

## 10 排序图解读 Visualizing multivariate data

展示排序分析结果最好的工具就是排序图。排序图能够用来近似模拟物种分布数据、样方之间距离的矩阵、物种之间的相异或相似矩阵。在带有环境因子的排序分析中，我们可以也用排序图来近似模拟环境因子数据内容、物种与环境因子之间的关系、环境因子之间的关系等等。下面两节将总结出如何从线性排序和单峰排序两种排序图获得有用的信息。

在讨论如何解读排序图之前，我们必须强调的是：排序对象(样方，物种和环境因子)在排序空间的坐标值 (scores) 一般来讲，只有相对意义，没有绝对意义。在解读排序图的，我们使用的是相对距离，相对方向和投影点的相对次序来解释对象之间的关系。

### 10.1 线性排序图的解读 (What we can infer from ordination diagrams: linear methods)

在线性排序图 (PCA或RDA) 中物种和数量型环境因子 (quantitative environmental variables) 用箭头表示，定性环境因子 (nominal) 和样方用符号 (Symbols) 表示 (质心)。表10-1总结我们能够得到的从线性排序图得到的信息 (Ter Braak 1994)。

在线性排序图中 (如 PCA 排序图)，如果我们从代表每个样方的点 (sample points) 投影到某一物种的箭头，投影点的相对位置可以代表该物种这些样方中多度值排序情况 (见图 10-1)。如果是在 RDA 图中，投影点代表的是拟合的多度值(即能够被排序模型所解释的部分，非观测值)的排序情况。在 PCA 图中，物种箭头的起始点的位置表示物种多度平均值位置，如果样方的投影点在箭头的反向延长线上，则表示该物种在此样方内多度小于平均值；反之，则大于平均值。

如果排序图是以样方的距离为标准 (Focus on sample distance)，在排序图内样方点之间的距离远近 (欧几里得距离) 可以代表样方之间的相异程度 (dissimilarity) (见图 10-2)

在线性排序图内，我们可以通过物种箭头之间的夹角来表示物种之间的相关性 (见图 10-3)。夹角越小，表示相关性越高，如果箭头同向，表示正相关；如果反向，表示负相关；如果夹角接近直角，表示相关性很小。用同样的规则，我们可以解读线性排序图内物种与数量型的环境因子关系 (见图 10-4)。比如，一个环境的因子的箭头方向与某一物种的箭头方向相同，可以预测该物种的多度是随着该环境因子的增加而递增。

在线性排序图内，从样方的点向数量环境因子的箭头做投影，投影点的位置可以近似表示该环境因子数值在这些样方内的排序。由于在进行排序模型拟合的时候，环境因子均已经标准化，所以环境因子箭头的起始点也可以认为是平均值的位点 (见图 10-5)。

数量型环境因子箭头之间的解读，跟物种箭头之间关系的解读是相似的 (见图 10-6)。但需要注意的是，这种关系不及直接拿环境因子作为原始数据进行 PCA 分析来得准确。也就是说，如果是简单考察环境因子之间的关系，可以直接将环境因子进行 PCA 分析。另外，在 RDA 排序图中，数量型的环境因子箭头的长短可以代表环境因子对于物种数据的影响程度 (解释量) 的大小，因为在环境因子进行分析之前，都已经被标准化。并且，约束排序图中环境因子的箭头长短不是来拟合环境因子在样方内值的排序，而是拟合环境因子与物种分布之间的关系。

在 CANOCO 的排序图中，对于哑变量(dummy environmental variables)的解读有很多

方法。由于哑变量在样方内数据是 0-1 数据，而在排序图中是以点来表示，这个点是该哑变量中所有为 1 的样方点的质心点 (Centroid)。正如第一、三章所提及那样，我们可以视每个哑变量为某一因子不同的水平 (等级)。关于哑变量与物种、样方、数量性环境变量之间的排序图关系，图 10-7、图 10-8、图 10-9 的解释基本可以说明清楚

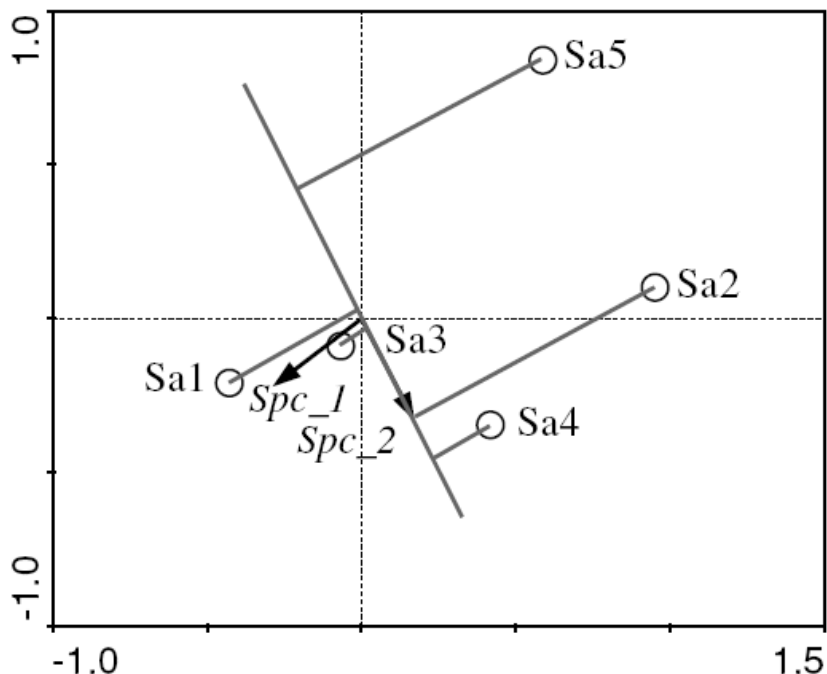


图10-1. 在线性排序图, 从样方的点在物种箭头及其延长线的投影点可以表示他们之间的关系。在图中, 我们可以预测, 物种 *Spc\_2* 的多度最大的样方应该是在样方 Sa4, 其次是 Sa2, 在样方 Sa3 和 Sa1 里多度接近平均值, Sa5 里面多度最小。

