机器人中的状态估计课后习题答案

完成人：高明

联系方式：

> 知乎：高明
> 微信：gaoming0901

2.5.1

假设u,v是相同维度向量, 请证明下面等式： $u^T v = tr(vu^T)$

solution:
$u = (x_1, x_2, ..., x_n)^T$

$v = (y_1, y_2, ..., y_n)^T$

$u^T v = x_1 y_1 + x_2 y_2 + ... + x_n y_n = \sum_{i=1}^{n} x_i y_i$

$$uv^T = \begin{bmatrix} x_1 y_1 & \cdots & \cdots & \cdots \\ \cdots & x_2 y_2 & \cdots & \cdots \\ \vdots & \vdots & \ddots & \vdots \\ \cdots & \cdots & \cdots & x_n y_n \end{bmatrix}$$

$tr(uv^T) = \sum_{i=1}^{n} x_i y_i = u^T v$

2.5.2

如果有两个相互独立的随机变量x,y,它们的联合分布为$p(x,y)$,请证明它们概率的香浓信息等于各自独立香浓信息的和:

$H(x, y) = H(x) + H(y)$

solution:

$H(x, y)$

$= -E_{(x,y)}(ln(f(x,y)))$

$= -\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f(x,y)ln(f(x,y))dxdy$

因为x,y独立

$$
\begin{aligned}
&H(x,y)\\
&= -\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f(x)f(y)[ln(f(x)) + ln(f(y))]dxdy\\
&= -[\int_{-\infty}^{\infty}(f(x)ln(f(x))dx] * \int_{-\infty}^{\infty} f(y)dy - [\int_{-\infty}^{\infty} f(y)ln(f(y))dy] * \int_{-\infty}^{\infty} f(x)dx\\
&= -\int_{-\infty}^{\infty}(f(x)ln(f(x))dx - \int_{-\infty}^{\infty} f(y)ln(f(y))dy\\
&= H(x) + H(y)
\end{aligned}
$$

2.5.3

对于高斯分布的随机变量，x~N(μ,Σ),请证明下面的等式:

$\mu = E[xx^T] = \Sigma + \mu\mu^T$

solution:

$$
\begin{aligned}
&\Sigma\\
&= E[(x-\mu)(x-\mu)^T]\\
&= E(xx^T - x\mu^T - \mu x^T + \mu\mu^T)\\
&= E(xx^T) - E(x)\mu^T - \mu E(x^T) + \mu\mu^T
\end{aligned}
$$

因为$E(x) = \mu$

$\Sigma = E(xx^T) - \mu\mu^T$

因此

$E(xx^T) = \Sigma + \mu\mu^T$

2.5.4

对于高斯分布的随机变量，x~N(μ,Σ),请证明下面的等式:

$\mu = E(x) = \int_{-\infty}^{\infty} xp(x)dx$

solution:

$E(x)$

$= \int_{-\infty}^{\infty} \frac{x}{\sqrt{(2\pi)^N det(\Sigma)}} exp(-\frac{1}{2}(x-\mu)^T\Sigma^{-1}(x-\mu))dx$

做变换:
$y = x - \mu$

可得:
$x = y + \mu$

$E(x)$

$= \int_{-\infty}^{\infty} \frac{y+\mu}{\sqrt{(2\pi)^N det(\Sigma)}} exp(-\frac{1}{2}y^T\Sigma^{-1}y)dy$

$= \int_{-\infty}^{\infty} \frac{y}{\sqrt{(2\pi)^N det(\Sigma)}} exp(-\frac{1}{2}y^T\Sigma^{-1}y)dy + \int_{-\infty}^{\infty} \frac{\mu}{\sqrt{(2\pi)^N det(\Sigma)}} exp(-\frac{1}{2}y^T\Sigma^{-1}y)dy$

上式第一项由于奇函数在关于0对称空间积分为0

上式第二项扣除μ满足概率归一化条件

$E(x) = \mu$

2.5.5

对于高斯分布的随机变量，x~$N(\mu, \Sigma)$,证明下式:

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

solution:

