Conditional compositional biplots: theory and application

J. Aitchison

Department of Statistics, University of Glasgow, UK
Email: john.aitchison@btinternet.com

K.W. Ng*

Department of Statistics and Actuarial Science, University of Hong Kong, People's Republic of China
Email: kaing@hku.hk

Abstract

The biplot has proved to be a powerful descriptive and analytical tool in many areas of applications of statistics. For compositional data the necessary theoretical adaptation has been provided, with illustrative applications, by Aitchison (1990) and Aitchison and Greenacre (2002). These papers were restricted to the interpretation of simple compositional data sets. In many situations the problem has to be described in some form of conditional modelling. For example, in a clinical trial where interest is in how patients' steroid metabolite compositions may change as a result of different treatment regimes, interest is in relating the compositions after treatment to the compositions before treatment and the nature of the treatments applied. To study this through a biplot technique requires the development of some form of conditional compositional biplot. This is the purpose of this paper. We choose as a motivating application an analysis of the 1992 US Presidential Election, where interest may be in how the three-part composition, the percentage division among the three candidates - Bush, Clinton and Perot - of the presidential vote in each state, depends on the ethnic composition and on the urban-rural composition of the state. The methodology of conditional compositional biplots is first developed and a detailed interpretation of the 1992 US Presidential Election provided. We use a second application involving the conditional variability of tektite mineral compositions with respect to major oxide compositions to demonstrate some hazards of simplistic interpretation of biplots. Finally we conjecture on further possible applications of conditional compositional biplots.

Keywords: Simplicial singular value decomposition; tektite compositions, US Presidential elections

---

*Address for correspondence: Department of Statistics and Actuarial Science, University of Hong Kong, Pokfulam Road, Hong Kong, People's Republic of China

1 Introduction

It is now well established that the analysis and interpretation of compositional data, vectors of proportions of parts of some whole, such as the major oxide [SiO_2, Al_2O_3, . . . , MnO] composition of rock specimens or the ethnic [White, Black, Hispanic, Native American, Other] composition of a US state, require statistical methods appropriate to a simplex sample space. The monograph (Aitchison, 1986) established appropriate methodologies for compositional data analysis based on logratio transformation techniques. The useful graphical technique, the biplot introduced by Gabriel (1971, 1981) for unconstrained multivariate data has already been adapted to the constrained nature of compositional data by Aitchison (1990) and Aitchison and Greenacre (2002), with illustrative applications to demonstrate the use of such compositional biplots as descriptive and analytical tools. The main purpose of this paper is to extend the compositional biplot from applications involving simple compositional data sets to those requiring conditional modelling. We shall illustrate the formation and interpretation of such conditional biplots in Section 5 and 6 to two data sets. First, in Section 5, we investigate the 1992 US Presidential election, in particular the dependence of the share of the vote in each state by the three candidates on the ethnic and urban-rural compositions of the state. Secondly, in Section 6, we examine the mineral and major-oxide compositions of tektites and attempt, with no geochemical pre-knowledge, to relate the variability of the mineral compositions to that of the major oxide compositions. For completeness we first summarise in Section 2 the essential features of the simplex sample space. We shall also need the basic ideas on unconditional biplots and we summarise these in Section 3 together with some specific biplots required later in the paper. Section 4 then builds on Sections 2 and 3 to provide a methodology for conditional compositional biplots and provides guidelines for interpretation. In the final Section 7 we discuss some possible extensions to a variety of conditional compositional situations,

2 The simplex sample space: operations and notation

A D-part composition x is a vector of positive components $[x_1, \ldots, x_D]$ where x_i $(i = 1, \ldots, D)$ are the proportions of the various parts of some whole unit. Such compositions thus belong to a unit simplex sample space

$$S^D = \{[x_i \ldots, x_D]: \ x_i > 0 \ \ (i = 1, \ldots, D), \ \ x_1 + \ldots + x_D = 1\}. \tag{2.1}$$

Our development requires a clear understanding of simple algebraic-geometric concepts and operations in this space and, for convenience, we provide a summary here and, at the same time, establish the notation of the paper.

Logratio transformations from S^D to real spaces. The early development of a methodology for compositional data analysis (Aitchison, 1982, 1986) relied on transformation techniques involving logratios of components of compositions. These mapped compositions into vectors in real spaces allowing, with care, the use of standard unconstrained multivariate analysis to resolve inference problems about the compositions. We shall require the most useful two of these transformations, the so called *additive* logratio transformation $alr: S^D \to R^{D-1}$, defined as follows:

$$alr(x) = [\log(x_1 / x_D), \ldots, \log(x_{D-1} / x_D)] = y \quad (x \in S^D, y \in R^{D-1}), \qquad (2.2)$$

with inverse

$$\begin{aligned} alr^{-1}(y) &= [\exp(y_1), \ldots, \exp(y_{D-1}), 1] / \{\exp(y_1) + \ldots + \exp(y_{D-1}) + 1\} \\ &= C[\exp(y_1), \ldots, \exp(y_{D-1}), 1], \end{aligned} \qquad (2.3)$$

where the operator C simply divides each component of the contained vector by their sum. The *centred* logratio transformation $clr: S^D \to U^{D-1}$, where U^{D-1} is the hyperplane $\{u \in R^D : u_1 + \ldots + u_D = 0\}$ in R^D, is defined as follows:

$$clr(x) = [\log\{x_1 / g(x)\}, \ldots, \log\{x_D / g(x)\}] = u \quad (x \in S^D, u \in U^{D-1}), \qquad (2,4)$$

where $g(x) = (x_1 x_2 \ldots x_D)^{1/D}$ is the geometric mean of the components of x. The corresponding inverse is given by

$$clr^{-1}(u) = C[\exp(u_1), \ldots, \exp(u_D)]. \qquad (2.5)$$

It is the 1-1 nature of these transformations which allows a satisfactory approach to compositional data analysis through the use of R^D-based multivariate statistical analysis.

Compositional covariance structures. These transformations provide appropriate covariance or dependence structures to describe the dependence of components within a composition x. These are the logratio covariance matrix $\Sigma(x) = \operatorname{cov}\{alr(x)\}$, the centred logratio covariance matrix $\Gamma(x) = \operatorname{cov}\{clr(x)\}$ and the variation matrix $\mathrm{T}(x) = [\operatorname{var}\{\log(x_i / x_j)\}]$. These three characteristics are equivalent in the sense that given any one, the others are automatically determined. All have a role to play in compositional data analysis and the choice is largely determined by the particular application. See Aitchison (1986) for fuller details..

Perturbation. The group operation corresponding to displacement or translation in real space R^D is a perturbation defined as follows. Corresponding to any two compositions $x, y \in S^D$ there is a perturbation $x \oplus y \in S^D$, defined as follows:

$$x \oplus y = [x_1 y_1, \ldots, x_D y_D] / (x_1 y_1 + \ldots + x_D y_D) = C[x_1 y_1, \ldots, x_D y_D]. \qquad (2.6)$$

This operation defines an Abelian group with identity $e = \{1/D\}[1, \ldots, 1]$ and with inverse operation $x \ominus y$ defined by $x \ominus y = C[x_1 / y_1, \ldots, x_D / y_D]$.

An important use of the concept of perturbation is in specifying change in compositions. If in part of a compositional process a composition *x* changes to a composition *X* then the change *p* is simply the perturbation $p = X \ominus x$. For a simple application of this concept in compositional data analysis see Aitchison and Ng (2005).

Powering. The second operation we require is that of powering, defined as follows. Corresponding to any composition $x \in S^D$ and any real number $a \in R^1$ there is a powered composition $a \otimes x \in S^D$, defined as follows:

$$a \otimes x = C[x_1^a, , , , , x_D^\backslash]. \qquad (2.7).$$

So far in the development of compositional data analysis this operation has had a very limited role. Later in this paper we shall see that in terms of process modelling and analysis it plays a central role. In terms of the geometry of the simplex the internal perturbation group operation and the external powering operation ensure a linear space structure.

Metric We can extend the structure of S^D to a metric space with the definition of a metric $\Delta : S^D \times S^D \to R^1_{\geq 0}$ defined as follows. Corresponding to every two compositions $x, y \in S^D$,

$$\begin{aligned} \Delta(x,y) &= (\{clr(x) - clr(y)\}\{clr(x) - clr(y)\}^T)^{1/2} \\ &= (\{alr(x) - alr(y)\} H^{-1} \{alr(x) - alr(y)\}^T)^{1/2}, \end{aligned} \qquad (2.8)$$

where $H = [h_{ij}]$ is of order $(D-1) \times (D-1)$ with $h_{ij} = 1\,(i \neq j), = 2\,(i = j)$.

