2.3 理论基础

2.3.1 SLAM概率问题建模

SLAM模型由一个运动方程和一个观测方程构成

$$\left\{\begin{array}{l} x_k=f\left(x_{k-1}, u_k\right)+w_k \\ z_{k, j}=h\left(y_j, x_k\right)+v_{k, j} \end{array}\right. \tag{2.1}$$

其中x_k为相机位姿，u为运动测量读数，y_j为路标点，$z_{i,j}$为观测到的像素点位置，w，v为噪声，两个噪声项满足零均值的高斯分布，$w_k \sim N(0, \boldsymbol{R_k})$，$v_k \sim N(0, \boldsymbol{Q_{k,j}})$，考虑某个状态$\boldsymbol{x}$，以及一次与该变量相关的观测 $\boldsymbol{z_i}$，由于噪声的存在，观测服从概率分布 $\mathbf{p}(\mathbf{z_i}|\mathbf{x})$。多次观测时，各个测量值相互独立，则多个测量 $\boldsymbol{r}=(\boldsymbol{z_1},...,\boldsymbol{z_n})^T$ 构建的似然概率为：

$$\boldsymbol{p}(\boldsymbol{z}|\boldsymbol{x})=\prod_i \boldsymbol{p}(\boldsymbol{z_i}|\boldsymbol{x}) \tag{2.2}$$

如果知道机器人状态的先验信息 $p(x)$，如 GPS, 车轮码盘信息等，则根据 Bayes 法则，有后验概率：

$$\boldsymbol{p}(\boldsymbol{x}|\boldsymbol{z})=\frac{\boldsymbol{p}(\boldsymbol{z}|\boldsymbol{x}) p(x)}{p(\boldsymbol{z})} \tag{2.3}$$

求解目标：给定观测的信息，求系统的最优状态，使得在该状态下后验概率最大化，等价于求解状态量x的最大似然估计$\boldsymbol{p}(\boldsymbol{x}|\boldsymbol{z})$。通俗地来说就是“在什么样的状态下，最可能产生现在观测到的数据”

$$P(\boldsymbol{x}|\boldsymbol{z})=\frac{P(\boldsymbol{z}|\boldsymbol{x}) P(\boldsymbol{x})}{P(\boldsymbol{z})} \propto P(\boldsymbol{z}|\boldsymbol{x}) P(\boldsymbol{x}) \tag{2.4}$$

由此可以得到状态量x的最大后验估计

$$\boldsymbol{x}_{MAP}^*=\arg\max P(\boldsymbol{x}|\boldsymbol{z})=\arg\max P(\boldsymbol{x}|\boldsymbol{z}) P(\boldsymbol{x}) \tag{2.5}$$

给定一组观测的信息，可以用最小化负对数的方式求一个高斯分布的最大后验

$$\boldsymbol{x}_{MAP}^*=\arg\max_x \prod_i p\left(\boldsymbol{z}_i \mid \boldsymbol{x}\right) p(\boldsymbol{x}) \tag{2.6}$$

$$\boldsymbol{x}_{MAP}^*=\arg\min_x\left[-\sum_i \ln p\left(\boldsymbol{z}_i \mid \boldsymbol{x}\right)-\ln p(\boldsymbol{x})\right] \tag{2.7}$$

已知高斯分布

$$P(x)=\frac{1}{\sqrt{(2\pi)^{\mathcal{N}} \det(\boldsymbol{\Sigma})}} \exp\left(-\frac{1}{2}(x-\mu)^T \boldsymbol{\Sigma}^{-1}(x-\mu)\right), x \sim N(\mu, \Sigma) \tag{2.8}$$

两边取负对数

$$-\ln(P(x))=\frac{1}{2} \ln\left((2\pi)^N \det(\boldsymbol{\Sigma})\right)+\frac{1}{2}(x-\mu)^T \boldsymbol{\Sigma}^{-1}(x-\mu) \tag{2.9}$$

等式右边略去与x无关的第一项，只要最小化右侧的二次型项，等式2.7代入式2.8后等价于

$$x_{\mathrm{MAP}}{}^*=\arg\min_x \sum_i \frac{1}{2}(z_i-w_i)^T \Sigma_i^{-1}\left(z_i-w_i\right)+\frac{1}{2}(x-v_x)^T \Sigma_x^{-1}\left(x-v_x\right) \tag{2.10}$$

写成马氏距离（ Σ 范数）的形式：

$$x_{\mathrm{MAP}} = \arg\min_{x} \sum_{i} \|z_i - w_i\|^2_{\boldsymbol{\Sigma}_i} + \|x - v_x\|^2_{\boldsymbol{\Sigma}x} \tag{2.11}$$

