RR – 85-09

Relations among k sets of variables, with geometrical representation, and an application to nominal variables.

John P. Van de Geer

Department of Data Theory
Faculty of Social Sciences
University of Leiden
The Netherlands

ABSTRACT

The paper discusses some possible solution for the analysis of relations among k sets of variables. The general approach is that for each of the k sets a linear compound is calculated. Relations among sets then can be described in terms of relations between k such compounds.

A first possibility is to base the analysis on sets P_i (i=1,..,k), where P_i is an orthogonal and unit-normalized basis of the i^{th} set. The alternative is to base the analysis on principal components of each set, written as $P_i Ø_i$.

Linear compounds are $P_i t_i$ or $P_i Ø_i t_i$. The second major consideration is whether weights t_i are restricted by $\Sigma t_i' t_i = k$, by $t_i' t_i = 1$, or by $T_i' T_i = I$ (where T_i collects different solutions for t_i in its columns).

The paper illustrates these different procedures in terms of their geometrical properties.

The paper then shows how these solutions work out for analysis of sets of nominal variables. This results in "optimal quantification", both of categories and of object scores (where object scores are defined as the centroid of those quantifications which apply to the object). The general conclusion is (1) that analysis of sets P_i tends to emphasize differences between infrequent categories, compared to analysis of sets $P_i Ø_i$; (2) that the choice for $\Sigma t_i' t_i = k$ may result in solutions where some of the k sets are dominating (with neglect of other sets), that the choice for $t_i' t_i = 1$ makes those contributions more equal (but with the risk that they are highly correlated), and that the choice for $T_i' T_i = I$ makes contributions more equal but at the same time counterbalances such high intercorrelations.

1. Formulation of the problem.

Data are given in a matrix X, with n rows (for objects) and as many columns as there are variables. Given is that X is partitioned, column-wise, in k sets: $X=(X_1,X_2,..,X_k)$. The problem is to identify optimal relations between the k sets. We assume here that columns of X have zero mean.

This problem can be approached by taking linear compounds X_iv_i (i=1,..,k) with v_i a vector of weights. We then have to select the v_i in such a way that relations between the k vectors X_iv_i are optimized. The criterion could be that the sum of the correlations between all pairs should have a maximum. But there are other possible criteria, as we shall see below.

In Van de Geer (1984) it was shown that there are a number of possible choices.

(a) The first basic choice is whether or not it is allowed to replace each X_i by an orthogonal and unit-normalized basis. Such a basis can be found by taking the singular value decomposition $X_i=P_i\emptyset_iQ_i'$. This solution satisfies

$P_i'P_i=I$

$Q_i'Q_i=I$

$\emptyset_i$ is a diagonal matrix, with positive diagonal elements, and with dimension equal to the rank of X_i.

Then P_i is an orthogonal and unit-normalized basis of X_i. Obviously, replacing each X_i by its basis P_i implies that all information about relations between variables within a set, is thrown away.

The alternative is to repace X_i by $P_i\emptyset_i$, in which case basic information abou relations within sets is retained.

(b) The second basic choice is related to requirements imposed on weights. Obviously, a solution for a linear compound X_iv_i will be equal to a solution for t_i such that $X_iv_i=P_it_i$, or such that $X_iv_i=P_i\emptyset_it_i$. The first possibility is that t_i is restricted by $t_i't_i=1$. The second possibility is that this restriction is softened to $t't=k$ (implying that the average value of $t_i't_i$ is equal to 1). The third possibility is the requirement $T_i'T_i=I$; it means that successive solutions for t_i are orthogonal (and unit-normalized).

Taken together, the two choices result in 6 different solutions. They are shown in Table 1. For each solution, the criterion is that the trace of $T'P'PT$ or of $T'\emptyset P'P\emptyset T$ must be maximized, where $PT=\Sigma P_iT_i$.

