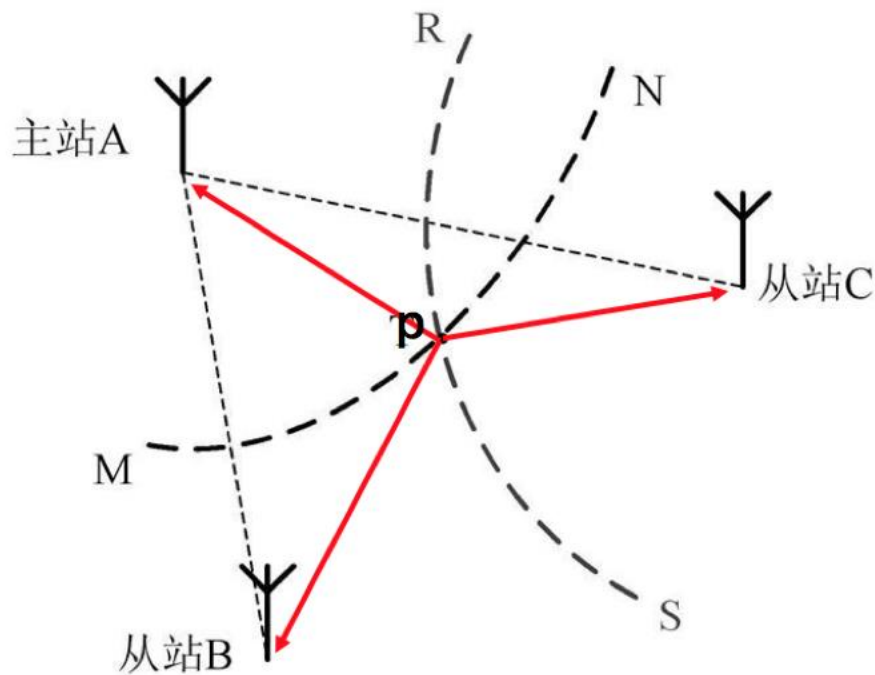


课程作业

背景知识

室内定位系统中, 一种常用的定位原理是 TDOA(Time-Difference-of-Arrival), 即到达时间差定位。以下图为例:



假设我们待定位的点是 P 点, 为了完成定位任务, 即确定 P 点的坐标, 首先需要在场地上部署多个基站。假设我们部署了 3 个基站, 分别为 A、B、C, 其中 A 是主基站, 它们的坐标已知。现在通过基站和 P 点标签之间的通信, 可以确定 P 点到主基站和到其他基站之间的距离差, 即已知 $PA-PB$ 和 $PA-PC$ 。显然, 通过 A、B 坐标和 $PA-PB$, 可以确定 P 在一条双曲线上, 同样地, 通过 A、C 坐标和 $PA-PC$ 可以确定另外一条双曲线, 两者求交点即可得到 P 点坐标。