$E[(x-\mu)(x-\mu)^T]$

$= \int_{-\infty}^{\infty} \frac{(x-\mu)(x-\mu)^T}{\sqrt{(2\pi)^N det(\Sigma)}} exp(-\frac{1}{2}(x-\mu)^T\Sigma^{-1}(x-\mu))dx$

做代换$y = x - \mu$

$E[(x-\mu)(x-\mu)^T]$

$= \int_{-\infty}^{\infty} \frac{yy^T}{\sqrt{(2\pi)^N det(\Sigma)}} exp(-\frac{1}{2}(y^T\Sigma^{-1}y))dy .... < 0 >$

下面参考文献【1】中公式(108)如下式:

$\frac{\partial}{\partial X}(X^TBX) = BX + B^TX$

上式中X是矩阵，向量算特殊矩阵，直接带入，向量表达式如下:

$\frac{d}{dx}(x^TBx) = Bx + B^Tx .... < 1 >$

由于协方差矩阵是对称矩阵，根据等式$< 1 >$:

$\frac{d}{dx}(x^T\Sigma^{-1}x) = \Sigma^{-1} * x + \Sigma^{-T} * x = 2 * \Sigma^{-1} * x ..... < 2 >$

对于<2>式变换:

$\frac{d}{dx}(-\frac{1}{2}x^T\Sigma^{-1}x) = (-\frac{1}{2})(\Sigma^{-1} * x + \Sigma^{-T} * x) = \Sigma^{-1} * x$

$= x^T\Sigma^{-1} .... < 3 >$

将<3>式带入<0>式:

$E[(x-\mu)(x-\mu)^T]$

$= \int_{-\infty}^{\infty} \frac{-y*\Sigma}{\sqrt{(2\pi)^N det(\Sigma)}} exp(-\frac{1}{2}(y^T\Sigma^{-1}y))d(-\frac{1}{2}(y^T\Sigma^{-1}y))$

$= \int_{-\infty}^{\infty} \frac{-y*\Sigma}{\sqrt{(2\pi)^N det(\Sigma)}} d(exp(-\frac{1}{2}(y^T\Sigma^{-1}y))$

分步积分法:

$E[(x-\mu)(x-\mu)^T]$

$= \frac{y*\Sigma}{\sqrt{(2\pi)^N det(\Sigma)}} * exp(-\frac{1}{2}(y^T\Sigma^{-1}y))|_{-\infty}^{+\infty} + \int_{-\infty}^{\infty} \frac{\Sigma}{\sqrt{(2\pi)^N det(\Sigma)}} exp(-\frac{1}{2}(y^T\Sigma^{-1}y))dy$

$= 0 + \Sigma$

$= \Sigma$

2.5.6

对于K个相互独立的高斯变量，x_k~$N(\mu_k, \Sigma_k)$,请证明它们的归一化积仍然是高斯分布:

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

其中:

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

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

且η归一化因子。

solution:

随机变量x_k的概率密度函数如下:

$f_k(x)=\frac{1}{\sqrt{(2\pi)^{N_k}det(\Sigma_k^{-1})}}exp(-\frac{1}{2}(x-\mu_k)^T\Sigma_k^{-1}(x-\mu_k))$

$f_1(x)*f_2(x)*...*f_K(x)$

$=\frac{1}{\sqrt{(2\pi)^{\sum_{k=1}^{K}N_k}\prod_{k=1}^{K}det(\Sigma_k)}}exp(-\frac{1}{2}\sum_{k=1}^{K}(x-\mu_k)^T\Sigma_k^{-1}(x-\mu))$

将指数部分的求和号展开:

$f_1(x)*f_2(x)*...*f_K(x)$

$=\frac{1}{\sqrt{(2\pi)^{\sum_{k=1}^{K}N_k}\prod_{k=1}^{K}det(\Sigma_k)}}exp(-\frac{1}{2}(x^T(\sum_{k=1}^{K}\Sigma_k^{-1})x-(\sum_{k=1}^{K}\mu_k^T\Sigma_k^{-1})x-x^T\sum_{k=1}^{K}\Sigma_i^{-1}\mu_i+\sum_{k=1}^{K}\mu_k^T\Sigma_k^{-1}\mu_k))\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots<0>$

