Applied Mathematics & Information Sciences **4**(1) (2010), 15–30
– An International Journal

Formal Analysis and Specification of the Triangular Format Graph

Michael D. Vose[1] and Donald W. Hastings[2]

[1]EECS Department, The University of Tennessee, 203 Claxton Complex,
Knoxville, TN 37996-3450, U. S. A.
Email Address: vose@eecs.utk.edu

[2]Sociology Department, 901 Mc Clung Tower, The University of Tennessee
Knoxville, TN 37996-0490, U. S. A.
Email Address: dhasting@utk.edu

Received July 25, 2008; Revised March 12, 2009

The triangular format graph was originally introduced as a means to visualize specialized data sets, most commonly arising in the social sciences. The graphing paradigm is defined, formalized, and extended to obtain a general method for displaying 4-dimensional data sets whose points may be approximated by a hyperplane (distortions introduced by the graphing process vanish as the approximation improves). The graphing methodology is also analyzed to determine what, geometrically speaking, the distortions introduced by the graphing process represent.

Keywords: Graphing, 4-dimensional data, geometric distortion.

1 Introduction

This paper explains the triangular format graph, stressing its mathematical underpinnings and considering its inherent distortions. The presentation takes a layered approach, where the graphing process is gradually revised and refined (over the following five sections) to finally take the form summarized in the conclusion (readers interested only in the ultimate formal specification can skip to the conclusion). The final result is a general graphing method suitable for arbitrary 4-dimensional data sets whose points may be approximated by a hyperplane (distortions introduced by the graphing process vanish as the approximation improves). The reason for leading the reader through a layered development is to motivate and explain what would otherwise seem to be the strange and arbitrary objects appearing in the formal definition of the graphing process.

This work was partially supported by National Institutes of Health grant "New approaches to the modeling of speciation" (GM56693).

The triangular format graph has historically been an aid to visualization in the social sciences — sometimes called quasi-scenic rendering [2, 3] — serving two purposes. First, images engage researchers and allow them to develop a gestalt understanding of how data are structured. Second, images provide the appropriate cues for researchers to select which family of functions may be applied to best interpret the data's structure, and where necessary assist "the researcher in the search for plausible decomposition of effects" [15].

The most common graphing paradigm is the traditional orthogonal x-y-z axis format, which is sometimes unsatisfactory; when displaying an age-period-cohort (APC) data set where a dependent variable is referenced by three time coordinates (age A, calendar year P, and birth year C) a format other than the orthogonal may be more cognitively appealing to the researcher. Such data sets naturally arise in demography, history, and developmental psychology. The traditional method would plot the dependent variable on the z axis and two of the three time referent coordinates on the x and y axes. Unfortunately, to discern a possible response effect for the third time coordinate the viewer must envision a family of curves in the plotted surface situated obliquely with respect to the x and y axes. Under the x-y-z orthogonal format, only two of the three time reference coordinates may be selected; each pair selected in effect rotates the third reference axis out of the user's direct field of observation. Thus three graphs are necessary to examine fully each of the three time referent profiles of the dependent variable. (For general discussions on the logic of cohort analysis see [9–11, 20], or [21, 22], and for discussions on graphing techniques of APC data sets see [5–7, 15–18, 24]).

As early as 1975, Norman Ryder [21] suggested that cohort data might better be presented employing an equilateral triangle whose sides represent age, period, and cohort, respectively. Informally, one side denotes cohorts C, another side marks periods P, and the triangle base references age A. Together the APC coordinates identify a point within the triangle above which the dependent variable is represented as a height. By rotating the displayed graphic, a user can directly observe each of the age, period, and cohort profiles for the dependent variable, because each reference axis or combination of axes is treated geometrically in an equivalent manner and all are in the user's direct field of observation.

More formally, a *cohort data set* may be defined as a collection of 4-tuples $\langle w, x, y, z\rangle$ which satisfy $w + x = y$. The triangular format graph of a cohort data set is obtained as follows (see figure 1.1). Point $\langle w', x', y', z'\rangle$ is plotted at a height of z' above a point $q(w', x', y')$ located in a equilateral triangle in the plane. One side of the triangle (the side with slope zero) represents the "w axis", another side of the triangle (the side with negative slope) represents the "x axis", and the remaining side (with positive slope) represents the "y axis". The point $q(w', x', y')$ is the point of intersection of the following three lines: the first line is parallel to the x axis and intersects the w axis at coordinate w', the second line is parallel to the y axis and intersects the x axis at coordinate x', the third line is parallel to the w axis and intersects the y axis at coordinate y'.

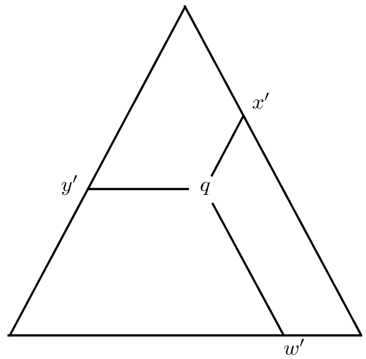

Figure 1.1: Point q with reference coordinates $\langle w', x', y' \rangle$.

