机器人学中的状态估计 第一课作业

边城量子 2020.4.19

1. 证明高斯分布积分为1

- 即如下分布函数积分为1:
  $p(\boldsymbol{x}|\boldsymbol{\mu},\boldsymbol{\Sigma})=\frac{1}{\sqrt{(2\pi)^N\det\Sigma}}\exp\left(-\frac{1}{2}(\boldsymbol{x}-\boldsymbol{\mu})^{\mathrm{T}}\boldsymbol{\Sigma}^{-1}(\boldsymbol{x}-\boldsymbol{\mu})\right)$

回答:

思路:

1. 先将多维高斯通过对称正定矩阵的性质，分解为多个一维高斯概率密度函数;
2. 证明并应用一维高斯概率密度函数积分为1的结论，从而证明上式1中的多维高斯密度函数;

已知信息和性质:

- 已知其中的 Σ 为协方差矩阵， 它是对称正定矩阵， 因此其逆矩阵必然存在，并且可以进行特征值分解。由此:
  - 可令 $\Sigma=\boldsymbol{U}^{\mathbf{T}}\Lambda\boldsymbol{U}$,
  - 由于对称正定矩阵的特征值大于0，所以 Λ 为对角阵， 即 $\Lambda=diag(\sigma_1^2\ldots\sigma_n^2)$, 其中 σ_i 是矩阵 Σ 特征值的开方值。 则 $\Lambda^{-1}=diag(\frac{1}{\sigma_1^2}\ldots\frac{1}{\sigma_n^2})$ (1)
  - 其中 U 是正交阵，可知:
    - $U^TU=U^{-1}U=I$ (2)
    - $\Sigma^{-1}=(U^T\Lambda U)^{-1}=U^{-1}\Lambda^{-1}(U^T)^{-1}=U^T\Lambda^{-1}U$ (3)
  - 而且，可也可知其行列式值 $|\Sigma^{-1}|=|\Lambda|=\frac{1}{|\Sigma|}$; (4)

基于以上信息和性质， 开始积分推导， 为简化计算，先不考虑指数前面的常数系数，针对指数部分积分如下:

$\int_{-\infty}^{\infty}\ldots\int_{-\infty}^{\infty}e^{-\frac{1}{2}(\boldsymbol{x}-\boldsymbol{\mu})^T\boldsymbol{\Sigma}^{-1}(\boldsymbol{x}-\boldsymbol{\mu})}dx_1\ldots dx_n$ (5)

令 $\boldsymbol{y}=\boldsymbol{x}-\boldsymbol{\mu}$ 上式变为:

$$\int_{-\infty}^{\infty}\ldots\int_{-\infty}^{\infty}e^{-\frac{1}{2}(\boldsymbol{x}-\boldsymbol{\mu})^T\boldsymbol{\Sigma}^{-1}(\boldsymbol{x}-\boldsymbol{\mu})}dx_1\ldots dx_n=\int_{-\infty}^{\infty}\ldots\int_{-\infty}^{\infty}e^{-\frac{1}{2}\boldsymbol{y}^T\boldsymbol{\Sigma}\boldsymbol{y}}dy_1\ldots dy_n$$

把 $\Sigma^{-1}=\boldsymbol{U}^{\mathbf{T}}\Lambda\boldsymbol{U}$ 代入， 并应用上面的已知信息和性质， 可得:

$$
\begin{aligned}
\int \dots \int_{-\infty}^{\infty} e^{-\frac{1}{2}\boldsymbol{y}^T(\boldsymbol{U}^T\Lambda\boldsymbol{U})^{-1}\boldsymbol{y}} dy_1 \dots dy_n &= \int \dots \int_{-\infty}^{\infty} e^{-\frac{1}{2}\boldsymbol{y}^T\boldsymbol{U}^T\Lambda^{-1}\boldsymbol{U}\boldsymbol{y}} dy_1 \dots dy_n \\
&= \int \dots \int_{-\infty}^{\infty} e^{-\frac{1}{2}(\boldsymbol{U}\boldsymbol{y})^T\Lambda^{-1}(\boldsymbol{U}\boldsymbol{y})} dy_1 \dots dy_n \\
&= \int \dots \int_{-\infty}^{\infty} e^{-\frac{1}{2}\boldsymbol{z}^T\Lambda^{-1}\boldsymbol{z}} |\boldsymbol{U}| dz_1 \dots dz_n \\
&= \int \dots \int_{-\infty}^{\infty} e^{-\frac{1}{2}[(\frac{z_1}{\sigma_1})^2+(\frac{z_2}{\sigma_2})^2+\dots+(\frac{z_n}{\sigma_n})^2]} |\boldsymbol{U}| dz_1 dz_2 \dots d_z n \\
&= \int_{-\infty}^{\infty} e^{-\frac{1}{2}(\frac{z_1}{\sigma_1})^2} dz_1 \dots \int_{-\infty}^{\infty} e^{-\frac{1}{2}(\frac{z_n}{\sigma_n})^2} dz_n \quad (6)
\end{aligned}
$$

