The front door to the fourth corner: variations on the sample unit × trait matrix in community ecology

B. McCune

Department of Botany and Plant Pathology, Oregon State University, Corvallis, OR 97331-2902 U.S.A.
E-mail: mccuneb@onid.orst.edu

Abstract: Calculating a sample unit × trait matrix provides a flexible first step in analyzing the relationships between species traits and explanatory variables. This matrix is obtained by multiplying a sample unit × species matrix by a species × trait matrix, but the content of the resulting matrix depends on whether and how traits are standardized and whether or not the multiplication is followed by a weighted averaging step. To maximize versatility of the SU × trait matrix, including comparability among traits, and usability with a wide range of distance measures, we recommend first standardizing traits by min-to-max, then calculating abundance-weighted trait averages in each sample unit.

Abbreviation: SU − Sample Unit.

Introduction

Linking species traits to the effects of environmental gradients, disturbances, and experimental treatments on ecological communities is a common goal in ecology. This goal can be formulated as understanding the relationships between three matrices: (1) a community matrix of sample units (SUs) × species (or taxa in general), (2) an explanatory matrix of SUs × environmental variables, disturbances measures, or experimental-design variables, and (3) a species × trait matrix containing values that represent one or more ecological traits of the species. Because three matrices are involved, many methodological choices must be made. The purpose of this paper is not to review those but to summarize a simple, effective preliminary step that can position the analyst to directly answer questions concerning those relationships.

The species × trait matrix itself is of considerable interest and worthy of analysis, for studying relationships among species, among traits, or both. Quite often, however, we focus on the relationship of this matrix to other matrices – primarily variation in communities in relation to environmental gradients or experimental factors. This can be done with univariate or bivariate analyses, or it can be done as multivariate analyses, considering multiple dependent variables simultaneously.

Broadly speaking two multivariate approaches have been used to relate the trait matrix to the community and environmental matrices. If we adopt parametric statistics, relationships between the variation in traits and environment can be modeled simultaneously, with the SU × species matrix providing the linkage between the trait matrix and the environmental matrix. Two methods for this are the fourth corner analysis (Dray and Legendre 2008, ter Braak et al. 2012) and RLQ analysis (Dray et al. 2003, Dray et al. 2014).

Alternatively, one can avoid the assumptions of those methods, instead opening the analysis to any functional form, by first calculating the SU × trait matrix. To do this we multiply the trait matrix by the community matrix, and then incorporate the resulting SU × trait matrix into other analyses (Feoli and Scimone 1984, Diaz et al. 1992, 1999, McCune and Grace 2002, pp. 9-10, Ozinga et al. 2004, and many others). For example, one might overlay the traits from the SU × trait as a joint plot on an ordination of SU in species space. Or one might represent the SU × trait matrix with one, two, or three axes with nonmetric multidimensional scaling, then overlay environmental factors with a smoothing function, showing the topography of an environmental factor in trait space. Alternatively, one might superimpose both environmental factors and traits on an ordination of SUs in species space—for example, the hilltop plots of Nelson et al. (2015).

My goal here is not to review the general approaches to the problem, but rather to focus on alternative ways to derive the SU × trait matrix. Derivation of this matrix is a critical step for analysts and, although it is a simple one, I have seen through teaching that it is often a point of confusion. Mistakes are commonly made such that the resulting matrix and subsequent analyses do not have the properties desired by the analyst. The purpose of this paper is, therefore, to summarize some of the most important options in calculating the SU × trait matrix, such that analysts can more quickly decide how best to calculate it, given their goals. The SU × trait matrix is not an end in itself; instead it is a vehicle to examine traits in relationship to environmental gradients, experimental factors, or disturbances. Because the calculation of that matrix is a critical first step, it is important that the matrix represents what is intended.

Definitions

Assume that the main community matrix ($\mathbf{A} = n$ SU $\times$ p species) contains abundances (e.g., counts of individuals).

Assume that the trait matrix ($\mathbf{S} = p$ species $\times q$ traits) contains either quantitative or binary traits or both.

