0. 说明

本 PDF 文档为自动生成，如有遗漏的格式错误但不影响阅读请见谅，若影响了阅读请告知！

问1：证明高斯分布积分为1

证： 高斯分布的概率密度函数为：$p(x) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right)$，所以我们要证明的是：
$\int_{-\infty}^{\infty} p(x)dx = 1 \quad (式\,1-1)$

上面这个式子显得不够简洁，我们首先使用 $y = \frac{x-\mu}{\sqrt{2}\sigma}$ 来进行**变量替换**化简可得：
$\int_{-\infty}^{\infty} \frac{1}{\sqrt{\pi}} \exp\left(-y^2\right)dy \quad (式\,1-2)$

要证明式子（1-2）等于 1就是要证明：$\int_{-\infty}^{\infty} \exp\left(-y^2\right)dy = \sqrt{\pi} \quad (式\,1-3)$

式子（1-3）也被称为高斯积分，那么接下来我们就只需要去证明**高斯积分**即可。

证明方法1：

令 $\mathcal{I} = \int_{-\infty}^{\infty} \exp\left(-y^2\right)dy \quad \longrightarrow \mathcal{I}^2 = \int_{-\infty}^{\infty} \exp\left(-y^2\right)dy \times \int_{-\infty}^{\infty} \exp\left(-y^2\right)dy \quad (式\,1-4)$

因为式子（1-4）中的两个积分是完全独立的，所以可以写成：
$\mathcal{I}^2 = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \exp\{-(x^2+y^2)\}dxdy$

看到 $x^2 + y^2$ 这种符号时可以考虑将坐标系换到**极坐标系**下进行计算：
$\mathcal{I}^2 = \int_0^{2\pi} \int_0^{\infty} \exp\left(-r^2\right)rdrd\theta \quad (式\,1-5)$

$\mathcal{I}^2 = \int_0^{2\pi}(\int_0^{\infty} \exp\left(-r^2\right)rdr)d\theta = \int_0^{2\pi} \frac{1}{2}d\theta = \pi \quad \longrightarrow \quad \mathcal{I} = \sqrt{\pi}$

证明方法2：

通过极限来进行计算，见链接
https://zh.wikipedia.org/wiki/%E9%AB%98%E6%96%AF%E7%A7%AF%E5%88%86

问2：假设 $\mathbf{u}, \mathbf{v}$ 是两个相同维度的列向量，请注明下面这个等式：

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

证： 令 $\boldsymbol{u} = [x_1, x_2, \cdots, x_n]^T; \quad \boldsymbol{v} = [y_1, y_2, \cdots, y_n]^T$

$\boldsymbol{u^T v} = x_1y_1 + x_2y_2 + \cdots + x_ny_n$

$$\boldsymbol{vu^T} = \begin{bmatrix} y_1x_1 & \cdot & \cdots & \cdot \\ \cdot & y_2x_2 & \cdot & \cdot \\ \vdots & \cdot & \ddots & \vdots \\ \cdot & \cdot & \cdots & y_nx_n \end{bmatrix}_{n\times n}$$

根据矩阵的迹的定义：一个$n\times n$矩阵A的主对角线上各个元素的总和被称为矩阵A 的**迹** 。

$tr(\boldsymbol{vu^T})=y_1x_1+y_2x_2+\cdots+y_nx_n=\boldsymbol{u^Tv}$

问3：对于高斯分布的随机变量，$\mathbf{x}\backsim\mathcal{N}(\boldsymbol{\mu},\boldsymbol{\Sigma})$,请证明下面这个等式：

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

证:

1）多维高斯的式子看起来比较复杂，我们首先从一维高斯开始证明，思路就是通过 $t=x-\mu$换元，将这个积分化为两个比较容易计算的积分之和，然后直接计算即可；

2）一维的情况证完了之后，考虑多维情况，如果随机变量 $\boldsymbol{x}$ 的各个分量之间是独立的，那么就只需要将 $\boldsymbol{x}-\boldsymbol{\mu}=[x_1-\mu_1,\ x_2-\mu_2,\ \cdots,x_n-\mu_n]^T$ 和 $\boldsymbol{\Sigma}=\begin{bmatrix}\sigma_1^2 & \cdot & \cdots & \cdot\\ \cdot & \sigma_2^2 & \cdot & \cdot\\ \vdots & \cdot & \ddots & \vdots\\ \cdot & \cdot & \cdots & \sigma_n^2\end{bmatrix}_{n\times n}$ 带入积分计算，然后利用一维的结论即可；