The historical roots described above are taken as the starting point for formalizing the triangular format graph using vector and matrix algebra. The mathematical formalization serves as the basis from which the graphing paradigm is extended to a more general and useful method for displaying 4-dimensional data sets (which are not necessarily cohort data sets), and allows the analysis of what distortions are introduced by the triangular format graphing process. The following sections incrementally revise, generalize, and refine the method, ultimately yielding a specification for computer implementation.

2 Formalization

Formalizing the triangular format graph begins with a representation for the equilateral triangle Λ above which points are plotted; it is the convex hull of the unit basis vectors in $\Re^3$

$$
\begin{aligned}
e_0 &= \langle 1, 0, 0 \rangle, \\
e_1 &= \langle 0, 1, 0 \rangle, \\
e_2 &= \langle 0, 0, 1 \rangle.
\end{aligned}
$$

The bold triangle (and its interior) in figure 2.2 depicts Λ (the diagram shows the coordinate axes which extend beyond e_0, e_1, e_2).

The point $q(v_0, v_1, v_2)$ with reference coordinates v_0, v_1, v_2 is given by the intersection of three lines l_0, l_1, l_2 defined as follows. Let p_0 on edge $e_1 - e_0$ be determined by reference coordinate v_0 according to

$$p_0 = (1 - v_0)e_0 + v_0 e_1$$

and define the line

$$l_0(t_0) = p_0 + t_0(e_2 - e_1).$$

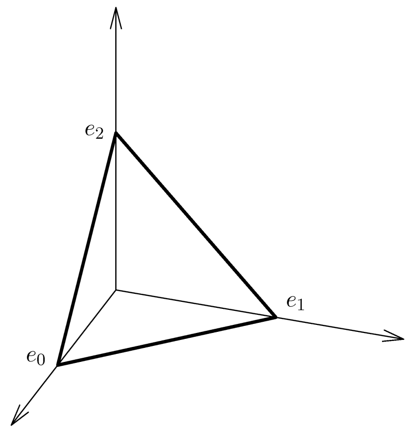

Figure 2.2: Domain for the triangular format graph.

Applying the following mapping to subscripts

$$i \mapsto i+1 \quad \text{modulo } 3$$

transforms the definitions of p_0 and $l_0(t_0)$ given above into definitions for p_1 and $l_1(t_1)$, a second application of the mapping yields the definitions for p_2 and $l_2(t_2)$. The intersection of l_0 with l_1 is determined by the equation

$$p_0 + t_0(e_2 - e_1) = p_1 + t_1(e_0 - e_2).$$

Applying the subscript mapping as before yields the other intersections, which leads to

$$\begin{aligned} l_0 \cap l_1 &= \langle \quad 1 - v_0 \quad , \quad 1 - v_1 \quad , \quad v_1 + v_0 - 1 \quad \rangle, \\ l_1 \cap l_2 &= \langle \quad v_2 + v_1 - 1 \quad , \quad 1 - v_1 \quad , \quad 1 - v_2 \quad \rangle, \\ l_2 \cap l_0 &= \langle \quad 1 - v_0 \quad , \quad v_0 + v_2 - 1 \quad , \quad 1 - v_2 \quad \rangle. \end{aligned}$$

Point $q(v)$ with reference coordinates v_0, v_1, v_2 is the common intersection displayed above. Using reference coordinates implicitly assumes $v_0 + v_1 + v_2 = 2$ since otherwise the lines would not have common intersection. It follows that

$$q(v) = \mathbf{1} - v,$$

where $\mathbf{1}$ is the vector $\langle 1, 1, 1 \rangle$.

The triangular graphing process as applied to a data point $\langle v, z \rangle = \langle w, x, y, z \rangle$ (hence $v = \langle w, x, y \rangle$) assumes that $l_0 \cap l_1 = l_1 \cap l_2 = l_2 \cap l_0$ which is necessary and sufficient for $q(v)$ to be well defined. Data point $\langle v, z \rangle$ is plotted as

$$q(v) + \mathbf{1}\frac{z}{\sqrt{3}}.$$

Note that this does not restrict q to Λ (it is inside if and only if $0 \le v_i \le 1$), and so a more general graphing paradigm than what was described in the introduction has been obtained. Adopt the convention that angle brackets $\langle \cdots \rangle$ are used to denote tuples which are regarded as column vectors. The assumption concerning the common intersection is equivalent to

$$\mathbf{1}^T v = 2.$$

Note, however, that cohort data *do not* satisfy this condition. The choice of reference coordinates used in the formalization above is evidently not compatible with cohort data. This "problem" is in some sense an illusion, as the next section will show.

