机器人学中的状态估计 - 作业 1

彭博

May 21, 2020

1. 证明高斯分布积分为 1.
   构造积分 I

$$
\begin{aligned}
I &= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \exp\{-\frac{x^2+y^2}{2\sigma^2}\}dxdy \\
&= \int_{0}^{2\pi}\int_{-\infty}^{\infty} \rho\exp\{-\frac{\rho^2}{2\sigma^2}\}d\rho d\theta \\
&= \pi\int_{-\infty}^{\infty} \exp\{-\frac{\rho^2}{2\sigma^2}\}d\rho^2 \\
&= 2\pi\sigma^2
\end{aligned} \tag{1}
$$

   分离变量 x 和 y，有

$$
\begin{aligned}
I &= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \exp\{-\frac{x^2+y^2}{2\sigma^2}\}dxdy \\
&= \int_{-\infty}^{\infty} \exp\{-\frac{x^2}{2\sigma^2}\}dx \int_{-\infty}^{\infty} \exp\{-\frac{y^2}{2\sigma^2}\}dy
\end{aligned} \tag{2}
$$

   因此

$$
\int_{-\infty}^{\infty} \exp\{-\frac{x^2}{2\sigma^2}\}dx = \sqrt{2\pi}\sigma \tag{3}
$$

   最后得到

$$
\begin{aligned}
\int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}\sigma}\exp\{-\frac{(x-\mu)^2}{2\sigma^2}\}dx &= \frac{1}{\sqrt{2\pi}\sigma}\int_{-\infty}^{\infty} \exp\{-\frac{x^2}{2\sigma^2}\}dx \\
&= 1
\end{aligned} \tag{4}
$$

2. Show for any two columns of the same length, $\mathbf{u}$ and $\mathbf{v}$, that

$$\mathbf{u}^{\mathbf{T}}\mathbf{v} = \mathrm{tr}(\mathbf{v}\mathbf{u}^{\mathbf{T}}) \tag{5}$$

设 $\mathbf{u}$ 和 $\mathbf{v}$ 为 N 维向量，则有

$$\mathbf{u}^{\mathbf{T}}\mathbf{v} = \sum_{i=1}^{N} u_i v_i \tag{6}$$

$$\begin{aligned}
\mathrm{tr}(\mathbf{v}\mathbf{u}^{\mathbf{T}}) &= \mathrm{tr}(\begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ v_N \end{bmatrix} \begin{bmatrix} u_1 & u_2 & \cdots & u_N \end{bmatrix}) \\
&= \mathrm{tr}(\begin{bmatrix} v_1u_1 & v_1u_2 & \cdots & v_1u_N \\ \vdots & \vdots & & \vdots \\ v_Nu_1 & v_Nu_2 & \cdots & v_Nu_N \end{bmatrix}) \\
&= v_1u_1 + v_2u_2 + ... + v_Nu_N \\
&= \sum_{i=1}^{N} u_i v_i
\end{aligned} \tag{7}$$

因此

$$\mathbf{u}^{\mathbf{T}}\mathbf{v} = \sum_{i=1}^{N} u_i v_i \tag{8}$$

3. For a Gaussian random variable, $\mathbf{x} \sim \mathcal{N}(\boldsymbol{\mu}, \boldsymbol{\Sigma})$, show directly that

$$\boldsymbol{\mu} = E[\mathbf{x}] = \int_{-\infty}^{\infty} \mathbf{x} p(\mathbf{x}) d\mathbf{x} \tag{9}$$

设 $\mathbf{x}$ 为 N 维随机变量，根据多元随机变量期望的定义有

$$\begin{aligned} E[\mathbf{x}] &= \int_{-\infty}^{\infty} \mathbf{x} p(\mathbf{x}) d\mathbf{x} = \int_{-\infty}^{\infty} \mathbf{x} \frac{1}{\sqrt{(2\pi)^N \det \boldsymbol{\Sigma}}} \exp\left(-\frac{1}{2}(\mathbf{x}-\boldsymbol{\mu})^T \boldsymbol{\Sigma}^{-1} (\mathbf{x}-\boldsymbol{\mu})\right) d\mathbf{x} \\ &= \frac{1}{\sqrt{(2\pi)^N \det \boldsymbol{\Sigma}}} \int_{-\infty}^{\infty} (\mathbf{u}+\boldsymbol{\mu}) \exp\left(-\frac{1}{2}\mathbf{u}^T \boldsymbol{\Sigma}^{-1} \mathbf{u}\right) d\mathbf{u} \\ &= \frac{1}{\sqrt{(2\pi)^N \det \boldsymbol{\Sigma}}} \int_{-\infty}^{\infty} \mathbf{u} \exp\left(-\frac{1}{2}\mathbf{u}^T \boldsymbol{\Sigma}^{-1} \mathbf{u}\right) d\mathbf{u} + \frac{1}{\sqrt{(2\pi)^N \det \boldsymbol{\Sigma}}} \int_{-\infty}^{\infty} \boldsymbol{\mu} \exp\left(-\frac{1}{2}\mathbf{u}^T \boldsymbol{\Sigma}^{-1} \mathbf{u}\right) d\mathbf{u} \end{aligned} \tag{10}$$

