



数据预处理

从RefTek130数采到SAC

作者：汪晟

版本：0.01

June 10, 2015

Contents

1	Ref格式数据的转换	2
2	SAC格式头段信息简介	3
3	补全SAC文件头段台站信息变量	4
4	补全SAC文件头段事件信息变量	5
5	时间序列预处理	6

1 Ref格式数据的转换

野外操作中，自130数采拷贝的原始数据是Ref格式，其单文件包含了多分量的数字记录以及部分台站与时间信息，可以使用PQL查看。此外，RefTek130数采亦生成log文件，实时记录了台站的状态信息，包括有：温度、GPS卫星数、电压等，可以使用newlogpeek.py查看。这些不同功能的软件均由PASSCAL () 提供。

根据需求，可以将Ref格式文件转换成其他格式的文件，相关程序均由PASSCAL提供。本文档依托项目为腾冲火山地区接受函数研究，需要数据转换成SAC格式。由Ref到SAC，需要借助中间文件格式，一般有两种方法较为常用：

Ref → miniSeed → SAC

Ref → Segy → SAC

本文档使用第二种方法，其命令行操作如下：

```
bash> ref2segy -f in.ref -l cas.file
bash> segy2sac in.ref
```

其中，ref2segy将Ref格式转换为Segy格式。-f参数声明Ref文件名或目录。-l参数声明130数采部分配置信息。一般情况下，原始Ref文件携有全部的数采配置信息，但是BUG时有发生，此时-l参数开启就很有必要。关于-l的详细介绍以及对应的130数采文件请参看ref2segy的manual，在本文档的后续章节中，亦有提及，可供参考。

转换得到的SAC格式文件头信息缺失，这里引出下一节。

2 SAC格式头段信息简介

除了时间序列记录之外，地震数据中的台站、事件等信息十分重要。本节承接上一节文件格式转换步骤，将简介SAC格式文件中的头段信息变量。

SAC头段信息变量包含了所有的数据参数信息，罗列接分类如下：

1. 基本变量 nvhdr nzyear nzjday nzhour nzmin nzsec nzmsec iztype iftype idep
2. 数据相关变量 npts delta odelta b e leven depmin depmax depmen scale xminimum xmaximum yminimum ymaximum nxsize nysize ical isynth
3. 事件相关变量 kevnv evla evlo evel evdp ievreg ievtyp mag imagsrc imagtyp gcarc dist az bz o ko khole
4. 台站相关变量 kstnm knetwk istreg stla stlo stel stdp cmpaz cmpinc kcmpnm kstcmp lpmpol
5. 震相相关变量 a ka f kf tn kn
6. 仪器相关变量 kinst iinst respn
7. 其他变量 usern kusern lovrok lcalda kdatrd

以上罗列中，×××变量为转换生产的SAC文件携有的有效头段信息。亦存在其他变量有值，但非有效，所以需要补全。

第一步需要补全的是台站相关变量。对任一数据记录，台站信息都是存在的且必须的。并非只有事件相关数据才可以用于研究，噪声数据的价值亦不可估量。

第二步需要补全的是事件相关变量。此操作只针对那些包含了时间记录的文件，所以需要先抽取出这些文件，再补全头信息。

第三步操作，需要补全部分震相等其他信息。

需要注意的是SAC头段中的时间类变量。由Ref转换得到的SAC文件携有的有效信息中有“参考时间”、“起始相对时间”、“结束相对时间”、“采样间隔”、“采样点数”。显然，这些时间信息倘若缺失，那么整个预处理都将失去意义。转换得到的SAC文件“参考时间”为数据首个采样点的绝对时间（伦敦时间），那么“起始相对时间”自然为0。在更为靠后的操作中，需要写入事件信息，并截断有效长度，此时“参考时间”会发生改变，相对类时间将会由SAC软件自动计算变化。总之，时间类变量的写入与更改，应当细心。

此外，SAC头段信息变量冗余，一些变量可由其他变量推导计算获得，这一特性减轻了补充头段信息变量的工作。

3 补全SAC文件头段台站信息变量

上一节提及了对所有的文件记录，台站信息都是存在且必需的。本节将描述补全这一类信息。

SAC程序提供了更改任意头段信息变量的命令，如下：

```
SAC> r eg1.sacS
SAC> ch ARGV1 VAL1 ARGV2 VAL2 ...
SAC> w over
```