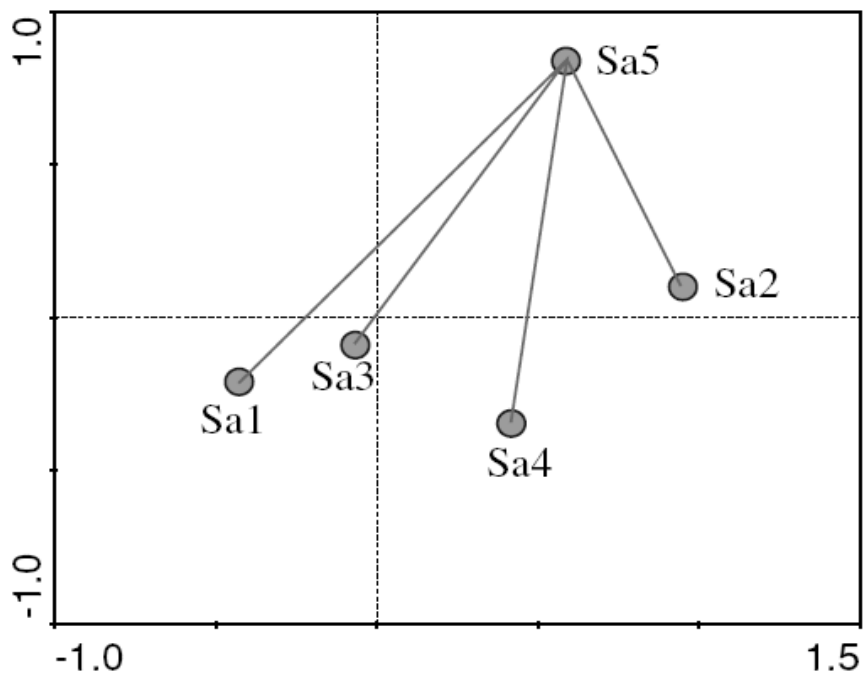


图10-2. 样方点之间的距离可以代表它们之间的关系。如图, 我们可以预测样方 Sa5 与 Sa2 相似度最高, 与 Sa3 和 Sa4 次之, 与 Sa1 相似度最小。

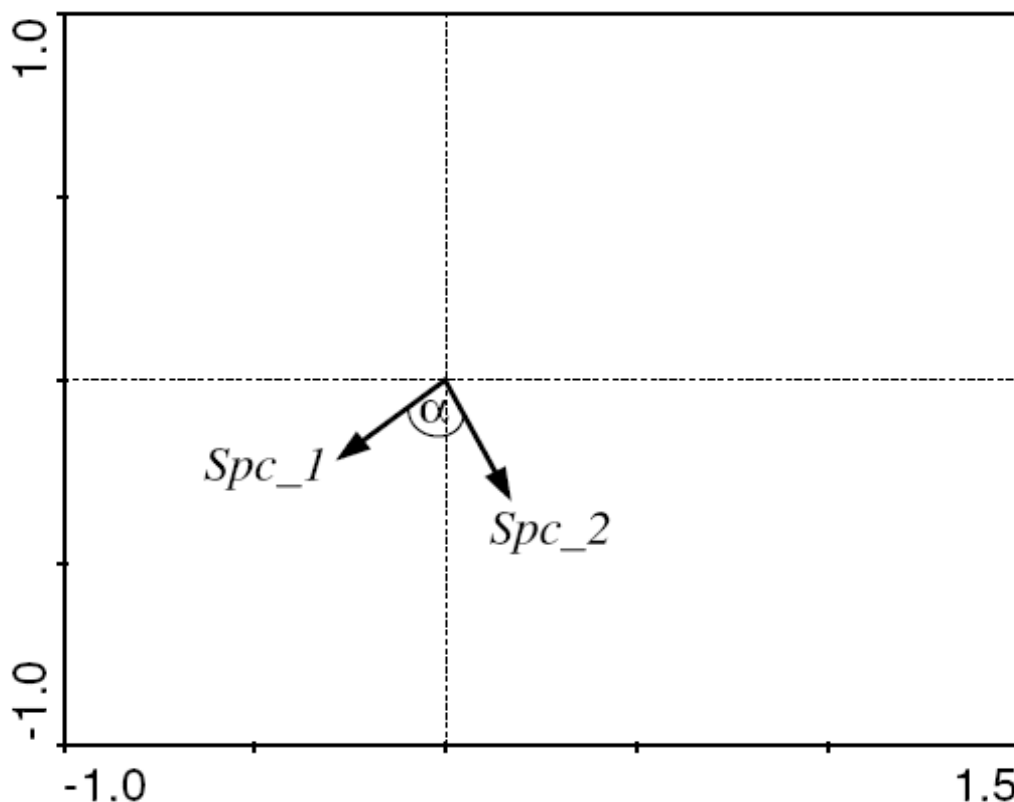


图10-3. 物种箭头夹角（余弦值）可以表示物种之间的相关性。如图中的物种 *Spc\_1* 和 *Spc\_2* 之间的箭头夹角接近90度，其余弦值接近于0，说明这两个物种的相关系数接近0，相关性很低。