For further details see Aitchison (2001), and Aitchison et al (2002). , who further define an associated inner product

$$\langle x, y \rangle = \sum_{i=1}^{D} \log \frac{x_i}{g(x)} \log \frac{y_i}{g(x)}, \qquad (2.9)$$

and norm

$$\| x \| = \left[\sum_{i=1}^{D} \left\{ \log \frac{x_i}{g(x)} \right\}^2 \right]^{1/2}, \qquad (2.10)$$

thus providing a Hilbert space in which to consider statistical modelling.

Distributional concepts in the simplex. For statistical modelling we have to consider distributions on the simplex and their characteristics. The well-established 'measure of central tendency' $x \in S^D$ which minimizes E(Δ(x, ξ)) is

$$x = cen(x) = C(\exp(E(\log x))), \qquad (2.11)$$

with simple proerties such as

$$cen(a \otimes x) = a \otimes cen(x) \text{ and } cen(x \oplus y) = cen(x) \oplus cen(y). \qquad (2.12)$$

This definition conforms with the covariance structure or dispersion definitions defined above Importantly these dispersion characteristics are consistent with the following properties for any of these matrices, say $dis(x)$:

for any real number a $dis(a \otimes x) = |a|^2\, dis(x)$;
for any constant perturbation p, $dis(x \oplus p) = dis(x)$;
for independent. $x, y \in S^{D,}$ $dis(x \oplus y) = dis(x) + dis(y)$.

Relevance to compositional data sets. There are substantial implications in the above development for the analysis of a $N \times D$ compositional data set $X = [x_1; \ldots; x_N]$. A main feature is that the estimate of the centre x is given by $\hat{x} = C[g_1, \ldots, g_D]$, where g_i is the geometric mean of the ith component of the N compositions. Measures of dispersion are simply estimated from the estimated variances of the appropriate logratios.

There is for the simplex a result, analogous to the singular value decomposition for data sets associated with the sample space R^d. This simplicial singular value decomposition plays a central role in the study of compositional variability, in particular in obtaing graphical approximations to the data set. Any $N \times D$ compositional data matrix X with nth row composition x_n can be decomposed in a power-perturbation form as follows

$$x_n = x \oplus (u_{n1} s_1 \otimes b_1) \oplus \ldots \oplus (u_{nR} s_R \otimes b_R), \qquad (2.13)$$

where ξ is the centre of the data set, the s's are positive 'singular values' in descending order of magnitude, the β's are compositions, R is a readily defined rank of the compositional data set and the u's are power components specific to each composition. In a way similar to that for data sets in R^d we may consider an approximation of order $r < R$ to the compositional data set given by

$$x_n^{(r)} = x \oplus (u_{n1} s_1 \otimes b_1) \oplus \ldots \oplus (u_{nr} s_r \otimes b_r) . \qquad (2.14)$$

Such an approximation retains a proportion

$$(s_1^2 + \ldots + s_r^2) / (s_1^2 + \ldots + s_R^2) \qquad (2.15)$$

of the total variability of the $N \times D$ compositional data matrix as measured by the trace of the estimate of the centered logratio covariance matrix $\Gamma(x)$ or equivalently in terms of total mutual distances as

$$(N(N-1))^{-1} \sum_{m<n}^{D} \Delta_S^2 (x_m, x_n) . \qquad (2.16)$$

We shall see in the next section that these measures of variation are related to the Frobenius norms of certain matrices which provide measures of the extent of the variability captured by biplots.

We may also note here that the power-perturbation expression of the singular value decomposition has exactly the same form as regression of a composition on some set of variables. The form is exactly what would be obtained if the logratio form of regression analysis in Aitchison (1986, Chapter 7) were transformed back into terms of the simplex.

3 Unconditional compositional biplots

As a step towards the constuction of conditional compositional biplots we summarise here the main aspects of unconditional compositional biplots. For full dtails of these and their interpretation see Aitchison (1990) and Aitchison and Greenacre (2002).

Compositional Singular Value Decomposition. Biplots are constructed on the basis of a singular value decompositions of data matrices, first introduced into statistical work by (Eckart and Young, 1936), (Whittle, 1952) and Good, 1969), and adapted for biplot construction by Gabriel (1971, 1981). The singular value decomposition property states that any N $\times$ D matrix Z of rank R can be expressed as a product

$$Z = U diag(s_1, \ldots, s_R) V^T, \tag{3.1}$$

where U and V are of orders $N \times R$ and $D \times R$, each with orthonormal columns, and the positive numbers $s_1, \ldots, s_R$ assumed here to be arranged in descending order of magnitude, are the singular values. For the application of this result to unconstrained multivariate data it is standard practice to centre the data at the mean vector by subtracting from each element of the data matrix its corresponding column average. The adjustment for a $N \times D$ compositional data matrix $X = [x_{ni}]$ with rows x_n $(n = 1, \ldots, N)$ the N D-part compositions. We first form $clr(X)$ ensuring that we are working symmetrically with logratios, and then centre the colums as for unconstrained data. the effect is to produce a centred logratio data matrix $Z = [z_{ni}]$ with

$$z_{nd} = \log x_{ni} - D^{-1} \sum_{d=1}^{D} \log x_{ni} - N^{-1} \sum_{n=1}^{N} \log x_{ni} + (ND)^{-1} \sum_{n=1}^{N} \sum_{d=1}^{D} \log x_{ni}, \tag{3.2}$$

with all row sums and column sums zero. This zero sum property carries over to the columns of U and of V in the singular value decomposition $Z = U diag(s_1, \ldots, s_R) V^T$ of *Z*. We shall see that what the singular value decomposition is achieving is the provision of a series of approximations to the data matrix *Z*.

Approximations to the centred logratio data matrix Z. The centred logratio data matrix Z will usually be of rank d = D - 1 so that the matrices U and V will be of order N $\times$ d and D $\times$ d. A hope of the decomposition is that the eigenvalues will decrease rapidly so that Z will be well approximated by $Z^{(r)} = U_r diag(s_1, \ldots, s_r) V_r^T \ (r < R)$, where U_r and V_r are the leading N $\times$ r and D $\times$ r submatrices of U and V respectively. The degree of approximation is usually measured in terms of the Frobenius norm of the difference between Z and $Z^{(r)}_r$ namely

$$\| Z - Z^{(r)} \|^2 = \sum_{n=1}^{N} \sum_{i=1}^{D} (x_{ni} - z_{ni}^r)^2 = s_{r+1}^2 + \ldots + s_R^2 . \qquad (3.3)$$

The optimizing property is that of all matrices of rank at most *r*, $Z^{(r)}$ is that which minimizes this Frobenius norm. As a measure of the quality of the approximation we can thus take the customary measure

$$1 - \frac{\| Z - Z^{(r)} \|^2}{\| Z \|^2} = \frac{s_1^2 + \ldots. + s_r^2}{s_1^1 + \ldots + s_R^2}, \qquad (3.4)$$

which is the proportion of the total variability of the compositional data set retained by $Z^{(r)}$ or equivalently by the first *r* principal logcontrast components. In order to obtain any useful graphical representation of the compositional data set we shall have to take *r* = 2 and in order to state the properties of the relative variation diagram clearly we shall assume for the moment that *R* = 2 so that the relationship $Z = U_2 diag(s_1, s_2) V_2^T$ is exact. Before we set out these properties we have first to describe the construction of this second order graphical approximation..

Construction of an unconditional compositional biplot

The relative variation diagram. Suppose that with origin O in a two-dimensional diagram (Fig. 3.1) we plot the D points

$$(s_1 v_{i1}, s_2 v_{i2}) / (N-1)^{1/2} \quad (i = 1, \ldots, D) \qquad (3.5)$$

and regard these as the *vertices* $i \ (i = 1, \ldots, D)$ of the relative variation part of the biplot for the compositional data matrix *X*. Each vertex then corresponds to a *part* of |the compositional data matrix *X*. We then term *Oi* a *ray* and the join of two vertices such as *ij* a *link* of the diagram. It can then be shown that the relative variation diagram, consisting of the set of vertices, rays and links, together with the various angles defined by rays and links, contains a complete quantitative picture of all the various covariance structures associated with the compositional data set .