第一行读SAC文件到内存中，第二行更改了内存中数据的头段变量值，第三行将内存中数据覆盖写入文件。本文档所有的SAC头段变量信息的赋值与更改均为以上操作。

本节中赋值更改的ARG范围列表如下： 其中，通道名与通道角度之间的关系为：

Table 1: 台站信息列表

变量名	解释	是否需要赋值
kstnm	台站名	是
knetwk	台网名	是
istreg	台站地理区域	
stla	台纬度	是
stlo	台经度	是
stel	台高度	是
stdp	台深度	是
cmpaz	通道方位角	是
cmpinc	通道倾角	是
kmpnm	通道名	是
kstcmp	辅助名	
lpspol	通道极性正负	是

Table 2: 通道信息表

通道名	cmpaz	cmpinc
N	0	90
E	90	90
U	0	0

显然NEU为左手坐标系，lpspol变量为TRUE表明通道的正极性与NEU方向相同，为FALSE则相反。

4 补全SAC文件头段事件信息变量

补全所有记录文件的台站信息之后，可挑选出包含有时间记录的文件。此时，当务之急是如何挑选出这类文件，并为这些文件补全事件类信息。

一种简单且可批量处理的方法是，给出台站与时间信息列表，组合成不同的“事件-台站”对，计算理论P波初至震相到时。由这些“事件-台站-到时”组合，挑选出包含了P波初至震相的记录文件。更为复杂的情况是，需要提取出P波初至震相前后一段时长内的记录。那么这些时间记录片可能会跨不同文件，甚至缺失某些时段的记录文件。对此，可以事先给出P波初至震相前后时间偏移长度，计算出有效时间片的“起始绝对时间”和“终止绝对时间”，随后提取出所有与此时间区间存在交集的记录文件，合并即可。此步得到的“起始绝对时间”和“终止绝对时间”亦将用到后续的数据截断中。

获取了包含有事件记录的文件后，需要补全对应事件信息，以及理论计算的P波初至震相到时。变量列表如下：

Table 3: 事件信息列表

变量名	解释	是否需要赋值
kevnrm	事件	是
evla	事件纬度	是
evlo	事件经度	是
evel	事件高程	是
evdp	事件深度	是
ievreg	事件地理区域	
ievtyp	事件类型	选
mag	震级	是
imagsrc	震级信息来源	是
imagtyp	震级类型	是
gcarc, dist	大圆路径长	当lcalda为TRUE时，将自动计算
az,bz	前后方位角	当lcalda为TRUE时，将自动计算
t1	理论P波初至相对时间	是

5 时间序列预处理

补全SAC文件头段信息之后，接下来的预处理操作将针对时间序列展开。与头段信息不同的是，时间序列的更改不可逆，且不同的处理手段及先后顺序将直接影响到后续的数据研究分析。随意随性的操作将会带来假象，所以先于预处理，应有数据备份。

时间序列的预处理包括有：1.去毛刺、2去仪器相应、3.时间截断、4.坐标旋转、5.去均值与线性趋势、6.尖灭、7.滤波。简介：

以上操作中，“去仪器相应”、“去均值与线性趋势”、“尖灭”三项均从根本上改变了采样点的数值，其执行需要严格遵守一定的先后顺序。在确定先后顺序中，需要深刻理解时间-频谱关系及仪器工作原理，否则将产生不可预知的假象和错误。一般预处理中，按照“去均值与线性趋势”→“尖灭”→“去仪器相应”的先后顺序进行。而“时间截断”和“坐标旋转”可以在后续操作中任意进行。“滤波”则针对研究目的的需要而进行。将预处理完的数据应用到研究中，如果涉及到频谱操作或（反）卷积操作，则需要再做一次“去均值与线性趋势”与“尖灭”。

当然，以上的步骤与顺序介绍并非放置四海而皆准。微小的预处理差异会带来截然不同的结果，所以需要备份数据，反复操作对比，并针对研究目的确定最合适的预处理流程和参数选择。

SAC软件预处理命令范例如下：

```
SAC> r eg.sac           #读入SAC文件
SAC> rmean              #去均值
SAC> rtr               #去线性趋势
SAC> taper ***          #尖灭
SAC> transfer           #去仪器响应
SAC> rotate to gcp      #旋转至大圆路径，即RTZ分量
SAC> cut t1 -50 200     #截取t1时间前50s到后200s
SAC> bandpass ***       #带通滤波
SAC> lowpass ***        #低通滤波
SAC> highpass ***       #高通滤波
SAC> bandrej ***        #带阻滤波
```