上式最后几步推导， 是令 $\boldsymbol{z} = \boldsymbol{U}\boldsymbol{y}$ 并代入后， 然后把 Λ 是对角阵的性质（1）代入，由于 z_i 分量相互独立， 因此可以分别积分再相乘。 由此， 就把公式（5）的积分推导， 转变为了公式（6）中对 z_i 的积分并相乘的形式。

然后， 需要用到一个结论： $\int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-\mu)^2}{2\sigma^2}} dx = 1$, 即一维高斯分布的积分为1。 这个结论稍后会进行证明，

此处先使用此结论，将 z 看作是此结论中的 $x - \mu$, 可以得到此结论的等价式：

$\int_{-\infty}^{\infty} e^{-\frac{1}{2}\frac{z^2}{\sigma^2}} dz = \sqrt{2\pi}\sigma$

代入到公式（6）， 则可以得到：

$$
\begin{aligned}
\int_{-\infty}^{\infty} e^{-\frac{1}{2}(\frac{z_1}{\sigma_1})^2} dz_1 \dots \int_{-\infty}^{\infty} e^{-\frac{1}{2}(\frac{z_n}{\sigma_n})^2} dz_n &= \sqrt{2\pi}\sigma_1\sqrt{2\pi}\sigma_2 \dots \sqrt{2\pi}\sigma_n \\
&= (2\pi)^{\frac{n}{2}} |\Sigma|^2 \quad (7)
\end{aligned}
$$

以上计算的都是高斯分布的指数部分的积分， 所以公式（7）乘以系数部分， 系数部分为 $\frac{1}{\sqrt{(2\pi)^n \det \Sigma}}$ ， 相乘以后可以看出， 两者乘积为1， 得证。

补充证明： 一维高斯分布的积分为1。

以下证明上述过程中用到的一个结论： 一维高斯分布的积分为1。

$$
\begin{aligned}
\int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-\mu)^2}{2\sigma^2}} dx &= \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}\sigma} e^{-(\frac{y}{\sqrt{2}\sigma})^2} dy \\
&= \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}\sigma} e^{-z^2} \sqrt{2}\sigma dz \\
&= \frac{1}{\sqrt{\pi}} \sqrt{\pi} \\
&= 1
\end{aligned}
$$

其中倒数第三步，详细过程如下：

$$
\begin{aligned}
\int\int_{-\infty}^{\infty} e^{-(x^2+y^2)} dxdy &= \int_0^{2\pi} d\theta \int_0^{\infty} e^{-r^2} rdr \\
&= \int_0^{2\pi} (-\frac{1}{2}e^{-r^2})|_0^{\infty} d\theta \\
&= \int_0^{2\pi} \frac{1}{2} d\theta \\
&= \pi \\
&= \int_{-\infty}^{\infty} e^{-x^2} dx \int_{-\infty}^{\infty} e^{-y^2} dy \\
&= (\int_{-\infty}^{\infty} e^{-x^2} dx)^2
\end{aligned}
$$

可以得出：$\int_{-\infty}^{\infty} e^{-x^2} dx = \sqrt{\pi}$, 下面基于这个结果，来推导 一维高斯积分为1这个结论：

$$
\begin{aligned}
\int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-\mu)^2}{2\sigma^2}} dx &= \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}\sigma} e^{-(\frac{y}{\sqrt{2}\sigma})^2} dy \\
&= \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}\sigma} e^{-z^2} \sqrt{2}\sigma dz \\
&= \int_{-\infty}^{\infty} \frac{1}{\sqrt{\pi}} e^{-z^2} dz \\
&= 1
\end{aligned}
$$

2.1 假设 $\boldsymbol{u}, \boldsymbol{v}$ 是两个相同维度的列向量，请证明下面这个等式：

$\boldsymbol{u}^T \boldsymbol{v} = tr(\boldsymbol{v}\boldsymbol{u}^T)$

回答:

2.1 习题：

设 u, v 是 n 维列向量，

$u = (u_1, u_2, \cdots, u_n)^T$

$v = (v_1, v_2, \cdots, v_n)^T$

则 $u^T v = (u_1, u_2, \cdots, u_n)\begin{pmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{pmatrix} = \sum_{i=1}^{n} u_i v_i$

$v u^T = \begin{pmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{pmatrix}(u_1, u_2, \cdots, u_n)$

$= \begin{bmatrix} v_1 u_1 & v_1 u_2 & \cdots & v_1 u_n \\ v_2 u_1 & v_2 u_2 & \cdots & v_2 u_n \\ \vdots & & \ddots & \vdots \\ v_n u_1 & v_n u_2 & \cdots & v_n u_n \end{bmatrix}$

$tr(v u^T) = \sum_{i=1}^{n} v_i u_i = \sum_{i=1}^{n} u_i v_i$

故 $u^T v = tr(v u^T)$

证毕.