Properties of the relative variation diagram

Property 1. The origin *O* is the centroid of the vertices 1, . . . , *D.*

Property 2. The squared lengths of the links represent the set of estimated relative variances:

$$| ij |^2 = \tau_{ij}^{\,2} = \text{estimate of var}\{\log(x_i/x_j)\}.$$

Property 3. Rays and inter-ray angles represent the centred logratio covariance matrix Γ:

$$| Oi |^2 = \gamma_{ii} = \text{estimate of var}[\log\{x_i/g(x)\}],$$

$$Oi\,.\,Oj = \gamma_{ij} = \text{estimate of cov}[\log\{x_i/g(x)\},\ \log\{x_j/g(x)\}],$$

so that

$$\cos(iOj) = \text{estimate of corr}[\log\{x_i/g(x)\},\ \log\{x_j/g(x)\}].$$

A generalization of Property 3 involving four parts i, j, k, l is easily established and proves extremely useful in the exploration of independence properties of compositional data sets.

Property 4. If the links ij and kl intersect at M then

$$\cos(iMk) = \text{estimate of corr}\{\log(x_i/x_j),\log(x_k/x_l)\}.$$

Interpretation of a relative variation diagrams

The following hints may prove useful in interpreting various aspects of compositional variability.

(a) *Coincident vertices and proportionality*: When two vertices i and j coincide or are close together then the length of the link ij and, from Property 2, the estimate τ_{ij} of var$\{\log(x_i/x_j)\}$ are zero or small and so components x_i and x_j are in constant proportion or nearly so. While this is obvious it is not unimportant, particularly when we realise that the whole covariance structure is most simply defined in terms of relative variances; and further that the concept of small relative variance with its associated high dependence of one component on another is essentially what is required to replace uninterpretable measures of dependence such as crude product-moment correlations corr(x_i, x_j).

(b) *Subsets of vertices and subcompositional analysis.* If we consider a subset, say of parts $1, \ldots, C$ of a D-part composition, then the concept of the subcomposition formed from these parts and defined by

$$(s_1, \ldots, s_C) = (x_1, \ldots, x_C) / (x_1 + \ldots + x_C) \qquad (3.6)$$

plays a central role in compositional data analysis. One of the main reasons for claiming that relative variances provide the simplest specification of compositional covariance structure is that

$$\text{var}\{\log(s_i/s_j)\} = \text{var}\{\log(x_i/x_j)\} \quad (i = 1, \ldots, C - 1;\ j = i + 1, \ldots, C) \qquad (3.7)$$

The fact that the relative variance of two parts is the same within a subcomposition and within the full composition means that the relative variation diagram for any subcomposition is simply the subdiagram formed by the links of parts, or equivalently by the selection of vertices, associated with the subcomposition. Moreover the centre of the subcompositional diagram is, by Property 1, at the centroid, say O^*, of the subcompositional vertices. It is thus very simple to inspect visually within the full diagram the nature of any subcompositional variability. Indeed we could go further and estimate the proportion of the total compositional variability (Aitchison, 1986, Section 8.6) retained by the subcomposition. For example for the subcomposition formed from parts 1, . . . , C, this proportion is, by Property 5,

$$\sum_{i=1}^{C} |O^* i|^2 / \sum_{i=1}^{D} |Oi|^2 . \qquad (3.8)$$

(c) *Collinear vertices and constant logcontrasts* If a subset, say 1, . . . C, of vertices is approximately collinear then we know that the associated subcomposition has a relative variation diagram which is one-dimensional. Remembering the nature of the singular value decomposition we see that if we were to perform a principal component analysis on the set of such subcompositions we would find that only one eigenvalue was appreciably non-zero. An immediate implication therefore is that the subcompositional variability is one-dimensional and the nature of the one-dimensionality can be expressed as the constancy of logcontrasts, of the form

$$a_1 \log x_1 + \ldots + a_D \log x_D \quad (a_1 + \ldots a_D = 0). \qquad (3.9)$$

We shall see that this can give substantial insight into the nature of compositional variability.

(d) *Orthogonal links and subcompositional independence*. If two links ij and kl intersect at right angles then we see from Property 4 that

$$\text{corr}\{\log(x_i/x_j), \log(x_k/x_l)\} .= 0$$

and so the ratios x_i/x_j and x_k/x_l) are uncorrelated and, within the context of additive logistic normality (Aitchison, 1986, Chapter 6), independent. Thus in exploratory compositional data analysis the search for ratios which vary independently is associated with detecting orthogonal links. We may note that in this search for independent ratios i, j, k, l need not be different. It is for example meaningful to ask whether x_i/x_D and x_j/x_D are independent and in the relative variation diagram this would be associated with iDj being right-angled.

There is also no need to confine this search to a pair of ratios. If two subsets of vertices lie on two lines at right angles then the associated subcompositions are independent (Aitchison, 1986,Section 10.3) while showing a highly dependent structure within each subcomposition, because of the collinearity of vertices.

Compositional markers

The relative variation diagram may be regarded as part of a compositional biplot. To complete the picture we require compositional markers which allow a visual inspection of the relationship of each composition to the covariance structure of the compositional data set. To do this we use the nth row of U to plot the point $(N-1)^{1/2}(u_{n1}, u_{n2})$ as the marker representing the nth composition ($n = 1, \ldots, N$), as in Fig.3.2. Such markers have the easily established property that $On.Oi$ represents the departure of $\log(x_{ni}/x_{nj})$ from the average of this logratio over all the replicates. Let P and P_n denote the projections of the centre O and the compositional marker n on the possibly extended link ji. Then $On.ji = \pm|PP_n||ji|$, where the positive sign is taken if the directions of PP_n and ji are the same, otherwise the negative sign is taken. A simple interpretation can be obtained as follows. Consider the extended line ji as divided into positive and negative parts by the point P_n the positive part being in the direction of ji from . If P_n falls on the positive (negative) side of this line then the logratio $\log(x_{ni}/x_{nj})$ of the nth composition exceeds (falls short of) the average value of this logratio over all replicates and the further P_n is from P the greater is this exceedance (shortfall); if P_n coincides with P then the compositional logratio coincides with the average. In Fig. 3.2 the nth composition clearly has a logratio $\log(x_{ni}/x_{nj})$ which falls short of the overall average of this logratio.

A similar form of interpretation can be obtained from the fact that $On.Oi$ represents the departure of the centred logratio $\log\{x_{ni}/g(x_n)\}$ of the nth composition x_n from the average of this centred logratio over all replicates. In Fig. 3.2 let Q_n be the projection of the composition marker n on the possibly extended ray Oi. Then $On.Oi = \pm|OQ_n||Oi|$, the positive or negative sign depending on whether Q_n and the vertex i lie on the same side or opposite sides from O. We then have the following simple interpretation. If Q_n lies on the same (opposite) side of the divided line as the vertex i then the centred logratio $\log\{x_{ni}/g(x_n)\}$of the nth composition exceeds (falls short of) the average of this logratio over all replicates, and so we can infer that the ith component of the nth composition is higher (lower) than average relative to the other components. Obviously also the further Q_n is from O the greater is the divergence from the average.

Although $|On|^2 = (N-1)(u_{n1}^2 + u_{n2}^2)$ provides the Mahalanobis distance of the nth composition when $s_3 = s_4 = \ldots = 0$ we have found it as simple and more reliable to indicate possible outliers on the diagram by using the complete singular value decomposition and the exact Mahalanobis distance $q_n = (N-1)(u_{n1}^2 + \ldots + u_{nd}^2)$. With this distance it is simple to compute the atypicality index of any composition (roughly the probability that a future composition will be more typical or have a smaller Mahalanobis distance than the considered composition). To avoid resubstitution bias the standard missing-one-out technique (Aitchison, 1986, p.175) is recommended and the atypicality index of a composition with Mahalanobis distance q computed from the singular value decomposition for the *full* data set can be shown to be

$$Be(qN/(N-1)^2, \tfrac{1}{2}(D-1), \tfrac{1}{2}(N-D)) \qquad (3.10)$$

where $\text{Be}(t;\, a,\, b)$ is the distribution function of a beta distribution of the first kind with parameters a and .b

Differences between unconstrained and compositional biplots

It must be clear from the above aspects of interpretation that the fundamental elements of a relative variation diagram are the links, not the rays as in the case of variation diagrams for unconstrained multivariate data. The complete set of links, by specifying all the relative variances, determines the compositional covariance structure and provides direct information about subcompositional variability and independence. It is also obvious that interpretation of the relative variation diagram is concerned with its internal geometry and would, for example, be unaffected by any rotation or indeed mirror-imaging of the diagram.