3 Affine Relation

Variables w, x, y are said to satisfy an *affine relation* if

$$\alpha w + \beta x + \gamma y = \varepsilon$$

for some constants $\alpha, \beta, \gamma, \varepsilon$. We are interested in the nontrivial case $\langle \alpha, \beta, \gamma \rangle \neq 0$. Hence in this section $\|\langle \alpha, \beta, \gamma \rangle\| = \|\mathbf{1}\|$ without loss of generality. Assume the displayed relation above, and consider the change of variables

$$v = M\langle w, x, y \rangle + \mathbf{1}\frac{2-\varepsilon}{3},$$

where M is a rotation matrix such that

$$M\langle \alpha, \beta, \gamma \rangle = \mathbf{1}.$$

In particular, $M^T = M^{-1}$ and $\mathbf{1}^T = \langle \alpha, \beta, \gamma \rangle^T M^T$. It follows that

$$\begin{aligned} \mathbf{1}^T v &= \mathbf{1}^T\Big(M\langle w, x, y \rangle + \mathbf{1}\frac{2-\varepsilon}{3}\Big) \\ &= \langle \alpha, \beta, \gamma \rangle^T M^T M \langle w, x, y \rangle + \mathbf{1}^T\mathbf{1}\frac{2-\varepsilon}{3} \\ &= \alpha w + \beta x + \gamma y + 2 - \varepsilon \\ &= 2. \end{aligned}$$

Therefore, the "transformed" data $\langle v, z \rangle$ satisfies the necessary and sufficient condition for triangular format graphing. When plotting $\langle w, x, y, z \rangle$ in the triangular format as the point

$$q(v) + \mathbf{1}\frac{z}{\sqrt{3}}$$

we may regard that point as the result of a graphing process which has either inputs v and z, or else inputs $\langle w, x, y \rangle$ and z. In the latter case, input $\langle w, x, y \rangle$ is called the vector of *coordinate variables*. In the former case, the point is said to result from plotting z *with respect to reference coordinates* v. In either case, input z is called the *dependent variable*.

Theorem 3.1. *The triangular format is suited to plotting dependent variable z with respect to reference coordinates v if the coordinate variables w, x, y satisfy an affine relation*

$$\alpha w + \beta x + \gamma y = \varepsilon$$

and

$$v = M\langle w, x, y\rangle + \mathbf{1}\frac{2-\varepsilon}{3},$$

where M is any rotation matrix such that

$$M\langle \alpha, \beta, \gamma\rangle = \mathbf{1}.$$

Theorem 3.1 completes the initial formalization of the triangular format graph. It appears to be a graphing paradigm more general than one which is applicable only to cohort data; it is suited to data sets whose first three coordinates satisfy any affine relation. A more general formulation is possible and is shown in the next section.

4 Affine Approximation

Let $\mathcal{D}$ be a 4-dimensional data set for which there exist constants $\alpha, \beta, \gamma, \delta, \varepsilon$ that minimize the following mean squared error

$$\mathcal{S} \;=\; \sum_{\langle d_0, d_1, d_2, d_3\rangle \,\in\, \mathcal{D}} (\alpha d_0 + \beta d_1 + \gamma d_2 + \delta d_3 - \varepsilon)^2,$$

subject to the constraint $\|\langle \alpha, \beta, \gamma, \delta\rangle\| = 1$. The best-fitting 4-dimensional affine relation to $\mathcal{D}$ (in the least squares sense) is therefore the hyperplane π with equation

$$\alpha w + \beta x + \gamma y + \delta z = \varepsilon.$$

This equation is "approximately" satisfied by points $d \in \mathcal{D}$ (with mean squared error $\mathcal{S}$) and is exactly satisfied when $\mathcal{S} = 0$. Assume $\mathcal{D} \subset \pi$, which is equivalent to assuming $\mathcal{S} = 0$, and consider the change of variables

$$u = P_k\langle d_0, d_1, d_2, d_3\rangle,$$

where P_k is a rotation matrix such that

$$P_k\langle \alpha, \beta, \gamma, \delta\rangle = e_k$$

and e_k is any of

$$\begin{aligned} e_0 &= \langle 1, 0, 0, 0\rangle, \\ e_1 &= \langle 0, 1, 0, 0\rangle, \\ e_2 &= \langle 0, 0, 1, 0\rangle, \\ e_3 &= \langle 0, 0, 0, 1\rangle. \end{aligned}$$

Note that

$$\begin{aligned} u^T e_k &= \langle d_0, d_1, d_2, d_3\rangle^T P_k^T P_k \langle \alpha, \beta, \gamma, \delta\rangle \\ &= \langle d_0, d_1, d_2, d_3\rangle^T \langle \alpha, \beta, \gamma, \delta\rangle \\ &= \varepsilon. \end{aligned}$$

Refer to k as the *pivot*, and refer to u_k as the *pivot variable*. What has been shown is that if $\mathcal{S} = 0$, then the transformed data $P\mathcal{D}$ (where $P = P_k$ for pivot k) represented by the vector variable u (ranging over that transformed data set) satisfies the following affine relation

Pivot	Pivot variable	Affine relation
0	u_0	$u_0 = \varepsilon$
1	u_1	$u_1 = \varepsilon$
2	u_2	$u_2 = \varepsilon$
3	u_3	$u_3 = \varepsilon$

The pivot variable together with any two other variables from $\{u_0, u_1, u_2, u_3\}$ satisfy a 3-dimensional affine relationship (the other variables simply have coefficients of 0). In other words, the results of the previous section apply, as summarized by the following.