This criterion has its own special interpretation for each cell of the table. In cell [P,t't=k] the criterion implies that a sumvector Pt has stationary value for its sum of squares, which comes to the same thing as requiring that the sum of the squared correlations between Pt and all P_it_i

This criterion has its own special interpretation for each of the six cells of Table 1. In cell [P,t't=k] there will be stationary value for the sum of the squared correlations between Pt and all columns of P, or between Pt and all columns $P_i t_i$. In cell [P,$t_i' t_i$=1] there is stationary value for the squared sum of all correlations betwee Pt and the vectors $P_i t_i$. For this analysis it is not required that columns of PT=$\Sigma P_i T_i$ form an orthogonal set. In cell [P,$T_i' T_i$=I] the interpretation of the stationary value remains the same, but under the condition that columns of PT are orthogonal. In cell [PØ,t't=k] it will be true that there is stationary value for the amount of variance in PØ (or in X itself) "explained" by PØt. This comes to the same thing as saying that there is stationary value for the sum of the squared covariances between all columns of PØ (or of X) and the unit-normalized vector PØt$(\mu k)^{-\frac{1}{2}}$. In cell [PØ,$t_i' t_i$=1] the stationary value refers not to the sum of the squared covariances, but to the squared sum of the covariances between all $P_i Ø_i t_i$ and the unit normalized vector PØt$(\Sigma\mu)^{-\frac{1}{2}}$. In cell [PØ,$T_i' T_i$=I] there is the additional requirement that columns of PØT form an orthogonal set.

2. Basic geometry.

The essentials of a geometric interpretation can be explained easiest by assuming that X_i has only two columns. Obviously, the two vectors x_{i1} and x_{i2} then are vectors locatec in a plane.

Suppose we take compounds $X_i v_i$, where $v_i' v_i$=1. Vectors $X_i v_i$ then will be located in the same plane as X_i. In addition, such vectors will appear as "radii" of an ellipse passing through x_{i1} and x_{i2}. There will be two solutions for v_i such that the compounds $X_i v_{i1}$ and $X_i v_{i2}$ will correspond to the two principal axes of this ellipse. These principal axes then also can be expressed as $p_{i1} ø_{i1}$ and $p_{i2} ø_{i2}$. Each other "radius" of the ellipse can be expressed as a vector $P_i Ø_i t_i$, with $t_i' t_i$=1.

Assume now that we have two solutions $t_{i(1)}$ and $t_{i(2)}$ such that $t_{i(1)}' t_{i(2)}$=0. The geometric implication is that the tangential to the ellipse at $P_i Ø_i t_{i(1)}$ is parallel to $P_i Ø_i t_{i(2)}$ - and vice versa: the tangential at $P_i Ø_i t_{i(2)}$ is parallel to $P_i Ø_i t_{i(1)}$. In other words: tangentials circumscribe the ellipse as a parallelogram, with sides parallel to $P_i Ø_i T_i$.

If we take a solution based on analysis of P, vectors $P_i t_i$ with $t_i' t_i$=1 will appear as radii of a circle. When $T_i' T_i$=I, the two radii will be orthogonal, so that the tangentials surround the circle as a square.

3. Applications of the geometrical interpretation.

3.1 Analysis of P.

Results will be illustrated with an example, with k=3 sets, each with m=2 variables. More details of this example will be given in the next section.

(i) Analysis of P with $t_i't_i=1$. These results are shown in Figure 2. The figure shows three ellipses; they are the projection of the circles defined by P_i (i=1,2,3) on the plane of $Pt_{(1)}$ and $Pt_{(2)}$. The essential characteristic of the figure is that the tangentials at the points $P_it_{i(1)}$ are orthogonal to $Pt_{(1)}$, and the tangentials at $P_it_{i(2)}$ orthogonal to $Pt_{(2)}$. The rationale is quite simple: we want $Pt_{(j)}$ (j=1,2) to be as long as possible. On the other hand $Pt_{(j)}=\Sigma P_it_{i(j)}$, so that we also want the projection of $P_it_{i(j)}$ on $Pt_{(j)}$ to be as long as possible. The latter implies that the tangential at $P_it_{i(j)}$ must be orthogonal to $Pt_{(j)}$. In this way the tangentials circumscribe each ellips in the form of a parallelogram, with sides ortogonal to PT.

Another characteristic of the figure is that the vectors PT are not orthogonal to each other. A third characteristic is that vectors P_iT_i are not orthogonal. The latter is shown best for P_3T_3, where the corresponding ellipse is rather narrow, and where $P_3t_{3(1)}$ is very close to $P_3t_{3(2)}$.

(ii) Analysis of P with $T_i'T_i=I$. This result is shown in Figure 3. The three ellipses now have about the same shape. Vectors $Pt_{(1)}$ and $Pt_{(2)}$ are orthogonal. However, the price one must pay is that tangentials at $P_it_{i(j)}$ no longer are orthogonal to $Pt_{(j)}$. Tangentials circumscribe ellipses as parallelograms which no longer have parallel sides. In stead, the tangentials are only roughly orthogonal to $Pt_{(1)}$ or $Pt_{(2)}$, in the sense that the projections of $P_it_{i(1)}$ on $Pt_{(2)}$ have zero average (and vice versa). A further characteristic is that vectors P_iT_i are uncorrelated (geometrically this implies that the tangential at $P_it_{i(1)}$ is parallel to $P_it_{i(2)}$, and the tangential at $P_it_{i(2)}$ parallel to $P_it_{i(1)}$.

