次のように表される ($D \times D$ 行列):

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

- $\boldsymbol{\Sigma}$ の (i, j) 成分は, $\mathrm{Cov}\left(x_i, x_j\right) = \mathbb{E}\left[\left(x_i - \mathbb{E}\left[x_i\right]\right)\left(x_j - \mathbb{E}\left[x_j\right]\right)\right]$.
- $\mathrm{Cov}\left(x_i, x_j\right) = \mathrm{Cov}\left(x_j, x_i\right)$ だから, $\boldsymbol{\Sigma}$ は対称.
- $\mathrm{Var}\left[x_i\right] = \mathrm{Cov}\left(x_i, x_i\right)$ に注意.
- $\boldsymbol{\Sigma}$ 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, $\boldsymbol{\Sigma}$ は対角行列となる ($i \neq j \rightarrow \mathrm{Cov}\left(x_i, x_j\right) = 0$).

標準正規分布 (多次元)

平均 $\boldsymbol{\mu} = \mathbf{0}$, 共分散 $\boldsymbol{\Sigma} = \mathbf{I}$ 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:

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

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.

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

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

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

多変量ガウス分布の共分散

- 共分散 $\boldsymbol{\Sigma}$ は, 正定値対称行列.
- このとき, 逆行列 $\boldsymbol{\Sigma}^{-1}$ も, 正定値対称行列になる.
- 対称性:

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

- 正定値性:

$$\begin{aligned} \boldsymbol{\Sigma}\text{ が正定値} &\rightarrow \boldsymbol{\Sigma}\text{ の固有値 } \{\lambda_i\} \text{ は全て正} \\ &\rightarrow \boldsymbol{\Sigma}^{-1}\text{ の固有値は } \left\{\lambda_i^{-1}\right\} \text{ であるが, それらも全て正} \\ &\rightarrow \boldsymbol{\Sigma}^{-1}\text{ は正定値} \end{aligned}$$

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

- $\mathbf{\Sigma}$ の固有値が $\{\lambda_1, \dots, \lambda_D\}$ であるとき, $\mathbf{\Sigma}^{-1}$ 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 $\mathbf{\Sigma}$ は, 直交行列 $\mathbf{U}$ で, $\mathbf{\Lambda} = \mathbf{U}^\top \mathbf{\Sigma} \mathbf{U}$ と対角化できる.
- $\mathbf{\Lambda}$ は, $\mathbf{\Sigma}$ の固有値 λ_i を対角成分にもった, 対角行列:

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

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

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

- $\mathbf{\Lambda}^{-1} = \mathbf{U}^\top \mathbf{\Sigma}^{-1} \mathbf{U}$ 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).
- $\mathbf{\Lambda}^{-1}$ の対角成分は, $\mathbf{\Sigma}^{-1}$ の固有値 (= $\mathbf{\Sigma}$ の固有値の逆数)

多変量ガウス分布の共分散

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

$$\mathbf{\Lambda} = \mathrm{diag}(\lambda_1, \ldots, \lambda_D)$$

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

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

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

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

多変量ガウス分布の共分散

- $\boldsymbol{\Sigma}, \boldsymbol{\Sigma}^{-1}$ は正定値行列なので，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，

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

- $\boldsymbol{\Lambda}$ は，$\mathbf{A}$ の固有値 λ_i を対角成分にもった，対角行列:

$$\boldsymbol{\Lambda} = \mathrm{diag}(\lambda_1, \ldots, \lambda_D), \quad \boldsymbol{\Lambda}^{-1} = \mathrm{diag}\left(\frac{1}{\lambda_1}, \ldots, \frac{1}{\lambda_D}\right)$$
$$\boldsymbol{\Lambda}^{\frac{1}{2}} = \mathrm{diag}\left(\sqrt{\lambda_1}, \ldots, \sqrt{\lambda_D}\right), \quad \boldsymbol{\Lambda}^{-\frac{1}{2}} = \mathrm{diag}\left(\frac{1}{\sqrt{\lambda_1}}, \ldots, \frac{1}{\sqrt{\lambda_D}}\right)$$

