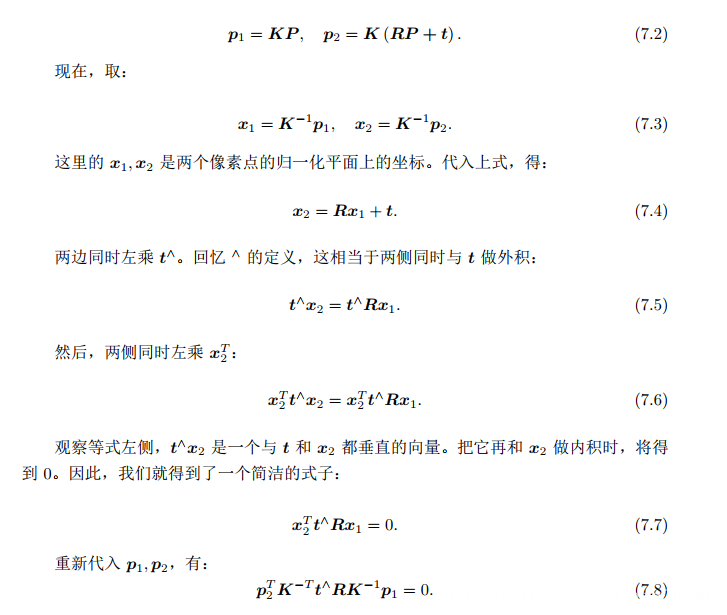
1. ORB-SLAM2讲解

ORB-SLAM2使用非常宽松的关键帧策略，然后再后端剔除冗余的关键帧

1. 优化方式，对极几何，PNP ICP BA

**1）对极几何**：估计相机运动



2D-2D，根据图相平面信息估计单目帧间运动相对位姿关系。

求解的Rt各有两组解，但是只有一组解能让P都具有正的深度，所以带入检验排除即可

E是由等于0的约束得到的，E=t^R，tR工六个自由度，而E没有尺度，所以只有5个自由度，只要5点就可，但是可以用多点（八点法），利用线性最小二乘或者RANSAC（一般是这个，因为经常会有误匹配）求解，然后再用SVD分解E得到t和R

单目估计的轨迹和位姿与真实地图相差一个尺度因子，对极约束得到的R是准的，但是t不具有真实尺度

单目初始化不能只有纯旋转，必须要有一定的平移，否则E矩阵为0，因为E=t^R，所以一般单目初始化都会左右平移

2）**三角测量**：估计地图点深度

根据前后两帧图像中匹配的特征像素坐标和对极几何得到的R、t，计算特征点的三维空间坐标，但是因为噪声的影响，两个匹配的像素坐标得到的特征点坐标有误差，所以一般用最小二乘

矛盾：平移小时，匹配的特征点像素误差会导致特征点的深度误差，平移太大会导致外观发生变化，特征提取和匹配变得困难。即平移增大，匹配失效，平移太小，三角化精度不够

3）**3D-2D：PnP，不需要使用对极约束**

当我们用前两帧利用对极几何得到帧间运动R1和t1之后，然后根据三角测量可以得到特征点的深度，求得其3D坐标。我们选择第一帧图像的相机坐标为世界坐标系，后续所有坐标变换都以此作为参考

第三帧和第二帧匹配的时候，第二帧的3D坐标和第三帧的2D坐标构成了PnP的3D-2D问题，然后就可以求得第三帧图像相对于第二帧图像的R2和t2，也可以求得第三帧相对于世界坐标系的位姿，后续所有计算均同理进行，便可实时跟踪相机位姿。

PnP的求解可以用P3P，直接线性优化（DLT）、EPnP，UPnP，**BA**。一般是先用P3P或EPnP，然后用BA进行优化

特征点在两个相机位姿下的归一化坐标为x1，x2，深度为s1，s2，则两点坐标为s1x1，s2x2，根据PnP求得R和t之后，s1x1=R(s2x2)+t，利用最小二乘求解坐标