Assume that qualitative traits (i.e., nominal or categorical traits) have been re-expressed as a series of binary traits, one such trait for each qualitative attribute. Binary traits that represent presence vs. absence, are simply and effectively coded as 1 = presence and 0 = absence. If a binary trait is in fact nominal (i.e., the states can be interchanged such as black = 0, white = 1 or black = 1, white = 0), then the results will depend on coding. This can be handled by creating two separate binary variables or by careful specification of the coding and its interpretation.

Binary and quantitative traits are each considered in the discussion that follows.

Multiply $\mathbf{A}$ by $\mathbf{S}$ to obtain the SU $\times$ trait matrix, $\mathbf{T} = \mathbf{AS}$ ($\mathbf{T} = n$ SU $\times q$ traits, Fig. 1).

Standardizing traits

Standardizing traits has the goal of placing equal emphasis on a variety of traits measured on a variety of scales. Trait data naturally tend to be mixtures of quantitative, binary, and qualitative values. Standardizing them improves comparability across traits, but for quantitative traits, sacrifices the original units of the traits. This loss may be unimportant for studies of the relative explanatory value of different traits, but if the goal is to use new trait measurements in a predictive model, then standardized traits will require that new measurements also receive the same standardization as the old—us-

ing the same relativization standard (e.g., the trait mean and standard deviation from the original data set).

Two simple and effective standardizations are to rescale by min-to-max (0 = minimum to 1 = maximum) or by standard deviates, such that each trait has a mean = 0 and standard deviation = 1.

Rescaling trait k for each species j in the $\mathbf{S}$ matrix (s_{jk}) by min-to-max is done by:

$$s_{jk}{}^* = (s_{jk} - s_{k\min})/(s_{k\max} - s_{k\min})$$

where

$s_{k\min}$ = minimum value for trait k

$s_{k\max}$ = maximum value for trait k

The domain of s_{jk} includes any real number and the range of $s_{jk}{}^*$ is 0 to 1. This standardization has the distinct advantage that it is concordant with Gower's (1971) distance measure. Gower's distance is often a good choice for the species by trait matrix, because it handles quantitative, qualitative, and binary variables (de Bello et al. 2013). A fundamental step in calculating Gower's distance or similarity is that individual variables are rescaled according to their ranges, similar to the formulation above.

Rescaling trait s_k by standard deviates is done by first calculating the mean (m_k) and standard deviation (sd_k) of each trait, k, across species, j, in the $\mathbf{S}$ matrix, then expressing each trait value as the number of standard deviations from that species' mean:

$$s_{jk}{}^* = (s_{jk} - m_k)/sd_k$$

The domain of s_{jk} includes any real number and range of $s_{jk}{}^*$ is $\pm\infty$, although in practice most values will fall between -3 and +3 standard deviations from the mean. This method has the advantage of being the most commonly used standardization in traditional parametric statistics. But it has the disadvantage of precluding the use of certain distance measures on