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

多変量ガウス分布の共分散

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

$$
\begin{aligned}
\boldsymbol{\Sigma}\ \text{が正定値} &\to \boldsymbol{\Sigma}\ \text{の固有値}\ \{\lambda_i\}\ \text{は全て正} \\
&\to \text{行列式は，全ての固有値の積なので正}
\end{aligned}
$$

多変量ガウス分布の無相関化

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

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

- $\mathbf{U}\mathbf{U}^\top = \mathbf{I}$ 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,

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

多変量ガウス分布の無相関化

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

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

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

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

多変量ガウス分布の無相関化

- 係数部分:

$$\det \boldsymbol{\Sigma} = \det \boldsymbol{\Lambda} = \prod_{i=1}^{D} \lambda_i \quad \longrightarrow \quad \frac{1}{\sqrt{\det \boldsymbol{\Sigma}}} = \prod_{i=1}^{D} \frac{1}{\sqrt{\lambda_i}}$$

$$\longrightarrow \quad \frac{1}{(2\pi)^{\frac{D}{2}} \sqrt{\det \boldsymbol{\Sigma}}} = \prod_{i=1}^{D} \frac{1}{\sqrt{2\pi\lambda_i}}$$

- $\exp(\cdot)$ の中身:

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

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

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.

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

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, $\mathrm{Cov}\left(y_i, y_j\right)$ である (対角成分は $\mathrm{Var}\left[y_i\right]$).
- $i \neq j$ であるとき, $\mathrm{Cov}\left(y_i, y_j\right) = 0$ 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:

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

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

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

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

多変量ガウス分布

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

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

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

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

- $\mathbf{U}$ 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,

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

- 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 $\mathcal{N}(\mathbf{x} \mid \boldsymbol{\mu}, \mathbf{\Sigma})$ 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:

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

ただし,

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

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

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

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, $\dfrac{\partial \mathbf{x}}{\partial \mathbf{x}} = \mathbf{I}$ 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,

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

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

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

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，

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

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

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

多変量ガウス分布の共分散

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,

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

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

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

多変量ガウス分布の共分散

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

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

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

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

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

多変量ガウス分布の共分散

また,

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

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:

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

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

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

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,

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

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

ヤコビ行列は $\mathbf{J} = \dfrac{\partial \mathbf{z}}{\partial \mathbf{y}} = \mathbf{U}$ 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,

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

以下の積分を示そう:

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

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,

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

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,

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

多変量ガウス分布の共分散

$i \neq j$ ならば,

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

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

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

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

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

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:

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

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

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

多変量ガウス分布の共分散

よって，

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

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

多変量ガウス分布の共分散

以上より，

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

ただし，

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

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,

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

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

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

多変量ガウス分布の共分散

ヤコビ行列は $\mathbf{J} = \dfrac{\partial \mathbf{x}}{\partial \mathbf{z}} = \mathbf{I}$, 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:

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

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,

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

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

- 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)

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, $\mathbb{E}[\mathbf{x}]$ 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, $\mathbb{E}\left[\mathbf{x}\mathbf{x}^\top\right]$ が得られる.

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

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

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):

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

- 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,

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

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:

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

モーメント母関数 (多次元)

- モーメント母関数:

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

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

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

モーメント母関数 (多次元)

- $\mathbf{t}$ で 1 回微分すると,

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

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

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

- $\mathbf{t} = \mathbf{0}$ を代入すれば, $\mathbb{E}\left[\mathbf{x}\mathbf{x}^\top\right]$ 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,