相关算法和优化思路

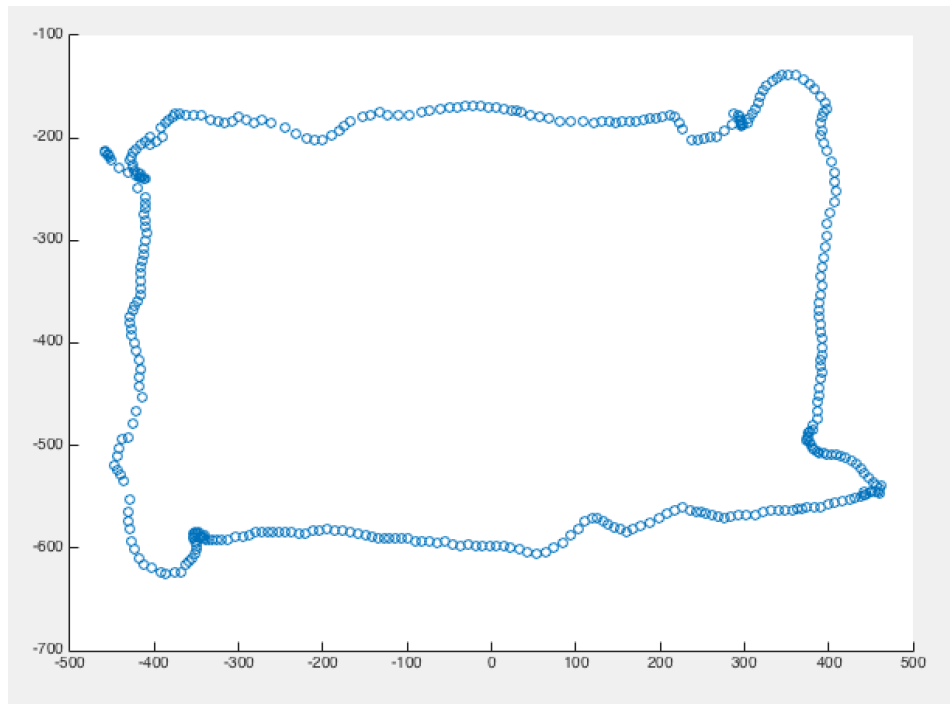
在实际任务中，由于环境中各种因素的影响，误差是普遍存在的，即距离差的测量是不准确的。为了提升系统的精度和鲁棒性，一般会部署多于 3 个基站，同样分为 1 个主基站和多个从基站。最简单的定位方式是选择主基站和任意 2 个从基站，通过上述双曲线求交点的方式进行定位，但这种方法在数据误差较大的情况下有明显的偏差。一种简单的优化策略是每次选择主基站和不同的 2 个从基站，按上述方式得到多个解，然后选择其中最合理的一个作为定位结果（理论上来讲，如果测量没有误差，所有的双曲线会交于同一个点，即 P 点，但实际环境下，误差会导致这些双曲线并不交于用一个点）。选点的策略有很多，比如根据上一时刻的位置选最近的，或是分析运动轨迹，选择位于轨迹上的点（定位系统中并不仅仅是对静止的点进行定位，而是需要给出一段时间内的运动轨迹）。

除了上述通过联立双曲线方程求解坐标的方法，目前还有许多解决 TDOA 定位的算法，其中比较经典的是文献[1]中提出的 Chan 算法和[2]中提出的 Taylor 序列方法，感兴趣的同学可以参考。

任务描述

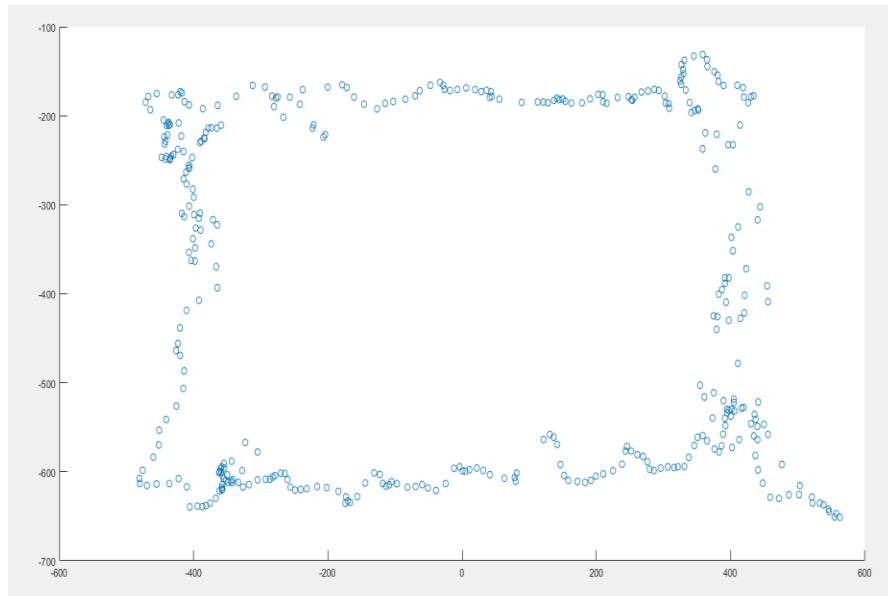
本次作业的主要任务是了解**最小二乘法拟合室内定位系统中的轨迹和对数数据进行合理的插值**。对第一个任务，我们给出一组 TDOA 数据，每一条数据对应一个时刻待定位点到各基站的距离差信息，可以计算出一个坐标，随后根据这些点的分布情况选择合适的直线或曲线函数进行最小二乘逼近，形成运动轨迹。