(iii) Analysis of P with $t't=k$. Figure 1 shows this solution. Vectors $Pt_{(1)}$ and $Pt_{(2)}$ now are orthogonal, and tangentials at $P_it_{i(1)}$ are orthogonal to $Pt_{(2)}$, and vice versa. However, vectors $P_it_{i(j)}$ will no longer be located on the ellipses, since $t_{i(j)}'t_{i(j)}\neq 1$. The vectors on the ellipses therefore are, in fact, the vectors $P_it_{i(j)}(t_{i(j)}'t_{i(j)})^{-\frac{1}{2}}$. We therefore may interpret the values of $(t_{i(j)}'t_{i(j)})^{\frac{1}{2}}$ as if they are weights: it the value is larger than 1, the corresponding vector will be located outside the ellipse, and if the value is smaller than 1, the vector is located in the inside of the ellipse. For the example, these values are 1.08, 1.20, and .64, respectively, for the first dimension; it shows that P_2 plays the most dominant role, whereas P_3 is relatively neglected. For the second dimension the weights are 1.17, .95, and .85.

3.2 Analysis of PØ.

In general, analysis of P does not pay any attention to the location of the principal components P_i. These principal components are two orthogonal radii of their circles, but for the analysis it does not matter at all where they are located. It follows that solutions for $P_i t_{i(j)}$ could be very well be close to the second principal component of P_i. Such a result is illustrated in Figure 1 for the solution of $P_2 T_2$: both vectors are close to the second principal component. This implies that the vectors $P_2 t_{2(j)}$ explain little of the variance of X_2.

Analysis of PØ has the effect that the analysis is based on the ellipses with the columns of $P_i Ø_i$ as principal axis. Such ellipses project on the plane of PØT again as ellipses, where, in general, the projection of the first principal component $p_{i1} ø_{i1}$ will be longer than that of the second principal component. As a result, solutions for $P_i Ø_i T_i$ will tend to be located closer to the projection of the first principal component.

For the example this result is illustrated in Table 2. This table gives for the principal components the percentage of explained variance, averaged over three sets. The result shows quite clearly that in analysis of PØ the percentage explained variance for the first component becomes much larger than the percentage explained for the second component, in comparison with the corresponding analysis of P.

(i) Analysis of PØ, t't=k. The result is shown in Figure 4. As in Figure 1, tangential circumscribe the ellipses as rectangles, with sides orthogonal to PT. For $P_2 Ø_2$ the projected ellipse is very narrow, with the further implication that $P_2 Ø_2 t_{2(1)}$ has almost perfect correlation with $p_2 ø_{21}$, whereas $P_2 Ø_2 t_{2(2)}$ has almost perfect negative correlation with the first principal component.

Again, in the analysis the contributions of each set have differential weight $(t'_{i(j)} t_{i(j)})^{\frac{1}{2}}$. For the first dimension these weights are 1.03. 1.28. and .56, respectively, showing that $P_2 Ø_2$ has dominant contribution, and $P_3 Ø_3$ is relatively neglected. For the second dimension the weights are .95, .22, and 1.43, so that now $P_3 Ø_3$ is dominant, and $P_2 Ø_2$ neglected.

(ii) Analysis of PØ, $t_i't_i=1$. The result is given in Figure 5. Ellipses are circumscribed by parallelograms with sides orthogonal to PØT. For the example $PØt_{(1)}$ has quite large correlation (r=.72) with $PØt_{(2)}$. The fact that the solution for $P_2Ø_2T_2$ is very much dominated by the first principal component, has already been mentioned above. The same is true, to less extent, for $P_1Ø_1T_1$.

(iii) Analysis of PØ with $T_i'T_i=I$. As in Figure 3, ellipses are circumscribed by tangential parallelograms of which the sides are only roughly orthogonal to PØT. Also, as in Figure 3, the tangential at $P_iØ_it_{i(1)}$ is parallel to $P_iØ_it_{i(2)}$ (and vice versa).

4. Application to the analysis of nominal variables.

4.1 Terminology.

A variable is said to be nominal if it sorts objects into a discrete number of categories which have no apriori quantification. Table 3 gives an example, with three nominal variables, each with three categories, for n=15 objects.