2.4 对于高斯分布的随机变量，$\boldsymbol{x} \sim \mathcal{N}(\boldsymbol{\mu}, \boldsymbol{\Sigma})$，请证明下面这个等式：

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

回答

方法(2):

$$E(x) = \int_{-\infty}^{\infty} x \cdot \frac{1}{\sqrt{(2\pi)^N}\,|\Sigma|^{1/2}} \exp\left(-\frac{1}{2}(x-\mu)^T \Sigma^{-1} (x-\mu)\right) dx$$

令 $y = x - \mu$, $H = \sqrt{(2\pi)^N}\,|\Sigma|^{1/2}$

$$= \int_{-\infty}^{\infty} (y+\mu) \frac{1}{H} \exp\left(-\frac{1}{2} y^T \Sigma^{-1} y\right) dy$$

$$= \underbrace{\int_{-\infty}^{\infty} y \cdot \frac{1}{H} \exp\left(-\frac{1}{2} y^T \Sigma^{-1} y\right) dy}_{\text{奇函数.积分为0}} + \mu \underbrace{\int_{-\infty}^{\infty} \frac{1}{H} \exp\left(-\frac{1}{2} y^T \Sigma^{-1} y\right) dy}_{\int p(y)dy = 1}$$

$$= \mu$$

得证#.

2.5 对于高斯分布的随机变量， $\boldsymbol{x} \sim \mathcal{N}(\boldsymbol{\mu}, \boldsymbol{\Sigma})$, 请证明下面这个等式：

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

回答

方法二：

$$\int (x-u)(x-u)^T \frac{1}{\sqrt{(2\pi)^N}|\Sigma|^{1/2}} exp\left(-\frac{1}{2}(x-u)^T\Sigma^{-1}(x-u)\right)dx$$

令 $H=\sqrt{(2\pi)^N}|\Sigma|^{1/2}$，令 $y=x-u$

$$=\int yy^T\cdot\frac{1}{H}\cdot exp\left(-\frac{1}{2}y^T\Sigma^{-1}y\right)dy$$

令 $\Sigma^{-1}=\sum_k^N \frac{1}{\lambda_k}u_k u_k^T$，令 $y=\sum_{i=1}^N z_i u_i$

$$=\int\sum_{i=1}^N\sum_{j=1}^N z_i u_i z_j u_j^T \frac{1}{H}\cdot exp\left(-\frac{1}{2}z_j u_j^T\sum_{k=1}^N\frac{1}{\lambda_k}u_k u_k^T z_i u_i\right)dz$$

$$=\frac{1}{H}\sum_i^N\sum_j^N\int z_i u_i z_j u_j^T exp\left(-\frac{1}{2}\sum_k^N\frac{z_i z_j}{\lambda_k}u_j^T u_k u_k^T u_i\right)dz$$

由于 Σ^{-1} 为对角阵，仅当 $i=j=k$ 时上式非0

$$=\frac{1}{H}\sum_i^N\sum_j^N\int z_i u_j z_j u_j^T exp\left(-\frac{1}{2}\sum_k^N\frac{z_k^2}{\lambda_k}\right)dz$$

当 $i\neq j$ 时，由奇偶性知上式为0

且 $\int z^2 exp\left(-\frac{1}{2}\frac{z^2}{\lambda}\right)dz=(2\pi\lambda)^{-\frac{1}{2}}\cdot\lambda$

$$=\frac{1}{H}\sum_i^N u_i u_i^T\lambda_i \underbrace{\prod_{k=1}^N(2\pi\lambda_k)^{1/2}}_{\text{消去}}$$

把H代入后，消去后部

$$=\sum_i^N\lambda_i u_i u_i^T=\Sigma \text{．（由定义）}$$

#证毕．

另一个方法，采用分部积分的方法：

2.5 习题：方法一.