注意到 $\mathbf{u} \exp(-\frac{1}{2}\mathbf{u}^T \boldsymbol{\Sigma}^{-1} \mathbf{u})$ 在任意维度上为奇函数，因此其在对称区间上积分为 0。故期望可化简为

$$\begin{aligned} E[\mathbf{x}] &= \frac{1}{\sqrt{(2\pi)^N \det \boldsymbol{\Sigma}}} \int_{-\infty}^{\infty} \boldsymbol{\mu} \exp\left(-\frac{1}{2}\mathbf{u}^T \boldsymbol{\Sigma}^{-1} \mathbf{u}\right) d\mathbf{u} \\ &= \boldsymbol{\mu} \frac{1}{\sqrt{(2\pi)^N \det \boldsymbol{\Sigma}}} \int_{-\infty}^{\infty} \exp\left(-\frac{1}{2}\mathbf{u}^T \boldsymbol{\Sigma}^{-1} \mathbf{u}\right) d\mathbf{u} \end{aligned} \tag{11}$$

其中 $\frac{1}{\sqrt{(2\pi)^N \det \boldsymbol{\Sigma}}} \int_{-\infty}^{\infty} \exp(-\frac{1}{2}\mathbf{u}^T \boldsymbol{\Sigma}^{-1} \mathbf{u}) d\mathbf{u}$ 为标准多元正态分布概率密度函数的积分，其值恒为 1。故最终期望可化简为

$$\begin{aligned} E[\mathbf{x}] &= \boldsymbol{\mu} \frac{1}{\sqrt{(2\pi)^N \det \boldsymbol{\Sigma}}} \int_{-\infty}^{\infty} \exp\left(-\frac{1}{2}\mathbf{u}^T \boldsymbol{\Sigma}^{-1} \mathbf{u}\right) d\mathbf{u} \\ &= \boldsymbol{\mu} \end{aligned} \tag{12}$$

4. For a Gaussian random variable, $\mathbf{x} \sim \mathcal{N}(\boldsymbol{\mu}, \boldsymbol{\Sigma})$, show directly that

$$\boldsymbol{\Sigma} = E[(\mathbf{x}-\boldsymbol{\mu})(\mathbf{x}-\boldsymbol{\mu})^T] = \int_{-\infty}^{\infty} (\mathbf{x}-\boldsymbol{\mu})(\mathbf{x}-\boldsymbol{\mu})^T p(\mathbf{x}) d\mathbf{x} \tag{13}$$

设 $\mathbf{x}$ 为 N 维随机变量，根据多元随机变量协方差的定义有

$$\begin{aligned} E[(\mathbf{x}-\boldsymbol{\mu})(\mathbf{x}-\boldsymbol{\mu})^T] &= \int_{-\infty}^{\infty} (\mathbf{x}-\boldsymbol{\mu})(\mathbf{x}-\boldsymbol{\mu})^T \frac{1}{\sqrt{(2\pi)^N \det\boldsymbol{\Sigma}}} \exp(-\frac{1}{2}(\mathbf{x}-\boldsymbol{\mu})^T \boldsymbol{\Sigma}^{-1}(\mathbf{x}-\boldsymbol{\mu})) d\mathbf{x} \\ &= \int_{-\infty}^{\infty} \mathbf{u}\mathbf{u}^T \frac{1}{\sqrt{(2\pi)^N \det\boldsymbol{\Sigma}}} \exp(-\frac{1}{2}\mathbf{u}^T \boldsymbol{\Sigma}^{-1}\mathbf{u}) d\mathbf{u} \\ &= \frac{1}{\sqrt{(2\pi)^N \det\boldsymbol{\Sigma}}} \int_{-\infty}^{\infty} \mathbf{u}\mathbf{u}^T \exp(-\frac{1}{2}\mathbf{u}^T \boldsymbol{\Sigma}^{-1}\mathbf{u}) d\mathbf{u} \end{aligned} \tag{14}$$