Another fundamental difference between the practice of biplots for unconstrained and compositional data is in the use of data scaling. For unconstrained data if there are substantial differences in the variances of the components, biplot approximation may concentrate its effort on capturing the nature of the variability of the most variable components and fail to provide any picture of the pattern of variability within the less variable components. Since such differences in variances may simply arise because of scales of measurement a common technique in such biplot applications is to apply some form of individual scaling to the components of the unconstrained vectors prior to application of the singular value decomposition. No such individual scaling is necessary for compositional data when the analysis involves logratio transformations. Indeed, since for any set of constants $(c_1, \ldots, c_D)$, we have

$$\text{cov}\{\log(c_i x_i / c_j x_j), \log(c_k x_k / c_l x_l)\} = \text{corr}\{\log(x_i / x_j), \log(x_k / x_l)\} \qquad (3.11)$$

it is obvious that the covariance structure and therefore the relative variation diagram are unchanged by any differential scaling or 'perturbation' (Aitchison, 1986, Section 2.8) of the parts. Only the centering process is affected by such differential scaling. Moreover any attempt at differential scaling of the *logratios* of the components would be equivalent to applying differential power transformations to the *components* of the compositions, a distortion which would prevent any compositional interpretation from the resulting diagram. It is perhaps worth pointing out here that, even for unconstrained multivariate data consisting of positive vectors, there is an advantage in the use of the logarithmic transformation, since biplots are then invariant to changes of scale because

$$\text{cov}\{\log(c_i x_i), \log(c_j x_j)\} = \text{cov}\{\log x_i, \log x_j\} \qquad (3.12)$$

for any constants c_i , c_j .

Biplots for bicompositions

In compositional data analysis a number of problems arise where the vector associated with the experimental unit may consist of two compositions. For example, we shall require in our study of two conditional compositional biplot applications in Sections 5 and 6 to consider two such bicompositions, in Section 5 the ethnic and urban-rural compositions of 50 US states, and in Section 6 the mineral and major-oxide compositions of 21 tektites. Later in this Section we shall display in Figs. 3.3 and 3.4 the bicompositional biplots for these two data sets.

As we have seen in earlier in this Section biplots for compositional variability are not beset by the problems of scaling encountered in unconstrained variability. For this reason it is worth considering whether it is possible to obtain within one biplot a picture of the individual variation of each composition and also their joint variability or association. The relative variation diagram may be readily adapted to such a data set, simply by the construction of row- and column-centred logratio data matrices Z_1 and Z_2 for each of the two compositional data sets, and then application of the singular value decomposition to the partitioned vector

$$Z = [Z_1 \;\; Z_2] = US[V_1^T \;\; V_2^T]. \tag{3.13}$$

The columns of V_1^T and V_2^T then refer to the parts of the first and second compositions and the construction of the relative variation diagram proceeds exactly as described earlier Moreover the fact that the column sums of V_1 and V_2 are both zero ensures that the centre O of the diagram is the centroid separately of each of the compositional part vectors. Indeed within this one diagram we have approximations to the separate relative variation diagrams for each of the two compositional data sets.

Ethnic and urban-rural compositions of US states

For each of 50 states of the USA the 5-part ethnic composition [white, black, native American, hispanic, others including asian], abbrecviated to [wh, bl, na, hi, ot],nd the 2-part (urban, rural), abbreviated to [ur, ru], compositions are available. This data set excludes Washington D.C. which is completely urban. Treating these as 5×2 bicompositions we obtain the joint biplot of Fig. 3.3 retaining 85.0 per cent of the total variability. The subplots for the ethnic and urban-rural compositions are virtually identical to biplots that would arise from separate constuction of biplots for each set of compositions..

A notable feature of the ethnic plot is the almost coincidence of the hispanic and other vertices hi and ot, indicating a low relative variation 0.47 compared with the largest relative variation 4.87 between native American and black, associated with the longest link na-bl. We may also note some links intersecting at approximately right angles, for example hi-wh (or ot-wh) and na-bl indicating almost zero correlation, namely 0.041 (or 0.042) between the logratios of hi/wh (or ot/wh) and na/bl. Similarly links na-ot and bl-ot intersect at approximately right angles again corresponding to near-zero correlation (0.023) between the associated logratios.

In considering the joint variability of the ethnic and urban-rural compositions we note that the hi-wh link makes an acute angle with the ur-ru link corresponding to high positive correlation (actually 0.74) between the logratios hi/wh and ur/ru. On the other hand the na-wh link is almost perpendicular to the ur-ru link corresponding to almost zero correlation (in fact 0.05) between the associated logratios.

The composition markers for the states fall into well-defined clusters according to ethnic ratios. For example, California (CA), Hawaii (HA), New Mexico (NM) and Texas (TX) with the highest hispanic and/or asian proportions are clearly associated with the hi and ot vertices, whereas Alabama (AL), Georgia (GE), Louisiana (LO), Mississippi (MI) and South Carolina (SC) with relatively the largest black populations are clearly associated with the bl vertex. Also the highly urban California is well out on the ru-ur link. The reader will be able to identify other interesting ethnic and urban-rural features of the biplot. Of the 50 states four have bicompositional atypicality indices greater than 0.95, Nebraska (NE) (0.998), Hawaii (0.996), New Mexico (0.97), Massachusetts (MA) (0.958). The reasons for these registering as atypical states are readily determined: Nebraska has the relatively lowest hispanic proportion despite its high urban-rural ratio, Hawaii has by far the highest other (asian) proportion, New Mexico has by far the highest hispanic proportion, and Massachusetts has the relatively lowest native American proportion.

Mineral and major-oxide compositions fo tektites

For an interesting comparison between bicompositional and conditional biplots we provide the bicomposition for mineral and major oxide compositions of 21 tektites in Fig. 3.4, See Section 6 for full discussion of this application.

4 Conditional biplots for bicompositions

In many studies involving compositions, interest may lie in the nature of the dependence of one composition (the response composition) on another (the covariate composition). For example, in the study of US Presidential elections interest may be in how the three-part composition, the percentage division among the three candidates, of the 1992 presidential vote in each state depends on the ethnic composition of the state.

Suppose that interest is in the dependence of a D_2-part composition on a D_1-part composition and that the corresponding compositional data sets based on N cases are the response compositional matrix X_2 and the covariate compositional matrix X_1. A variety of biplot techniques related to reduced rank regression, canonical correlation analysis and redundancy analysis designed to bring out conditional features of unconstrained data sets are well established. Of these, we adapt here the version of Ter Braak (1990, 1991), which is intermediate between redundancy analysis and canonical correlation analysis. We first summarise the necessary adaptation of the computations and plotting to compositional data sets. In these computations Moore-Penrose pseudo-inverses are used.

Computations

1 Compute the centred logratio data matrices Z_1, Z_2 corresponding to X_1, X_2 as defined in (3.2).

2 Compute the estimated centred logratio covariance matrices

$$\Gamma_{ij} = Z_i^T Z_j \,/\, (N-1) \quad (i, j = 1, 2).$$

3 Obtain the singular value decomposition of $\Gamma_{11}^{-1/2}\Gamma_{12}\Gamma_{22}^{-1/2}$ as PSQ^T.

4 Evaluate $U = \Gamma_{11}^{1/2}P.\ V = \Gamma_{22}^{1/2}Q$.

5 Evaluate $A = \Gamma_{11}^{-1}U$.

6 Evaluate $M = Z_1 A$.

We shall find the following easily proved relationships involving the above constructs important in the interpretation of the various biplots we shall construct.

(a) $UU^T = \Gamma_{11}$,

(b) $MU^T = Z$,

(c) $USV^T = \Gamma_{12}$,

(d) $MSV^T = \tilde{Z}_1$, the fitted values of the centred logratio responses in a regression on the covariate centred logratios.

(e) $ASV^T = \tilde{\Theta}$, the regression coefficients in the regression described in (d).

As a basis for applications we first describe here what may be termed a full conditional biplot which uses the order 2 approximations from the above computations, in which case relationships (a)-(e) take an approximate form, as in unconditional biplots,.

Construction of the standard conditional biplot

Refer to Fig 4.1.

