深蓝学院
shenlanxueyuan.com

机器人学中的状态估计

第一次作业讲评

主讲人 顾津铭

习题2.5.1

深蓝学院 shenlanxueyuan.com

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

$$\mathbf{u}^T\mathbf{v} = \operatorname{tr}(\mathbf{v}\mathbf{u}^T).$$

作业概况： 完成情况很好；

证明思路： 按照矩阵的迹的定义进行证明即可。

解答：

已知 $\mathrm{tr}\boldsymbol{A}=\sum_{i=1}^{n}a_{ii}$

设 $\boldsymbol{u}=[u_1,u_2,\ldots,u_n]^{\mathrm{T}}$， $\boldsymbol{v}=[v_1,v_2,\ldots,v_n]^{\mathrm{T}}$

则 $\mathrm{tr}(\boldsymbol{v}\boldsymbol{u}^{\mathrm{T}})=\mathrm{tr}\begin{bmatrix} v_1u_1 & \cdots & v_1u_n \\ \vdots & \ddots & \vdots \\ v_nu_1 & \cdots & v_nu_n \end{bmatrix}=\sum_{i=1}^{n}v_iu_i=\boldsymbol{u}^{\mathrm{T}}\boldsymbol{v}$

补充题

证明高斯分布积分为1，即： $1=\int_{-\infty}^{+\infty}\frac{1}{\sqrt{2\pi\sigma^2}}\exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right)dx$

作业概况： 完成情况很好；

证明思路： 将直角坐标系转化到极坐标系，再进行积分计算。

解答：

作线性变换 $t=\dfrac{x-\mu}{\sigma}$

则 $x=\sigma t+\mu$， $\int_a^b f(x)\mathrm{d}x=\int_\alpha^\beta f[x(t)]x'(t)\mathrm{d}t$

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

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

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

令 $I=\dfrac{1}{\sqrt{2\pi}}\displaystyle\int_{-\infty}^{+\infty}\exp\left(-\frac{t^2}{2}\right)\mathrm{d}t$

$$I^2=\frac{1}{2\pi}\int_{-\infty}^{+\infty}\exp\left(-\frac{t^2}{2}\right)\mathrm{d}t\int_{-\infty}^{+\infty}\exp\left(-\frac{u^2}{2}\right)\mathrm{d}u$$

$$=\frac{1}{2\pi}\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}\exp\left(-\frac{t^2+u^2}{2}\right)\mathrm{d}t\mathrm{d}u$$

因为极坐标 $t=\rho\cos\theta$， $u=\rho\sin\theta$，所以上式转化为

$$I^2=\frac{1}{2\pi}\int_0^{2\pi}\int_0^{+\infty}\exp\left(-\frac{\rho^2}{2}\right)\rho\mathrm{d}\rho\mathrm{d}\theta$$

$$=\frac{1}{2\pi}\int_0^{2\pi}\mathrm{d}\theta\int_0^{+\infty}\exp\left(-\frac{\rho^2}{2}\right)\rho\mathrm{d}\rho$$

$$=\frac{1}{2\pi}\times 2\pi\times\left[-\exp\left(-\frac{\rho^2}{2}\right)\right]\Bigg|_0^{+\infty}$$

$$=1$$

而 $I>0$，故有 $I=1$

p维高斯积分为1的证明

解答：

$$p(\boldsymbol{x})=\frac{1}{(2\pi)^{p/2}|\boldsymbol{\Sigma}|^{1/2}}\exp\left(-\frac{1}{2}(\boldsymbol{x}-\boldsymbol{\mu})^{\mathrm{T}}\boldsymbol{\Sigma}^{-1}(\boldsymbol{x}-\boldsymbol{\mu})\right)$$

协方差矩阵$\boldsymbol{\Sigma}$是一个实对称矩阵，设$\lambda_i\ (i=1,2,\dots,p)$为实数，

$\boldsymbol{u}_i=\left(u_1,u_2,\dots,u_p\right)^{\mathrm{T}}$为非零向量，则关系式$\boldsymbol{\Sigma}\boldsymbol{u}_i=\lambda_i\boldsymbol{u}_i$成立.