4）**Bundle Adjustment（BA）**，最小化重投影误差

将相机位姿和特征点空间位置一起优化，也可以对ICP和PnP的结果进行优化，最小化重投影误差，即像素坐标和3D点按照当前估计的位姿进行投影得到的位置相比较的误差

5）ICP：3D-3D，

已知对应点求解，不知对应点时用最近点作为对应点，所以旋转大于30°之后就容易失效。ICP可能有多解，有唯一解时

6）卡尔曼滤波推导：

马尔科夫性：k时刻状态只与k-1时刻相关，与之前的状态无关

7）EM（Expectation Maximization）：

比如200个人（但是不知道是男是女，只知道身高），男的和女的各服从一个N(μ1，σ1^2), N(μ2，σ2^2).要估计这四个参数。

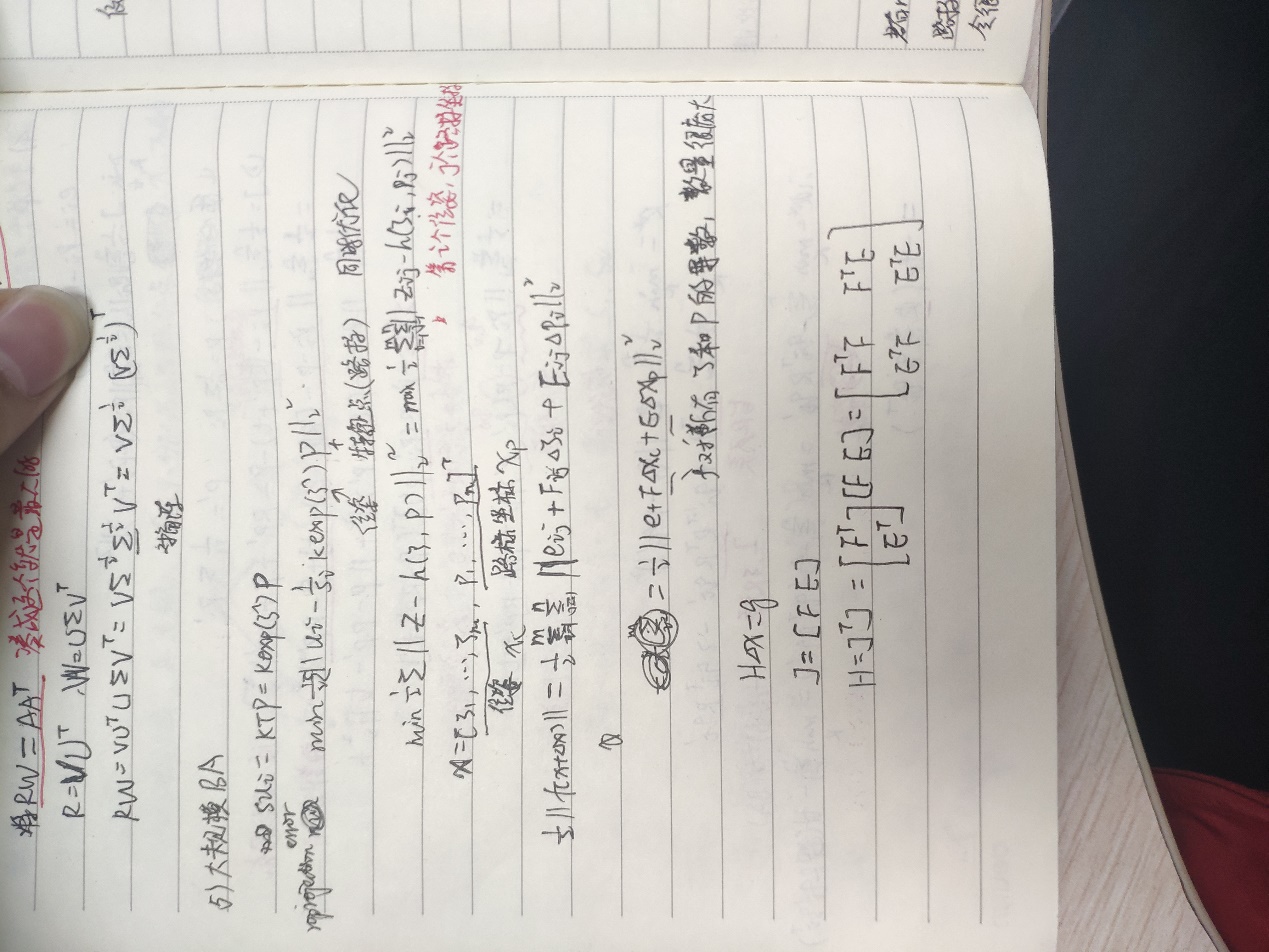
则可以首先假设两个分布N(μ1=172，σ1^2=5^2), N(μ2=162，σ2^2=5^2)。然后计算每个身高知道某个人属于男还是女，这就是Expectation