1 Corresponding to each of the parts of the covariate composition, construct a *vertex i.* (*i = 1, . . . , D_1*) with coordinates (u_{i1}, u_{i2}).

2 Corresponding to each of the parts of the response compositions construct a *vertex k* (k = 1, . . . , D_2) with coordinates $(s_1 v_{k1}, s_2 v_{k2})$.

3 Corresponding to each case or bicomposition construct a case or bicompositional marker n (n = 1, . . . , N) with coordinates (m_{n1}, m_{n2}) from the first two columns of M.

4 Corresponding to each of the parts of the covariate composition construct a *secondary vertex I* (I = 1, . . . , D_1) with coordinates (a_{I1}, a_{I2}) from the first two columns of A. .

First, the relationships (a) and (b) mean that the vertices associated with the parts of the covariate composition and the markers provide a complete biplot for the covariate compositions of the kind we have already considered in Section 3. Secondly relationship (c) gives the opportunity of investigating the dependence of the response composition on the covariate composition within the conditional biplot. For example, with indices i, j referring to the covariate and indices k, l referring to the response compositions, the scalar product $\vec{ij}.\vec{kl}$ and $\cos(iLj)$ provide estimates of $\mathrm{cov}\{\log(x_{1i}/x_{1j}),\log(x_{2k}/x_{2l})\}$ and $corr\{\log(x_{1i}/x_{1j}),\log(x_{2k}/x_{2l})\}$, respectively, where L is the intersection of ij and kl. Thirdly from relationship (d) we have that $\vec{On}.\vec{Ok}$ provides an estimate of the amount by which the fitted value of the centred response logratio $\log\{x_{2k}/g(x_2)\}$ exceeds the average over all the cases of the fitted values of that centred logratio. A simpler and more useful form of this is in terms of simple logratios with $\vec{On}.\vec{kl}$ estimating the amount by which the fitted value of the logratio $\log(x_{2k}/x_{2l})$ exceeds the average over all the cases of the fitted values of that logratio. We emphasize here that this conditional biplot does not allow the reconstruction of the response composition but only of the fitted response composition. Fourthly from (e) we see that the scalar products $\vec{LI}.\vec{Ok}$ provide approximations to the regression coefficients in a regression of the centred logratio responses on the covariate centred logratios. Finally we note that since this conditional biplot is based on the singular value decomposition of the mixed covariance matrix Γ_{12} the proportion of total variation retained by the biplot is related to this covariance matrix.

The application of these results to the following examples will help to illustrate their use.

.

5 Application to the 1992 US Presidential Election

For each of the 50 states of the US the proportions of votes cast for the three Presidential candidates Bush, Clinton and Perot form a three-part composition, abbreviated to [bu, cl, pe]. The question we now pose is whether we can obtain insights into the possible dependence of this voting pattern on the ethnic and urban-rural compositions of the states through the construction of the conditional biplot as described above. In this construction the covariate is the bicomposition consisting of the ethnic and urban-rural compositions and so consists of the two centred logratio matrices as described in Section 3.

There is a danger in conditional biplot use of crowding too many features into a single diagram, such as the full conditional biplot. We shall avoid this here by using two different graphical representation to bring out features of the dependence of the voting composition on the ethnic and urban-rural compositions. In Fig. 4.2 we plot a vertex for each of the voting parts as in Step 2 of the construction of the standard biplot above. Then for each of the covariate parts we plot a supplementary vertex as indicated in Step 4. Then a simplified version of relationship (e) allows us to interpret any scalar product of a voting link and an ethnic or urban-rural link as the relevant regression coefficient. As indicated in Section 3 logratio transformations have a stabilising effect on variability and in this example the ranges of values of all the logratios are fairly similar. This means that the effect of any of the ethnic and urban-rural logratios depends substantially on the appropriate regression coefficients. For example, the scalar product of the Bush-Perot link bu-pe and the white-black link wh-bl is -|wh-bl| × (projection of bu-pe on wh-bl) since the angle of intersection of bu-pe and wh-bl is obtuse and this is obviously large relative to other such scalar products. The inference therefore is that the larger the black-white ratio is in a state the greater the advantage is likely to be towards Bush away from Perot. Similarly consideration of the bu-pe and ur-ru links shows that Perot's better performances relative to Bush are in states with higher urban to rural ratios, whereas the almost perpendicular bu-cl and ur-ru links show that the relative Clinton-Bush vote is hardly affected by the urban-rural composition.

We can report here that comparison of the signs and magnitudes of all the 33 possible scalar products with those which can be computed by standard compositional regression methods (Aitchison, 1986 Chapter 12) shows agreement in all except one, the bu-pe and wh-ot links where the scalar product is negative whereas the true regression coefficient is slightly positive. The overall general picture is that Perot should do relatively badly compared with Bush and Clinton in states with higher black proportions but should fare better in states with a higher urban-rural ratio. Moreover Clinton fares better relative to Bush in states with higher proportions in the category other. Readers may easily work out other implications for themselves.

Other forms of the conditional biplot can be used to bring out different aspects of the conditional modelling but will not be pursued here.

The quality of the biplot can be measured in terms of the Frobenius norm variation retained by the second order approximation used by the biplot. For the centred logratio covariance matrix Γ_{12} on which the singular value decomposition is based and which has rank 2 the proportion retained is clearly 1; similarly for the fitted response matrix Z_2 the proportion retained is 1. For the covariate centred data matrix Z_1 and the covariate centred logratio covariance matrix Γ_{11} the proportions are 0.40 and 0.61, respectively. .

This departure from overall pattern raises the question of how we may handle the equivalent of examination of the errors in a standard unconstrained multivariate regression analysis. The answer is to be found in terms of the perturbation which measures the change between actual composition, say x, and a fitted composition, say $\tilde{x}$. This can be expressed as the perturbation

$$p = \tilde{x} \ominus x = C[\tilde{x}_1 / x_1, \ldots, \tilde{x}_D / x_D]. \qquad (4.1)$$

The set of these perturbations then plays a role similar to the error vectors in unconstrained multivariate regression analysis. If we then construct in Fig. 4.3 a simple compositional biplot of these perturbations along the lines of Sections 3 we can see in comparison with Fig.. 4.2 the direction of the departures from the general pattern. For example, Utah is atypical in favouring Bush despite having small black/white and large urban-rural ratios, which according to the general pattern as seen in Fig.4.2 would favour Perot. The Hawaii perturbation indicates that the fitted pattern [0.33, 0.46, 0.21] is incorrect despite Hawaii's dominant other proportion and high urban/rural ratio and requires adjustment in favour of Bush and Clinton to arrive at the actual vote [0.37, 0.49, 0.14]. In a similar way the Tennessee perturbation takes a fitted vote [0.41, 0.44, 0.15] to an actual vote [0.43, 0.47, 0.10]. We may note also the substantial departure of Clinton's home state Arkansas from patterned support for Bush towards the home candidate.

6 Application relating tektite mineral compositions to major oxide compositions

As a second example to illustrate conditional compositional biplot techniques and to provide some unusual features which require care in interpretation we consider a data set for 21 tektites (Chao, 1963; Miesch et al, 1966), for which the two compositions are 8-part major-oxide compositions and 8-part mineral compositions. These are subcompositions of the original data set, this reduction being adopted for the sake of simpler exposition. While experimentally these two types of compositions are determined by completely different processes they are obviously chemically related since the minerals are themselves more complicated major oxide compounds. The challenge of the conditional biplot of Fig. 4.4 with mineral composition as the response and major-oxide composition as the covariate, is whether it can at least identify these relationships from the compositional data alone, without any additional information about the chemical formulae of the minerals, and hopefully provide other meaningful interpretations of the data.

Table 4.1. Oxides and associated minerals in the tektite study

Oxide	Mineral	Abbreviation	Formula
SiO_2	quartz	qu	SiO_2
K_2O	orthoclase	or	$KAl Si_3O_8$
Na_2O	albite	al	$NaAlSi_3O_8$
CaO	anorthite	an	$CaAl_2Si_3O_8$
MgO	enstatitie	en	$MgSiO_3$
Fe_2O_3	magnetite	ma	Fe_3SiO_4
Ti	ilmenite	il	$FeTiO_3$
P_2O_5	apatite	ap	$Ca_3(Fe,Cl)(PO_4)_3$

A striking feature of the diagram is that it is indeed successful in identifying which oxides are associated with which minerals. From Table 4.1, which provides the chemical association between minerals and major oxides, we see that, apart from SiO_2, each of the other seven major oxides is associated with only one of the minerals, for example MgO is contained only in enstatite. In the biplot diagram each of these seven major oxide vertices is close to its corresponding mineral vertex. This means that the link associated with any two of these major oxides is nearly parallel to the link of the corresponding minerals and so the mineral logratios are all highly correlated with the corresponding major oxide logratios. It is in this sense that the conditional biplot identifies the chemical relationships. Moreover even SiO_2, which is a constituent of all eight minerals is nevertheless primarily identified with quartz which is simply its oxide self.