为了降低作业的难度，我们事先给出实验数据对应的大致真实轨迹，如下图所示：



1. 第一个任务就是拟合形成类似于上图的轨迹，这一轨迹类似于矩形，你可以使用最简单的函数来拟合，也就是 4 个一次函数（表示四条边），也可以在一些有明显弯曲的地方用二次函数这类曲线来拟合，或是采用一些更复杂的函数，展示效果时你要讲数据投影到你所拟合的曲线上。
2. 我们建议在左下角(第 67-83 条数据)和右上角(第 269-295 条数据)使用二次函数拟合，如果你完成了这一步，请在报告中写明拟合结果，不要求做投影。

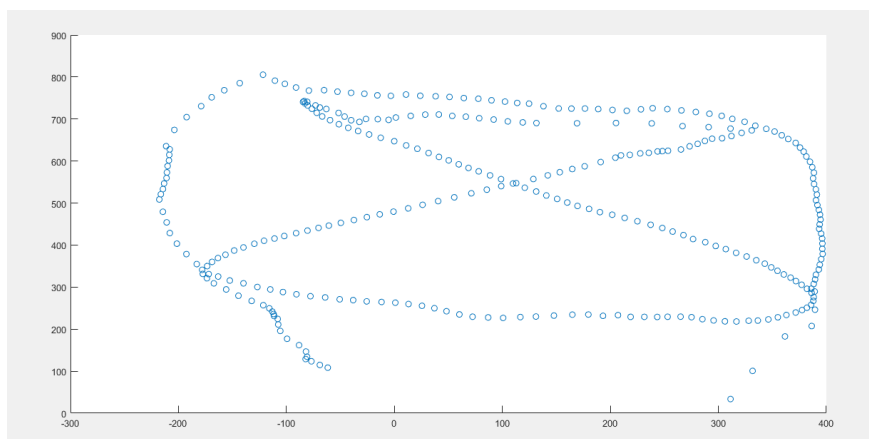
我们也给出了实验数据用 Chan 算法计算出的点，并进行了一些简单处理，其散点图如下：



不难看出，Chan 算法得到的点已经有了真实轨迹的大致形状。你可以不用我们给出的 TDOA 数据去逐条计算坐标，而直接用我们给出的这些已经计算出来的点进行拟合。

数据中为少量点被添加了随机误差，但忽视它依然能够顺利完成此任务。，如果你能设计方法对异常点进行处理，你将获得额外的附加分数。

2.第二个任务要求掌握基本的插值方法，我们给出一组真实环境下 tdoa 的解算结果，这是一组记录人行走的结果，每条数据都包含了准确的时刻，这组结果中包含时间条，解算结果，以及指出了哪些数据为缺失结果（不是时间条上每一个点都能得到解算坐标），你的任务是根据已解算出的坐标插值出部分缺失结果，未插值前的数据如下图：



你的任务是要不通过已有的插值库自行实现“**分段线性插值**”和“**三次样条插值**”，“**多项式插值**”也是工业界常用的插值方法，在这里我们不做硬性要求，通过查阅资料实现“**多项式插值**”可以获得一定的奖励分数，或者实现我们上述以外的插值算法并进行对比分析也可以得到**奖励分数**。完成这项作业我们给出以下几条建议：

1. 选择丢失点前后一定范围内的解算坐标作为插值的依据，范围可以自行给定，可做不同范围的对比实验，需在实验报告中给出。
2. 不是所有的缺失结果都能被很好的插值，你需要对得到的插值结果进行一个简单的评估（根据已有坐标）来决定是否保留它，最后会根据保留的插值结果数量和结果的合理性来进行评分。
3. 插值时只使用给出的坐标，不使用自己已经插值出的结果，当然你也可以都进行尝试并对比，我们在这里指出后者的效果没有前者好。
4. 我们在数据中设置了一些较长的连续缺失数据，全部合理的插值出来显然有些难度，你可以只对这些数据的前一部分和后一部分进行插值。
5. 在你的实验报告中，需要展现插值数量和插值效果，以及对“**分段线性插值**”和“**三次样条插值**”，两种方法的对比和分析，你的实验结果可以和一些现有的插值函数进行对比，比如 matlab 的 `interp1()`或 `spline()`函数等。

