VIO lesson 4

Yin Wu

2021 年 12 月

目录

1 Task 2

假设我们有一个高斯随机变量 $\boldsymbol{\theta}$，其均值为 $\boldsymbol{\theta}^*$，协方差矩阵为 $\boldsymbol{\Sigma_\theta}$，所以其概率密度函数可以写为

$$p(\boldsymbol{\theta}) = (2\pi)^{-\frac{N_\theta}{2}}|\boldsymbol{\Sigma_\theta}|^{-\frac{1}{2}}exp\left[-\frac{1}{2}(\boldsymbol{\theta}-\boldsymbol{\theta}^*)^T\boldsymbol{\Sigma}_{\boldsymbol{\theta}}^{-1}(\boldsymbol{\theta}-\boldsymbol{\theta}^*)\right] \tag{1}$$

将目标函数定义为概率密度函数的负对数形式

$$\boldsymbol{J(\theta)} = -\mathrm{ln}p(\boldsymbol{\theta}) = \frac{N_{\boldsymbol{\theta}}}{2}\mathrm{ln}2\pi + \frac{1}{2}|\boldsymbol{\Sigma_\theta}| + \frac{1}{2}(\boldsymbol{\theta}-\boldsymbol{\theta}^*)^T\boldsymbol{\Sigma}_{\boldsymbol{\theta}}^{-1}(\boldsymbol{\theta}-\boldsymbol{\theta}^*) \tag{2}$$

其形式是一个二次型的方程，我们令它对 θ_l 和 θ_l' 进行求偏微分导数，从而得到 Hessian 矩阵

$$H^{l,l'}(\boldsymbol{\theta}^*) = \frac{\partial^2\boldsymbol{J(\theta)}}{\partial\theta_l\partial\theta_l'}\bigg|_{\boldsymbol{\theta}=\boldsymbol{\theta}^*} = (\boldsymbol{\Sigma}_{\boldsymbol{\theta}}^{-1})^(l,l') \tag{3}$$

所以我们可以得出，Hessian 矩阵与协方差矩阵的逆相等，即

$$H(\boldsymbol{\theta}^*) = \boldsymbol{\Sigma}_{\boldsymbol{\theta}}^{-1} \tag{4}$$

对于高斯随机变量而言，由于目标方程是一个二次型方程，所以其二阶导数得到的 Hessian 对于所有 $\boldsymbol{\theta}$ 都是一个不变的常数。因此，求解 Hessian 矩阵不需要知晓均值 $\boldsymbol{\theta}^*$。

Hessian 矩阵中的元素保存着随机向量的条件信息，因为随机向量是通过固定其他参数而获得的。对角元素是目标方程的曲率。对角元素的倒数则是 $\boldsymbol{\theta}$ 中不确定参数的条件方差。然而，对角元素在协方差矩阵 $\boldsymbol{\Sigma_\theta}$ 中是参数中的微小量。

在很多应用中，目标函数并不是直接已知的，因此 Hessian 矩阵实际上需要只能数值计算，比容通过有限微分方法，对角元素可以通过以下公式计算

$$H^{l,l'}(\boldsymbol{\theta}^*) \approx \frac{1}{4\Delta\theta_l\Delta\theta_{l'}}[\boldsymbol{J}(\boldsymbol{\theta}^*+\Delta\boldsymbol{\theta}_l+\Delta\boldsymbol{\theta}_{l'}) - \boldsymbol{J}(\boldsymbol{\theta}^*+\Delta\boldsymbol{\theta}_l-\Delta\boldsymbol{\theta}_{l'}) - \boldsymbol{J}(\boldsymbol{\theta}^*-\Delta\boldsymbol{\theta}_l+\Delta\boldsymbol{\theta}_{l'}) + \boldsymbol{J}(\boldsymbol{\theta}^*-\Delta\boldsymbol{\theta}_l-\Delta\boldsymbol{\theta}_{l'})] \tag{5}$$

其中 $\Delta\boldsymbol{\theta}_l$ 和 $\Delta\boldsymbol{\theta}_{l'}$ 是只有第 l 和 l' 项不为 0，其余全为 0 的向量。

综上证明作业中的问题：对于高斯分布的随机变量而言，其求解最大概率而构建的目标函数的 Hessian 矩阵与高斯分布的协方差矩阵的逆相等，即信息矩阵等于 Hessian 矩阵