由于 $\boldsymbol{\Sigma}$ 为正定矩阵，其逆阵仍为正定矩阵。故可对其进行分解:

$$\boldsymbol{\Sigma}^{-1} = P^T P \tag{15}$$

带入积分项有

$$\begin{aligned} \int_{-\infty}^{\infty} \mathbf{u}\mathbf{u}^T \exp(-\frac{1}{2}\mathbf{u}^T \boldsymbol{\Sigma}^{-1}\mathbf{u}) d\mathbf{u} &= \int_{-\infty}^{\infty} \mathbf{u}\mathbf{u}^T \exp(-\frac{1}{2}(P\mathbf{u})^T (P\mathbf{u})) d\mathbf{u} \\ &= |\det P^{-1}| P^{-1} \int_{-\infty}^{\infty} \mathbf{v}\mathbf{v}^T \exp(-\frac{1}{2}\mathbf{v}^T \mathbf{v}) d\mathbf{v} P^{-T} \end{aligned} \tag{16}$$

注意到

$$d\mathbf{v} \exp(-\frac{1}{2}\mathbf{v}^T\mathbf{v}) = \exp(-\frac{1}{2}\mathbf{v}^T\mathbf{v})(I_N - \mathbf{v}\mathbf{v}^T) d\mathbf{v} \tag{17}$$

$$\mathbf{v}\mathbf{v}^T \exp(-\frac{1}{2}\mathbf{v}^T\mathbf{v}) d\mathbf{v} = I_N \exp(-\frac{1}{2}\mathbf{v}^T\mathbf{v}) d\mathbf{v} - d\mathbf{v} \exp(-\frac{1}{2}\mathbf{v}^T\mathbf{v}) \tag{18}$$

两边同时积分得到

$$\begin{aligned} \int_{-\infty}^{\infty} \mathbf{v}\mathbf{v}^T \exp(-\frac{1}{2}\mathbf{v}^T\mathbf{v}) d\mathbf{v} &= \int_{-\infty}^{\infty} I_N \exp(-\frac{1}{2}\mathbf{v}^T\mathbf{v}) d\mathbf{v} - \int_{-\infty}^{\infty} d\mathbf{v} \exp(-\frac{1}{2}\mathbf{v}^T\mathbf{v}) \\ &= \int_{-\infty}^{\infty} I_N \exp(-\frac{1}{2}\mathbf{v}^T\mathbf{v}) d\mathbf{v} \\ &= \sqrt{(2\pi)^N} I_N \end{aligned} \tag{19}$$

带回原始积分得到

$$\begin{aligned} E[(\mathbf{x}-\boldsymbol{\mu})(\mathbf{x}-\boldsymbol{\mu})^T] &= \frac{1}{\sqrt{(2\pi)^N \det\boldsymbol{\Sigma}}} \int_{-\infty}^{\infty} \mathbf{u}\mathbf{u}^T \exp(-\frac{1}{2}\mathbf{u}^T \boldsymbol{\Sigma}^{-1}\mathbf{u}) d\mathbf{u} \\ &= \frac{1}{\sqrt{(2\pi)^N \det\boldsymbol{\Sigma}}} |\det P^{-1}| P^{-1} \int_{-\infty}^{\infty} \mathbf{v}\mathbf{v}^T \exp(-\frac{1}{2}\mathbf{v}^T\mathbf{v}) d\mathbf{v} P^{-T} \\ &= \frac{1}{\sqrt{(2\pi)^N \det\boldsymbol{\Sigma}}} |\det P^{-1}| P^{-1} \sqrt{(2\pi)^N} I_N P^{-T} \\ &= P^{-1} P^{-T} \\ &= (P^T P)^{-1} \\ &= \boldsymbol{\Sigma} \end{aligned} \tag{20}$$