6. 如果你想完成“多项式插值”，我们给出的建议是不要超过三次多项式。

数据说明

给出的数据文件为“homework.mat”，这是 matlab 下的标准数据文件，在 matlab 下双击或调用 load(“homework.mat”)命令即可读取该文件。读取后可以在 workspace 中看到该文件包含了四组数据，“tdoa_path”、“path_chan”和“config”用于完成第一个任务，“coord”用于完成第二个任务。

“tdoa_path”：这是原始的 TDOA 数据文件，按时间的先后顺序采集，共有 385 条。对于每一条数据，第 1 维表示该数据中基站的个数；假设基站有 k 个，则随后的 k 维数据是这 k 个基站的标号，其中第 1 个是主基站；再后面的 k 维数据是待定位的点到主基站和到这 k 个基站的距离之差，显然，这 k 维数据中的第 1 个数据必然为 0，因为待定位点到主基站和到第 1 个基站（就是主基站）的距离相同。最后 1 维数据在本次任务中没有作用，可以忽略。

5	25	16	21	172	230	0	-495.7700	215.7600	-364.4400	-43.1500
---	----	----	----	-----	-----	---	-----------	----------	-----------	----------

以上面的一条数据为例，5 表示这条数据中有 5 个基站，其中主基站标号为 25，另外的基站标号是 16、21、172 和 230。我们需要定位的点到 25 号基站和到 25、16、21、172、230 号基站的距离之差分别为 0、-495.77、215.76、-364.44、-43.15。最后 1 维数据没有截图。

“path_chan”：这是“tdoa_path”数据通过 Chan 算法得出的点，每一条 TDOA 数据对应一个点，共 385 个。这些点形成的散点图上面已做过展示。

“config”：这是基站的位置数据，第 1 维是基站标号，后面 3 维是该基站的 x、y、z 轴坐标。注意，我们现在只考虑二维形式，因此只需用到(x, y)。

“coord”：这是用于插值任务的数据，每一行是一条单独的数据，数据的第 1

维是时间轴上时刻值，第 2 维和第 3 维表示了该时刻值得坐标(x,y)，最后一维表示数据是否缺失，缺失为 0，未缺失为 1，你的任务是根据缺失数据的时刻值插值坐标。

评分标准

本次作业需要提交 matlab 代码（建议，但其他语言也行）和实验报告，报告中需给出拟合用的函数、优化策略（如果有的话）、坐标的散点图（如果自己计算坐标的话）和得到的轨迹图像。本次作业最简单的完成方式是在我们给出的 Chan 算法计算结果的基础上，用 4 条直线进行拟合，并在第二组数据上简单实现两种插值方法，你的实验报告里要包含你运行的结果，这样可以得到基本的分数。如果你想要更高的分数，可以（但不限于）：

1. 自己设计算法（或者查阅相关文献）通过 tdoa 数据计算坐标
2. 用二次函数或其他更复杂的函数进行拟合
3. 完成插值部分的奖励题
4. 检测异常点并剔除

上述几点不做强制要求，你可以完成 1 点或几点，也可以忽略，我们会依据你采用的算法、策略和得出的结果进行评分。我们强烈鼓励提出自己的想法，即使这些想法很简单或算出的坐标、轨迹不够理想，只要你的想法有启发性，同样可以得到高分。

参考文献

- 【1】 Chan Y T, Ho K C. A simple and efficient estimator for hyperbolic location[J].

IEEE Transactions on Signal Processing, 2002, 42(8):1905-1915.

【2】 Foy W H. Position-Location Solutions by Taylor-Series Estimation[J]. IEEE Transactions on Aerospace & Electronic Systems, 2007, AES-12(2):187-194.