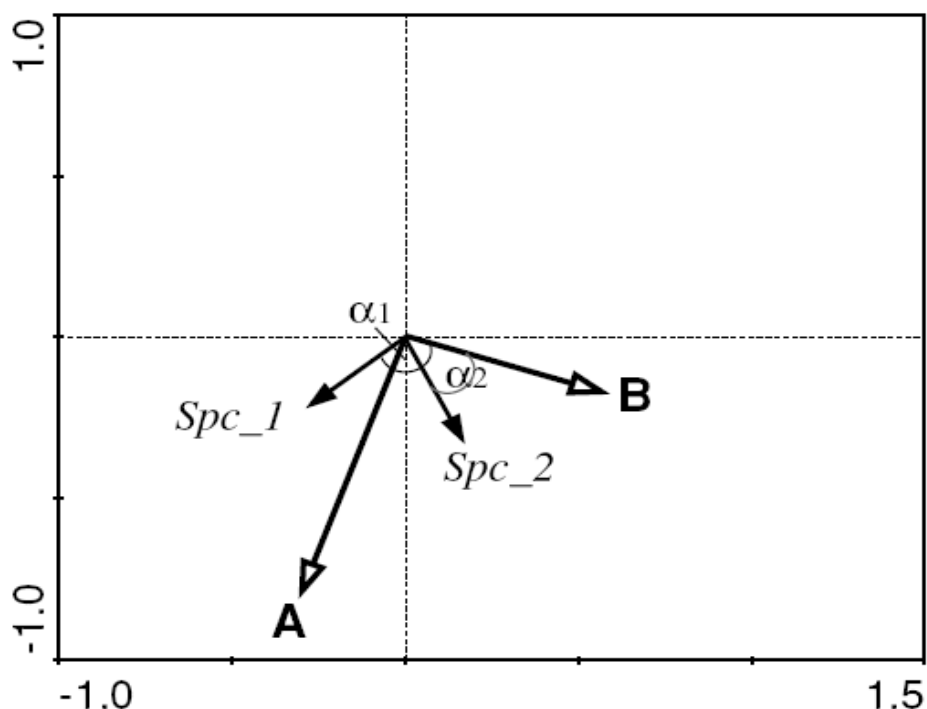


图10-4. 环境因子的箭头与物种因子之间的箭头的夹角也可以表示物种与环境因子之间的相关性。如图中，环境因子B与物种 *Spc\_2* 之间的夹角小于90°，表示之间正相关，即 *Spc\_2* 的多度会随环境因子B值的增加而增加。相反，环境因子B与物种 *Spc\_1* 之间的夹角大于90°，表示环境因子B与物种 *Spc\_1* 之间负相关。

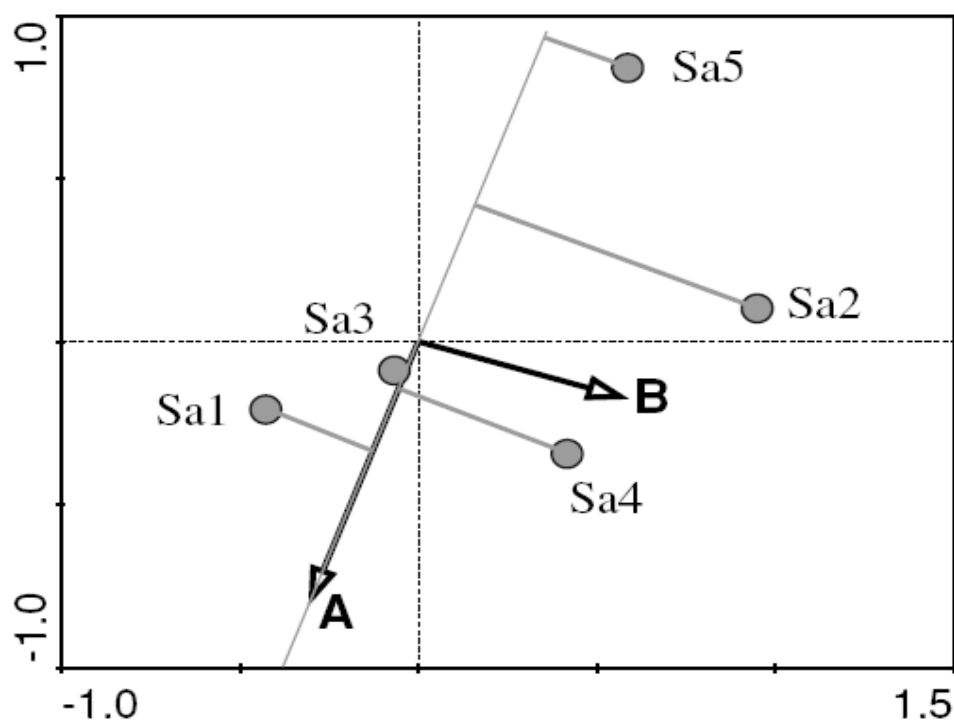


图10-5 从样方点到数量型环境因子的箭头投影点,可以表示该环境因子的值在样方内的情况。如图中,尽管样方Sa4和Sa3的点到环境因子A的箭头距离不一样,但是A在这两个样方内的值应该上差不多,因为它们的投影点很接近。