All of this seems splendid until the quality of the approximation is investigated. The proportion of the covariance matrix Σ_{12} which is retained by the biplot is only 0.204. The reason for this is to be found in Step 3 of the computations. The singular value decomposition has singular values 1.00, 1.00, 1.00. 0.999, 0.994. 0.868, 0,060 and it would require a fourth order approximation and a four-dimensional biplot to raise the quality to a reasonable 0.911 proportion retained. The reason for this disappointing quality is easily determined. It lies in the fact that within the constraints of compositional data each mineral is almost independently related to its major oxide, in the sense that each mineral logratio is almost perfectly linearly related to the corresponding major-oxide ratio. An analogous situation with unconstrained data would be the assemblage of independent univariate regressions, each with a different response and different covariate, into a multivariate regression. The apparent success of the conditional biplot lies more in the strength of the individual logratio regressions than in the quality of the biplot. It is important here to distinguish between the quality of the biplot and the reliability of the logratio regression of mineral on major oxide composition. The proportion of the mineral variability explained by the regression can be shown to be 0.983.

In such circumstances it seems worth considering how the bicompositional biplot of Section 3 fares. Fig. 3.4 provides such a biplot, with the centred logratio data matrix $[Z_1\ Z_2]$ having 0.861 of its variability retained. The corresponding proportions retained by the second order approximation are for Γ_{11} 0.860, for Γ_{12} 0.874, for Γ_{22} 0.887. It is also clear that Fig. 3.4 picks out the chemical relationships between minerals and major oxides as firmly as in Fig. 4.4. It certainly seems in this case that rather than place the emphasis of the biplot approximation on the conditional aspects it is as effective to consider the joint biplot as the means of investigating dependence of response composition on covariate composition.

Fig.3.4 can be used to investigate a theory put forward by Miesch et al (1966) of the formation of tektites which they identify as the independence of the set of three ratios (Fe_2O_3, MgO, P_2O_5) / SiO_2 from the set of three ratios (CaO, Na_2O, K_2O) / SiO_2. This in turn implies that the subcompositions

$$[Fe_2O_3, MgO, P_2O_5],\ [CaO, Na_2O, K_2O]$$

must be independent and we would then be disappointed in failing to identify the necessary approximate right angles within Fig.3.4.. For example, the Fe_2O - P_2O_5 and CaO - Na_2O links are approximately parallel instead of orthogonal, throwing considerable doubt on the theory. Rejection of the theory can, of course, be confirmed by a full statistical analysis testing the hypothesis of independence of the above two subcompositions by the procedure described in Aitchison (1986, Section 10.3) for which the significance probability is 0.04.

7 Discussion

We have not treated many of the finer points of biplots which appear in the ever-growing literature on the subject. We have preferred to keep the diagrams simple in the hope that those faced with compositional data analysis may see them as a convincing case for the use of the logratio form of analysis. Although we have not given any examples it is clear that conditional compositional biplots may be readily adapted to situations where either the response or the covariate consists of a multivariate vector of unconstrained measurements. All that is required is that such vectors be centred as for standard unconstrained biplots.

The question of what constitutes a satisfactory degree of approximation in compositional biplot use is, as with unconstrained biplots, a difficult one to answer with any objectivity. Our own view is that biplots are no substitute for relevant statistical analysis but often play an important role in allowing a simple visual exposition of the findings of the analysis. For this purpose we would be reluctant to use a biplot which retained less than a proportion 0.75 of the total variability of the data set. A lesser degree of approximation, as in the conditional biplot of Fig. 4.4, is an indication that a higher order approximation is required. In such a situation there is the option of attempting a three-dimensional representation or the use of the (1, 2) order biplot together with the (1, 3) order biplot, as is commonly reported in the use of the first three principal components in a principal component analysis.

Two main distinctions between compositional and unconstrained biplots should be reemphasized. First, in order to ensure simple valid ratio comparisons links rather than rays are the basic components of biplots. Secondly, scaling transformations have no role to play in the construction of biplots after the basic logratio transformation has been performed. Indeed the logratio transformation has a stabilising effect on component variability which ensures that the singular value decomposition, unlike its action in unconstrained variability, does not concentrate on some aspects of the variability at the expense of others.

There are obviously many other situations where conditional biplots may prove useful if only for expository purposes. For example it would be of interest to compare analyses of later US Presidential Elections with that of 1992 discussed in Section 5. Also in the recent 2005 UK elections there was considerable interest in the nature of the ‘swing’ from party to party within constituencies, relative to the 2001 election, particularly in the main parties [Conservative, Labour, Liberal]. Such swings are essentially perturbations and so interest might be directed towards interpretation of how a conditional biplot relates these perturbations to other covariates descriptive of the constituencies.

Acknowledgement

We are grateful to Diane Lightey, who as an M. Sc student at the University of Virginia under the supervision of one of us performed an early and less detailed analysis of the 1992 US Presidential election data.

References

Aitchison, J. (1982). The statistical analysis of compositional data (with discussion) *Journal of the Royal Statistical Society* **B44**, 139-177.

Aitchison, J. (1986). *The statistical analysis of compositional data.* Chapman and Hall: London, 416 p. Reprinted in 2003 with extensions by The Blackburn Press, Caldwell, NJ, USA..

Aitchison, J. (1990). Relative variation diagrams for describing patterns of variability of compositional data. *Mathematical Geology* **22**, 487-512.

Aitchison, J. (2001). Simplicial inference. In *Algebraic Methods in Statistics and Probability* (M. A. G. Viana and D. St. P. Richards, eds), pp. 1-22. Contemporary Mathematics Series 287. Providence, Rhode Island: American Mathematical Society.

Aitchison, J., Barceló-Vidal, C., Ecozcue, J. J. and Pawlowsky-Glahn, V. (2002). concise guide to the algebraic-geometric structure of the simplex, the sample space for compositional data analysis: *Terra Nostra, sp*ecial issue: Eighth Annual Conference of the International Association for Mathematical Geology **3**, 387-392.

Aitchison, J. and Greenacre, M. (2002). Biplots for compositional data. *Applied Statistics*. **51**, 375-382.

Aitchison, J. and Ng, K.W. (2005). The role of perturbation in compositional data analysis. *Statistical Modelling* **5**, 173-185.

Chao, E.C.T. (1963). The petrographic and chemical charactersitics of tektites. In *Tektites* (J.A. O'Keefe, ed.), pp. 51-64. Chicago University Press.

Eckart, C. and Young, G. (1936). The approximation of one matrix by another of lower rank. *Psychometrika* **1**, 211-218.

Miesch, A.T., Chao, E.C.T. and Cuttitta, F. (1966). Multivariate analysis of geochemical data on tektites. *Journal of Geology* **74**, 673-691.

Gabriel, K.R. (1971).. The biplot-graphic display of matrices with application to principal component analysis. *Biometrika* **58**, 453-467.

Gabriel, K.R. (1981). Biplot display of multivariate matrices for inspection of data and diagnosis. In *Interpreting Multivariate Data* (V. Barnett, ed.), pp. 147-173. New York: Wiley.

Good, I.J. (1969). Some applications of the singular decomposition of a matrix. . *Technometrics* **11**, 828-831.

Ter Braak, C.J.F. (1990). Interpreting canonical correlation analysis through biplots of structural correlations and weights. *Psychometrika* **55**, 519-531.

Ter Braak, C.J.F. (1991). Biplots in reduced-rank regression. *Biometrical Journal* **8**, 983-1003.

Whittle, P. (1952). On principal component analysis and least square methods of factor analysis. *Skandinavisk Aktuaritidskrift* **35**, 223-239.