Theorem 4.1. *Using the notation above, if $\mathcal{S} = 0$, then the triangular format graph of the data set $P\mathcal{D}$ can be constructed based on the pivot variable and any two other variables as the triplet of coordinate variables; the remaining variable (not chosen as a coordinate variable) can be graphed as the dependent variable.*

Theorem 4.1 restores symmetry to the graphing process in the sense that no particular variable is *necessarily* singled out as being the dependent variable; any one can be graphed with respect to the remaining three. Note also that the condition of Theorem 4.1 is less strict than that of Theorem 3.1; any 3-dimensional affine relation between any three variables implies $\mathcal{S} = 0$. Moreover, a 4-dimensional affine relation may exist when no 3-dimensional one does.

Note that the pivot merely specifies one of three coordinate variables; any of the three could be the pivot variable. It would be more natural to specify the choice of *dependent* variable. That formulation will be presented in the next section.

The next step to further generalization is to include the situation where $\mathcal{S} > 0$. The difficulty in that case is in the graphing process where the point $q \in \Lambda$ is not well defined (the three lines do not intersect in a single point), all else remains the same: P is determined by minimizing $\mathcal{S}$ and choosing coordinate variables $\langle u_i, u_j, u_k\rangle$, then the remaining variable can be graphed as the dependent variable with respect to reference coordinates as determined by Theorem 3.1.

It is natural to let q be the centroid of the triangle with vertices $\{l_0 \cap l_1, l_1 \cap l_2, l_2 \cap l_0\}$ (the centroid is the average of the vertices). This collapses to the previous definition when the triangle is degenerate (i.e., a single point). Therefore, extend the definition of q to the case in which the three lines no not necessarily intersect,

$$q(v) \;=\; \frac{l_0 \cap l_1 \;+\; l_1 \cap l_2 \;+\; l_2 \cap l_0}{3}$$

This nearly completes the generalization of the triangular graphing process; later we discuss why and how the function q should be altered further to minimize geometric distortions introduced by the graphing process. An analysis of what those distortions are is the subject of the next section.

5 Distortion

There are two potential sources of distortion. The first is from the change of variables Theorems 3.1 and 4.1 entail. Let u_i be the dependent variable, let $\alpha, \beta, \gamma, \delta, \varepsilon$ be determined by minimizing $\mathcal{S}$, and let P_k correspond to pivot $k \;= i+1$ modulo 4 so that if $\mathcal{S} = 0$ then

$$\sqrt{3}u_{i+1} + 0u_{i+2} + 0u_{i+3} = \sqrt{3}\varepsilon,$$

where subscripts are interpreted modulo 4. Let the rotation matrix $R_k \;=\; (R')^k (R'')^k$ be determined by the orthogonal matrices

$$R' \;=\; \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}, \qquad R'' \;=\; \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \end{pmatrix},$$

and let M be any rotation matrix such that

$$M\langle\sqrt{3}, 0, 0\rangle = \mathbf{1}.$$

The combined change of variable involved in the graphing process may be taken to be

$$\Theta(d) = \begin{pmatrix} M & 0 \\ 0 & 1 \end{pmatrix} R_k P_k \langle d_0, d_1, d_2, d_3\rangle + \langle \mathbf{1}, 0\rangle \frac{2-\sqrt{3}\varepsilon}{3},$$

which is easily verified as follows. Let $\bar{i}$ denote 1 if i is odd, and 0 otherwise. By definition, the right hand side above simplifies to

$$\begin{aligned} & \begin{pmatrix} M & 0 \\ 0 & 1 \end{pmatrix} \langle u_{i+1}, u_{i+3-\bar{i}}, u_{i+2+\bar{i}}, u_i\rangle + \langle \mathbf{1}, 0\rangle \frac{2-\sqrt{3}\varepsilon}{3} \\ = \quad & \langle M\langle u_{i+1}, u_{i+3-\bar{i}}, u_{i+2+\bar{i}}\rangle + \mathbf{1}\frac{2-\sqrt{3}\varepsilon}{3}, u_i\rangle. \end{aligned}$$

This corresponds to the reference coordinates as determined by Theorem 3.1 (first partition above – first three components of $\Theta(d)$) for graphing dependent variable u_i (second partition above – last component of $\Theta(d)$) of the transformed data set $\Theta(\mathcal{D})$.

The transformation Θ is a rigid motion of four dimensional space; because only rotation and translation are involved, no distortions are introduced (the transformed data set $\Theta(\mathcal{D})$ and the original data set $\mathcal{D}$ are not only congruent, orientation is preserved).

The second potential source of distortion is the plotting of points which takes place after the change of variable. Consider first the case where $\mathcal{S} = 0$. Then the (partitioned) point $\langle v, z \rangle = \Theta(d) \in \Theta(\mathcal{D})$ is plotted as the point

$$\mathbf{1} - v + \mathbf{1}\frac{z}{\sqrt{3}}.$$

Note the central inversion of the domain over which the dependent variable is graphed caused by

$$q(v) = \mathbf{1} - v$$

(obviously translation by $\mathbf{1}$ is also involved, but that is not the focus of concern). Thus orientation is reversed in that domain. However, the map

$$\phi(v) = \mathbf{1}\frac{4}{3} - v$$

not only introduces an inversion, but also leaves $\mathbf{1}^T v = 2$ invariant. Therefore, the inversions in q and ϕ cancel out in the composition $q \circ \phi$, and the inclusion of ϕ leaves the necessary and sufficient condition $\mathbf{1}^T v = 2$ undisturbed.