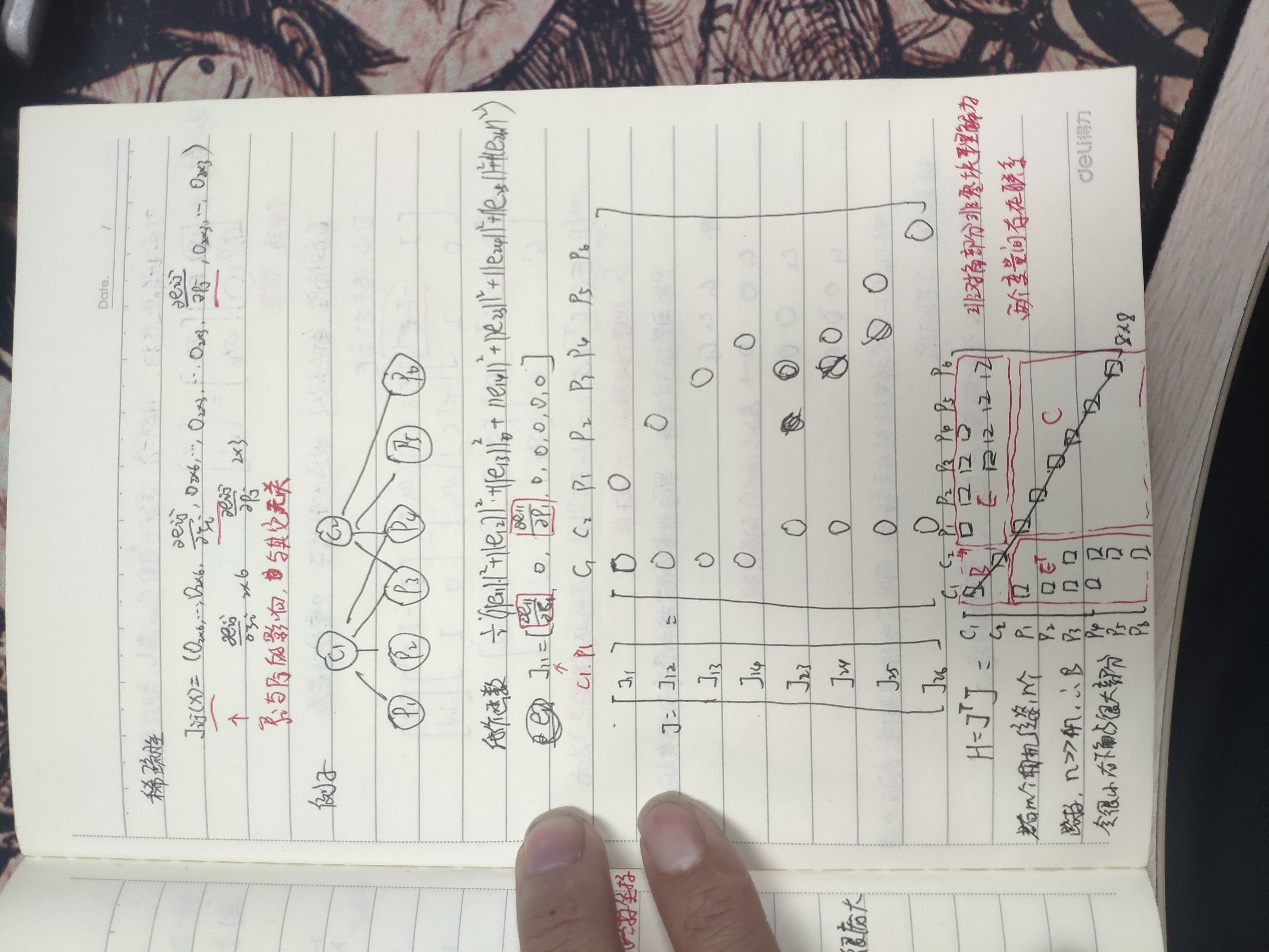
当我们分出男女后，又可以根据分类后每人的身高进行极大似然估计，得到男女的分布参数，这步交Maximization