对协方差矩阵$\boldsymbol{\Sigma}$进行特征值分解，可得$\boldsymbol{\Sigma}=\boldsymbol{U}\boldsymbol{\Lambda}\boldsymbol{U}^{\mathrm{T}}$，

其中矩阵$\boldsymbol{U}=\left(\boldsymbol{u}_1,\boldsymbol{u}_2,\dots,\boldsymbol{u}_p\right)_{p\times p}$是正交矩阵，$\boldsymbol{U}\boldsymbol{U}^{\mathrm{T}}=\boldsymbol{U}^{\mathrm{T}}\boldsymbol{U}=\boldsymbol{I}$，$\boldsymbol{U}^{\mathrm{T}}=\boldsymbol{U}^{-1}$，

$\boldsymbol{\Lambda}$是对角矩阵，即$\boldsymbol{\Lambda}=diag(\lambda_i)\ \ i=1,2,\dots,p$.

$$\boldsymbol{\Sigma}=\boldsymbol{U}\boldsymbol{\Lambda}\boldsymbol{U}^{\mathrm{T}}$$

$$=\left(\boldsymbol{u}_1,\boldsymbol{u}_2,\dots,\boldsymbol{u}_p\right)\begin{pmatrix}\lambda_1&&&\\&\lambda_2&&\\&&\ddots&\\&&&\lambda_p\end{pmatrix}\begin{pmatrix}\boldsymbol{u}_1^{\mathrm{T}}\\\boldsymbol{u}_2^{\mathrm{T}}\\\vdots\\\boldsymbol{u}_p^{\mathrm{T}}\end{pmatrix}$$

$$=\left(\lambda_1\boldsymbol{u}_1,\lambda_2\boldsymbol{u}_2,\dots,\lambda_p\boldsymbol{u}_p\right)\begin{pmatrix}\boldsymbol{u}_1^{\mathrm{T}}\\\boldsymbol{u}_2^{\mathrm{T}}\\\vdots\\\boldsymbol{u}_p^{\mathrm{T}}\end{pmatrix}$$

$$=\boxed{\sum_{i=1}^{p}\lambda_i\boldsymbol{u}_i\boldsymbol{u}_i^{\mathrm{T}}}$$

$$\boldsymbol{\Sigma}^{-1}=\left(\boldsymbol{U}\boldsymbol{\Lambda}\boldsymbol{U}^{\mathrm{T}}\right)^{-1}=\left(\boldsymbol{U}^{\mathrm{T}}\right)^{-1}\boldsymbol{\Lambda}^{-1}\boldsymbol{U}^{-1}=\boldsymbol{U}\boldsymbol{\Lambda}^{-1}\boldsymbol{U}^{\mathrm{T}},$$

其中$\boldsymbol{\Lambda}^{-1}=diag(\frac{1}{\lambda_i})$，所以$\boxed{\boldsymbol{\Sigma}^{-1}=\sum_{i=1}^{p}\frac{1}{\lambda_i}\boldsymbol{u}_i\boldsymbol{u}_i^{\mathrm{T}}}$.

p维高斯积分为1的证明

深蓝学院
shenlanxueyuan.com

将上式带入到概率密度函数 $p(\boldsymbol{x})$ 中，可得

$$
\begin{aligned}
p(\boldsymbol{x}) &= \frac{1}{(2\pi)^{p/2}|\boldsymbol{\Sigma}|^{1/2}}\exp\left(-\frac{1}{2}(\boldsymbol{x}-\boldsymbol{\mu})^{\mathrm{T}}\boldsymbol{\Sigma}^{-1}(\boldsymbol{x}-\boldsymbol{\mu})\right) \\
&= \frac{1}{(2\pi)^{p/2}|\boldsymbol{\Sigma}|^{1/2}}\exp\left(-\frac{1}{2}(\boldsymbol{x}-\boldsymbol{\mu})^{\mathrm{T}}\sum_{i=1}^{p}\frac{1}{\lambda_i}\boldsymbol{u}_i\boldsymbol{u}_i^{\mathrm{T}}(\boldsymbol{x}-\boldsymbol{\mu})\right) \\
&= \frac{1}{(2\pi)^{p/2}|\boldsymbol{\Sigma}|^{1/2}}\exp\left(\sum_{i=1}^{p}-\frac{1}{2}(\boldsymbol{x}-\boldsymbol{\mu})^{\mathrm{T}}\boldsymbol{u}_i\frac{1}{\lambda_i}\boldsymbol{u}_i^{\mathrm{T}}(\boldsymbol{x}-\boldsymbol{\mu})\right)
\end{aligned}
$$

定义 $\boldsymbol{y}=\boldsymbol{U}(\boldsymbol{x}-\boldsymbol{\mu})$，向量 $\boldsymbol{y}=\left(y_1,y_2,\ldots,y_p\right)^{\mathrm{T}}$，