Fig 3.1 The basic elements of a compositional biplot

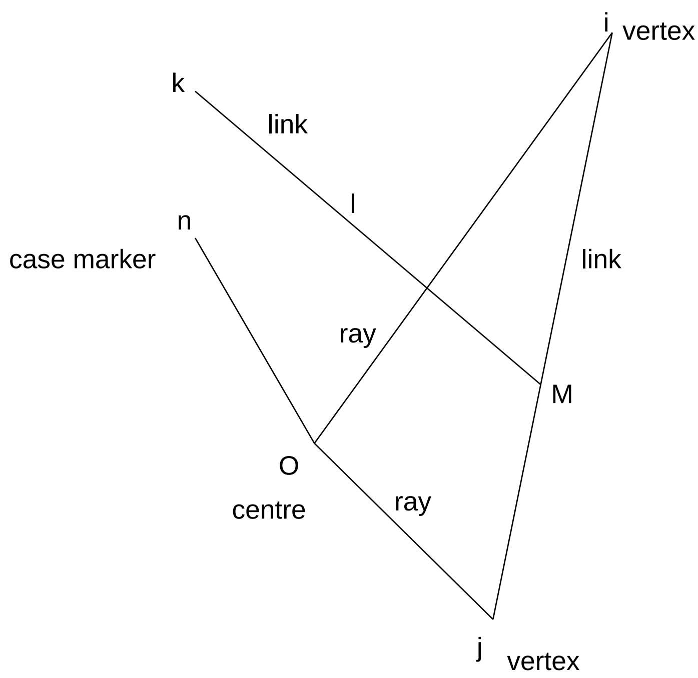

Fig. 3.2 Interpretation of case markers in a compositional biplot

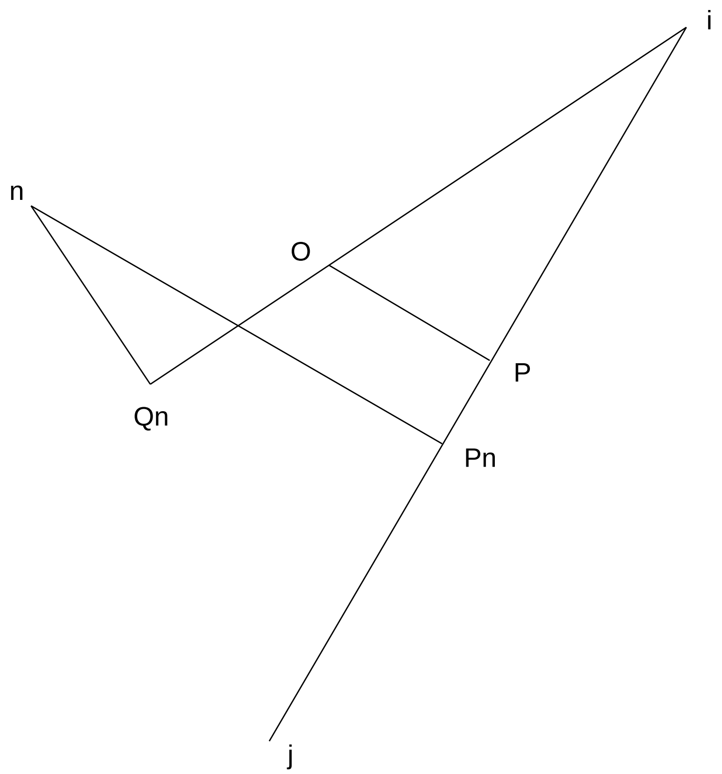

Fig 3.3 Bicompositional biplot of ethnic and urban−rural compositions of 50 US states

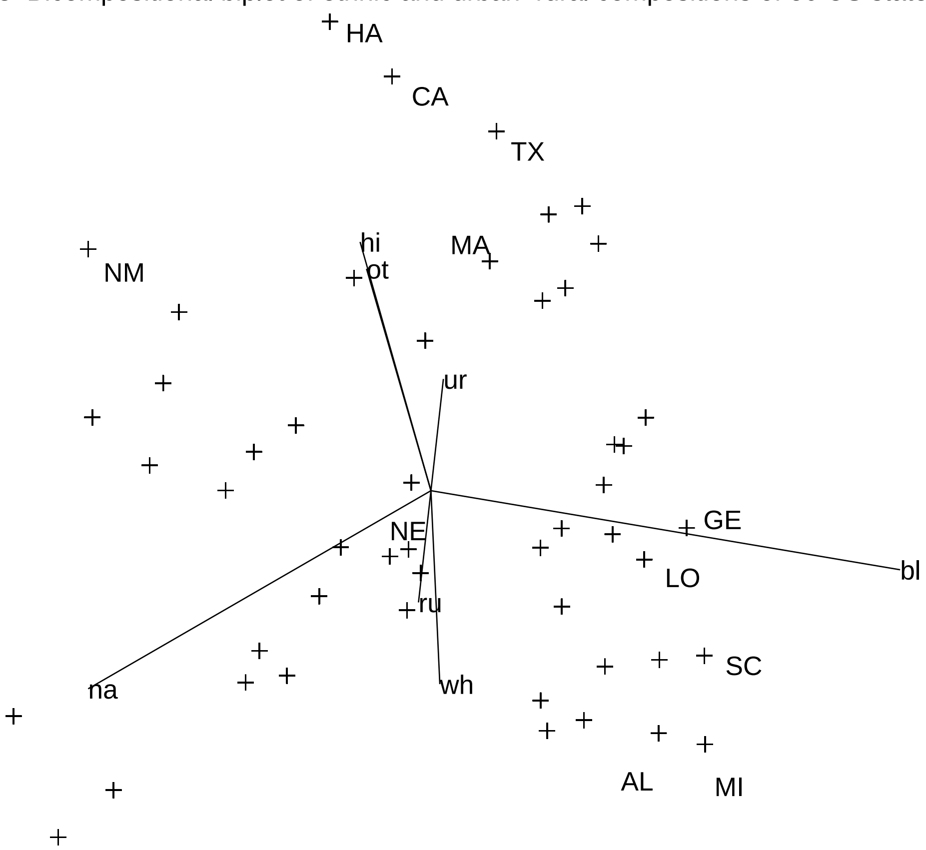

Fig. 3.4 .Bicompositional biplot of mineral and major−oxide compositions of 21 tektites