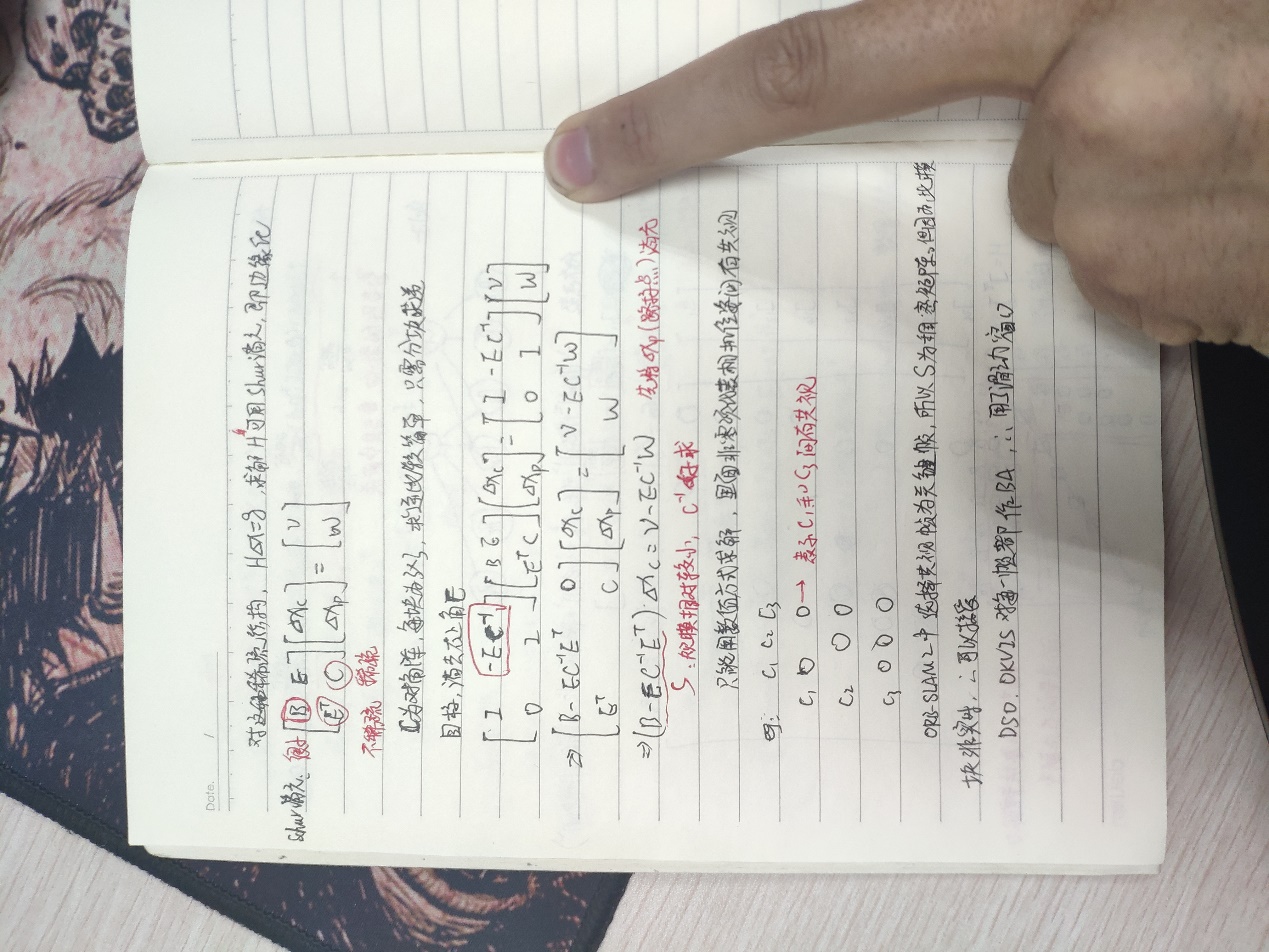
当我们调整分布的时候，属于男女的人又变了，则需要调整E

如此往复，直到参数基本不变或达到最大迭代次数

8）BA：每个特征反射的光束，镜柜特征位姿调整和相机姿态调整，最终汇聚于光心的过程







ORB-SLAM2中选择的都是有共视的关键帧，所以S矩阵是稠密矩阵，但因为这个线程处理非实时，所以也能接受

DSO、OKVIS用的所有关键帧做BA，所以需要用滑动窗口来降低S矩阵维数

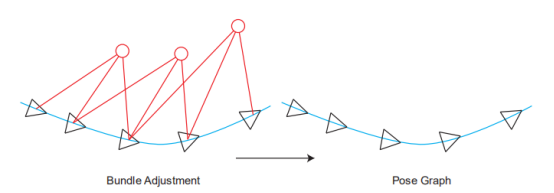
9）位姿图（Pose Graph）

BA和图优化，是将位姿点和路标点一起优化。特征点一般非常多，机器人轨迹越长，特征点增长很快，H矩阵很大

所以对路标点和位姿优化几次后，就舍弃路标点的优化，只保留位姿优化

滑动窗口：丢弃历史数据

位姿图：只保留Pose的边



表示估计得到的相机位姿的绝对坐标，表示ij个位姿之间的相对位姿变换，可以用直接法或者特征点法求得。但是因为误差的存在，等号不会严格成立，有误差，所以需要优化误差。





10）因子图优化

11）后端优化方式的差异

A.卡尔曼滤波器：从k-1时刻后验推k时刻先验，从k时刻先验推k时刻后验；

B.扩展卡尔曼滤波器：对卡尔曼滤波器进行修正，针对不是线性的情况，采用一阶泰勒展开近似线性。

C.BA优化：把一路上的所有坐标点与位姿整体放在一起作为自变量进行非线性优化