Figure 1. Calculating the sample unit (SU) $\times$ trait matrix. Initially, we wish to relate three matrices to one another: the community matrix of SU $\times$ spp ($\mathbf{A}$), the species traits matrix of spp $\times$ traits ($\mathbf{S}$), and an explanatory matrix of SU $\times$ explanatory variables ($\mathbf{E}$, which may include design variables such as factors in experimental designs or levels in a sampling design). As a first step, the SU $\times$ trait matrix ($\mathbf{T}$, see variants in Table 1) can be calculated by multiplying $\mathbf{A} \times \mathbf{S}$.
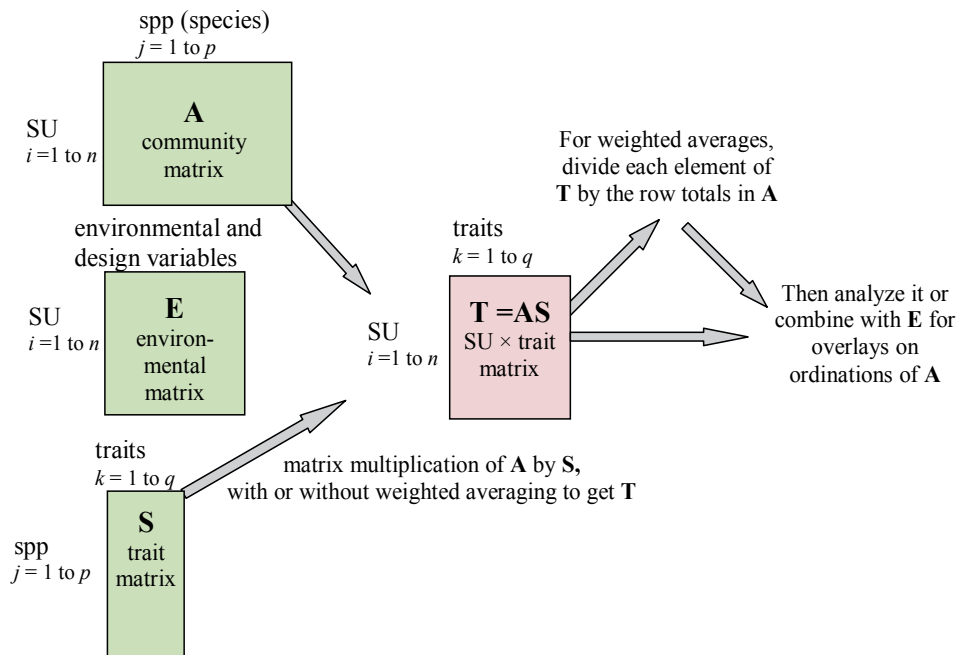
Table 1. Variations on the sample unit × trait matrix produced by choices of standardization and weighting. "Ranges are comparable" refers to whether traits with different scales or ranges are scaled to equal footing with each other. The range of resulting values is compared for quantitative and binary traits. Compatibility of the resulting matrix with proportional city-block distance measures (e.g., Sørensen or Bray-Curtis, Jaccard) is also indicated. The shaded row is recommended unless a specific research objective dictates otherwise.

Matrix type	Traits standardized by	Resulting elements are...	Are ranges of resulting variables directly comparable with each other?	Range of result, if quantitative trait	Range of result, if binary trait	Compatible with proportional distance measures
Abundance-weighted totals						
1	none	abundance-weighted trait total (sum of trait×abundance products)	no (vary by trait scale and total abundance in SU)	0 to product of row total in **A** times trait total in **S**	0 to row total in **A**	yes
2	0-1 (min to max)	abundance-weighted proportional trait total (sum of proportional trait×abundance products)	no (vary by total abundance)	0 to row total in **A**	0 to p	yes
3	mean = 0, sd = 1	abundance-weighted total deviation from trait mean	no (vary by total abundance)	not pre-defined	$\pm \infty$	no
Abundance-weighted averages						
4	none	abundance-weighted trait average	no (vary by trait scale)	$s_{k\min}$ to $s_{k\max}$	0 to 1	yes
5	0-1 (min to max)	abundance-weighted proportional trait average	yes	0 to 1	0 to 1	yes
6	mean = 0, sd = 1	abundance-weighted average deviation from trait mean	yes	theoretically $\pm\infty$, but practically ± 3	theoretically $\pm\infty$, but practically ± 3	no

the resulting matrix, specifically the popular proportionalized city-block distances such as Sørensen and Jaccard. These are particularly useful for zero-based measures of abundance, and have a domain of non-negative numbers. Rescaling the traits by standard deviates yields both positive and negative numbers, precluding those distance measures.

Weighted totals vs. weighted averages

If a community matrix, **A**, that contains abundance data (i.e., counts, frequency, cover, etc.) is multiplied by a trait matrix, **S**, the resulting matrix, **T,** will contain abundance-weighted totals, simply by the process of matrix multiplication. Each element of **T** (t_{ik}) is obtained by summing the products of a species' abundance (a_{ij}) by its corresponding trait (s_{jk}), across all species ($j = 1$ to p) in a sample unit:

$$t_k = \sum_{j=1}^{p} a_j s_k$$

Abundance-weighted averaging follows this by dividing that total by the summed abundance of species in the sample unit, i.