因为协方差矩阵是对称矩阵，<0>式中 $(\sum_{k=1}^{K}\mu_k^T\Sigma_k^{-1})x=x^T\sum_{k=1}^{K}\Sigma_i^{-1}\mu_i$

因此:

$f_1(x)*f_2(x)*...*f_K(x)$

$=\frac{1}{\sqrt{(2\pi)^{\sum_{k=1}^{K}N_k}\prod_{k=1}^{K}det(\Sigma_k)}}exp(-\frac{1}{2}(x^T(\sum_{k=1}^{K}\Sigma_k^{-1})x-2x^T\sum_{k=1}^{K}\Sigma_i^{-1}\mu_i+\sum_{k=1}^{K}\mu_k^T\Sigma_k^{-1}\mu_k))\ldots\ldots\ldots\ldots<1>$

在式<1>中：
$x^T(\sum_{k=1}^K \Sigma_k^{-1})x$为二次项
$2x^T\sum_{k=1}^K \Sigma_i^{-1}\mu_i$为一次项
可以凑出“完全平方形式”

$f_1(x)*f_2(x)*...*f_K(x)$
$= \frac{1}{\sqrt{(2\pi)^{\sum_{k=1}^K N_k}} \prod_{k=1}^K det(\Sigma_k)} exp(-\frac{1}{2}(x-\mu)^T\Sigma^{-1}(x-\mu)+M)$

上式中M为一个常数;

$f_1(x)*f_2(x)*...*f_K(x)$
$= \frac{1}{\sqrt{(2\pi)^{\sum_{k=1}^K N_k}} \prod_{k=1}^K det(\Sigma_k)} exp(-\frac{1}{2}(x-\mu)^T\Sigma^{-1}(x-\mu))*$
$exp(M)..<2>$

根据上式二次项一次项对应参数，可以得到：

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

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

为了满足归一化条件，需要将变量指数项外的其他常数项归一到η中，也即证明K个独立正态分布随机变量概率密度相乘归一化之后仍为正态分布

2.5.7

假设有K个互相独立的随机变量x_k,它们通过加权组成一个新的随机变量：

$$x = \sum_{k=1}^K \omega_k x_k$$

其中$\sum_{k=1}^K \omega_k = 1$且$\omega_k \geq 0$,它们的期望表示为：

$$\mu=\sum_{k=1}^{K}\omega_k\mu_k$$

其中μ_k是x_{k}的均值，请定义出一个计算方差的表达式，注意，这些随机变量并没有假设服从高斯分布

solution:

2.5.8

当K维随机变量x服从标准正态分布，即x~N$(0,1)$,则随机变量：

$$y=x^Tx$$

服从自由度为K的卡方分布，请证明该随机变量的均值为K，方差为2K(**题目条件暗含每一维度随机变量独立同分布假设，远书为准确提及**)

solution:

$y=x_1^2+x_2^2+...+x_K^2$

$E(y)=E(x_1^2+x_2^2+...+x_K^2)$

$=E(x_1^2)+E(x_2^2)+...+E(x_K^2)$

根据统计学：

$E(X^2)=D(X)+(E(X))^2$

因此对于任意$1\leq i\leq K$：

$E(x_k^2)=1+0=1$

$E(y)=K$

根据Isserlis定理：
$E\left[x_ix_jx_kx_\ell\right]=E\left[x_ix_j\right]E\left[x_kx_\ell\right]+E\left[x_ix_k\right]E\left[x_jx_\ell\right]+$
$E\left[x_ix_\ell\right]E\left[x_jx_k\right]$.. $<0>$

方差：

$D(y)$

$= E((x_1^2 + x_2^2 + ...x_K^2)^2)$

$= E(\sum_{i=1}^{K} x_i^4) + 2E(\sum_{i=1,j=1,i\neq j}^{K} E(x_i^2 x_j^2))............ < 1 >$

根据<0>,其中:

$E(\sum_{i=1}^{K} x_i^4)$

$= \sum_{i=1}^{K} E(x_i^4)$

$= 3K$

$E(\sum_{i=1,j=1,i\neq j}^{K} E(x_i^2 x_j^2))$

$= \frac{K(K-1)}{2}$

将上述两式带入<0>:

$D(y) = 2K$

参考文献:

1.Matrix Cookbook---Kaare Brandt Petersen