



深蓝学院
shenlanxueyuon.com

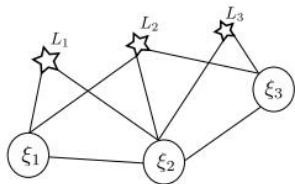
第十三章作业提示

主讲人 会打篮球的猫



第一题

- ① 某时刻，SLAM 系统中相机和路标点的观测关系如下图所示，其中 ξ 表示相机姿态， L 表示观测到的路标点。当路标点 L 表示在世界坐标系下时，第 k 个路标被第 i 时刻的相机观测到，重投影误差为 $r(\xi_i, L_k)$ 。另外，相邻相机之间存在运动约束，如 IMU 或者轮速计等约束。



- 1 请绘制上述系统的信息矩阵 Λ 。
- 2 请绘制相机 ξ_1 被 marg 以后的信息矩阵 Λ' 。

$$\mathbf{J}_2 = \frac{\partial \mathbf{r}_{13}}{\partial \xi} = \begin{bmatrix} \frac{\partial \mathbf{r}_{13}}{\partial \xi_1} & 0 & \frac{\partial \mathbf{r}_{13}}{\partial \xi_3} & 0 & 0 & 0 \end{bmatrix}$$
$$\mathbf{\Lambda}_2 = \mathbf{J}_2^\top \Sigma_2^{-1} \mathbf{J} = \begin{bmatrix} \left(\frac{\partial \mathbf{r}_{13}}{\partial \xi_1}\right)^\top \Sigma_2^{-1} \frac{\partial \mathbf{r}_{13}}{\partial \xi_1} & 0 & \left(\frac{\partial \mathbf{r}_{13}}{\partial \xi_1}\right)^\top \Sigma_2^{-1} \frac{\partial \mathbf{r}_{13}}{\partial \xi_3} & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ \left(\frac{\partial \mathbf{r}_{13}}{\partial \xi_3}\right)^\top \Sigma_2^{-1} \frac{\partial \mathbf{r}_{13}}{\partial \xi_1} & 0 & \left(\frac{\partial \mathbf{r}_{13}}{\partial \xi_3}\right)^\top \Sigma_2^{-1} \frac{\partial \mathbf{r}_{13}}{\partial \xi_3} & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix} \quad (46)$$

其实就是将系统所有约束对应的小信息矩阵累加的一个过程，只有该约束相连接的变量所对应位置才有值，其余为 0。

第一题

Λ

	ξ_1	ξ_2	ξ_3	L_1	L_2	L_3
ξ_1	■	■	□	■	■	□
ξ_2	■	■	■	■	■	■
ξ_3	□	■	■	□	■	■
L_1	■	■	□	■	□	□
L_2	■	■	■	□	■	□
L_3	□	■	■	□	□	■

将 ξ_1 进行边缘化，根据原来的系统，与 ξ_1 相连的有 ξ_1 、 L_1 和 L_2 ，边缘化操作后，这三者之间便会两两相连。

第一题

$$\begin{array}{c} \Lambda \end{array} \xrightarrow{\Lambda_{\alpha\alpha} - \Lambda_{\alpha\beta} \Lambda_{\beta\beta}^{-1} \Lambda_{\beta\alpha}} \begin{array}{c} \Lambda_{\alpha\alpha} \end{array} - \begin{array}{c} \Lambda_{\alpha\beta} \end{array} \begin{array}{c} \Lambda_{\beta\beta}^{-1} \end{array} \begin{array}{c} \Lambda_{\beta\alpha} \end{array} = \begin{array}{c} \Lambda_{\alpha\alpha} - \Lambda_{\alpha\beta} \Lambda_{\beta\beta}^{-1} \Lambda_{\beta\alpha} \end{array}$$

Diagram illustrating the Schur complement operation on a block matrix Λ .