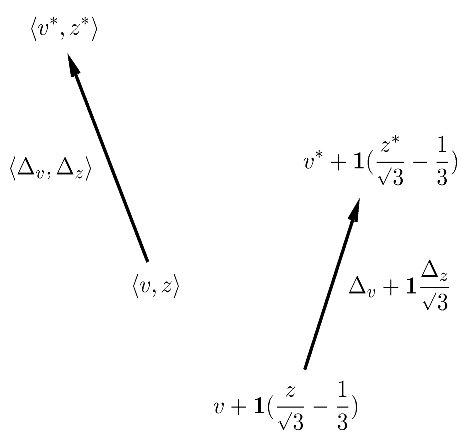

Figure 5.3: Corresponding vectors in $\Theta(\mathcal{D})$ and the triangular format graph.

This motivates a *redefinition* of the graphing process to include precomposition with ϕ so that orientation is preserved. In the general case ($\mathcal{S} \geq 0$) replace q by

$$q'(v) = v + \frac{\mathbf{1}}{3} - \frac{\mathbf{1}\mathbf{1}^T v}{3}.$$

Because we are considering the case $\mathcal{S} = 0$ (and thus $\mathbf{1}^T v = 2$), the point $\langle v, z\rangle = \Theta(d)$ would therefore be plotted as

$$v + \mathbf{1}\Big(\frac{z}{\surd 3} - \frac{1}{3}\Big).$$

Now consider a pair of corresponding vectors, the first $\langle \Delta_v, \Delta_z\rangle$ is between two points of $\Theta(\mathcal{D})$ and the the second $\Delta_v + \mathbf{1}\Delta_z/\surd 3$ is between those points plotted in the triangular format graph (see figure 5.3). Let $\langle \Delta_v', \Delta_z'\rangle$ and $\Delta_v' + \mathbf{1}\Delta_z'/\surd 3$ be another pair of corresponding vectors, and consider the following inner product between the vectors in the triangular format graph

$$\Big(\Delta_v + \mathbf{1}\frac{\Delta_z}{\surd 3}\Big)^T \Big(\Delta_{v'} + \mathbf{1}\frac{\Delta_{z'}}{\surd 3}\Big).$$

Expanding and using $\langle v, z\rangle \in \Theta(\mathcal{D}) \Rightarrow \mathbf{1}^T v = 2$ yields

$$\begin{aligned}
& \Delta_v^T\Delta_{v'} + \Delta_z\Delta_{z'} + \Delta_v^T\mathbf{1}\frac{\Delta_{z'}}{\surd 3} + \frac{\Delta_z}{\surd 3}\mathbf{1}^T\Delta_{v'} \\
= \;& \langle \Delta_v, \Delta_z\rangle^T\langle \Delta_{v'}, \Delta_{z'}\rangle + \mathbf{1}^T(v^* - v)\frac{\Delta_{z'}}{\surd 3} + \mathbf{1}^T(v^{*\prime} - v')\frac{\Delta_z}{\surd 3} \\
= \;& \langle \Delta_v, \Delta_z\rangle^T\langle \Delta_{v'}, \Delta_{z'}\rangle + (2-2)\frac{\Delta_{z'}}{\surd 3} + (2-2)\frac{\Delta_z}{\surd 3} \\
= \;& \langle \Delta_v, \Delta_z\rangle^T\langle \Delta_{v'}, \Delta_{z'}\rangle.
\end{aligned}$$

The result matches the inner product between vectors *before* graphing. Because the inner product determines angles and distances, they are preserved by graphing, no distortions are introduced. Let I denote the 3×3 identity matrix. The partitioned matrix $\Psi = (I\ \mathbf{1}/\surd 3)$ relates differences $\langle \Delta_v, \Delta_z\rangle$ before graphing to differences $\Delta_v + \mathbf{1}\Delta_z/\surd 3$ after graphing,

$$\Delta_v + \mathbf{1}\Delta_z/\surd 3 = (I\ \ \mathbf{1}/\surd 3)\langle \Delta_v, \Delta_z\rangle.$$

We have seen Ψ is an isometry (between its domain and range) if its domain is taken to be spanned by differences of vectors within $\Theta(\mathcal{D})$. Moreover, that partitioned matrix preserves orientation when mapping that domain. The previous observations are summarized by the following.

