



1. 信息矩阵相关

1.1 绘制上述系统的信息矩阵 Λ

我们可以先将系统所有约束对应的小信息矩阵求出，然后进行累加。

在求小信息矩阵时，只有与约束（即图中的边）相连接的变量（即图中的节点）所对应位置才有值，其余为0。

根据如图观测约束关系，有如下9个误差函数（边的数量）

$$r_1(\xi_1, L_1) \quad r_2(\xi_1, L_2) \quad r_3(\xi_1, \xi_2)$$

$$r_4(\xi_2, L_1) \quad r_5(\xi_2, L_2) \quad r_6(\xi_2, L_3) \quad r_7(\xi_2, \xi_3)$$

$$r_8(\xi_3, L_2) \quad r_9(\xi_3, L_3)$$

以 $r_1(\xi_1, L_1)$ 为例，有 $J_1 = \frac{\partial r_1(\xi_1, L_1)}{\partial \underline{x}} = \begin{bmatrix} \frac{\partial r_1}{\partial \xi_1} & 0 & 0 & \frac{\partial r_1}{\partial L_1} & 0 & 0 \end{bmatrix}$

状态量 $(\xi_1^T, \xi_2^T, \xi_3^T, L_1^T, L_2^T, L_3^T)^T$

那么 $\Lambda_1 = J_1^T \Sigma_1^{-1} J_1$

$$= \begin{bmatrix} \left(\frac{\partial r_1}{\partial \xi_1}\right)^T \Sigma_1^{-1} \frac{\partial r_1}{\partial \xi_1} & 0 & 0 & \left(\frac{\partial r_1}{\partial \xi_1}\right)^T \Sigma_1^{-1} \frac{\partial r_1}{\partial L_1} & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ \left(\frac{\partial r_1}{\partial L_1}\right)^T \Sigma_1^{-1} \frac{\partial r_1}{\partial \xi_1} & 0 & 0 & \left(\frac{\partial r_1}{\partial L_1}\right)^T \Sigma_1^{-1} \frac{\partial r_1}{\partial L_1} & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