The relation with analysis of k sets becomes visible by re-coding the nominal data in the format of an indicator matrix G. Table 4 shows what this means for the example. The indicator matrix has n=15 rows, and as many columns (9) as there are different categories in all variables. In the row for a particular object, an entry 1 is placed in the columns of the categories which apply to the object, and an entry 0 otherwise.

Obviously, columns of G form k sets, where k is the number of nominal variables. In the example, G can be partitioned into $G=(G_1,G_2,G_3)$, with three binary column vectors for each of the three sets.

We then may replace G by a matrix of deviations from column means. Call this matrix X, then X is also partitioned into k sets: $X=(X_1,X_2,X_3)$. The matrix X_i can be expressed as $X_i=G_i-uu'D_i/n$, where u is a vector with all elements equal to 1, and where D_i is the diagonal matrix $D_i=G_i'G_i$. Note that elements of D_i correspond to the marginal frequencies of the categories of the i^{th} nominal variable.

Rows of X_i will add up to 0, and it follows that X_i has rank m_i-1, where m_i is the number of categories in variable i.

4.2 Analysis in terms of k sets.

Given the matrix X as defined in section 4.1, procedures for analysis of k sets can be applied to it. These procedures will be based on the singular value decomposition $X_i=P_i\emptyset_iQ_i'$, where P_i has m_i-1 columns. In the example, each P_i (i=1,2,3) will have two columns, and therefore each $P_i\emptyset_i$ also has two columns.

To illustrate, take the analysis based on P, with t't=k. This results in solutions for Pt, where Pt has sum of squares equal to $m\mu$. We then may re-normalize Pt such that after re-normalization the sum of squares is equal to n (in other words, Pt is standardized to unit variance). For the example this means that we should take $Pt(n/k\mu)^{\frac{1}{2}}$. This result is called a vector of object scores. At the same time, we can quantify each nominal variable by taking $P_it_i(n/k\mu)^{\frac{1}{2}}.k$. This means that in Table 3 the category labels are replaced by numerical values. It then follows that object scores are the average of the category quantifications of categories which apply to the object. The reason is that $Pt=\Sigma P_it_i$, so that Pt becomes the average of the values in columns $P_i t_i.k$. For other procedures similar adaptations must be made; they will be formulated in the discussion of these various procedures, below (Table 5).

In general, however, it can be said that categories of the nominal variables are quantified in such a way that after quantification the criterion value for the solution is maximized. E.g., in the analysis of P with t't=k, the solution for category quantification in the first dimension will be such that this category quantification ensures largest possible value of $\mu_{(1)}$; any other quantification would result in a lower value of $\mu_{(1)}$.

4.3 Numerical illustration.

Figures 1-6 show for each of the six solutions the 15 points for objects, and the 9 points for quantified categories. Table 5 shows how coordinates of these points are defined for each solution.

For all solutions it is true that object points have sum of squares equal to n=15, both in the horizontal and the vertical direction, and that the sum of cross-products is zero. Note that points for identical objects 1 4 7, or 6 9, or 8 10, do coincide. Note also that in each graph an object point is the centroid of the categories which apply to the object. E.g., object 1 is plotted as the centroid of categories a p u, object 2 is the centroid of b q v, etc.

The solution for P with t't=k is known as the HOMALS solution. This solution has the special property that categories also could have been quantified by taking the average of the object scores of objects in the category; such a quantification would be strictly proportional to the quantification plotted in Figure 1. This special property does not apply to the other five solutions.

A possible disadvantage of the HOMALS solution is that it tends to become dominated by categories with low marginal frequency. This will be true for all solutions based on analysis of P. Analysis of PØ tends to suppress the influence of infrequent categories. This is related to the discussion in section 3.2, and illustrated in Table 2. In fact, in an example with nominal variables, columns of X_i have larger variance to the extent that the marginal frequency comes closer to .5, whereas columns for categories with frequency close to 0 have small variance. The first principal component $p_{i1}\phi_{i1}$ therefore will be correlated more with columns of X_i for categories with intermediate frequency than with columns for categories with low frequency. This is illustrated in Table 6 in which the solutions for $P_i\emptyset_i$ are given. Clearly, $p_{11}\phi_{11}$ contrasts category a with the other two categories, whereas $p_{12}\phi_{12}$ contrasts the two infrequent categories b and c. Also, $p_{21}\phi_{21}$ makes a contrast between the most frequent category p and the other two, whereas $p_{22}\phi_{22}$ makes a contrast between q and r. Finally, $p_{31}\phi_{31}$ gives the contrast between the two most frequent categories v and u, whereas $p_{32}\phi_{32}$ gives the contrast between the infrequent category w and the other two. It follows directly that a solution (as for PØ) which gives more attention to first principal components, therefore must be less dominated by differences between infrequent categories.