Theorem 5.1. *If $\mathcal{S} = 0$, then the original data set $\mathcal{D}$ and the triangular format graph of $\mathcal{D}$ are congruent and have the same orientation (they may be regarded as alternate views of the same points from different perspectives in space).*

Let T_q denote translation by q. When the data set $\mathcal{D}$ lies in the hyperplane π, points are plotted in the triangular format graph by the *plotting function* $\mathcal{T} : \Re^4 \to \Re^3$ defined by

$$\begin{aligned}
\mathcal{T}(d) &= T_{-\mathbf{1}/3}\Psi\Theta(d) \\
&= \Psi\langle v, z\rangle - \mathbf{1}/3
\end{aligned}$$

$$= \quad v + \mathbf{1}z/\surd 3 - \mathbf{1}/3.$$

According to Theorem 5.1, $\mathcal{T}$ preserves the orientation and geometry of the data set, which suggests (since π is mapped into $\Re^3$) that the plotting function $\mathcal{T}$ is realized by:

1. Translate and rotate π into the three dimensional subspace $\Re \times \Re \times \Re \times \{0\}$.
2. Project to $\Re \times \Re \times \Re$.
3. Rotate and translate the result.

However, $\mathcal{T}$ *cannot* be realized in that way. If it were, then there would exist rotation matrices A and B and translation vectors r and s for which $\mathcal{T} = T_s B(I\ 0)AT_r$ which leads to a contradiction as follows:

$$T_{-\mathbf{1}/3}\Psi\Theta = \mathcal{T} = T_s B(I\ 0)AT_r \implies \Psi = T_{\mathbf{1}/3+s}B(I\ 0)AT_r\Theta^{-1}.$$

Because Θ is a rotation followed by a translation, there must therefore exist a rotation matrix A' and translation vectors a, b, c such that for all v

$$\Psi v \;=\; a + B(I\ 0)A(b + A'v) \;=\; c + B(I\ 0)AA'v.$$

Since Ψ is linear, the translation on the right hand side above can be ignored (linear maps preserve the origin), yielding an expression for Ψ which has norm one (contradiction).

What is going on is that even though $\mathcal{T}$ maps π to $\Re^3$ appropriately, it nevertheless transforms points of $\Re^4 \setminus \pi$ in a non-intuitive manner. In some sense that is due to a missing projection in the definition of $\mathcal{T}$.

In the general case $\mathcal{S} \geq 0$, the plotting function $\mathcal{G} : \Re^4 \to \Re^3$ is given by

$$\mathcal{G}(d) \quad = \quad T_{\mathbf{1}/3}\Psi'\Theta(d),$$

where Ψ' is the partitioned matrix $(I - \mathbf{1}\mathbf{1}^T/3\ \ \mathbf{1}/\surd 3)$. If $\mathcal{S} > 0$ then $\mathcal{D}$ is not contained in a hyperplane and angles and distances cannot possibly be preserved ($\mathcal{G}$ reduces dimension).

Let $\rho : \Re^4 \to \Re^4$ be the orthogonal projection to $\langle \mathbf{1}, 0\rangle^\perp$ given by

$$\begin{pmatrix} I - \mathbf{1}\mathbf{1}^T/3 & 0 \\ 0 & 1 \end{pmatrix}.$$

Note that $\Psi' = \Psi\rho$. Therefore, the difference between $\mathcal{T}$ and $\mathcal{G}$ — aside from a compensating difference in translation ($T_{-\mathbf{1}/3}$ vs $T_{\mathbf{1}/3}$) — is the inclusion of the projection ρ in Ψ' compared to its omission in Ψ. The presence of the projection is the influence of the centroid calculation in the definition of q', it was omitted in $\mathcal{T}$ (when $\mathcal{S} = 0$ the centroid calculation is irrelevant) but it is included in $\mathcal{G}$.

The next theorem extends Theorem 5.1 to the general case, summarizing the distortions that are introduced by the triangular format graph. They are in some sense minimal, since the projection involved is to the best-fitting hyperplane.

Theorem 5.2. *Let $\mathcal{D}_\pi$ be the orthogonal projection of the data set $\mathcal{D}$ to the hyperplane π. Then $\mathcal{D}_\pi$ and the triangular format graph of $\mathcal{D}$ are congruent and have the same orientation (they may be regarded as alternate views of the same points from different perspectives in space).*

Proof. Given $d \in \mathcal{D}$, let

$$d = d_\pi + d_\perp,$$

where $d_\pi \in \pi$ and $d_\perp$ is orthogonal to π; hence $d_\perp = \|d_\perp\|\langle \alpha, \beta, \gamma, \delta\rangle$. By definition, $\mathcal{D}_\pi = \{d_\pi : d \in \mathcal{D}\}$, and, since $\mathcal{S} = 0$ for the data set $\mathcal{D}_\pi$,

$$\langle \mathbf{1}, 0\rangle^T \Theta(d_\pi) = 2.$$

Therefore

$$\Theta(d_\pi) = \frac{2}{3}\langle \mathbf{1}, 0\rangle + \rho\,\Theta(d_\pi)$$

and

$$\rho\,\Theta(d_\pi) = \Theta(d_\pi) - \frac{2}{3}\langle \mathbf{1}, 0\rangle.$$