3）多维高斯分布一般情况下的证明：

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

$\quad=\int_{-\infty}^{\infty}\frac{\boldsymbol{(y+\mu)}}{\sqrt{(2\pi)^N\det\boldsymbol{\Sigma}}}\exp\left(-\frac{1}{2}\boldsymbol{y^T\Sigma^{-1}y}\right)d\boldsymbol{y}\qquad \boldsymbol{y}=\boldsymbol{x}-\boldsymbol{\mu}$ 换元

$$=\int_{-\infty}^{\infty}\frac{\boldsymbol{y}}{\sqrt{(2\pi)^N\det\boldsymbol{\Sigma}}}\exp\left(-\frac{1}{2}\boldsymbol{y^T\Sigma^{-1}y}\right)d\boldsymbol{y}+\int_{-\infty}^{\infty}\frac{\boldsymbol{\mu}}{\sqrt{(2\pi)^N\det\boldsymbol{\Sigma}}}\exp\left(-\frac{1}{2}\boldsymbol{y^T\Sigma^{-1}y}\right)d\boldsymbol{y}\quad(\text{式}\,3-1)$$

$\quad=\mathcal{I}_1+\mathcal{I}_2$

因为$\boldsymbol{y}=\boldsymbol{x}-\boldsymbol{\mu}$ 变换后的随机变量 $\boldsymbol{y}$ 也是服从高斯分布的，所以根据**全概率公理**可知
$\mathcal{I}_2=\boldsymbol{\mu}*1=\boldsymbol{\mu}$

那么接下来我们只需要证明 $\mathcal{I}_1=0$ 即可（这里我卡住了，哈哈，最后还是去谷歌了一下）

因为被积部分 $\boldsymbol{yf(y)}$ 是一个奇函数，所以在 $\int_{-\infty}^{\infty}$ 上等于 0。

注意： 其实这里不能直接说$\boldsymbol{yf(y)}$是一个奇函数，只能说每一个分量在其定义域上是奇函数。

问4：对于高斯分布的随机变量，$\mathbf{x}\backsim\mathcal{N}(\boldsymbol{\mu},\boldsymbol{\Sigma})$,请证明下面这个等式：

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

证： 先通过$\boldsymbol{y}=\boldsymbol{x}-\boldsymbol{\mu}$ 换元得到：

$E[(\boldsymbol{x}-\boldsymbol{\mu})(\boldsymbol{x}-\boldsymbol{\mu})^T]=E[\boldsymbol{yy}^T]=\int_{-\infty}^{\infty}\frac{\boldsymbol{yy^T}}{\sqrt{(2\pi)^N\det\boldsymbol{\Sigma}}}\exp\left(-\frac{1}{2}\boldsymbol{y^T\Sigma^{-1}y}\right)d\boldsymbol{y}\quad(\text{式}\,4-1)$

(这里卡住了，去谷歌了一下)，后续的推导主要用到了 $d(-\frac{1}{2}\boldsymbol{y}^{\boldsymbol{T}}\boldsymbol{\Sigma}^{-\mathbf{1}}\boldsymbol{y}) = -\boldsymbol{\Sigma}^{-\mathbf{1}}\boldsymbol{y}d\boldsymbol{y} \quad (\text{式}\,4-2)$

式子（4-2）的推导主要用到了公式 $\frac{d\boldsymbol{x}^{\boldsymbol{T}}\boldsymbol{B}\boldsymbol{x}}{d\boldsymbol{x}} = (\boldsymbol{B}+\boldsymbol{B}^{\boldsymbol{T}})\boldsymbol{x}$ 参考资料http://matrixcookbook.com