所以有 $y_i=\boldsymbol{u}_i^{\mathrm{T}}(\boldsymbol{x}-\boldsymbol{\mu})$，$\boxed{|\boldsymbol{\Sigma}|^{\frac{1}{2}}=\prod_{i=1}^{p}\lambda_i^{\frac{1}{2}}}$，则上式转化为

$$
\begin{aligned}
p(\boldsymbol{y}) &= \frac{1}{(2\pi)^{p/2}|\boldsymbol{\Sigma}|^{1/2}}\exp\left(\sum_{i=1}^{p}-\frac{1}{2}y_i^{\mathrm{T}}\frac{1}{\lambda_i}y_i\right) \\
&= \frac{1}{(2\pi)^{p/2}|\boldsymbol{\Sigma}|^{1/2}}\exp\left(\sum_{i=1}^{p}-\frac{y_i^2}{2\lambda_i}\right) \\
&= \frac{1}{(2\pi)^{p/2}|\boldsymbol{\Sigma}|^{1/2}}\prod_{i=1}^{p}\exp(-\frac{y_i^2}{2\lambda_i}) \\
&= \prod_{i=1}^{p}\frac{1}{(2\pi\lambda_i)^{1/2}}\exp(-\frac{y_i^2}{2\lambda_i})
\end{aligned}
$$

$$
\begin{aligned}
&\int_{-\infty}^{+\infty}p(\boldsymbol{x})\mathrm{d}\boldsymbol{x} \\
&= \int_{-\infty}^{+\infty}\frac{1}{(2\pi)^{p/2}|\boldsymbol{\Sigma}|^{1/2}}\exp\left(-\frac{1}{2}(\boldsymbol{x}-\boldsymbol{\mu})^{\mathrm{T}}\boldsymbol{\Sigma}^{-1}(\boldsymbol{x}-\boldsymbol{\mu})\right)\mathrm{d}\boldsymbol{x} \\
&= \prod_{i=1}^{p}\int_{-\infty}^{+\infty}\frac{1}{(2\pi\lambda_i)^{1/2}}\exp(-\frac{y_i^2}{2\lambda_i})\mathrm{d}y_i \\
&= 1
\end{aligned}
$$

习题2.5.4

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

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

作业概况：完成情况较好，部分作业只证明了1维高斯的情况；

证明思路：将积分拆解开，一部分是奇函数在对称区间积分为0，另一部分是μ乘以一维高斯分布的积分。

深蓝学院
shenlanxueyuan.com

解答：

$$
\begin{aligned}
E[\boldsymbol{x}] &= \int_{-\infty}^{+\infty} \boldsymbol{x}p(\boldsymbol{x})\mathrm{d}\boldsymbol{x} \\
&= \int_{-\infty}^{+\infty} \frac{\boldsymbol{x}}{\sqrt{(2\pi)^N \det \boldsymbol{\Sigma}}} \exp\left(-\frac{1}{2}(\boldsymbol{x}-\boldsymbol{\mu})^{\mathrm{T}}\boldsymbol{\Sigma}^{-1}(\boldsymbol{x}-\boldsymbol{\mu})\right)\mathrm{d}\boldsymbol{x} \\
&= \int_{-\infty}^{+\infty} \frac{\boldsymbol{x}-\boldsymbol{\mu}+\boldsymbol{\mu}}{\sqrt{(2\pi)^N \det \boldsymbol{\Sigma}}} \exp\left(-\frac{1}{2}(\boldsymbol{x}-\boldsymbol{\mu})^{\mathrm{T}}\boldsymbol{\Sigma}^{-1}(\boldsymbol{x}-\boldsymbol{\mu})\right)\mathrm{d}\boldsymbol{x} \\
&= \int_{-\infty}^{+\infty} \frac{\boldsymbol{x}-\boldsymbol{\mu}}{\sqrt{(2\pi)^N \det \boldsymbol{\Sigma}}} \exp\left(-\frac{1}{2}(\boldsymbol{x}-\boldsymbol{\mu})^{\mathrm{T}}\boldsymbol{\Sigma}^{-1}(\boldsymbol{x}-\boldsymbol{\mu})\right)\mathrm{d}\boldsymbol{x} + \int_{-\infty}^{+\infty} \frac{\boldsymbol{\mu}}{\sqrt{(2\pi)^N \det \boldsymbol{\Sigma}}} \exp\left(-\frac{1}{2}(\boldsymbol{x}-\boldsymbol{\mu})^{\mathrm{T}}\boldsymbol{\Sigma}^{-1}(\boldsymbol{x}-\boldsymbol{\mu})\right)\mathrm{d}\boldsymbol{x} \\
&= \underbrace{\int_{-\infty}^{+\infty} \frac{\boldsymbol{x}-\boldsymbol{\mu}}{\sqrt{(2\pi)^N \det \boldsymbol{\Sigma}}} \exp\left(-\frac{1}{2}(\boldsymbol{x}-\boldsymbol{\mu})^{\mathrm{T}}\boldsymbol{\Sigma}^{-1}(\boldsymbol{x}-\boldsymbol{\mu})\right)\mathrm{d}\boldsymbol{x}}_{\boldsymbol{0}} + \boldsymbol{\mu}\underbrace{\int_{-\infty}^{+\infty} \frac{1}{\sqrt{(2\pi)^N \det \boldsymbol{\Sigma}}} \exp\left(-\frac{1}{2}(\boldsymbol{x}-\boldsymbol{\mu})^{\mathrm{T}}\boldsymbol{\Sigma}^{-1}(\boldsymbol{x}-\boldsymbol{\mu})\right)\mathrm{d}\boldsymbol{x}}_{1} \\
&= \boldsymbol{\mu}
\end{aligned}
$$