This can be seen in Figure 1, where categories b and c are far apart, whereas in Figure 4 the distance between these categories is much smaller. Similarly for categories q and r, or for the extreme position of category w in Figure 1.

Similar changes are found when comparing figures 2 and 5. A comparison of figures 3 and 6 shows that changes are much less drastic, but this was already shown in Table 2.

5. Conclusions.

5.1 Analysis of P. Each set X_i is replaced by an orthogonal basis P_i. The analysis focusses on correlations between sets, and ignores the correlation structure within sets. With t't=k one obtains orthogonal vectors $Pt_{(j)}$, but these vectors are weighted sums of compounds $P_i t_{i(j)}(t'_{i(j)} t_{i(j)})^{-\frac{1}{2}}$. The effect can be that the solution is dominated by only some of the sets, with neglect of other sets. With $t'_i t_i=1$ the vector $Pt_{(j)}$ is a direct sum of the compounds $P_{i(j)} t_{i(j)}$. The contribution of individual sets therefore becomes more balanced. On the other hand, columns of PT are no longer orthogonal. With $T'_i T_i=I$ individual contributions are even more balanced; vectors of PT are orthogonal, and, in addition , vectors in $P_i T_i$ are orthogonal.

5.2 Analysis of PØ. These solutions do not ignore the covariance structure within sets. The general effect will be that for good solutions the contribution of the first principal component(s) of each set will be increased, whereas the contribution of principal components with small eigenvalue is decreased.

5.3 Geometry. Properties of the solutions can be described in terms of orthogonality of tangentials (or tangential hyperplanes) at $P_i t_{i(j)}$ (or $P_i \emptyset_i t_{i(j)}$) and vectors $Pt_{(s)}$ (or $P\emptyset t_{(s)}$) $(s \neq j)$.

5.4 Nominal variables. The solutions can be applied to nominal variables and then result in varieties of "non-linear" analysis. Analysis of P with t't=k then becomes identical to a HOMALS analysis. Analysis of P with $T'_i T_i$ results in orthogonal category quantifications. In all three procedures based on analysis of P the risk is that results are dominated by categories with low marginal frequency. This effect is counterbalanced in the procedures based on analysis of PØ.

References

Gifi, A. Non-linear multivariate analysis. Leiden, DSWO Press (in press).

Van de Geer, J.P. (1984), Relations among k sets of variables. Psychometrika,49,79-9

Van de Geer, J.P. Introduction to linear multivariate data analysis. (In press)

TABLE 1

	$t't=k$	$t_i't_i=1$	$T_i'T_i=I$
P	$P_i'Pt=t_i\mu$ $t'tP'Pt=\mu k$	$P_i'Pt=t_i\mu_i$ $t'P'Pt=\Sigma\mu_i$	$P_i'PT=T_iS_i$ $tr(T'P'PT)=tr(\Sigma S_i)$
PØ	$\emptyset_iP_i'P\emptyset t=t_i\mu$ $t'\emptyset P'P\emptyset t=\mu k$	$\emptyset_iP_i'P\emptyset t=t_i\mu_i$ $t'\emptyset P'P\emptyset t=\Sigma\mu_i$	$\emptyset_iP_i'P\emptyset T=T_iS_i$ $tr(T'\emptyset P'P\emptyset T)=tr(\Sigma S_i)$

Table 1. Six solutions for relations among k sets of variables. Top line in each cell gives the stationary equation; bottom line defines the criterion. In solutions with $T_i'T_i=I$ there is the additional requirement that $\Sigma S_i=\Psi^2$ must be a diagonal matrix.

TABLE 2

		1^{st} princ.comp.	2^{nd} princ. comp.
P	$t't=k$	.377	.545
PØ	$t't=k$	.731	.156
P	$t_i't_i=1$	.350	.586
PØ	$t_i't_i=1$	.687	.177
P	$T_i'T_i=I$	.470	.453
PØ	$T_i'T_i=I$	.544	.375

Table 2. Percentage of explained variance of first and second principal component, averaged over the three sets.