式子（4-1）可变换为：

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

$= \frac{-\boldsymbol{y}\boldsymbol{\Sigma}}{\sqrt{(2\pi)^N \det \boldsymbol{\Sigma}}} \exp\left(-\frac{1}{2}\boldsymbol{y}^{\boldsymbol{T}}\boldsymbol{\Sigma}^{-\mathbf{1}}\boldsymbol{y}\right)\Big|_{-\infty}^{\infty} + \int_{-\infty}^{\infty} \frac{\boldsymbol{\Sigma}}{\sqrt{(2\pi)^N \det \boldsymbol{\Sigma}}} \exp\left(-\frac{1}{2}\boldsymbol{y}^{\boldsymbol{T}}\boldsymbol{\Sigma}^{-\mathbf{1}}\boldsymbol{y}\right)d\boldsymbol{y}$ **(分部积分法)**

$= \mathbf{0} + \boldsymbol{\Sigma} = \boldsymbol{\Sigma}$

注意： 上面的证明过程有一定问题，对于重积分，不能直接应用分部积分法，而是应该先化为分量，再用分部积分法求解。

问5：对于 K 个相互独立的高斯变量， $\boldsymbol{x_k} \backsim \mathcal{N}(\boldsymbol{\mu_k}, \boldsymbol{\Sigma_k})$ ，请证明它们的归一化积任然是高斯分布：

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

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

证：

首先，这 K 个高斯变量之间相互独立，那么我们可以将等式写成：

$$\exp\left(-\frac{1}{2}(\boldsymbol{x}-\boldsymbol{\mu})\right)^T \boldsymbol{\Sigma}^{-1}(\boldsymbol{x}-\boldsymbol{\mu})) \equiv \eta \prod_{k=1}^{K} \exp\left(-\frac{1}{2}(\boldsymbol{x}-\boldsymbol{\mu_k})^T \boldsymbol{\Sigma}_k^{-1}(\boldsymbol{x}-\boldsymbol{\mu_k})\right) \quad (\text{式}\,5-1)$$

- $K = 1$ 的时候，式子（5-1）显然成立；
- $K = 2$ 的时候
  $\prod_{k=1}^{2} \exp\left(-\frac{1}{2}(\boldsymbol{x}-\boldsymbol{\mu_k})^T \boldsymbol{\Sigma}_k^{-1}(\boldsymbol{x}-\boldsymbol{\mu_k})\right)$
  $= \exp\left(-\frac{1}{2}\left(\sum_{k=1}^{2} \boldsymbol{x}^{\boldsymbol{T}}\boldsymbol{\Sigma_k^{-1}}\boldsymbol{x} - 2\sum_{k=1}^{2} \boldsymbol{x}^{\boldsymbol{T}}\boldsymbol{\Sigma_k^{-1}}\boldsymbol{\mu_k} + \sum_{k=1}^{2} \boldsymbol{\mu_k^T}\boldsymbol{\Sigma_k^{-1}}\boldsymbol{\mu_k}\right)\right)$ ，($\boldsymbol{x}^{\boldsymbol{T}}\boldsymbol{\Sigma_k^{-1}}\boldsymbol{\mu_k} = \boldsymbol{\mu_k^T}\boldsymbol{\Sigma_k^{-1}}\boldsymbol{x}$)
  $= \exp\left(-\frac{1}{2}(\boldsymbol{x}^{\boldsymbol{T}}\boldsymbol{\Sigma}^{-\mathbf{1}}\boldsymbol{x} - 2\boldsymbol{x}^{\boldsymbol{T}}\boldsymbol{\Sigma}^{-\mathbf{1}}\boldsymbol{\mu} + \boldsymbol{\mu}^{\boldsymbol{T}}\boldsymbol{\Sigma}^{-\mathbf{1}}\boldsymbol{\mu} + \mathcal{C})\right)$ ，(为了凑出第三项需要添加一些常数项，用 $\mathcal{C}$ 来表示)
  $= \exp\left(-\frac{1}{2}(\boldsymbol{x}-\boldsymbol{\mu})\right)^T \boldsymbol{\Sigma}^{-1}(\boldsymbol{x}-\boldsymbol{\mu})) \cdot \exp(\mathcal{C})$
  将 $\exp(\mathcal{C})$ 这一项纳入归一化因子即可证明式子（5-1）成立。
- $K \geq 2$ 的时候证明过程类似，只不过需要把 $\sum_{k=1}^{2}$ 换成 $\sum_{k=1}^{K}$ 即可。