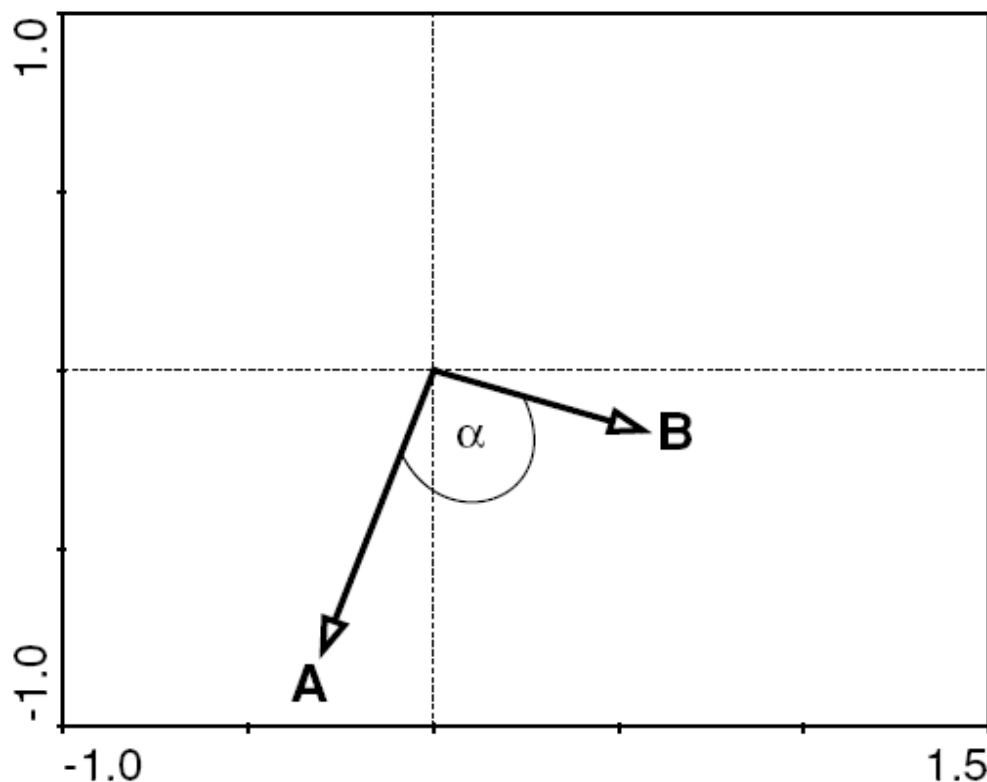


图10-6. 数量型环境因子箭头之间夹角可以表示环境因子之间的相关性。如图中的环境因子A和B之间的夹角接近90度,说明这两个环境因子几乎不相关。

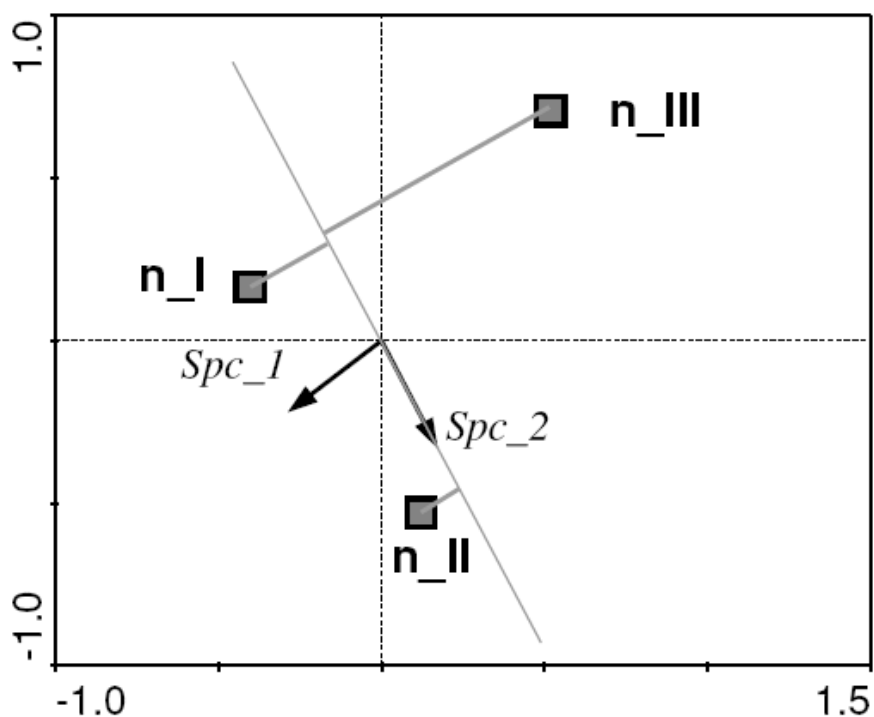


图10-7. 从哑环境变量点到物种的箭头投影点,可以看出该物种与环境变量之间的关系。如图中,我们可以预测物种Spc\_2在属于n\_II级样方内的多度平均值要比在属于n\_I和n\_III的样方内的多度平均值高。

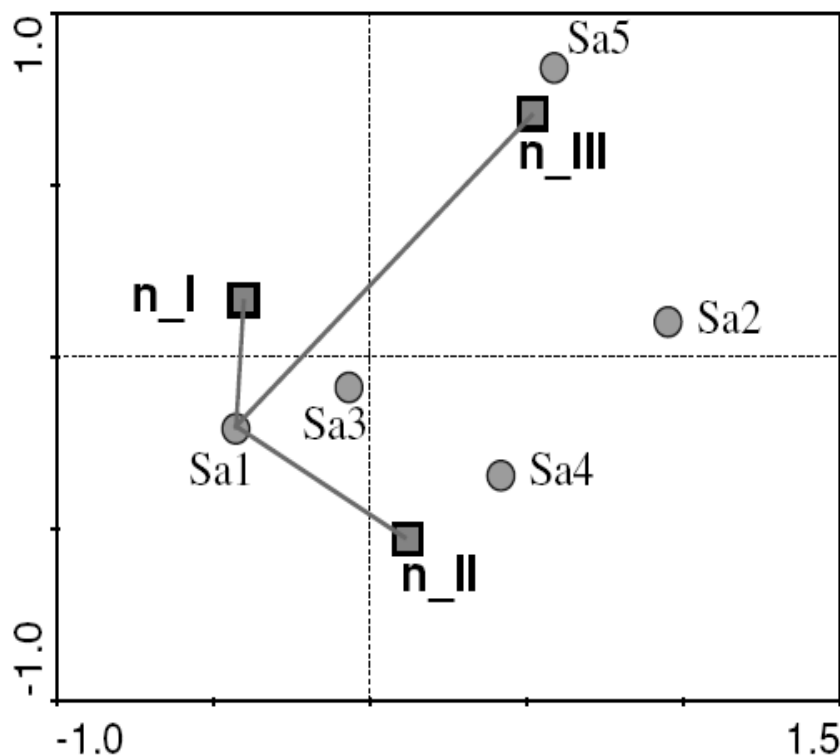


图10-8. 哑环境变量点与样方点之间的距离可以表示它们之间关系。如图中,我们可以预测样方Sa1属于n\_1的级的概率非常高,而属于n\_III的可能性很低。

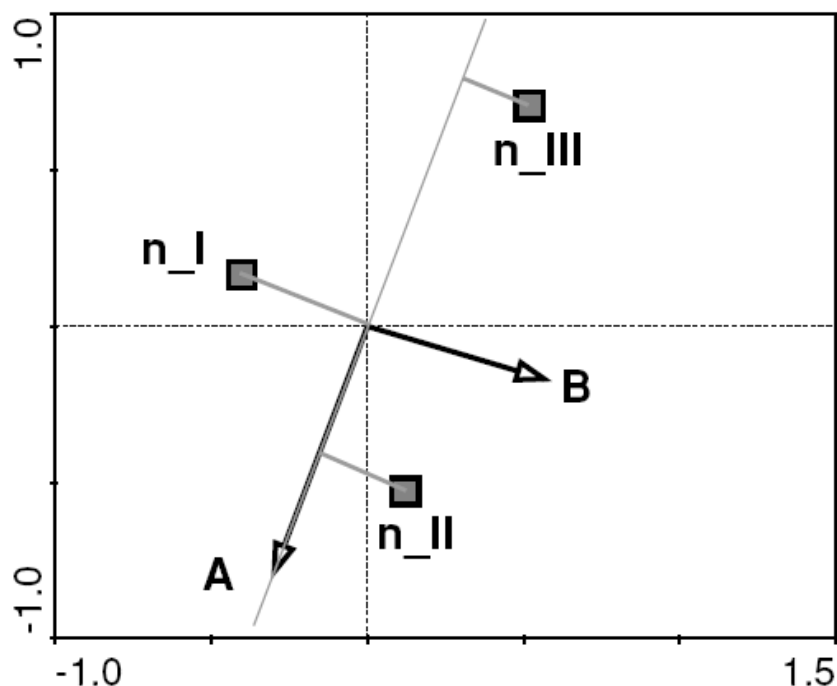


图10-9. 从哑环境变量点投影到数量环境变量的箭头，可以看出该哑变量与环境变量之间的关系。如图中，我们可以预测属于n\_II的样方环境因子A的平均值最高，其次是属于n\_I的样方，n\_III的样方的A的平均值最低。