D.PoseGraph优化：先通过一路递推方式算出的各点位姿，通过数学方式计算得到一个位姿的变换A，再通过单独拿出两张图像来算出一个位姿变换B，争取让B=A

E.因子图优化：保留中间结果，每加入一个点，对不需要重新计算的就直接用之前的中间结果，需要重新计算的再去计算，从而避免冗余计算。

12）回环检测：

可以建立全局一致地图，降低累积误差。可以用于重定位

需要计算图像间的相似性，s(A,B)=||A-B||.但是这种方式不好，准确率和召回率都很差，所以用BoW（k叉树，kmeans聚类建立），然后计算两张图片的IDF和TF，权重w=IDF\*TF

词袋的局限性：只依赖外观，没有用几何信息，所以外观相近的可能会被当成回环，所以需要验证。（建立回环缓存机制，多次回环才算；空间一致性检验，回环检测的两个帧进行特征匹配，贵相机运动，再把运动放到Pose Graph中，检查估计出入是否很大）

1. 卡尔曼滤波推导

1）卡尔曼的本质是预测和观测两个高斯分布融合，求解其中的重合部分，得到新的高斯分布（分布函数有一个缩放因子，概率密度积分不为一，但是方差均值性质不变），其为状态的后验分布。

2）两个高斯分布相乘仍为高斯分布的推导

|  |  |
| --- | --- |
| F:\Users\gxf\AppData\Local\Temp\WeChat Files\bc3234c8845e3d26e30f1d13ffd58b9.jpg | F:\Users\gxf\AppData\Local\Temp\WeChat Files\e4b818ef5232760b687aea1a7f79cf5.jpg |

