Report of Operational Statistics for SAR Imagery

Gao Sicheng 高斯骋 "Mr Tall"

2019 年 10 月 6 日

目录

1 Introduction of Report

本报告是用于《合成孔径雷达图像的操作统计》这门课的大作业。在两周的课程内，我们学习了基于R语言的数据处理,图像生成，参数估计等相关内容。本次报告将基于Frery教授提供的原始数据，一些通过SAR得到的城市的数据，抽取一部分数据作为样本，去估计样本中的显著目标的个数。

This report is a major assignment for the course– Operational Statistics for SAR Imagery.During the two-week course,We have learned about data processing , image generation , parameter estimation and something related based on R language.This report will be based on the original data provided by Professor Frery.These data are obtained through SAR scanning of the urban.This report will extract part of the data as a sample,and estimate the number of salient targets in the sample.

2   Programming Environment and Document Tools

本报告基于TeX写成，版本号为MiKTeX2.9。TeX是一种便捷的文本排版系统，在图像表格，数学公式等科学领域的排版工作有易用性的优势。MiKTeX是用于Windows系统的TeX工具。

This report is written on TeX. The version is MiKTeX 2.9. TeX is a convenient text typesetting system. It has the advantage of ease of use in the typesetting of image tables, mathematical formulas and other scientific fields.MiKTeX is a TeX tool for Windows systems.

本报告的程序基于R语言的环境编写，版本号是R x64 3.6.1。R语言是用于统计分析、绘图的语言和操作环境。R语言是一个自由、免费、源代码开放的软件，它是一个用于统计计算和统计制图的优秀工具。

The codes for this report is written in an environment based on R language. The version is R x64 3.6.1. R language is a language and operating environment for statistical analysis and drawing. R language is a free and open source software. It is an excellent tool for statistical calculation and image generation.

3　Overview for Sample

　　本项目的数据来源于https://github.com/acfrery/Statistics-SAR-Intensity.git。该网址由Frery教授提供。原始数据为一个1599×4000的矩阵，大小为97.6MB。原始数据的数据量过于庞大，不便于操作。本项目作为课程练习，旨在掌握数据处理的方法，所以本项目将抽取第500行至第699行与第500列至第699列的数据，共计20000个数据点，作为样本，进行处理与分析。这样的样本减小了计算机的操作量与操作时间，但是对于整体的操作流程并没有任何精简，任然保持完整的数据处理流程。

The data for this project comes from the web site: https://github.com /acfrery/Statistics-SAR-Intensity.git. The website is provided by Professor Frery. The raw data is a matrix of 1599 x 4000, its size is 97.6MB.The amount of raw data is too large to operate easily. As a course exercise, this project aims at mastering the method of data processing. So this project will extract data from Line 500 to 699 and from Column 500 to 699. A total of 20,000 data points are processed and analyzed as a sample. Such a sample reduces the amount and time of operation of the computer, but there is no simplification of the overall operation process, still maintaining a complete data processing process.

　　利用R语言的代码，将20000个样本点输出为PNG图片，即图1。所使用的核心代码在源文件imagematrix.R中。

Using R language codes, 20,000 sample points are output as a PNG picture, as shown in Figure 1. The codes used are in the source file imagematrix.R.
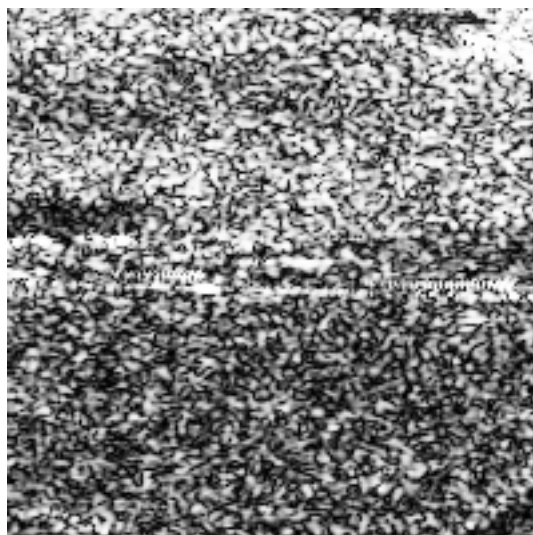
　　利用R语言，亦可得到样本的基础统计数据，其中：