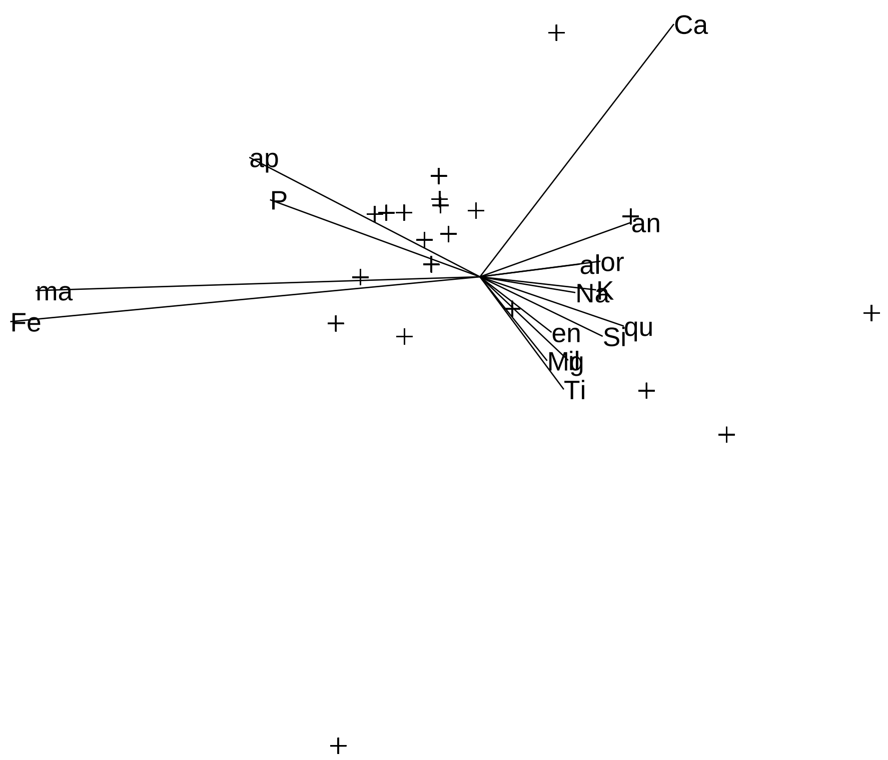

Fig. 4.1 Basic features of the stndarcd conditional biplot

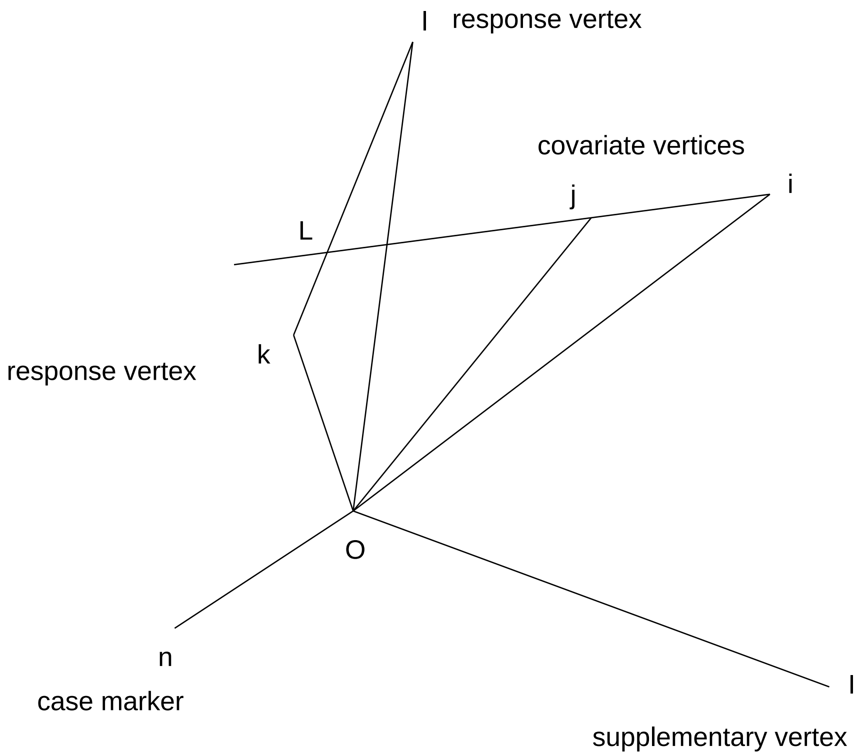

Fig. 4.2 Conditional biplot showing dependence of vote on ethnic and urban−rural compositions

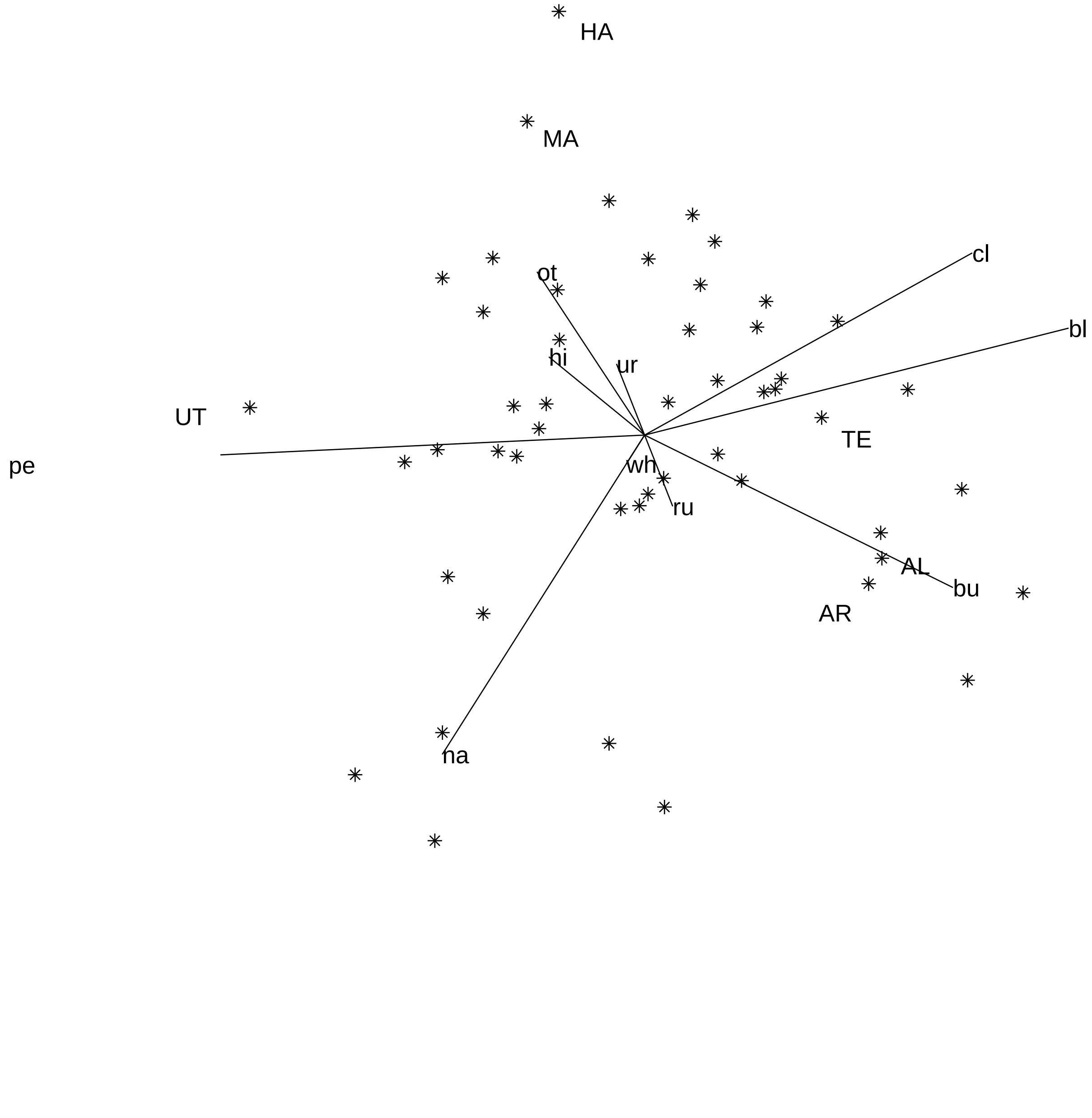

Fig 4.3 Biplot of conditional error perturbations for US 1992 Presidential election

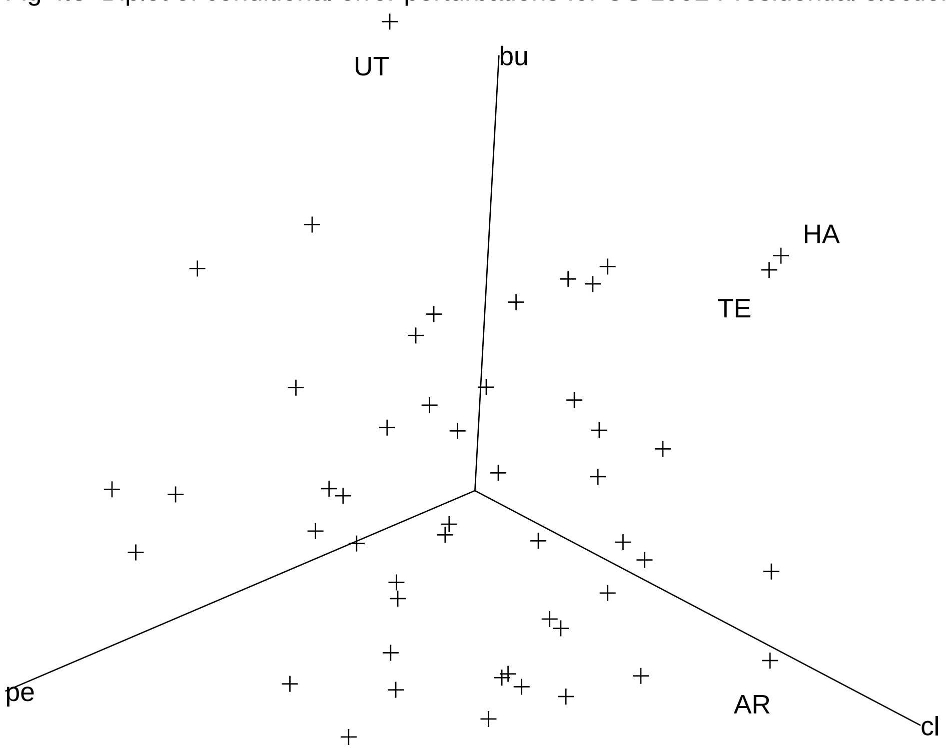

Fig. 4.4 Conditional biplot of mineral on major oxide compositions for 21 tektites