在仅有视觉传感器的情况下，忽略运动方程，代入SLAM观测方程(2.1)后可得

$$\boldsymbol{x}^* = \arg\min\,[\boldsymbol{z}_{k,j} - h(\boldsymbol{y}_j, \boldsymbol{x}_k)]^T \boldsymbol{\Sigma}^{-1}_{k,j} [\boldsymbol{z}_{k,j} - h(\boldsymbol{y}_j, \boldsymbol{x}_k)] \tag{2.12}$$

其中 Σ 为观测协方差，$\boldsymbol{z}_{k,j} - h(\boldsymbol{y}_j, \boldsymbol{x}_k)$为观测残差。在实际应用中，通常构建如下的最小二乘问题，通过不断地迭代求解状态量地增量 $\Delta\boldsymbol{x}$，并更新状态: $\boldsymbol{x} = \boldsymbol{x}' + \boldsymbol{\Delta x}$ ，将更新后的状态$\boldsymbol{x}$代入目标函数， $f(\boldsymbol{x}) = \boldsymbol{z}_{k,j} - h(\boldsymbol{y}_j, \boldsymbol{x}_k)$ 即观测残差，使得目标函数的 Σ 范数达到最小。

$$\Delta\boldsymbol{x}^* = \arg\min_{\Delta\boldsymbol{x}} \frac{1}{2} \parallel \boldsymbol{f}(\boldsymbol{x}) \parallel_{\Sigma} \tag{2.13}$$

状态的增量可由高斯-牛顿（Gauss-Newton）法求解，每次迭代需要求解如下的线性方程组

$$\boldsymbol{J}(\boldsymbol{x})^T \Sigma^{-1} \boldsymbol{J}(\boldsymbol{x}) \Delta\boldsymbol{x} = -\boldsymbol{J}(\boldsymbol{x})^T \Sigma^{-1} f(\boldsymbol{x}) \tag{2.14}$$

$\boldsymbol{J}(\boldsymbol{x})$表示目标函数对状态量的雅可比矩阵，$\Sigma^{-1}$表示状态量的之间的协方差矩阵，方程可简写为$\boldsymbol{H\Delta x = b}$，称为正规方程。

2.3.2 Bundle Adjustment

Bundle Adjustment中文译作光束平差法，是指从视觉重建中提炼出最优的3D模型和相机参数（内参和外参）。从每个特征点反射出来的几束光（Bundles of Light Rays），把相机姿态和特征点的位置做出最优的调整（Adjustment）最后收束到光心的这个过程，简称BA。

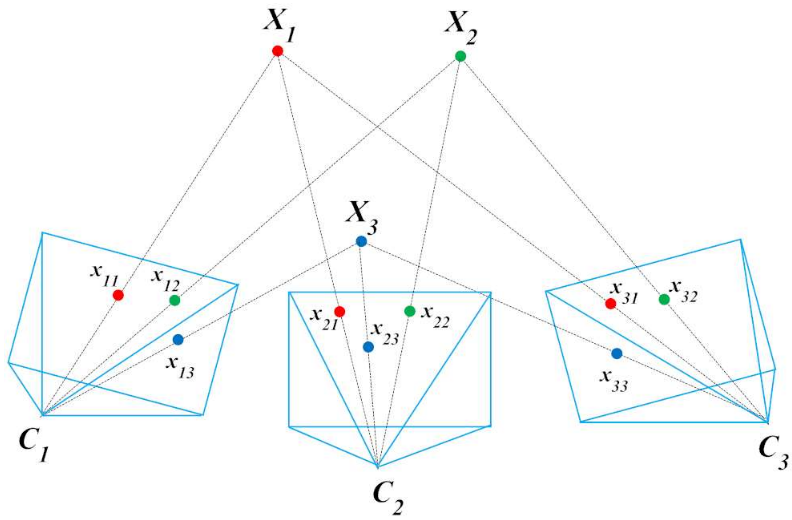

图 2-1 Bundle Adjustment示意图

对于BA，通过2.2.1中提到的概率建模方法，可利用重投影误差构建式（2.14）所示的最小二乘问题:

$$e(\boldsymbol{X}, \boldsymbol{T}) = \sum_{\mathrm{i}=1}^{n} \left\| \boldsymbol{u_i} - \frac{1}{s_i} \boldsymbol{K} \cdot \boldsymbol{T} \cdot \boldsymbol{X_i} \right\|_{\Sigma} \tag{2.15}$$