The initial matrix Λ is a 6x6 matrix with rows and columns labeled $\xi_1, \xi_2, \xi_3, L_1, L_2, L_3$. The matrix is partitioned into four blocks: $\Lambda_{\alpha\alpha}$ (top-left 3x3), $\Lambda_{\alpha\beta}$ (top-right 3x3), $\Lambda_{\beta\alpha}$ (bottom-left 3x3), and $\Lambda_{\beta\beta}$ (bottom-right 3x3).

The operation performed is the Schur complement calculation: $\Lambda_{\alpha\alpha} - \Lambda_{\alpha\beta} \Lambda_{\beta\beta}^{-1} \Lambda_{\beta\alpha}$.

The result is shown as a 3x3 matrix, which is the Schur complement of $\Lambda_{\beta\beta}$ in Λ .

第一题

$$\Lambda'$$

	ξ_2	ξ_3	L_1	L_2	L_3
ξ_2	Dark Gray	Light Gray	Dark Gray	Dark Gray	Light Gray
ξ_3	Light Gray	Light Gray	White	Light Gray	Light Gray
L_1	Dark Gray	White	Dark Gray	Dark Gray	White
L_2	Dark Gray	Light Gray	Dark Gray	Dark Gray	White
L_3	Light Gray	Light Gray	White	White	Light Gray

可以发现，和 $\Lambda_{\alpha\alpha}$ 相比，增加了 L_1 和 L_2 之间的关联，信息矩阵更稠密了。

第二题

- ② 阅读《Relationship between the Hessian and Covariance Matrix for Gaussian Random Variables》. 证明信息矩阵和协方差的逆之间的关系。

- Hessian和协方差逆的关系

一个多元高斯随机变量的概率密度函数的负对数函数的Hessian矩阵, 和其协方差矩阵的逆是相等的。

$$\mathbf{C}\delta\mathbf{x} = \delta\mathbf{z} \sim \mathcal{N}(\mathbf{0}, \Sigma_z)$$

$$p(\mathbf{C}\delta\mathbf{x}) = \alpha \exp(\delta\mathbf{x}^T \mathbf{C}^T \Sigma_z^{-1} \mathbf{C} \delta\mathbf{x}) = \alpha \exp(\delta\mathbf{x}^T \Sigma_x^{-1} \delta\mathbf{x})$$

Consider a Gaussian random vector θ with mean θ^* and covariance matrix Σ_θ so its joint probability density function (PDF) is given by:

$$p(\theta) = (2\pi)^{-\frac{N_\theta}{2}} |\Sigma_\theta|^{-\frac{1}{2}} \exp\left[-\frac{1}{2}(\theta - \theta^*)^T \Sigma_\theta^{-1} (\theta - \theta^*)\right] \quad (\text{A.1})$$

The objective function can be defined as its negative logarithm:

$$J(\theta) \equiv -\ln p(\theta) = \frac{N_\theta}{2} \ln 2\pi + \frac{1}{2} \ln |\Sigma_\theta| + \frac{1}{2} (\theta - \theta^*)^T \Sigma_\theta^{-1} (\theta - \theta^*) \quad (\text{A.2})$$

which is a quadratic function of the components in θ . By taking partial differentiations with respect to θ_l and $\theta_{l'}$, the (l, l') component of the Hessian matrix can be obtained:

$$\mathcal{H}^{(l,l')}(\theta^*) = \left. \frac{\partial^2 J(\theta)}{\partial \theta_l \partial \theta_{l'}} \right|_{\theta=\theta^*} = (\Sigma_\theta^{-1})^{(l,l')} \quad (\text{A.3})$$

so the Hessian matrix is equal to the inverse of the covariance matrix:

$$\mathcal{H}(\theta^*) = \Sigma_\theta^{-1} \quad (\text{A.4})$$

For Gaussian random variables, the second derivatives of the objective function are constant for all θ because the objective function is a quadratic function of θ . Therefore, the Hessian matrix can be computed without obtaining the mean vector θ^* .