图 1: 样本 the Sample

Basic statistical data of samples can also be obtained by using R language, as shown:

Min. : 1
1st Qu.: 7778
Median : 20732
Mean : 47775
3rd Qu.: 48570
Max. :6395295

利用R语言的代码，得到样本的直方图分布，输出为PDF图片，即图2。所使用的核心代码在源文件1006.R中的14行至22行。

Using the code of R language, we get the histogram of the sample and output them as a PDF picture, that is, Figure 2. The codes
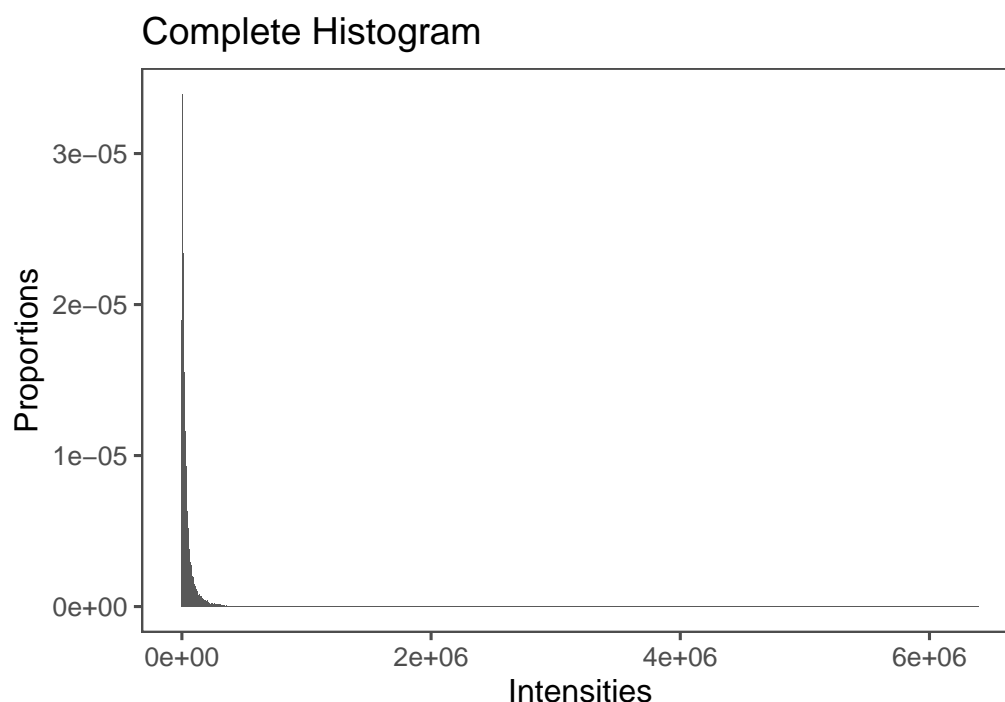
Complete Histogram

图 2: 样本直方图 Histogram of the Sample

used range from Line 14 to 22 in the source file 1006.R.

  由图2可以看出，当横轴大于150000后，数据量就显著减少。而最大值6395295相比于其他数据过分庞大，导致直方图并不直观。我们取横轴上限为200000，舍弃超限的数据，再次绘制直方图，可以让样本的数据分布情况更加直观，具体情况如图3所示。所使用的核心代码在源文件1006.R中的24行至33行。