利用性质 $(fg)' = f'g + fg'$ $\therefore fg' = (fg)' - f'g$

令 $f = (x-u)$. $g = \exp(-\frac{1}{2}(x-u)^T \Sigma^{-1}(x-u))$

则 $\therefore fg' = (x-u)\exp'(-\frac{1}{2}(x-u)^T\Sigma^{-1}(x-u))$,

$$= \left((x-u)\exp(-\frac{1}{2}(x-u)^T\Sigma^{-1}(x-u))\right)' - (x-u)'\cdot \exp(-\frac{1}{2}(x-u)^T\Sigma^{-1}(x-u))$$

$$E[(x-u)(x-u)^T] = \int_{-\infty}^{\infty}(x-u)(x-u)^T \frac{1}{\sqrt{(2\pi)^N}|\Sigma|^{1/2}} \exp(-\frac{1}{2}(x-u)^T\Sigma^{-1}(x-u))dx$$

令 $H = \sqrt{(2\pi)^N}|\Sigma|^{1/2}$, $y = x-u$

$$= \Sigma\int_{-\infty}^{\infty}\frac{1}{H}\cdot \underbrace{y\cdot \exp(-\frac{1}{2}y^T\Sigma^{-1}y)\, d(-\frac{1}{2}y^T\Sigma^{-1}y)}_{\text{此式为 } fg'\text{，用 } (fg)' - f'g \text{ 代替}}$$

而 $\int (fg)' = fg\Big|_{-\infty}^{\infty}$

$$= \Sigma\cdot\frac{1}{H}\underbrace{\left(y\cdot \exp(-\frac{1}{2}y^T\Sigma^{-1}y)\right)\Big|_{-\infty}^{+\infty}}_{\text{奇函数，积分为0}} - \underbrace{\frac{\Sigma}{H}\int_{-\infty}^{\infty}\exp(-\frac{1}{2}y^T\Sigma^{-1}y)\,dy}_{\text{全概率}}.$$

$= \Sigma$.

证毕.

2.6 高斯分布归一化积

2.5.6 对于 K 个相互独立的高斯概率密度，$\boldsymbol{x}_k \sim \mathcal{N}(\boldsymbol{\mu}_k, \boldsymbol{\Sigma}_k)$，请证明它们的归一化积仍然是高斯分布：

$$\exp\left(-\frac{1}{2}(\boldsymbol{x}-\boldsymbol{\mu})^{\mathrm{T}}\boldsymbol{\Sigma}^{-1}(\boldsymbol{x}-\boldsymbol{\mu})\right) \equiv \eta\prod_{k=1}^{K}\exp\left(-\frac{1}{2}(\boldsymbol{x}_k-\boldsymbol{\mu}_k)^{\mathrm{T}}\boldsymbol{\Sigma}_k^{-1}(\boldsymbol{x}_k-\boldsymbol{\mu}_k)\right)$$

其中：

$$\boldsymbol{\Sigma}^{-1} = \sum_{k=1}^{K}\boldsymbol{\Sigma}_k^{-1}, \qquad \boldsymbol{\Sigma}^{-1}\boldsymbol{\mu} = \sum_{k=1}^{K}\boldsymbol{\Sigma}_k^{-1}\boldsymbol{\mu}_k$$

回答

2.6 高斯分布归一化系数

$$\exp\left(-\frac{1}{2}(x-\mu)^T\Sigma^{-1}(x-\mu)\right) \equiv \eta \prod_{k=1}^{K} \exp\left(-\frac{1}{2}(x_k-\mu_k)^T\Sigma_k^{-1}(x_k-\mu_k)\right)$$

η 为归一化因子，其中 $\Sigma^{-1} = \sum_{k=1}^{K}\Sigma_k^{-1}$，$\Sigma^{-1}\mu = \sum_{k=1}^{K}\Sigma_k^{-1}\mu_k$

两侧取 $\ln$

$$-\frac{1}{2}(x-\mu)^T\Sigma^{-1}(x-\mu) = \sum_{k=1}^{K}\left(-\frac{1}{2}(x_k-\mu_k)^T\Sigma_k^{-1}(x_k-\mu_k)\right)$$

$$x^T\Sigma^{-1}x - 2\mu^T\Sigma^{-1}x + \mu^T\Sigma^{-1}\mu = \sum_{k=1}^{K} x_k^T\Sigma_k^{-1}x - 2\sum_{k=1}^{K}\mu_k^T\Sigma_k^{-1}x_k + \sum_{k=1}^{k}\mu_k^T\Sigma_k^{-1}\mu_k$$

由上式可知

$$\Sigma^{-1} = \sum_{k=1}^{K}\Sigma_k^{-1}$$

$$\Sigma^{-1}\mu = \sum_{k=1}^{K}\Sigma_k^{-1}\mu_k$$

证毕.