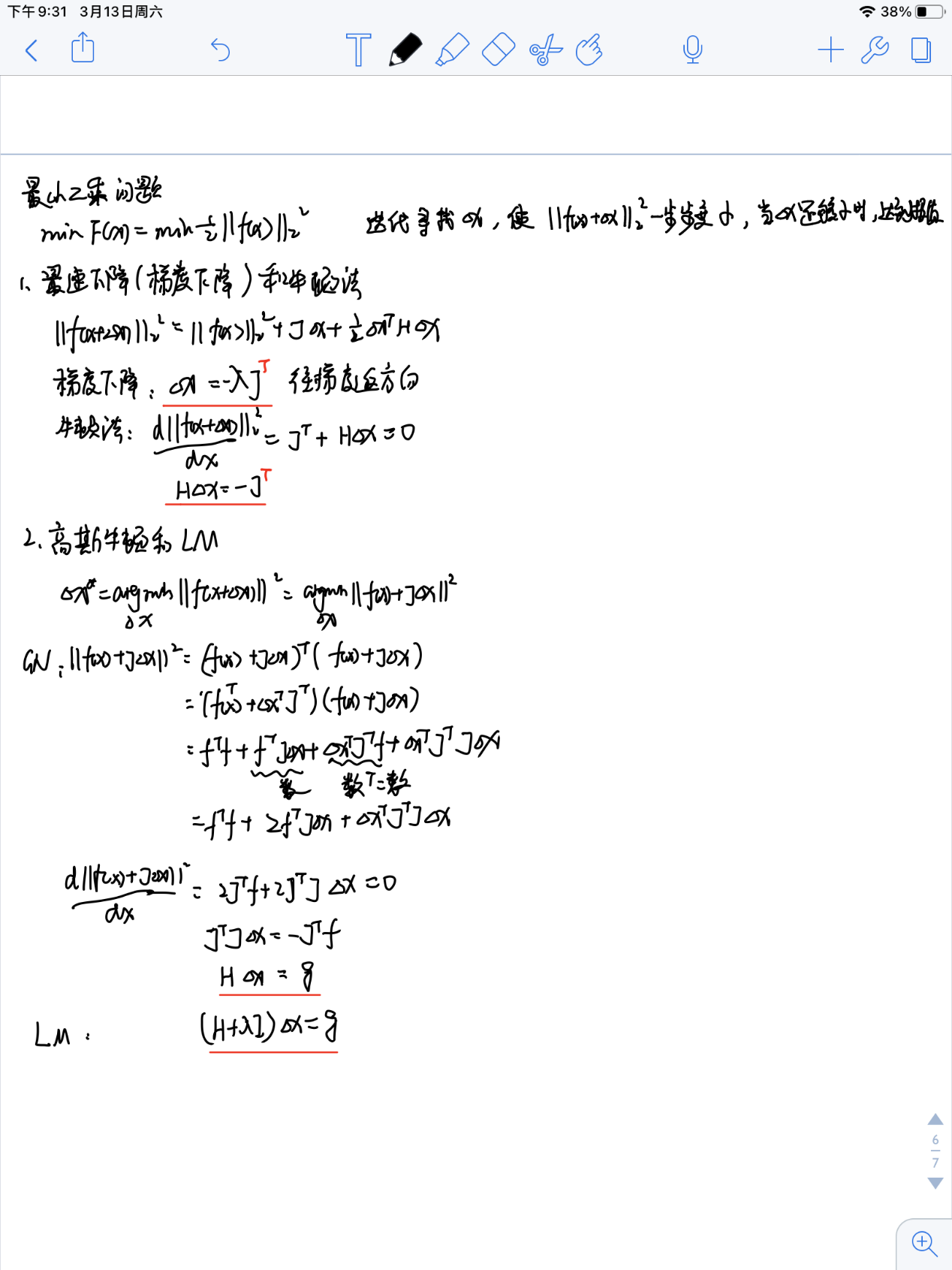
1. Cartographer
2. 前端：local SLAM，建立维护submaps

后端：global SLAM，主要进行loop closure，采用Branch-and-Bound Approach(BBA)。3D情况下还根据IMU寻找重力方向。

分支定界：将搜索窗口表达为一个树，每一个子节点形成了对父节点的所有可能性的划分。父节点代表了该子树下的所有解的可能性。每个节点都有一个得分score(c)，该分值是该子树所有分值的上界。所以如果得分阈值比score(c)大，则此子树不可能满足要求，则直接被剪枝，从而大大加速。实际问题中，大多数分枝（不是最优解的）得分都非常低，所以一个阈值就可以剪掉很多子树。所以相比于全局搜索更快。但是存储效率降低，score需要预先计算好子图的每个pixel的得分值，直接存储下来。当输入一个节点时，直接lookup即可。即存储空间换时间效率。

每一个分辨率对应一个栅格地图，构造了一个栅格地图的金字塔，0.05 0.1 0.2 0.4（上层）

5.梯度下降、牛顿法、高斯牛顿、LM方法



1. 最速下降
   1. 优点：快
   2. 缺点：过于贪心，容易走出锯齿状，增加迭代次数
2. 牛顿法
   1. 优点：快
   2. 缺点：规模大时H求解计算量大；需要方程二阶可导；非凸函数H逆可能不存在
3. 高斯牛顿
   1. 优点：避免求H，减少计算量
   2. 缺点：JJT半正定，奇异时求逆算法不收敛；若△x太大，可能局部近似不精确，严重时不收敛；容易陷入锯齿状
4. LM
   1. 优点：一定程度上避免JJT非奇异和病态，得到更稳定精确的解
   2. 缺点：收敛速度·比GN慢