第二题

② 阅读《Relationship between the Hessian and Covariance Matrix for Gaussian Random Variables》. 证明信息矩阵和协方差的逆之间的关系。

• Hessian和信息矩阵的关系

We can then see it as an information. The covariance of score function above is the definition of Fisher Information. As we assume θ is a vector, the Fisher Information is in a matrix form, called Fisher Information Matrix:

$$\mathbf{F} = \mathbb{E}_{p(x|\theta)} [\nabla \log p(x|\theta) \nabla \log p(x|\theta)^T].$$

However, usually our likelihood function is complicated and computing the expectation is intractable. We can approximate the expectation in \mathbf{F} using empirical distribution $\hat{q}(x)$, which is given by our training data $X = \{x_1, x_2, \dots, x_N\}$. In this form, \mathbf{F} is called Empirical Fisher:

$$\mathbf{F} = \frac{1}{N} \sum_{i=1}^N \nabla \log p(x_i|\theta) \nabla \log p(x_i|\theta)^T.$$

Claim: The negative expected Hessian of log likelihood is equal to the Fisher Information Matrix \mathbf{F} .

Proof. The Hessian of the log likelihood is given by the Jacobian of its gradient:

$$\begin{aligned} \mathbf{H}_{\log p(x|\theta)} &= \mathbf{J} \left(\frac{\nabla p(x|\theta)}{p(x|\theta)} \right) \\ &= \frac{\mathbf{H}_{p(x|\theta)} p(x|\theta) - \nabla p(x|\theta) \nabla p(x|\theta)^T}{p(x|\theta) p(x|\theta)} \\ &= \frac{\mathbf{H}_{p(x|\theta)} p(x|\theta)}{p(x|\theta) p(x|\theta)} - \frac{\nabla p(x|\theta) \nabla p(x|\theta)^T}{p(x|\theta) p(x|\theta)} \\ &= \frac{\mathbf{H}_{p(x|\theta)}}{p(x|\theta)} - \left(\frac{\nabla p(x|\theta)}{p(x|\theta)} \right) \left(\frac{\nabla p(x|\theta)}{p(x|\theta)} \right)^T, \end{aligned}$$

第二题

- ② 阅读《Relationship between the Hessian and Covariance Matrix for Gaussian Random Variables》. 证明信息矩阵和协方差的逆之间的关系。

• Hessian和信息矩阵的关系

where the second line is a result of applying quotient rule of derivative. Taking expectation wrt. our model, we have:

$$\begin{aligned}\mathbb{E}_{p(x|\theta)} [\mathbf{H}_{\log p(x|\theta)}] &= \mathbb{E}_{p(x|\theta)} \left[\frac{\mathbf{H}_{p(x|\theta)}}{p(x|\theta)} - \left(\frac{\nabla p(x|\theta)}{p(x|\theta)} \right) \left(\frac{\nabla p(x|\theta)}{p(x|\theta)} \right)^T \right] \\&= \mathbb{E}_{p(x|\theta)} \left[\frac{\mathbf{H}_{p(x|\theta)}}{p(x|\theta)} \right] - \mathbb{E}_{p(x|\theta)} \left[\left(\frac{\nabla p(x|\theta)}{p(x|\theta)} \right) \left(\frac{\nabla p(x|\theta)}{p(x|\theta)} \right)^T \right] \\&= \int \frac{\mathbf{H}_{p(x|\theta)}}{p(x|\theta)} p(x|\theta) dx - \mathbb{E}_{p(x|\theta)} [\nabla \log p(x|\theta) \nabla \log p(x|\theta)^T] \\&= \mathbf{H}_{\int p(x|\theta) dx} - \mathbf{F} \\&= \mathbf{H}_1 - \mathbf{F} \\&= -\mathbf{F}.\end{aligned}$$

Thus we have $\mathbf{F} = -\mathbb{E}_{p(x|\theta)} [\mathbf{H}_{\log p(x|\theta)}]$.

