

R 语言程序包 HalfsibBV 使用说明

王德源 童春发

2015/05/10

HalfsibBV 是以 R 3.1.1 软件为平台编写的 R 语言程序包，包括 Windows 和 Linux 两个版本，分别为 zip 和 tar.gz 格式，可在个人电脑 Windows XP/Vista/7/8 及 Linux 操作系统上使用，要求内存 1G 以上，主要针对以下两个遗传统计分析模型。

多地点模型：设有 f 个家系，布置在 s 个地点进行试验，每个地点设置 b 个区组，每小区若干个单株，可得单株育种值计算分析的线性模型为：

$$y_{ijkl} = \mu + S_i + SB_{ij} + A_{ijkl} + SF_{ik} + SBF_{ijk} + e_{ijkl}$$

其中， A_{ijkl} 即为第 i 个地点下第 j 个区组里第 k 个家系内第 l 个单株的育种值。 y_{ijkl} 为第 i 个地点第 j 区组中第 k 个家系的第 l 个个体的数量性状观测值， μ 为总体平均值， S_i 为第 i 个地点的效应 ($i=1,2,\dots,s$)， SB_{ij} 为第 i 个地点下第 j 个区组的效应 ($j=1,2,\dots,b$)， SF_{ik} 为第 i 个地点与第 k 个家系的交互效应， SBF_{ijk} 为第 i 个地点内第 j 个区组与第 k 个家系的交互效应， e_{ijkl} 为随机误差效应 ($l=1,2,\dots,n_{ijk}$)，一般假定 $e_{ijkl} \sim N(0, \sigma_e^2)$ 。用矩阵形式表示为：

$$\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \mathbf{A} + \mathbf{U}_2\mathbf{SF} + \mathbf{U}_3\mathbf{SBF} + \mathbf{e}$$

其中， \mathbf{y} 为观测值向量， $\boldsymbol{\beta} = (\mu, S_1, \dots, S_s, SB_{11}, SB_{12}, \dots, SB_{sb})'$ 为固定效应向量， $\mathbf{A} = (A_1, \dots, A_n)'$ 、 $\mathbf{SF} = (SF_{11}, \dots, SF_{sk})'$ 、 $\mathbf{SBF} = (SBF_{111}, \dots, SBF_{sbk})'$ 为随机效应向量， \mathbf{e} 为随机误差向量，且

$A \sim N(0, G\sigma_A^2)$ (G 为子代间亲缘关系矩阵)、 $SF \sim N(0, I_{sf}\sigma_{SF}^2)$ 、 $SBF \sim N(0, I_{sbf}\sigma_{SBF}^2)$ 、 $e \sim N(0, I_n\sigma_e^2)$ ， X 、 U_2 、 U_3 分别为各效应的系数矩阵。

单地点模型：设有 f 个家系，在某个地点进行试验，设置 b 个区组，每小区若干个单株，则单株育种值计算分析的线性模型为：

$$y_{ijk} = \mu + B_i + A_{ijk} + BF_{ij} + e_{ijk}$$

用矩阵形式表示为：

$$y = X\beta + A + U_2BF + e$$

其中 A_{ijk} 即为第 i 个区组里第 j 个家系内第 k 个单株的育种值，其他各符号的含义与多地点模型类似。

针对以上两种半同胞子代测定遗传模型，HalfsibBV 函数根据读入的数据文件类型选择相应的计算分析程序。首先应用传统的方差分析法计算得到各遗传方差分量的估计值；然后以这个估计值作为初始值，用 AI 算法进行迭代计算获得各方差分量的限制性极大似然估计 (REML)，进而计算遗传力及其标准误；最后采用混合模型方程组方法对育种值进行最佳线性无偏预测，获得家系及子代个体育种值的 BLUP 估计及其标准误。对于多个数量性状，计算了任意两个性状在各效应水平上的遗传相关系数及其标准误。

HalfsibBV 程序包使用的数据文件为 csv 格式，两种模型的输入数据文件格式可分别参考示例数据文件“example_for_mutil_site.csv”、和“example_for_single_site.csv”。在 Windows 版本的 R 语言软件上使用的程序包为 zip 格式，其安装和加载有鼠标操作和命令执行两种

实现方式，首先点击 R 语言界面上的“Packages”菜单项，在其下拉选项中选择“Install package(s) from local zip files”，在弹出窗口中找到并选中程序包压缩文件 `HalfSibBV.zip`，点击“打开”后便可自动安装到 R 软件上，或者在 R 语言命令窗口输入“`install.packages("程序包文件的完整路径",repos=NULL)`”按“enter”键进行安装，如 `install.packages("E:/package/HalfSibBV_1.0.zip",repos=NULL)`。然后再点击 R 语言“Packages”菜单项下的“Load package”选项，在弹出的窗口中选择程序包 `HalfSibBV`，点击“OK”按钮完成加载，或者在 R 语言命令窗口输入“`library(HalfSibBV)`”按“enter”键进行加载。在 Linux 版本的 R 语言软件上，使用的程序包为 `tar.gz` 格式，需要在终端进入 R 后使用命令形式安装和加载程序包。

程序包只需要安装一次，但是每次启动 R 软件都需要加载才能使用。加载完成后，在 R 语言命令窗口输入“`HalfSibBV("input_file")`”按“enter”键将准备好的数据文件提交并开始分析计算，其中“`input_file`”代表数据文件的名称及其所在的完整文件路径，如 `HalfSibBV("d:/data/example_for_mutl_site.csv")`。计算完成后，结果会显示在一个新的窗口里，同时当前工作目录（可通过命令“`getwd()`”来查看）下生成计算结果文件 `Result.txt`。