## 10.2 单峰模型排序图的解读 (What we can infer from ordination diagrams: unimodal methods)

表 10-2 总结我们能够得到的从单峰模型排序图得到的信息 (Ter Braak 1994)。单峰模型排序图和线性排序图的解读有很多相同的地方。但有几点的区别还是比较明显。最主要的不同的是，在线性排序图中，物种是用箭头表示，而在单峰模型排序图中，物种是用点表示。因为，在线性排序图中，我们的前提假设是物种的多度是沿着环境梯度单调变化；而在单峰模型中，物种所在点的位置是该物种在梯度上的最适值对应的位置 (optimum position)，物种点的位置坐标也是通过加权平均算出来的。

还有另外一个重要的不同的是，在单峰排序图中，样方中的相异度 (dissimilarity) 是以卡方距离 (chi-square) 依据，也就是说，如果两个样方内各个物种的相对多度一样，那么这两个样方在图上的位置是一样的 (打个比方，样方 sam1 和 sam2 均有三个物种 a、b、c，这三个物种在 sam1 的数量为 1、2、1；在 sam2 的数量分别为 10、20、10，那么这两个样方在单峰排序图上的位置应该是在一起的)。同样，物种在分布差异也可以用同样的方法算出，如果两个物种在各个样方内多度比例是一致的，那么它们在图上位置应该是在一起的。

物种与样方之间的相对距离可以用来判断物种在样方内相对多度的变化。如图 10-10，物种与样方之间的点越近，代表该物种在此样方内的相对多度越大 (这里的相对多度该物种的样方内数量比上该样方内所有个体数量，而不是物种绝对多度的比例)。同样，样方距离某个物种的距离长短代表该物种在这些样方内相对多度变化，距离越短，相对多度越大，距离越长，相对多度越少。

当然，如果梯度小于 2SD，进行 CCA 分析（或 DCCA），可以用类似与线性排序的规则来解读物种点与样方点的关系。如图 10-11，从排序图的坐标原点到物种点引一个箭头。样方到该箭头的投影点的位置次序可以代表这些样方内该物种绝对多度的排序。

同线性排序一样，单峰排序图内，样方之间的距离也是可以代表样方之间的差异程度。

在单峰排序图内，物种点之间的距离是卡方（Chi-square distance），可以代表不同物种空间分布差异，如图 10-12。

在单峰排序图内，从物种点到数量型环境因子箭头的投影点的位置次序可以代表这些物种在该环境因子最适值（optima）的排序，如图 10-13。

在单峰排序图内，数量型环境因子的关系如同线性排序图一样解读。

在单峰排序图内，哑环境变量点与物种点之间的距离可以表示它们之间关系，如图 10-14。

在单峰排序图内，哑环境变量点之间的关系及与样方点关系解释如同线性排序图解释。

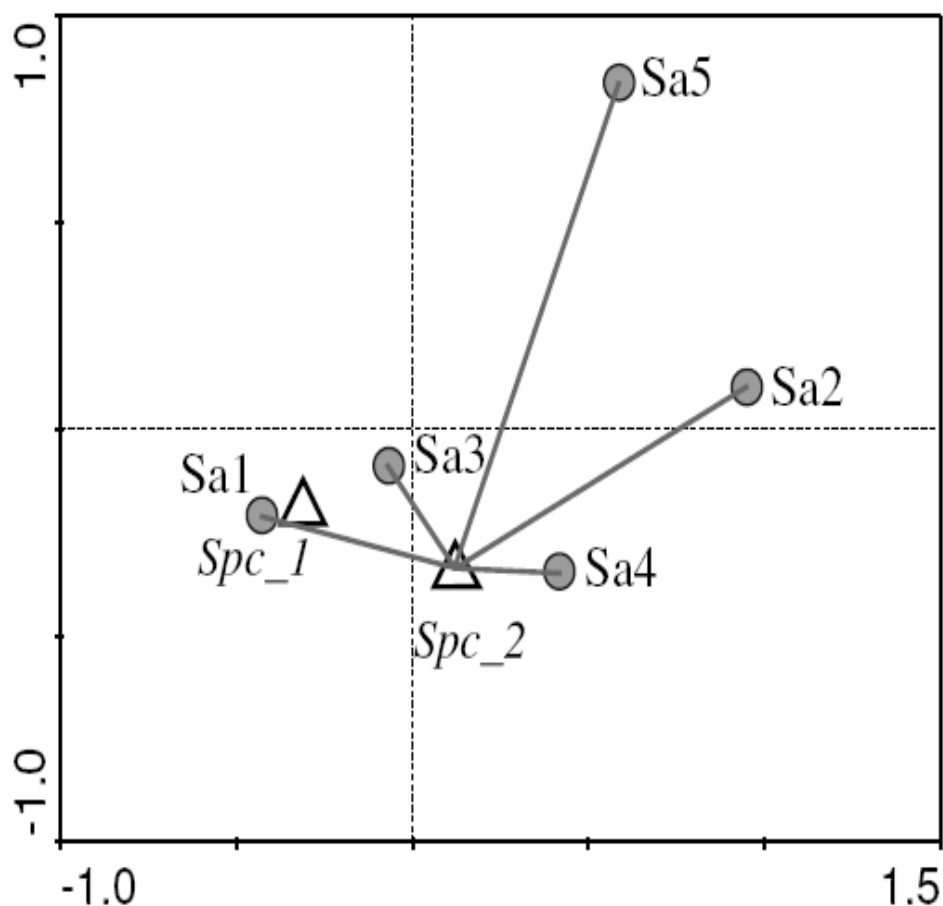


图10-10. 物种点与样方点之间的距离可以表示它们之间的关系。如图中，我们可以预测物种 *Spc\_2* 在样方 Sa4 和 Sa3 内相对多度最高，其次是在 Sa1，在 Sa5 的相对多度最小。需要注意的是，这里的均指相对多度，非绝对多度。