5. Show that the direct product of K statistically independent Gaussian PDFs, $\mathbf{x_k} \sim \mathcal{N}(\boldsymbol{\mu}_\mathbf{k}, \boldsymbol{\Sigma}_\mathbf{k})$ is also a Gaussian PDF:

$$\exp\Big(-\frac{1}{2}(\mathbf{x}-\boldsymbol{\mu})^T\boldsymbol{\Sigma}^{-1}(\mathbf{x}-\boldsymbol{\mu})\Big) = \eta\prod_{k=1}^{K}\exp\Big(-\frac{1}{2}(\mathbf{x}-\boldsymbol{\mu}_k)^T\boldsymbol{\Sigma}_k^{-1}(\mathbf{x}-\boldsymbol{\mu}_k)\Big) \tag{21}$$

where

$$\boldsymbol{\Sigma}^{-1} = \sum_{k=1}^{K}\boldsymbol{\Sigma}_k^{-1} \tag{22}$$

$$\boldsymbol{\Sigma}^{-1}\boldsymbol{\mu} = \sum_{k=1}^{K}\boldsymbol{\Sigma}_k^{-1}\boldsymbol{\mu}_k \tag{23}$$

and η is a normalization constant to enforce the axiom of total probability.

$(\mathbf{x}_1, ..., \mathbf{x}_K)$ 的联合概率密度函数为

$$P = \eta\prod_{k=1}^{K}\exp\Big(-\frac{1}{2}(\mathbf{x}-\boldsymbol{\mu}_k)^T\boldsymbol{\Sigma}_k^{-1}(\mathbf{x}-\boldsymbol{\mu}_k)\Big) \tag{24}$$

对联合概率密度函数 P 取对数有

$$\log P = \log\eta - \frac{1}{2}\sum_{k=1}^{K}(\mathbf{x}-\boldsymbol{\mu}_k)^T\boldsymbol{\Sigma}_k^{-1}(\mathbf{x}-\boldsymbol{\mu}_k) \tag{25}$$

其中求和的每一项均为二次型，求和后仍然为二次型。将求和展开有

$$\begin{aligned}\sum_{k=1}^{K}(\mathbf{x}-\boldsymbol{\mu}_k)^T\boldsymbol{\Sigma}_k^{-1}(\mathbf{x}-\boldsymbol{\mu}_k) &= \sum_{k=1}^{K}\mathbf{x}^T\boldsymbol{\Sigma}_k^{-1}\mathbf{x} - 2\boldsymbol{\mu}_k^T\boldsymbol{\Sigma}_k^{-1}\mathbf{x} + \boldsymbol{\mu}_k^T\boldsymbol{\Sigma}_k^{-1}\boldsymbol{\mu}_k \\ &= \mathbf{x}^T\sum_{k=1}^{K}\boldsymbol{\Sigma}_k^{-1}\mathbf{x} - 2\sum_{k=1}^{K}\boldsymbol{\mu}_k^T\boldsymbol{\Sigma}_k^{-1}\mathbf{x} + \sum_{k=1}^{K}\boldsymbol{\mu}_k^T\boldsymbol{\Sigma}_k^{-1}\boldsymbol{\mu}_k\end{aligned} \tag{26}$$

忽略掉常数项 $\sum_{k=1}^{K}\boldsymbol{\mu}_k^T\boldsymbol{\Sigma}_k^{-1}\boldsymbol{\mu}_k$ 并记为 C，引入系数

$$\boldsymbol{\Sigma}^{-1} = \sum_{k=1}^{K}\boldsymbol{\Sigma}_k^{-1} \tag{27}$$

$$\boldsymbol{\Sigma}^{-1}\boldsymbol{\mu} = \sum_{k=1}^{K}\boldsymbol{\Sigma}_k^{-1}\boldsymbol{\mu}_k \tag{28}$$

则求和可简记为

$$\begin{aligned}\mathbf{x}^T\sum_{k=1}^{K}\boldsymbol{\Sigma}_k^{-1}\mathbf{x} - 2\sum_{k=1}^{K}\boldsymbol{\mu}_k^T\boldsymbol{\Sigma}_k^{-1}\mathbf{x} + C &= \mathbf{x}^T\boldsymbol{\Sigma}^{-1}\mathbf{x} + 2\boldsymbol{\mu}^T\boldsymbol{\Sigma}^{-1}\mathbf{x} + C \\ &= (\mathbf{x}-\boldsymbol{\mu})^T\boldsymbol{\Sigma}^{-1}(\mathbf{x}-\boldsymbol{\mu})\end{aligned} \tag{29}$$

其中常数项 C 不足或多余的部分可由归一化常数 $\log\eta$ 补齐。因此联合概率 P 仍然是一个正态分布。