Note that

$$\begin{aligned}
\Theta(d) &= \Theta(d_\pi) + \begin{pmatrix} M & 0 \\ 0 & 1 \end{pmatrix} R^{i+1} P\, d_\perp \\
&= \Theta(d_\pi) + \|d_\perp\| \begin{pmatrix} M & 0 \\ 0 & 1 \end{pmatrix} R^{i+1} P \langle \alpha, \beta, \gamma, \delta\rangle \\
&= \Theta(d_\pi) + \|d_\perp\| \begin{pmatrix} M & 0 \\ 0 & 1 \end{pmatrix} R^{i+1} e_{i+1} \\
&= \Theta(d_\pi) + \|d_\perp\| \begin{pmatrix} M & 0 \\ 0 & 1 \end{pmatrix} e_0 \\
&= \Theta(d_\pi) + \|d_\perp\| \langle \mathbf{1}, 0\rangle / \sqrt{3}.
\end{aligned}$$

Using the fact that $\Psi' = \Psi\rho$, it follows from the above that

$$\begin{aligned}
\mathcal{G}(d) &= T_{\mathbf{1}/3}\Psi'\Theta(d) \\
&= T_{\mathbf{1}/3}\Psi\rho\big(\Theta(d_\pi) + \|d_\perp\|\langle \mathbf{1}, 0\rangle/\sqrt{3}\big) \\
&= T_{\mathbf{1}/3}\Psi\rho\,\Theta(d_\pi) \\
&= T_{\mathbf{1}/3}\Psi\big(\Theta(d_\pi) - \frac{2}{3}\langle \mathbf{1}, 0\rangle\big) \\
&= T_{\mathbf{1}/3}\Psi\Theta(d_\pi) - \frac{2}{3}\Psi\langle \mathbf{1}, 0\rangle \\
&= T_{\mathbf{1}/3}\Psi\Theta(d_\pi) - \frac{2}{3}\mathbf{1}
\end{aligned}$$

$$= T_{-\mathbf{1}/3}\Psi\Theta(d_\pi)$$
$$= \mathcal{T}(d_\pi).$$

The proof is completed by appealing to Theorem 5.1. □

6 Conclusion

Theorem 5.2 of the previous section summarizes the meaning of the triangular graphing process. That process is defined by the following specification:

- **Input:** A 4-dimensional data set $\mathcal{D}$ for which there exist constants $\alpha, \beta, \gamma, \delta, \varepsilon$ that minimize

$$\mathcal{S} = \sum_{\langle d_0,d_1,d_2,d_3\rangle \in \mathcal{D}} (\alpha d_0 + \beta d_1 + \gamma d_2 + \delta d_3 - \varepsilon)^2$$

subject to the constraint $\|\langle \alpha, \beta, \gamma, \delta\rangle\| = 1$.

- **Output:** The set of points

$$\mathcal{G}(\mathcal{D}) = \{T_{\mathbf{1}/3}\Psi'\Theta(d) : d \in \mathcal{D}\},$$

where

$$T_{\mathbf{1}/3}(x) = x + \frac{1}{3}\langle 1,1,1\rangle,$$

$$\Psi' = \frac{1}{3}\begin{pmatrix} 2 & -1 & -1 & \surd 3 \\ -1 & 2 & -1 & \surd 3 \\ -1 & -1 & 2 & \surd 3 \end{pmatrix},$$

$$R' = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix},$$

$$R'' = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \end{pmatrix},$$

$$\Theta(d) = \begin{pmatrix} M & 0 \\ 0 & 1 \end{pmatrix} (R')^{i+1}(R'')^{i+1}Pd + \langle 1,1,1,0\rangle \frac{2-\sqrt{3}\varepsilon}{3}.$$

Here M is any rotation matrix such that $M\langle\sqrt{3},0,0\rangle = \langle 1,1,1\rangle$, the (arbitrary) rotation matrix P must satisfy $P\langle\alpha,\beta,\gamma,\delta\rangle = \langle[i=0],[i=1],[i=2],[i=3]\rangle$ where $[i=j]$ is 1 if $i=j$ and is 0 otherwise, and i is a parameter interpreted as selecting the dependent variable. The coordinate axes are graphed as the following four parametric lines

$$\begin{aligned}
&t\,\langle 1,1,1\rangle,\\
&\langle 1,0,0\rangle \;+\; t\,\langle 1,0,-1\rangle,\\
&\langle 0,1,0\rangle \;+\; t\,\langle -1,1,0\rangle,\\
&\langle 0,0,1\rangle \;+\; t\,\langle 0,-1,1\rangle.
\end{aligned}$$

The advantage of the triangular format graph for visual presentation stems from keeping all four coordinate axes in the field of view while simultaneously minimizing distortions. By so doing, it facilitates understanding of the underlying structure of the data, and it facilitates the researcher's selection of which statistical solutions might be applied to decompose additional variable effects on the dependent variable.

Returning to the initial motivating example of APC data sets discussed in the introduction (where triplet reference coordinates are linearly dependent), the user's examination of each of the age, period, and cohort profiles for the dependent variable z would suggest, then, how age, period, and cohort effects operate and may be best interpreted. This approach is more facile than other strategies suggested in the APC literature by [1,4,8,9,12–14,19,23].