依次可以求出 $\Lambda_2 \sim \Lambda_9$ ，然后叠加得 Λ

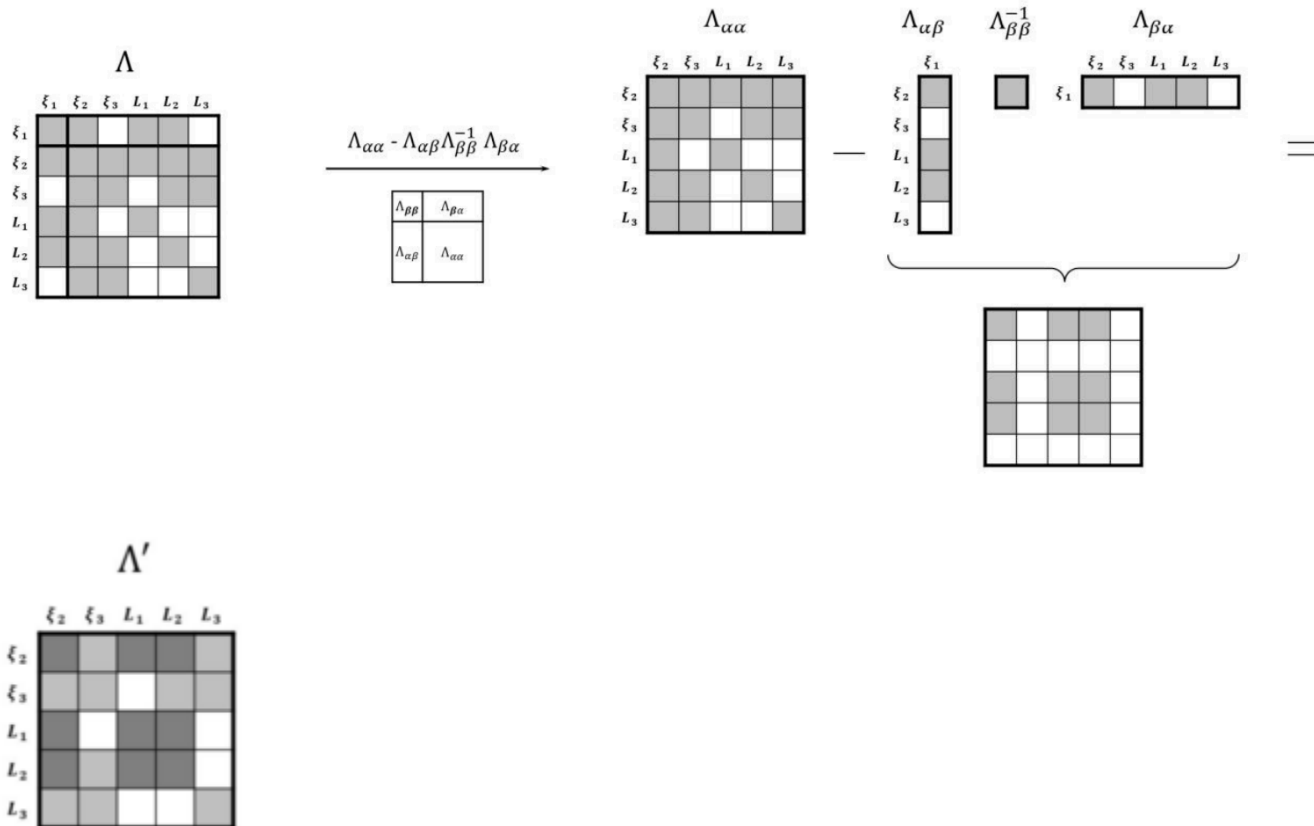
| | ξ_1 | ξ_2 | ξ_3 | L_1 | L_2 | L_3 |
|---------|---------|---------|---------|-------|-------|-------|
| ξ_1 | 3 | 1 | | 1 | 1 | |
| ξ_2 | 1 | 5 | 1 | 1 | 1 | 1 |
| ξ_3 | | 1 | 3 | | | 1 |
| L_1 | 1 | 1 | | 2 | | |
| L_2 | 1 | 1 | 1 | | 3 | |
| L_3 | | 1 | 1 | | | 2 |

总结一下也不难发现，非对角线上如果顶点之间存在一条边，该处对应累加元素个数为1，对角线上累加元素个数为与该顶点相连的边的个数。

1.2 绘制相机 ξ_1 被marg后的信息矩阵 Λ'

将 ξ_1 进行边缘化，根据原来的系统，与 ξ_1 相连的有 ξ_2 ， L_1 ， L_2 ，边缘化操作后，这三者之间

便会两两相连。



观察可见，和 $\Lambda_{\alpha\alpha}$ 相比，由于增加了 ξ_2, L_1, L_2 之间的关联，信息矩阵变得更加稠密了。

2.证明信息矩阵与协方差的逆之间的关系

1.Hessian矩阵和协方差逆的关系

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。

Consider a Gaussian random vector $\boldsymbol{\theta}$ with mean $\boldsymbol{\theta}^*$ and covariance matrix $\boldsymbol{\Sigma}_{\boldsymbol{\theta}}$ so its joint probability density function (PDF) is given by:

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

The objective function can be defined as its negative logarithm:

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

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

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

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

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

2. Hessian矩阵和信息矩阵的关系

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。

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.

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.

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

3. 零空间

补充代码部分需要我们回忆起重投影误差关于特征点坐标和相机位姿的雅可比如何写，并将其填入对应的位置即可。

源代码中的jacobian_Pj表示重投影误差对特征点的雅可比，计算其的理论依据参见《SLAM十四讲》P.187的公式（7.48）；jacobian_Ti表示重投影误差对相机位姿的雅可比，可以参见《SLAM十四讲》P.187的公式（7.46）。

以下给出相关代码实现：

```

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

```

```
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
xwl@xwl-Inspiron-15-7000-Gaming:~/Documents/VSLAM-f
```

奇异值最后7维接近0，即零空间维度为7。在单目视觉SLAM中，信息矩阵存在一个7维零空间，对应7自由度的不可观量。