习题2.5.5

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

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

作业概况： 完成情况一般，普遍错误是分部积分之后，误将被积函数认为是奇函数，从而出现证明不严密；

证明思路： 利用之前得出的协方差矩阵特征值分解的结论，将p维高斯化简为1维高斯后再进行证明。

解答：

作线性变换 $y_i = \boldsymbol{u}_i^{\mathrm{T}}(\boldsymbol{x}-\boldsymbol{\mu}) \quad i=1,2,\dots,p$，则 $\boldsymbol{x}-\boldsymbol{\mu} = \sum_{i=1}^{p} y_i \boldsymbol{u}_i$

$$
\begin{aligned}
E[(\boldsymbol{x}-\boldsymbol{\mu})(\boldsymbol{x}-\boldsymbol{\mu})^{\mathrm{T}}] &= \int_{-\infty}^{+\infty} p(\boldsymbol{x})(\boldsymbol{x}-\boldsymbol{\mu})(\boldsymbol{x}-\boldsymbol{\mu})^{\mathrm{T}}\mathrm{d}\boldsymbol{x} \\
&= \frac{1}{(2\pi)^{p/2}|\Sigma|^{1/2}} \int_{-\infty}^{+\infty} \exp\left(-\frac{1}{2}(\boldsymbol{x}-\boldsymbol{\mu})^{\mathrm{T}} \boldsymbol{\Sigma}^{-1}(\boldsymbol{x}-\boldsymbol{\mu})\right)(\boldsymbol{x}-\boldsymbol{\mu})(\boldsymbol{x}-\boldsymbol{\mu})^{\mathrm{T}}\mathrm{d}\boldsymbol{x} \\
&= \frac{1}{(2\pi)^{p/2}|\Sigma|^{1/2}} \int_{-\infty}^{+\infty} \exp\left(\sum_{k=1}^{p} -\frac{y_k^{\,2}}{2\lambda_k}\right) \sum_{i=1}^{p} y_i \boldsymbol{u}_i \left(\sum_{j=1}^{p} y_j \boldsymbol{u}_j\right)^{\mathrm{T}} \mathrm{d}\boldsymbol{y} \\
&= \frac{1}{(2\pi)^{p/2}|\Sigma|^{1/2}} \sum_{i=1}^{p}\sum_{j=1}^{p} \boldsymbol{u}_i \boldsymbol{u}_j^{\mathrm{T}} \int_{-\infty}^{+\infty} \exp\left(\sum_{k=1}^{p} -\frac{y_k^{\,2}}{2\lambda_k}\right) y_i y_j \mathrm{d}\boldsymbol{y} \\
&= \sum_{i=1}^{p} \boldsymbol{u}_i \boldsymbol{u}_i^{\mathrm{T}} \int_{-\infty}^{+\infty} \frac{1}{(2\pi)^{p/2}|\Sigma|^{1/2}} \exp\left(\sum_{k=1}^{p} -\frac{y_k^{\,2}}{2\lambda_k}\right) y_i^{\,2} \mathrm{d}\boldsymbol{y} \\
&= \sum_{i=1}^{p} \boldsymbol{u}_i \boldsymbol{u}_i^{\mathrm{T}} \int_{-\infty}^{+\infty} \prod_{k=1}^{p} \frac{1}{(2\pi\lambda_k)^{1/2}} \exp\left(-\frac{y_k^{\,2}}{2\lambda_k}\right) y_i^{\,2} \mathrm{d}y_k \\
&= \sum_{i=1}^{p} \lambda_i \boldsymbol{u}_i \boldsymbol{u}_i^{\mathrm{T}} \\
&= \boldsymbol{\Sigma}
\end{aligned}
$$