More generally, the triangular format graph is a visualization paradigm suited to arbitrary 4-dimensional data sets. It minimizes distortions – when distortions are unavoidable – and otherwise introduces no graphing artifacts (coordinate axes undergo distortion, but orientation, distances, and angles between data points are preserved). Moreover, it manages all that while simultaneously keeping all four coordinate axes in the field of view.

Acknowledgements

We would like to thank the anonymous reviewers. Their careful attention to detail has significantly improved the exposition of this paper.

References

[1] P. B. Baltes, Longitudinal and cross-sectional sequences in the study of age and generation effects. *Human Development* **II** (1968), 145–147.

[2] R. A. Becker and W. S. Cleveland, Viewing multivariate scattered data, *Pixel.* **2** (1991), 36–41.

[3] R. A. Becker and W. S. Cleveland, Taking the broader view of scientific visualization, *Pixel.* **2** (1991), 42–44.

[4] H. M. Blalock, The identification problem and theory building: The case of status inconsistency, *American Sociological Review.* **31** (1966), 52–56.

[5] R. D. Blanchard, J. B. Bunker, M. Wachs, Distinguishing aging, period, and cohort effects in longitudinal studies of elderly populations, *Socio-Economic Planning Sciences* **11** (1977), 137–146.

[6] R. A. Case, Cohort analysis on mortality rates as an historical or narrative technique, *British Journal of Preventive and Social Medicine.* **II** (1956), 159–71

[7] G. Caselli and E. Lombardo, Graphiques et analyse demographique: Quelques elements d' historie et d' actualite, *Population* **2** (1990), 399–414.

[8] S. E. Fienberg and W. M. Mason, Identification and estimation of age-period-cohort models in the analysis of discrete archival data, in: Schuessler KF (ed) *Sociological Methodology.* Jossey-Bass, San Francisco, 1979, 1–67.

[9] N. D. Glenn, Cohort analysts' futile quest: Statistical attempts to separate age, period, and cohort effects, *American Sociological Review* **41** (1976), 900–905.

[10] D. W. Hastings and L. G. Berry, *Cohort Analysis: A Collection of Interdisciplinary Readings*, Scripps Foundation for Research in Population Problems, Oxford, Ohio, 1979.

[11] D. W. Hastings, L. G. Berry, J. M. Rodriguez, El analisis cohorte-su logica y utilidad, *Revista Mexicana de Sociologia* **40**, 3 (Extraordinaria E) (1978) 299–331.

[12] K. O. Mason, W. H. Mason, H. H. Winsborough, W. K. Poole, Some methodological issues in cohort analysis of archival data, *American Sociological Review* **38** (1973), 242–258.

[13] W. H. Mason and H. L. Smith, Age-period-cohort analysis and the study of deaths from pulmonary tuberculosis, in: Mason WM, Fienberg SE (eds) *Cohort Analysis: Beyond the Identification Problem.* Springer-Verlag, New York, 1985, 151–227.

[14] E. Palmore, When can age, period, and cohort be separated?, *Social Forces* **57** (1978), 282–295.

[15] D. R. Ploch and D. W. Hastings, Cohort surfaces: A graphic approach to teaching social stability and change, *Teaching Sociology* **20** (1992), 192–200.

[16] D. R. Ploch and D. W. Hastings, Graphic presentations of church attendance using General Social Survey Data, *Journal for the Scientific Study of Religion* **33** (1994), 16–33.

[17] D. R. Ploch and D. W. Hastings, Some church; Some don't, *Journal for the Scientific Study of Religion* **34** (1995), 507–515.

[18] R. Pressat, Demographic Analysis: Methods, Results, and Applications, translated by Judah Matras. Aldine-Atherton, Chicago, 1972.

[19] T. W. Pullum, Parametrizing age, period, and cohort effects: An application to U.S. delinquency rates, 1964–73, in: Schuessler KF (ed) *Sociological Methodology.* Jossey-Bass, San Francisco, 1978, 116–140.

[20] M. W. Riley, M. Johnson, A. Foner, *Aging and Society, III.- A Sociology of Age Stratification*, Russell Sage Foundation, New York, 1972.

[21] N. B. Ryder, Fertility measurement through cross-sectional surveys, *Social Forces* **54** (1975), 7–35.
[22] N. B. Ryder, The cohort as a concept in the study of social change, *American Sociological Review* **30** (1965), 843–861.
[23] K. W. Schaie, A general model for the study of developmental problems, *Psychological Bulletin* **64** (1965), 92–107.
[24] C. M. Young, A cohort approach to the study of population growth, *Journal of Biosocial Science* **4** (1974), 75–90.

M. D. Vose has doctorates in Mathematics and Computer Science, from the University of Texas at Austin. He is currently an Associate Professor in Electrical Engineering and Computer Science at the University of Tennessee, Knoxville. He is especially well known for seminal contributions to the foundations of genetic algorithms.

Donald W. Hastings, PhD Sociology, University of Massachusetts (Amherst), is Professor Emeritus at the University of Tennessee, Knoxville. His major areas of publication include demography, sociology of religion and sport.