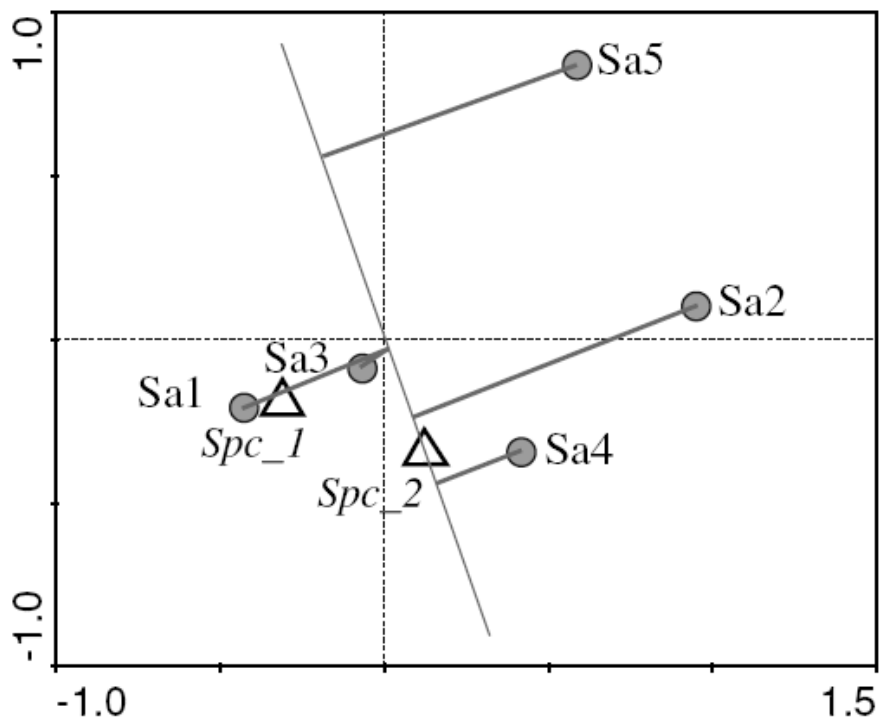


图10-11. 如果梯度小于2SD, 可以线性排序规则解读物种与样方之间的关系, 如图, 物种Spc\_2 被预测在样方Sa4有最大的绝对多度, 其次是在Sa2内, 而在Sa3中的多度接近平均值, 在Sa5内的多度最小。

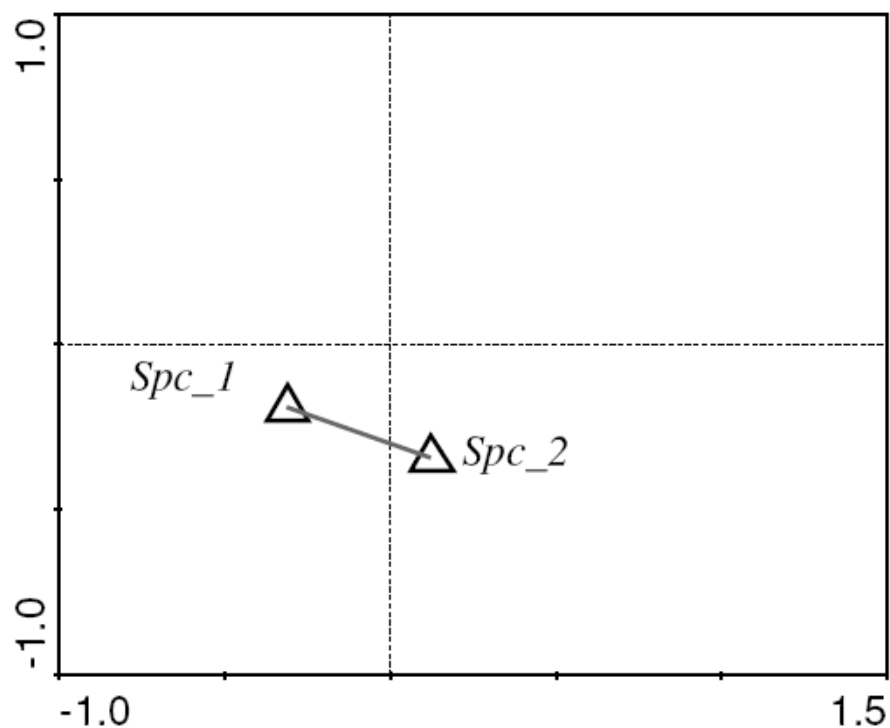


图10-12. 物种点之间的距离可以代表分布差异程度



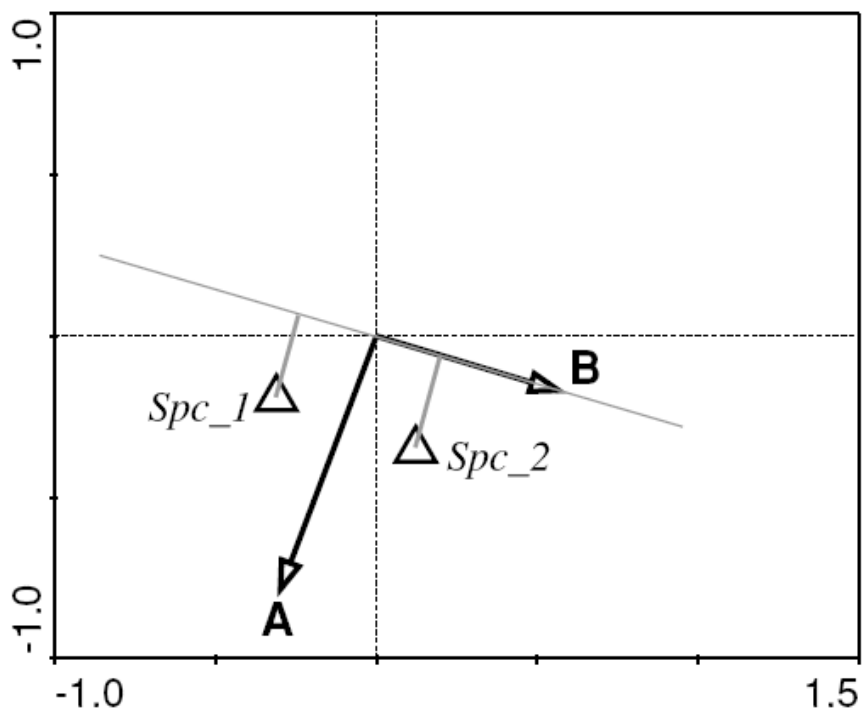


图10-13. 从物种点到数量型环境因子箭头的投影点的位置次序可以代表这些物种在该环境因子最适值 (optima) 的排序。图中在B梯度上, Spc\_2的最适值比Spc\_1高。