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

ただし,

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

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

$$
\begin{aligned}
&-\frac{1}{2}\left(\mathbf{x}^\top\boldsymbol{\Sigma}^{-1}\mathbf{x} - \boldsymbol{\mu}^\top\boldsymbol{\Sigma}^{-1}\mathbf{x} - \mathbf{t}^\top\mathbf{x} - \mathbf{x}^\top\boldsymbol{\Sigma}^{-1}\boldsymbol{\mu} - \mathbf{x}^\top\mathbf{t} + \boldsymbol{\mu}^\top\boldsymbol{\Sigma}^{-1}\boldsymbol{\mu}\right) \\
&= -\frac{1}{2}\Bigg(\mathbf{x}^\top\boldsymbol{\Sigma}^{-\frac{1}{2}}\boldsymbol{\Sigma}^{-\frac{1}{2}}\mathbf{x} - \left(\boldsymbol{\mu}^\top\boldsymbol{\Sigma}^{-\frac{1}{2}} + \mathbf{t}^\top\boldsymbol{\Sigma}^{\frac{1}{2}}\right)\boldsymbol{\Sigma}^{-\frac{1}{2}}\mathbf{x} \\
&\qquad - \mathbf{x}^\top\boldsymbol{\Sigma}^{-\frac{1}{2}}\left(\boldsymbol{\Sigma}^{-\frac{1}{2}}\boldsymbol{\mu} + \boldsymbol{\Sigma}^{\frac{1}{2}}\mathbf{t}\right) + \boldsymbol{\mu}^\top\boldsymbol{\Sigma}^{-1}\boldsymbol{\mu}\Bigg) \\
&= -\frac{1}{2}\Bigg(\left(\boldsymbol{\Sigma}^{-\frac{1}{2}}\mathbf{x}\right)^\top\boldsymbol{\Sigma}^{-\frac{1}{2}}\mathbf{x} - \left(\boldsymbol{\Sigma}^{-\frac{1}{2}}\boldsymbol{\mu} + \boldsymbol{\Sigma}^{\frac{1}{2}}\mathbf{t}\right)^\top\boldsymbol{\Sigma}^{-\frac{1}{2}}\mathbf{x} \\
&\qquad - \left(\boldsymbol{\Sigma}^{-\frac{1}{2}}\mathbf{x}\right)^\top\left(\boldsymbol{\Sigma}^{-\frac{1}{2}}\boldsymbol{\mu} + \boldsymbol{\Sigma}^{\frac{1}{2}}\mathbf{t}\right) + \boldsymbol{\mu}^\top\boldsymbol{\Sigma}^{-1}\boldsymbol{\mu}\Bigg) \quad (\because \text{対称性}) \\
&= -\frac{1}{2}\Bigg(\left(\boldsymbol{\Sigma}^{-\frac{1}{2}}\mathbf{x} - \left(\boldsymbol{\Sigma}^{-\frac{1}{2}}\boldsymbol{\mu} + \boldsymbol{\Sigma}^{\frac{1}{2}}\mathbf{t}\right)\right)^\top\left(\boldsymbol{\Sigma}^{-\frac{1}{2}}\mathbf{x} - \left(\boldsymbol{\Sigma}^{-\frac{1}{2}}\boldsymbol{\mu} + \boldsymbol{\Sigma}^{\frac{1}{2}}\mathbf{t}\right)\right) \\
&\qquad - \boldsymbol{\mu}^\top\mathbf{t} - \mathbf{t}^\top\boldsymbol{\mu} - \mathbf{t}^\top\boldsymbol{\Sigma}\mathbf{t}\Bigg)
\end{aligned}
$$