□

Indeed knowing this result, we can see the role of \mathbf{F} as a measure of curvature of the log likelihood function.

似然分布的概率密度函数的负对数函数的Hessian矩阵的期望等于信息矩阵。

第二题

② 阅读《Relationship between the Hessian and Covariance Matrix for Gaussian Random Variables》. 证明信息矩阵和协方差的逆之间的关系。

- Hessian和信息矩阵的关系

综上所述，负对数最大似然估计的 Hessian 矩阵就是信息矩阵，同时也等于协方差的逆。所以平时往往就直接将协方差矩阵的逆当做信息矩阵。

第三题

- ③ 请补充作业代码中单目 Bundle Adjustment 信息矩阵的计算，并输出正确的结果。正确的结果为：奇异值最后 7 维接近于 0，表明零空间的维度为 7。

```
69 H.block(i*6,i*6,6,6) += jacobian_Ti.transpose() * jacobian_Ti;
70 /// 请补充完整作业信息矩阵块的计算
71 H.block(poseNums*6+j*3, i*6, 3, 6) += jacobian_Pj.transpose() * jacobian_Ti;
72 H.block(i*6, poseNums*6+j*3, 6, 3) += jacobian_Ti.transpose() * jacobian_Pj;
73 H.block(poseNums*6+j*3, poseNums*6+j*3, 3, 3) += jacobian_Pj.transpose() * jacobian_Pj;
```

$$\mathbf{J}_2 = \frac{\partial \mathbf{r}_{13}}{\partial \boldsymbol{\xi}} = \begin{bmatrix} \frac{\partial \mathbf{r}_{13}}{\partial \xi_1} & 0 & \frac{\partial \mathbf{r}_{13}}{\partial \xi_3} & 0 & 0 & 0 \end{bmatrix}$$
$$\mathbf{\Lambda}_2 = \mathbf{J}_2^\top \Sigma_2^{-1} \mathbf{J} = \begin{bmatrix} \left(\frac{\partial \mathbf{r}_{13}}{\partial \xi_1}\right)^\top \Sigma_2^{-1} \frac{\partial \mathbf{r}_{13}}{\partial \xi_1} & 0 & \left(\frac{\partial \mathbf{r}_{13}}{\partial \xi_1}\right)^\top \Sigma_2^{-1} \frac{\partial \mathbf{r}_{13}}{\partial \xi_3} & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ \left(\frac{\partial \mathbf{r}_{13}}{\partial \xi_3}\right)^\top \Sigma_2^{-1} \frac{\partial \mathbf{r}_{13}}{\partial \xi_1} & 0 & \left(\frac{\partial \mathbf{r}_{13}}{\partial \xi_3}\right)^\top \Sigma_2^{-1} \frac{\partial \mathbf{r}_{13}}{\partial \xi_3} & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

(46)

第三题

- ③ 请补充作业代码中单目 Bundle Adjustment 信息矩阵的计算，并输出正确的结果。正确的结果为：奇异值最后 7 维接近于 0，表明零空间的维度为 7。

```
PROBLEMS 4 OUTPUT TERMINAL DEBUG CONSOLE JUPYTER

0.00547299
0.0053236
0.00520788
0.00502341
0.0048434
0.00451083
0.0042627
0.00386223
0.00351651
0.00302963
0.00253459
0.00230246
0.00172459
0.000422374
3.21708e-17
2.06732e-17
1.43188e-17
7.66992e-18
6.08423e-18
6.05715e-18
3.94363e-18
lcx@lcx:~/Desktop/vio_homework/ch4/course4/nullspace_test/build$
```

在单目视觉SLAM中，信息矩阵存在一个7维零空间，对应7自由度的不客观量。

感谢各位聆听 !
Thanks for Listening