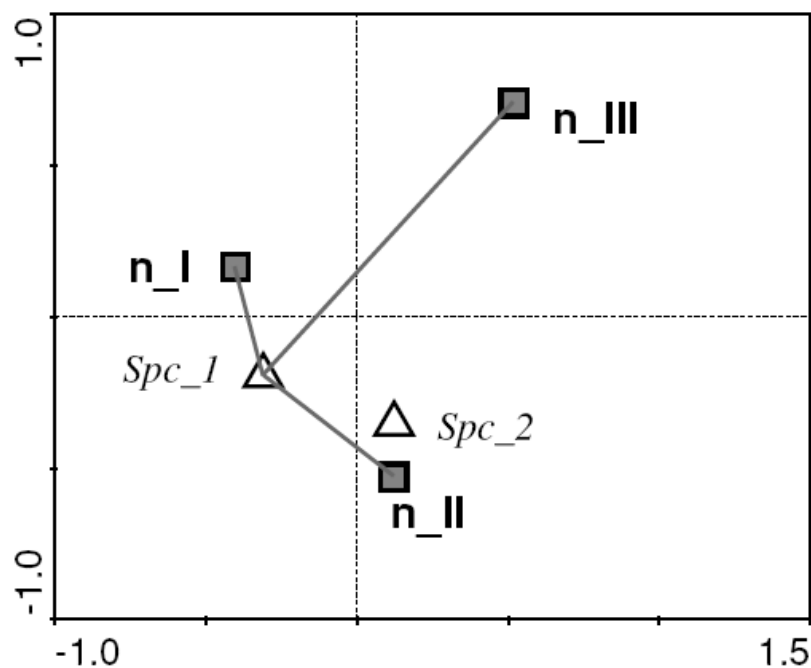


图10-14. 哑环境变量点与物种点之间的距离可以表示它们之间关系。如图,物种 *Spc\_1* 在属于 *n\_I* 级的样方内的平均多度要比属于 *n\_II* 类样方平均多度要大, 在 *n\_III* 级的样方内平均多度最小。

### 10.3 排序图内统计量图的解读 (Visualizing ordination results with statistical models)

Canodraw 不仅可以做出几个基本对象 (物种、样方、环境因子) 的排序结果图, 也可以研究同类变量之间的关系, 以及样方很多指标与环境变量及环境轴之间的关系, 比如样方的物种多度、多样性指数等等与环境因子的关系。这些关系, 有些是建立在线型回归模型的基础上, 这些内容在本书的 8.7 节都已经介绍。

比如, 你能够将一个环境因子的量值在排序空间内展示出来。该环境因子在每个样方内的值可以在样方在排序图内所在位置用半径不同的符号表示出来, 如图 10-15。但是这类图往往不容易看出变量之间关系, 有时可以用回归模型来表示变量之间关系, 具体详见第 8 章。

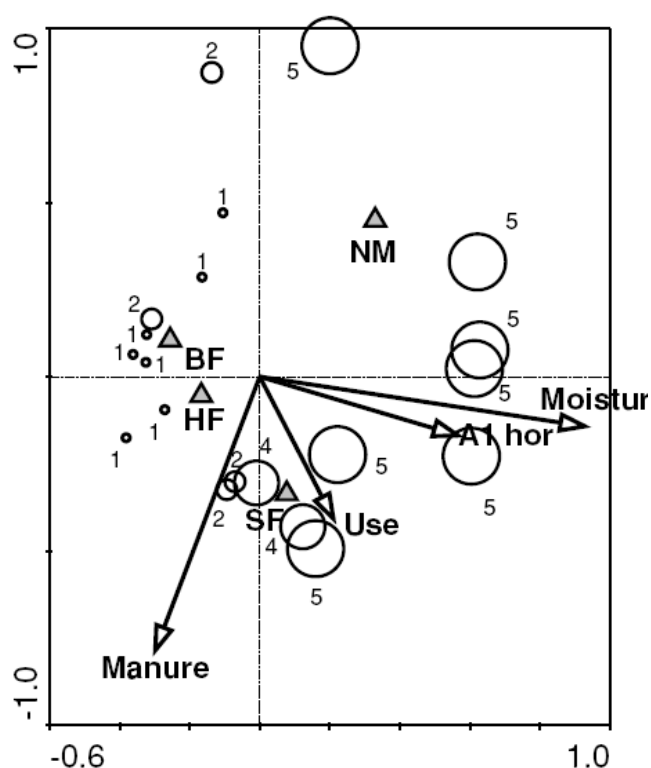


图10-15. 环境因子量值排序空间分布图。图中展示的是湿度 (Moisture) 在排序空间内的量值变化情况。BF, NM, SF和HF均为属性变量, 同事Use, Alhor, Moisture 和Manure 均为半数量变量。

### 10.4 排序诊断 (Ordination diagnostics)

Canodraw 还能画出对象的属性图 (attribute plots), 这些图对我们的研究的对象关系很大的帮助。比如, 要看物种沿着排序轴梯度上的响应曲线, 我们可以用物种在每个样方内拟合值与排序轴的值之间的建立回归关系, 然后画出回归曲线。在 Canodraw 里面, 可以从 Create > Attribute Plots > Species response curves 选项里面做类似的图。在这个选项内, 有广义线性模(GLM), 线性(GAM)可以选择, 具体操作, 可以参考本书第 8.7

节。

在 CANOCO 的使用手册中，有“Ordination diagnostics”这么一个内容，这个是指在排序图中展示单个样方或单个物种能够多少比例被排序轴的百分比例（拟合度）（详见 Ter Braak and Smilauer 2002, Section 6.3.11.2）。这个拟合度在 Canoco 的输出文件内的 Cfit 这个栏目内可以找到。在 CanoDraw 的画图中，经常是用这个拟合度来排除一些低拟合度的物种显示在排序图中。如图 10-16 中，我们可以在排序图中将这些物种的拟合度用符号图表示出来。