As can be seen from Figure 2, when the horizontal axis is more than 150000, the amount of data decreases significantly. Compared with other data, the maximum value 6395295 is too large, which
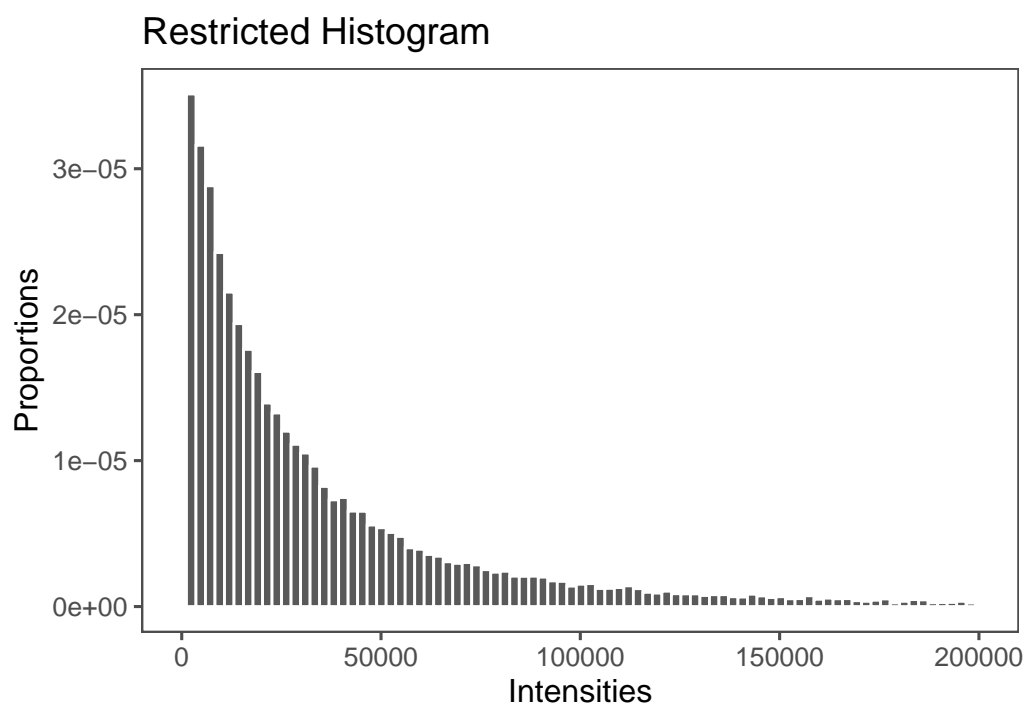
图 3: 有上限的样本直方图 Histogram of the Restricted Sample

makes the histogram not intuitive. We take the horizontal axis limit of 200,000, abandon the data beyond the limit, and draw the histogram again, which can make the data distribution of the sample more intuitive. It is shown as Figure 3. The codes used range from Line 24 to 33 in the source file 1006.R.

4   Data Analysis

至此，我们已经得到了样本的基本情况。根据直方图的分布与经验上的判定，我们用GI0分布去拟合这次的样本。对于一个GI0的随机分布而言，主要参数为 α ， γ 与L。对于实际问题中的参数L通常为已知量，这里取L值为1。而 α 与 γ 为未知量，需要我们去估计。在本课程中，我们使用两种不同的参数估计方法，分别为矩估计法与最大似然估计法。

So far, we have got the basic information of the sample. According to histogram distribution and empirical judgment, we use GI0 distribution to fit this sample. For a random distribution of GI0, the main parameters are alpha, gamma and L. The parameter L in practical problems is usually a known constant, and this time the value of L is 1.But alpha and gamma are unknown, which need to be estimated. In this course, we use two different parameter estimation methods, namely moment estimation and maximum likelihood estimation.

我们将先使用矩估计法。通过编写R语言代码，使用样本的一阶矩与二阶矩的数据，去估算两个未知参数。所使用的核心代码在源文件1006.R中的34行至45行。至此，我们得到矩估计法的结果， α 的估计量为-2.237618， γ 的估计量为106903。

We will use the moment estimation first. By compiling R language codes, two unknown parameters are estimated using the data of the first and second moments of the sample. The codes used range from Line 34 to 45 in the source file 1006.R. So far, we have obtained the result of moment estimation . The estimation of alpha is -2.237618. The estimate of gamma is 106903.