If a_{i+} = total for row (SU) i in matrix **A**, then the abundance-weighted average is

$$\bar{t}_k = \frac{1}{a_{i+}} \sum_{j=1}^{p} a_j s_k$$

This is the community-weighted mean (CWM) of Lavorel et al. (2008). Be sure that each sum is divided by the total abundance of species, not p, the number of terms in the summation. The latter would make the average sensitive to the number of absent species and would not be a weighted average.

Using the abundance-weighted average, rather than the total, makes matrix **T** more interpretable, because each trait in **T** has the same potential range as the trait in **S**. For example, if longevity ranges from 1 to 10 years in **S**, longevity in **T** has a minimum possible value of 1 year and a maximum possible value of 10 years. Or, if longevity had previously been standardized to a range of 0 to 1, then the values in **T** have that same potential range.

To illustrate the importance of standardizing traits, consider a quantitative trait for seed weight that ranged from 1 to 500 mg. If unstandardized and combined with a set of binary traits, seed weight would completely overwhelm the other traits in an analysis that has no built-in standardization of variables, such as nonmetric multidimensional scaling. Furthermore, the results would be sensitive to any linear rescaling of the trait. For example, if seed weight was expressed in grams rather than milligrams, the values in that column of the SU × trait matrix would shrink by 1/1000, relative to the other variables, greatly influencing the results. Standardizing traits by min-to-max, however, has the useful effect of rendering the units for quantitative traits irrelevant.

Abundance-weighted totals are, however, a simple, effective representation of the SU × trait matrix if the traits are all binary (or qualitative and recoded as binary traits). Each

element of the resulting matrix is, in that case, the total abundance of species with a particular trait in a given SU.

Variants of the SU × Trait Matrix

Combining the choices for standardization with the choice of abundance-weighted average vs. totals yields six possible outcomes for the SU × trait matrix (Table 1). Each of

these is expressed somewhat differently for binary and quantitative traits (Fig. 2).

Example of weighted averaging with traits standardized by min to max

Row (sample unit) from **A**: plot 4 has maple = 0, oak = 3, pine = 4

A	maple	oak	pine
Plot1	4	0	0
Plot2	2	1	1
Plot3	1	5	2
Plot4	0	3	4

S		Standardization of **S**					
	Raw data			By min-to-max		By standard deviates	
	ShadeTol	Hardwood		ShadeTol	Hardwood	ShadeTol	Hardwood
maple	10	1	maple	1	1	maple	1.407	0.707
oak	2	1	oak	0.111	1	oak	-0.579	0.707
pine	1	0	pine	0	0	pine	-0.828	-1.414
min	1	0		0	0		-0.579	-1.414
max	10	1		1	1		1.407	0.707
mean	4.333	0.667		0.370	0.667		0.000	0.000
sd	4.028	0.471		0.448	0.471		1.000	1.000

T = AS   Calculation of the SU × traits matrix

As abundance weighted totals

1	Not standardized		2	Min-to-max		3	Standard deviates	
	ShadeTol	Hardwood		ShadeTol	Hardwood		ShadeTol	Hardwood
Plot1	40	4	Plot1	4	4	Plot1	5.628	2.828
Plot2	23	3	Plot2	2.111	3	Plot2	1.407	0.707
Plot3	22	6	Plot3	1.556	6	Plot3	-3.145	1.414
Plot4	10	3	Plot4	0.333	3	Plot4	-5.048	-3.536

T = AS

As abundance-weighted averages

4	Not standardized		5	Min-to-max		6	Standard deviates	
	ShadeTol	Hardwood		ShadeTol	Hardwood		ShadeTol	Hardwood
Plot1	10	1	Plot1	1	1	Plot1	1.407	0.707
Plot2	5.750	0.750	Plot2	0.528	0.750	Plot2	0.352	0.177
Plot3	2.750	0.750	Plot3	0.194	0.750	Plot3	-0.393	0.177
Plot4	1.429	0.429	Plot4	0.048	0.429	Plot4	-0.721	-0.505

Figure 2. Example calculation of the six types of SU × trait matrices representing combinations of standardization and weighting choices. Table 1 describes the properties of these six matrices.

Column (trait 1) from **S**: shade tolerance is maple = 10, oak = 2, pine = 1.