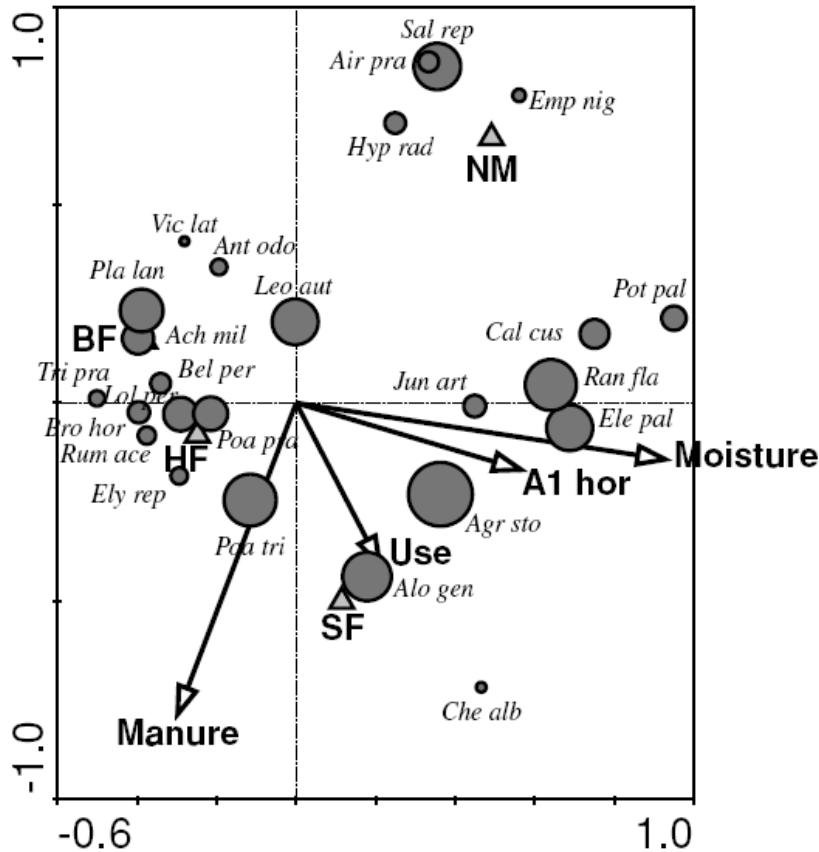


图10-16. 能显示物种拟合度（能被排序图解释的百分比例）的属性排序图（Attribution plot）。图中物种圈圈的大小代表该物种方差（变化量）能够被排序模型解释的比例，这些比例的具体情况可以在Canoco输出文件（.sol 为后缀的文件）Cfit statistics 栏中可以找到。

## 10.5 t-value 双序图（t-value biplot interpretation）

t-value 双序图包含了物种的箭头和环境因子的箭头和圆圈符号的排序图。首要的任务是揭示成对的物种与环境因子的统计显著关系（比如物种依赖环境因子的程度）。t-value 来自于多重回归的回归系数。这个多重回归是以该物种为因变量，以观测的环境因子为自变量的回归模型。

T-values 的 Van Dobben circles 是根据物种与环境因子多重回归系数画的，专门用来表现物种分布与环境因子变量的相关关系排序图。例图中的箭头长度与方向代表物种与此环境因子的典范相关关系。如图 10-17，如果某物种的箭头完全掉在湿度（Moisture）

正相关的 Van Dobben circles (例图中红色区域), 就意味着此种与此湿度显著正相关, 即物种多度随湿度大增大而增加; 如果物种的箭头完全掉在湿度 (Moisture) 负相关的 Van Dobben circles (例图中蓝色区域), 就意味着物种与此湿度显著负相关, 即物种多度随湿度大增大而减少。t-value 双序图可以在 CanoDraw 里面菜单 Create > Biplots and Joint Plots >t-values biplots 里面可以生成。

关于 t-values 更详细内容, 可以参考 Canoco 使用手册 6.3.1.2 节 Ter Braak and Smilauer 2002, Section 6.3.1.2

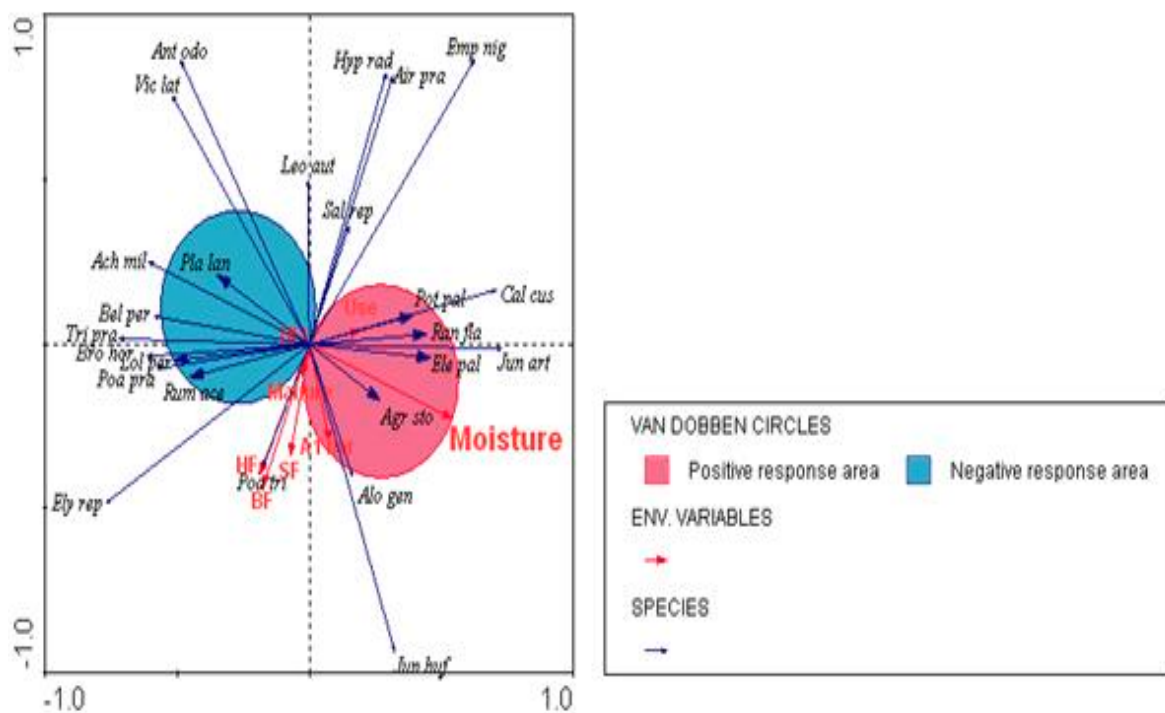


图 10-17. 如果某物种的箭头完全掉在湿度 (Moisture) 正相关的 Van Dobben circles (例图中红色区域), 就意味着此种与此湿度显著正相关, 即物种多度随湿度大增大而增加; 如果物种的箭头完全掉在湿度 (Moisture) 负相关的 Van Dobben circles (例图中蓝色区域), 就意味着物种与此湿度显著负相关, 即物种多度随湿度大增大而减少。