优秀作业代表

$$E[(x-\mu)(x-\mu)^T] = \int_{-\infty}^{+\infty}(x-\mu)(x-\mu)^T p(x)\,dx$$

$$= \int_{-\infty}^{+\infty}(x-\mu)(x-\mu)^T e^{-\frac{1}{2}((x-\mu)^T\Sigma^{-1}(x-\mu))}\,dx$$

设 $H^TH=\Sigma^{-1}$，令 $y=H(x-\mu)$

$$y^Ty = (x-\mu)^TH^TH(x-\mu) = (x-\mu)^T\Sigma^{-1}(x-\mu)$$

故 $|H||\Sigma|^{\frac{1}{2}}=1$

$$p(y) = \frac{1}{|H|}p(x) = \frac{1}{H}\frac{1}{\sqrt{(2\pi)^N\Sigma}}e^{(-\frac{1}{2}(x-\mu)^T\Sigma^{-1}(x-\mu))}$$

$$= \frac{1}{\sqrt{(2\pi)^N}}e^{-\frac{1}{2}y^Ty}$$

$\therefore y \sim N(0, I)$

$$I = \mathrm{cov}\,y = \mathrm{cov}(H(x-\mu)) = H\,\mathrm{cov}\,x\,H^T$$

$$\mathrm{cov}(x) = H^{-1}(H^{-1})^T = (H^TH)^{-1} = \Sigma$$

习题2.5.6

2.5.6 Show that the direct product of K statistically independent Gaussian PDFs, $\mathbf{x}_k \sim \mathcal{N}(\boldsymbol{\mu}_k, \boldsymbol{\Sigma}_k)$, is also a Gaussian PDF:

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

where

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

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

作业概况： 完成情况很好；

证明思路： 将连乘转化成指数部分的连加，等式两边的对应项相等即可证明。

$$\text{左边}=\exp\left(-\frac{1}{2}(x-\mu)^T\Sigma^{-1}(x-\mu)\right)$$

$$=\exp\left(-\frac{1}{2}\left(x^T\Sigma^{-1}x-x^T\Sigma^{-1}\mu-\mu^T\Sigma^{-1}x+\mu^T\Sigma^{-1}\mu\right)\right)$$

$$=\exp\left(-\frac{1}{2}\left(x^T\Sigma^{-1}x-2x^T\Sigma^{-1}\mu+\mu^T\Sigma^{-1}\mu\right)\right)$$

$$\text{右边}=\eta\prod_{k=1}^{K}\exp\left(-\frac{1}{2}(x-\mu_k)^T\Sigma^{-1}(x-\mu_k)\right)$$

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

$$=\eta\exp\left(-\frac{1}{2}\sum_{k=1}^{K}\left(x^T\Sigma_k^{-1}x-2x^T\Sigma_k^{-1}\mu_k+\mu_k^T\Sigma_k^{-1}\mu_k\right)\right)$$

$$=\eta\exp\left(-\frac{1}{2}\left(x^T\sum_{k=1}^{K}\Sigma_k^{-1}x-2x^T\sum_{k=1}^{K}\Sigma_k^{-1}\mu_k+\mu^T\sum_{k=1}^{K}\Sigma_k^{-1}\mu_k+C\right)\right)$$

$$=\eta\exp\left(-\frac{1}{2}\left(x^T\Sigma^{-1}x-2x^T\Sigma^{-1}\mu+\mu^T\Sigma^{-1}\mu+C\right)\right)$$

$$=\eta\exp\left(-\frac{1}{2}C\right)\exp\left(-\frac{1}{2}\left(x^T\Sigma^{-1}x-2x^T\Sigma^{-1}\mu+\mu^T\Sigma^{-1}\mu+C\right)\right)$$

令 $\eta\exp\left(-\frac{1}{2}C\right)=1$，此时左边=右边.

参考文献

深蓝学院
shenlanxueyuan.com

[1] Bishop C . Pattern Recognition and Machine Learning[M]// Stat Sci. 2006.

[2]高惠璇编著. 应用多元统计分析[M]. 北京：北京大学出版社, 2005.01.

[3] https://www.bilibili.com/video/BV1aE411o7qd?p=5

[4] https://www.bilibili.com/video/BV1QV411o7EQ

[5] https://math.stackexchange.com/questions/1905977/deriving-the-covariance-of-multivariate-gaussian

深蓝学院
shenlanxueyuan.com

感谢各位聆听！

Thanks for Listening