时间序列数据的符号化表示及可视展现

已知时间数据集（来源及格式说明参考文档“BriefingDocument2018.pdf（Title: UCR Time Series Classification Archive）”），给出其波形图。

要求：**<1>将其标准化处理（均值为零，标准差为1），并根据所给出的w和a值，对标准化处理之后的时间序列进行符号化表示，需要注意的是，对a的划分不再采用简单的a段均分，而是根据下表中a的取值所划分的breakpoint进行划分（高斯分布）。**

**eg:若a=3，则-0.43和0.43为两个breakpoint，将字符序列划分为三个取值区间，按取值由小到大分别对应“a”、”b”、”c”三个字符。（见[1]Table 3）。**

**<2>对标准化处理后的时间序列进行可视化展示。（示例如[1]Fig 5）**

[3]对选定数据集的不同分类在同一显示界面或多个窗体中分别进行可视化展现，以便于对比其形状与分类的关系。

[4]可以选择对于任意数据集或者某个固定数据集（如ECG数据集）实现本实验。

参考文献及资料：

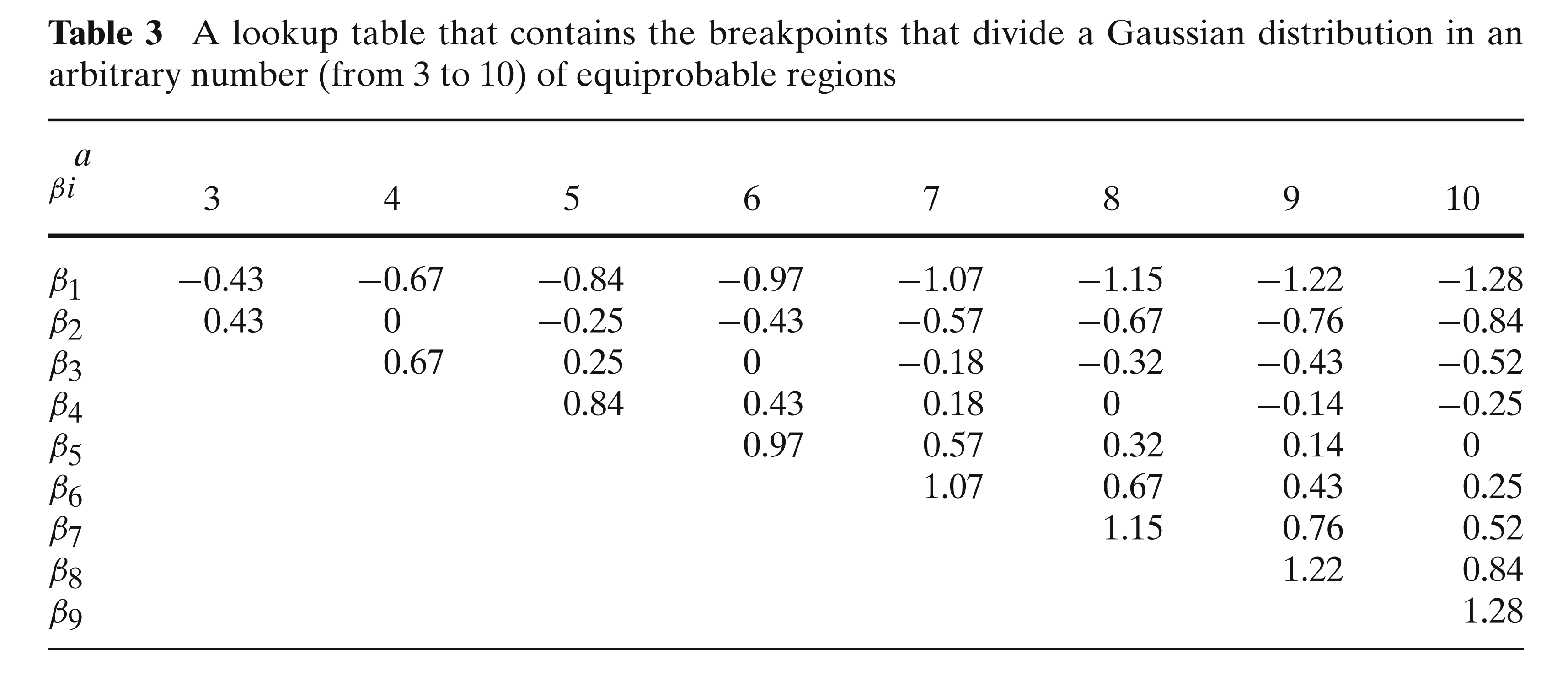
[1]Experiencing SAX\_ a novel symbolic representation of time series.pdf(第3部分3.1及3.2)