Standardized trait: shade tolerances are $(s - s_{j\min})/(s_{j\max} - s_{j\min})$:

maple = (10–1)/(10–1) = 1
oak = (2–1)/(10–1) = 0.111
pine = (1–1)/(10–1) = 0

Weighted total for plot 4 for trait 1, shade tolerance:
$t_{4,1} = \Sigma_{j=1}^{3} a_{4,j} \, s_{j,1} = 0\cdot1 + 3\cdot0.111 + 4\cdot0 = 0.333$

To convert weighted total to weighted average, divide by the row total from **A**:

$\bar{t}_{4,1} = t_{4,1} / a_{4+} = 0.333 / 7 = 0.0476.$

Conclusions

For the sake of transparency and reproducibility of results, authors should always report the exact method used to calculate a SU × trait matrix. Many readers will also appreciate a statement of the meaning of an element in the matrix. For example, stating that they are "abundance-weighted proportional trait averages" conveys that the elements are weighted averages and that traits have been proportionalized by their ranges. This then implies that the traits in that matrix are all on scales that are directly comparable with each other.

To maximize versatility of the SU × trait matrix, including comparability among traits, and usability with a wide range of distance measures, we recommend first standardizing traits by min-to-max, then calculating abundance-weighted trait averages in each sample unit. This is "matrix type 5" in Table 1 and Figure 2. This is only a general recommendation; other formulations may be appropriate for specific needs.

References

de Bello, F., C. P. Carmona, N. W. H. Mason, M. Sebastià and Lepš, J. 2013. Which trait dissimilarity for functional diversity: trait means or trait overlap? *J. Veg. Sci.* 24: 807-819.

Diaz, S., A. Acosta and M. Cabido. 1992. Morphological analysis of herbaceous communities under different grazing regimes. *J. Veg. Sci.* 3: 689-696.

Diaz, S., M. Cabido, M. Zak, E. Martinez Carretero and J. Araniba. 1999. Plant functional traits, ecosystem structure and land-use history along a climatic gradient in central-western Argentina. *J. Veg. Sci.* 10: 651-660.

Dray, S., D. Chessel and J. Thioulouse. 2003. Co-inertia analysis and the linking of ecological data tables. *Ecology* 84: 3078-3089.

Dray, S., P. Choler, S. Dolédec, P. R. Peres-Neto, W. Thuiller, S. Pavoine and C. J. F. ter Braak. 2014. Combining the fourth-corner and the RLQ methods for assessing trait responses to environmental variation. *Ecology* 95: 14–21.

Dray, S. and P. Legendre, P. 2008. Testing the species traits environment relationships: the fourth-corner problem revisited. *Ecology* 89: 3400–3412.

Feoli, E. and M. Scimone. 1984. A quantitative view of textural analysis of vegetation and examples of application of some methods. *Arch. Bot. Biogeogr. Ital.* 60: 72-94.

Gower, J. C. 1971. A general coefficient of similarity and some of its properties. *Biometrics* 23: 623-637.

Lavorel, S., K. Grigulis, S. McIntyre, N. S. G. Williams, D. Garden, J. Dorrough, S. Berman, F. Quétier, A. Thebault and A. Bonis. 2008. Assessing functional diversity in the field – methodology matters! *Funct. Ecol.* 22: 134-147.

McCune, B. and J. B. Grace. 2002. *Analysis of Ecological Communities*. MjM Software, Gleneden Beach, Oregon, U.S.A.

Nelson, P. R., B. McCune, C. Roland and S. Stehn. 2015. Nonparametric methods reveal nonlinear functional trait variation along environmental and fire age gradients. *J. Veg. Sci.* 26: (in press). DOI: 10.1111/jvs.12286

Ozinga, W. A., R. M. Bekker, J. H. J. Schaminée and J. M. van Groenendael. 2004. Dispersal potential in plant communities depends on environmental conditions. *J. Ecol.* 92: 767-777.

ter Braak, C. J. F., A. Cormont and S. Dray. 2012. Improved testing of species traits–environment relationships in the fourth-corner problem. *Ecology* 93: 1525–1526.