# assignment1

xiaoma

2022.09.06

# 1 第一章

## 1.1

$$
\begin{aligned}
& \frac{\partial \ln det(\mathbf{A})}{\partial x} \\
= \; & \frac{1}{det(\mathbf{A})} \sum_{i,j} \; \frac{\partial det(\mathbf{A})}{\partial A_{ij}} \frac{\partial \mathbf{A}_{ij}}{\partial x} \\
= \; & \frac{1}{det(\mathbf{A})} \sum_{i,j} \; adj^T(\mathbf{A})_{ij} \frac{\partial \mathbf{A}_{ij}}{\partial x} \\
= \; & \frac{1}{det(\mathbf{A})} \sum_{i,j} \; adj(\mathbf{A})_{ji} \frac{\partial \mathbf{A}_{ij}}{\partial x} \\
= \; & \frac{1}{det(\mathbf{A})} \sum_{j=1}^{n} \; (A^* \frac{\partial \mathbf{A}}{\partial x})_{jj} \\
= \; & tr(\mathbf{A}^{-1} \frac{\partial \mathbf{A}}{\partial x})
\end{aligned} \tag{1}
$$

## 1.2

假设空间一共有 $3 * 4 * 4 + 1 = 49$ 种假设，若不考虑空集，则有 48 种假设：

1. 全部泛化：1

2. 全部不泛化：$2 * 3 * 3 = 18$

3. 泛化一种属性：$3 * 3 + 2 * 3 + 2 * 3 = 21$

4. 泛化两种属性：$2 + 3 + 3 = 8$

用这 48 种假设进行组合，共有 $\sum_{i=1}^{k} \mathbf{C}_{48}^{i}$ 种，去重以后，在不泛化的 18 种中进行选择，共有 $2^{18} - 1$ 种假设。

## 1.3

$$
\begin{aligned}
& \mathbf{P}(x_1) \\
= & \frac{1}{2\pi\sqrt{|\sum|}} \int exp[-\frac{1}{2}(\vec{x}-\mu)^{\mathsf{T}} \sum{}^{-1} (\vec{x}-\mu)] dx_2 \\
& \mathsf{P}(x_1|x_2) \\
= & \frac{\mathsf{P}(x_1,x_2)}{\mathsf{P}(x_2)} \\
= & \frac{1}{\sqrt{2\pi|\frac{\sum}{\sum_{11}}|}} exp[-\frac{1}{2}(x_1-\mu_1)^T(\frac{\sum}{\sum_{11}})^{-1}(x_1-\mu_1-\sum{}_{12}^{\mathsf{T}}\sum{}_{11}^{-1}(\vec{x}-\mu)_2)]
\end{aligned}
\tag{2}
$$

## 1.4

证明：

$$
\begin{aligned}
&f(x)=\|x\|_p=\sqrt[p]{\sum_{i=1}^{n}x_i^p}\\
&\nabla f(x)=(\sum_{j=1}^{n}x_j^p)^{\frac{1}{p}-1}[x_1^{p-1},x_2^{p-1},...,x_n^{p-1}]^{\mathsf{T}}\\
&\begin{cases}\frac{\partial^2 f}{\partial x_i^2}=(1-p)(\sum_{j=1}^{n}x_j^p)^{\frac{1}{p}}\frac{1}{x_i^2}(\frac{z_i^p}{\sum_{j=1}^{n}x_j^p}-1)\frac{x_i^p}{\sum_{j=1}^{n}x_j^p}\\ \frac{\partial^2 f}{\partial x_i\partial x_k}=(1-p)(\sum_{j=1}^{n}x_j^p)^{\frac{1}{p}}\frac{1}{x_ix_k}(\frac{z_i^p}{\sum_{j=1}^{n}x_j^p})(\frac{z_k^p}{\sum_{j=1}^{n}x_j^p})\end{cases}
\end{aligned}
\tag{3}
$$

记

$$
\begin{aligned}
&\vec{z}=(x_1^p,x_2^p,...,x_n^p)^{\mathbf{T}},A=diag(\frac{1}{x_1},\frac{1}{x_2},...,\frac{1}{x_n})\\
&\nabla^2 f=(1-p)f\mathbf{A}^{\mathsf{T}}(\vec{z}\vec{z}^{\mathsf{T}}-\vec{z}diag(\vec{z}))\mathsf{A}\frac{1}{\vec{z}^2}
\end{aligned}
\tag{4}
$$

当 $p<1$ 且 $p\neq 0$ 时，要证 f(x) 为凹函数，只需要证

$$\forall y\in\mathbf{R}^n.|y^{\mathsf{T}}\nabla^2\vec{y}\leqslant 0$$

$$\vec{y}^{\mathsf{T}}\nabla^2\vec{y}=\frac{(1-p)f}{\vec{z}^2}[(\sum_{i=1}^{n}(\mathbf{A}\vec{y})_i\vec{z}_i)^2-(\sum_{i=1}^{n}\vec{z}_i)(\sum_{i=1}^{n}(\mathbf{A}\vec{y})_i^2\vec{z}_i)]\leqslant 0 \tag{5}$$

当 $p\geqslant 1$ 时，$\|x\|_p$ 为凸函数 $\sharp$

## 1.5

充分性：

令 $a=1-t$,$a\in[0,1]$

$$
\begin{aligned}
&f(tx+(1-t)y)\leqslant tf(x)+(1-t)(f(y))\\
&\Rightarrow a(f(y))\geqslant f[(1-a)x+ay]-(1-a)f(x)=f[x+a(y-x)]+af(x)-f(x)\\
&\Rightarrow f(y)\geqslant f(x)+\lim_{a\to 0}\frac{f[x+a(y-x)]-f(x)}{a(y-x)}(y-x)\\
&\Rightarrow f(y)\geqslant f(x)+\nabla f(x)^{\mathbf{T}}(y-x)
\end{aligned}
\tag{6}
$$

必要性：

令 $z = tx + (1-t)y$

$$
\begin{aligned}
& f(y) \geqslant f(z) + \nabla f(z)^{\mathbf{T}}(y-z) \\
\Rightarrow\ & f(x) \geqslant f(z) + \nabla f(z)^{\mathbf{T}}(yx-z) \\
\Rightarrow\ & tf(x) + (1-t)f(y) \geqslant f(z) + \nabla f(z)^{\mathbf{T}}[t(1-t)(x-y) + t(1-t)(y-x)] = f(z) \\
\Rightarrow\ & f(tx + (1-t)y) \leqslant tf(x) + (1-t)f(y)
\end{aligned}
$$

$\sharp$

(7)