视觉定位与感知作业（EE382）作业5

Q1:基础矩阵基础矩阵F与本征矩阵E有什么区别？两者之间的关系是什么？（20分）

基础矩阵F和本征矩阵E的区别有：

属性	基础矩阵	本征矩阵
坐标系	图像坐标系	相机坐标系
依赖参数	相机的内参和外参	只有相机外参
自由度	7	5

两者之间是可以互相转换的：我们可以来看看推导过程：

完整推导过程如下：

$$\begin{aligned} p_2^T K^{-T}[t]_\times R K^{-1} p_1 &= 0 x_2^T [t]_\times R x_1 = 0 \\ E &= [t]_\times R \\ F &= K^{'-T} E K^{-1} \\ x_2^T E x_1 &= 0 \\ p_2^T F p_1 &= 0 \end{aligned} \tag{1}$$

推导过程中，p_1, p_2是两个像素点，而x_1, x_2为上述两个像素点转化到归一化平面上的坐标。前两式除去p_1, p_2, x_1, x_2的部分我们称之为基础矩阵和本征矩阵。后面的几个式子我们也推导出了两个矩阵的转换关系

Q2:为什么本征矩阵只有两个非零奇异值，而且都为1？（20分）

根据上一问的结论我们写出本征矩阵的表达式：

$$E = [t]_\times R \tag{2}$$

在这个表达式里面，R是旋转矩阵，$[t]_x$表示叉乘矩阵（平移向量），是反对称矩阵。我们对这个叉乘矩阵做SVD分解，首先根据性质：

$$[t]_x = \lambda U \begin{bmatrix} 0 & 1 & 0 \\ -1 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} U^T = \lambda U \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} 0 & 1 & 0 \\ -1 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} U^T \tag{3}$$

根据SVD特征分解的性质，首先我们设矩阵A：

$$A = \begin{bmatrix} 0 & 1 & 0 \\ -1 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \tag{4}$$

矩阵A满足条件

$$AA^T = 0 \tag{5}$$

因此我们可以把一个本征矩阵做以下分解：

$$E = [t]_\times R = U \begin{bmatrix} 0 & 1 & 0 \\ -1 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} AU^T R \tag{6}$$

然后，我们可以根据矩阵W和旋转矩阵R的固有性质，得到证明：

$$(AU^T R)^T (AU^T R) = R^T U A^T AU^T R = I \tag{7}$$

由（7）可知，（$WU^T R$）是一个单位正交矩阵，因此（6）式中后三项的乘积是一个对角矩阵，因此（6）式符合SVD分解的条件。而根据（6）式中间对角矩阵的值可知，本征矩阵只有两个非零奇异值而且都是1。

Q3:若使用RANSAC求解P3P相机位姿估算问题，若要达到>99.9%的outlier剔除概率，那么需要多少次采样？（20分）

我们假设每次操作的采样点共有k个，并将outlier剔除概率简写为p，同时将inlier的比例计作q：

采样次数n与剔除概率p的关系式可写作：

$$p(n) = 1 - (1 - u^M)^N \tag{8}$$

根据上述关系我们可以得到不等式：

$$p(n_{min}) = 1 - (1 - u^M)^N > 0.999 \tag{9}$$

求解该不等式：

$$N \geq \frac{log(1 - p(n_{min}))}{log(1 - u^M)} \tag{10}$$

已知$M = 3, p(n_{min} = 0.999)$可求得：

$$N \geq \frac{-3}{log(1 - u^3)} \tag{11}$$

由于N必须为整数，所以我们取N为大于（11）式右边的最小整数

Q4：在求解Structure-from-motion问题上，基于矩阵分解的方法有什么局限性？（20分）

首先我们给出矩阵分解的方法：

$$
\begin{aligned}
&M = \Pi S \\
&(M \in R^{2M\times N}, \Pi \in R^{2M\times 3}, S \in R^{3\times N})
\end{aligned} \tag{12}
$$

我们将其局限性概括为以下三点：

1、观测矩阵如果规模很大的话，其分解将产生巨大的复杂度，进而导致很长的计算时间。

2、对M矩阵做SVD分解时，根据性质，对角矩阵的秩必须保证为3才行。但是由于各种因素的影响（如误差，噪声等），有可能M的秩大于3 。此时是没有办法进行无损失分解的。

3、求解变换矩阵A时，需要对Π矩阵进行正交性近似操作。既带来误差，也导致计算时间继续增加。

Q5:Structure-from-motion增量式求解过程中为什么需要时不时调用Bundle adjustment? （20分）

求解Structure-from-motion问题时，我们可以将BA问题看成一个优化目标为使重投影误差最小的优化问题。重投影误差的表达式如式（13）：

$$
z_{ij} = m_{ij} - P(\theta_i, X_j) \tag{13}
$$

我们知道，Bundle adjustment是一种非线性优化方式。调用Bundle adjustment在优化重投影误差的同时优化相机位姿和空间点位置，不断调用的过程可以使最终的结果更加准确。