其次，使用最大似然估计法。通过编写R语言代码，构造似然函数，去估算两个未知参数。所使用的核心代码在源文件1006.R中的46行至64行。至此，我们得到最大似然估计法的结果，α 的估计量为-3.527616，γ 的估计量为106922.134905。

Secondly, the maximum likelihood estimation is used. By compiling R language codes and constructing likelihood function, two unknown parameters are estimated. The codes used range from Line 46 to 64 in the source file 1006.R. So far, we have obtained the result of maximum likelihood estimation . The estimation of alpha is -3.527616. The estimate of gamma is 106922.134905.

5   Result Discussion

　　由上一章的数据分析中，我们可以得到两组不同的估计量。在矩估计中，α 的估计量为-2.237618，这意味着我们认为在样本中有2个显著的目标。在城市的数据样本中，我们通常认为这两个目标为2栋较为高大的建筑。而在最大似然估计中，α 的估计量为-3.527616，这意味着我们认为在样本中有3个显著的目标，即有3栋较为高大的建筑。这两种方法各有优劣，矩估计的计算难度较小，而最大似然估计的结果更为可靠。如果我们想进一步探索样本中的目标数，还需要更多的手段进行下一步处理。这里就得出初步的结论，即显著的目标数为2或者3。

From the data analysis in the previous chapter, we can get two sets of different estimators. In moment estimation, the estimate of alpha is -2.237618, which means that we think there are two significant targets in the sample. In urban data samples, we usually consider these two targets to be 2 taller buildings. However, in the maximum likelihood estimation, the estimate of alpha is -3.527616, which means that we think there are three significant targets in the sample, namely, 3 taller buildings. Both methods have their advantages and disadvantages. The calculation of moment estimation is less difficult, and the result of maximum likelihood estimation is more reliable. If we want to further explore the target number in the sample, we need more means for further processing. Here we make a preliminary conclusion that the significant target number is 2 or 3.

6   Summary and Evaluation

到现在，我们已经完成了教授授予的课程，也完成了所布置的作业。对我一个普通的学生而言，这门课是与众不同的。虽然只有短短的四次课程，历时两周的时间，我觉得我受益匪浅。首先是收获了知识。不论是关于TeX还是R语言，外语交流还是统计学知识，我都获得了很大的收获。除过知识外的收获，我和结识了新的老师与同学，得到了新的友谊。这相比于知识而言，是更大的收获。非常感谢西电有这样的课程!

So far, we have completed the course given by the professor and the assignments. For me, an ordinary student of mine, this course is out of the ordinary. Although only four courses lasted for two weeks, I think I benefited a lot. First of all, I got the knowledge. Whether it's about TeX or R, foreign language communication or statistical knowledge, I have gained a lot. In addition to knowledge, I made new acquaintances with teachers and classmates, and gained new friendship. This is a greater gain than knowledge. Thank Xidian University very much for this course!

至于说对该课程的评价，我觉得首先，我们的Frery教授是一个好老师，认真负责，而且非常友善。课程难度设置上也符合同学们的水平，即有一定挑战性，但都可以完成任务。同学也非常友好，乐于助人。我希望这样的课程可以多多益善。

As for the evaluation of the course, I think first of all, Professor Frery is a good teacher, conscientious and responsible, and very friendly. The difficulty of the course is also in line with the level of students, that is, it is challenging, but it can be completed The students are also very friendly and helpful. I hope we can have more courses like this.

7   Thanks

感谢Frery 教授！你带给了我们一堂完美的课程。

Thank you Professor Frery! You brought us a perfect course.

感谢杨晨同学！你在语言与知识上，给予了我很大帮助。

Thank you, Yang Chen! You have helped me a lot in language and knowledge.

感谢李天瑞同学！你在资料上给予了我支持。

Thank you, Li Tianrui. You have given me support in the information.

感谢昵称为"猫画桥"的同学，虽然你从来没告诉过我你的名字，但是你是一个很可爱的女生，感谢你的帮助。

Thank you to the students nicknamed "Cat Painting Bridge". Although you never told me your name, you are a very lovely girl. Thank you for your help.