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，

$$
\begin{aligned}
&-\frac{1}{2}(\mathbf{x}-\boldsymbol{\mu})^\top \boldsymbol{\Sigma}^{-1}(\mathbf{x}-\boldsymbol{\mu})+\mathbf{t}^\top\mathbf{x}\\
&=-\frac{1}{2}\Bigg(\left(\boldsymbol{\Sigma}^{-\frac{1}{2}}\mathbf{x}-\left(\boldsymbol{\Sigma}^{-\frac{1}{2}}\boldsymbol{\mu}+\boldsymbol{\Sigma}^{\frac{1}{2}}\mathbf{t}\right)\right)^\top\left(\boldsymbol{\Sigma}^{-\frac{1}{2}}\mathbf{x}-\left(\boldsymbol{\Sigma}^{-\frac{1}{2}}\boldsymbol{\mu}+\boldsymbol{\Sigma}^{\frac{1}{2}}\mathbf{t}\right)\right)\\
&\quad-\boldsymbol{\mu}^\top\mathbf{t}-\mathbf{t}^\top\boldsymbol{\mu}-\mathbf{t}^\top\boldsymbol{\Sigma}\mathbf{t}\Bigg)\\
&=-\frac{1}{2}(\mathbf{x}-(\boldsymbol{\mu}+\boldsymbol{\Sigma}\mathbf{t}))^\top\boldsymbol{\Sigma}^{-\frac{1}{2}}\boldsymbol{\Sigma}^{-\frac{1}{2}}(\mathbf{x}-(\boldsymbol{\mu}+\boldsymbol{\Sigma}\mathbf{t}))\quad(\boldsymbol{\Sigma}^{-\frac{1}{2}}\text{ をくくり出す})\\
&\quad+\boldsymbol{\mu}^\top\mathbf{t}+\frac{1}{2}\mathbf{t}^\top\boldsymbol{\Sigma}\mathbf{t}\quad(\because\ \boldsymbol{\mu}^\top\mathbf{t}=\mathbf{t}^\top\boldsymbol{\mu})\\
&=-\frac{1}{2}(\mathbf{x}-(\boldsymbol{\mu}+\boldsymbol{\Sigma}\mathbf{t}))^\top\boldsymbol{\Sigma}^{-1}(\mathbf{x}-(\boldsymbol{\mu}+\boldsymbol{\Sigma}\mathbf{t}))+\boldsymbol{\mu}^\top\mathbf{t}+\frac{1}{2}\mathbf{t}^\top\boldsymbol{\Sigma}\mathbf{t}
\end{aligned}
$$

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，

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

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,

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

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

$$
\begin{aligned}
\mathbb{E}\left[\mathbf{x}\mathbf{x}^\top\right] &= \frac{\mathrm{d}^2}{\mathrm{d}\mathbf{t}^\top\,\mathrm{d}\mathbf{t}} \exp\left(\boldsymbol{\mu}^\top \mathbf{t} + \frac{1}{2}\mathbf{t}^\top \boldsymbol{\Sigma}\mathbf{t}\right)\bigg|_{\mathbf{t}=\mathbf{0}} \\
&= \frac{\partial}{\partial \mathbf{t}^\top} \exp\left(\boldsymbol{\mu}^\top \mathbf{t} + \frac{1}{2}\mathbf{t}^\top \boldsymbol{\Sigma}\mathbf{t}\right)\left(\boldsymbol{\mu}^\top + \mathbf{t}^\top \boldsymbol{\Sigma}\right) \\
&\quad + \exp\left(\boldsymbol{\mu}^\top \mathbf{t} + \frac{1}{2}\mathbf{t}^\top \boldsymbol{\Sigma}\mathbf{t}\right)\frac{\partial}{\partial \mathbf{t}^\top}\left(\boldsymbol{\mu}^\top + \mathbf{t}^\top \boldsymbol{\Sigma}\right)\bigg|_{\mathbf{t}=\mathbf{0}} \\
&= \exp\left(\boldsymbol{\mu}^\top \mathbf{t} + \frac{1}{2}\mathbf{t}^\top \boldsymbol{\Sigma}\mathbf{t}\right)^2 (\boldsymbol{\mu} + \boldsymbol{\Sigma}\mathbf{t})\left(\boldsymbol{\mu}^\top + \mathbf{t}^\top \boldsymbol{\Sigma}\right) \\
&\quad + \exp\left(\boldsymbol{\mu}^\top \mathbf{t} + \frac{1}{2}\mathbf{t}^\top \boldsymbol{\Sigma}\mathbf{t}\right)\boldsymbol{\Sigma}\bigg|_{\mathbf{t}=\mathbf{0}} = \boldsymbol{\mu}\boldsymbol{\mu}^\top + \boldsymbol{\Sigma}
\end{aligned}
$$

$\mathbf{t}^\top$ による微分は, 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, $\mathbb{E}\left[\mathbf{x}\mathbf{x}^\top\right]$, $\mathbb{E}\left[\mathbf{x}\right]$ を使えば簡単に求められる:

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

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