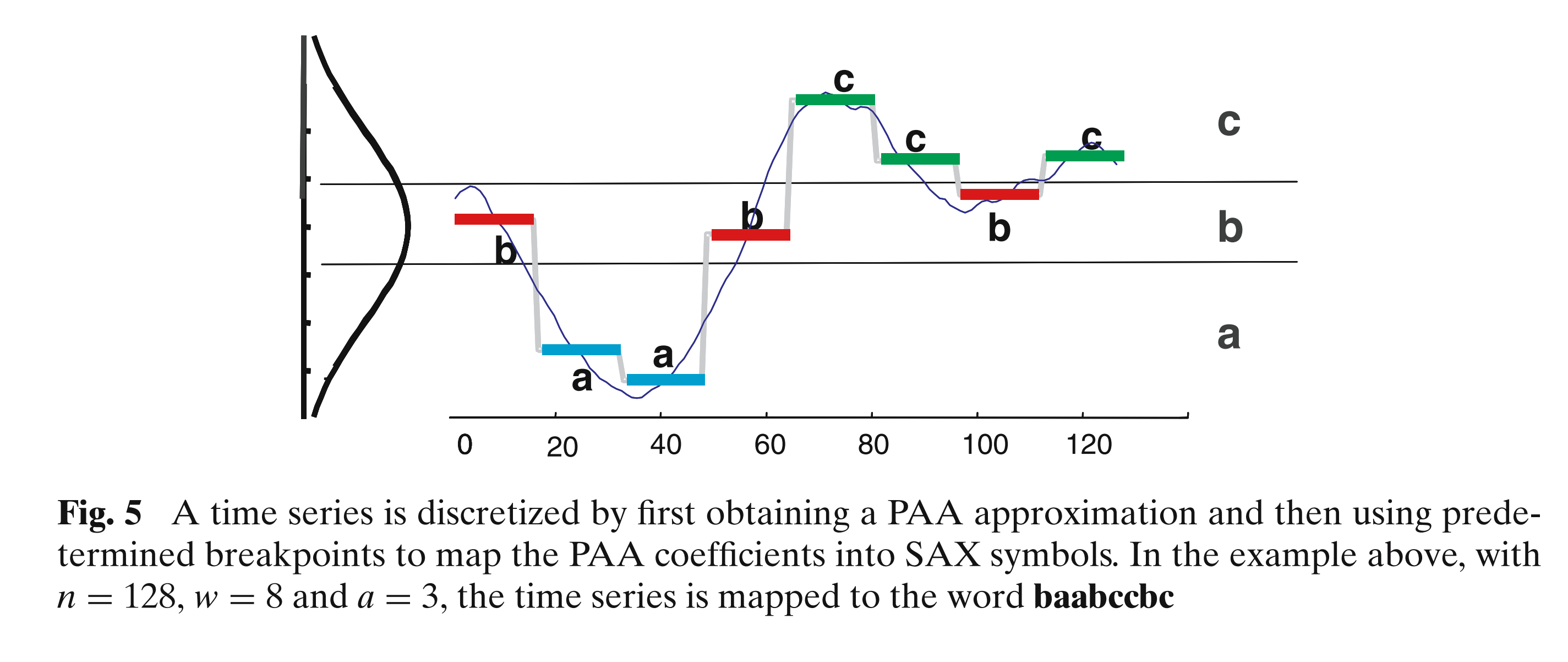
[2] BriefingDocument2018.pdf (对数据集内容、格式、密码等进行简要介绍)

[3] The UCR Time Series Archive.pdf （对数据集的发展、扩充及相关实验应用进行介绍 密码）

[4] DataSummary.csv (Class列为该数据集的分类数)

[5] UCRArchive\_2018\_Beta.zip





|  |
| --- |
| 提交内容： |
| 1）源程序、可执行程序、实验报告电子版 |
| 2）实验报告纸质版 |
| Deadline：2018/12/04 10: 00am |

**建议考虑申请**软件著作权**，可以向老师咨询并寻求帮助。**