TABLE 3

objects		variables 1	2	3
	1	a	p	u
	2	b	q	v
	3	a	r	v
	4	a	p	u
	5	b	p	v
	6	c	p	v
	7	a	p	u
objects	8	a	p	v
	9	c	p	v
	10	a	p	v
	11	a	q	w
	12	b	r	w
	13	c	p	w
	14	b	q	u
	15	c	r	u

Table 3. Numerical example of three nominal variable with three categories in each variable.

TABLE 4

	1			2			3		
	a	b	c	p	q	r	u	v	w
1	1	0	0	1	0	0	1	0	0
2	0	1	0	0	1	0	0	1	0
3	1	0	0	0	0	1	0	1	0
4	1	0	0	1	0	0	1	0	0
5	0	1	0	1	0	0	0	1	0
6	0	0	1	1	0	0	0	1	0
7	1	0	0	1	0	0	1	0	0
8	1	0	0	1	0	0	0	1	0
9	0	0	1	1	0	0	0	1	0
10	1	0	0	1	0	0	0	1	0
11	1	0	0	0	1	0	0	0	1
12	0	1	0	0	0	1	0	0	1
13	0	0	1	1	0	0	0	0	1
14	0	1	0	0	1	0	1	0	0
15	0	0	1	0	0	1	1	0	0
marg. freq.	7	4	4	9	3	3	5	7	3

Table 4. Indicator matrix derived from Table 3.

TABLE 5.

		object scores	category quantifications
	$t't=k$	$Pt_{(j)}(n/k\mu_{(j)})^{\frac{1}{2}}$	$P_i t_{i(j)}(nk/\mu_{(j)})^{\frac{1}{2}}$
P	$t_i't_i=1$	$Pt_{(j)}(n/\Sigma\mu_{i(j)})^{\frac{1}{2}}$	$P_i t_{i(j)}k(n/\Sigma\mu_{i(j)})^{\frac{1}{2}}$
	$T_i'T_i=I$	$PT\Psi^{-1}n^{\frac{1}{2}}$	$P_i T_i \Psi^{-1} kn^{\frac{1}{2}}$
	$t't=k$	$P\emptyset t_{(j)}(n/k\mu_{(j)})^{\frac{1}{2}}$	$P_i \emptyset_i t_{i(j)}(nk/\mu_{(j)})^{\frac{1}{2}}$
P$\emptyset$	$t_i't_i=1$	$P\emptyset t_{(j)}(n/\Sigma\mu_{i(j)})^{\frac{1}{2}}$	$P_i \emptyset_i t_{i(j)}k(n/\Sigma\mu_{i(j)})^{\frac{1}{2}}$
	$T_i'T_i=I$	$P\emptyset T\Psi^{-1}n^{\frac{1}{2}}$	$P_i \emptyset_i T_i \Psi^{-1} kn^{\frac{1}{2}}$

Table 5. Coordinates of object points and points for quantified categories.

TABLE 6

	$P_1\emptyset_1$		$P_2\emptyset_2$		$P_3\emptyset_3$	
1	.65	0	.49	0	-.76	.36
2	-.57	.71	-.73	.71	.64	.15
3	.65	0	-.73	-.71	.64	.15
4	.65	0	.49	0	-.76	.36
5	-.57	.71	.49	0	.64	.15
6	-.57	-.71	.49	0	.64	.15
7	.65	0	.49	0	-.76	.36
8	.65	0	.49	0	.64	.15
9	-.57	-.71	.49	0	.64	.15
10	.65	0	.49	0	.64	.15
11	.65	0	-.73	.71	-.24	-.96
12	-.57	.71	-.73	-.71	-.24	-.96
13	-.57	-.71	.49	0	-.24	-.96
14	-.57	.71	-.73	.71	-.76	.36
15	-.57	-.71	-.73	-.71	-.76	.36
$\emptyset^2$	5.60	4.00	5.40	3.00	5.92	3.55

Table 6. Solution for PØ.

FIGURE CAPTIONS

Figure 1. Solution for P with $t't=k$. Ellipses are drawn on scale 8/3. Of vectors $Pt_{(j)}$ only the directions are indicated (not the proper length). Vectors $P_i t_{i(j)}$ are labelled as t_{ij}; they are connected by their corresponding points on the P_i-ellipse by dotted lines. Objects are labelled with their number of Table 3. Objects 1 4 7 are identical (label 1), 6 9 are identical (label 6) and 8 10 (label 8).

The essential geometrical property is that $Pt_{(1)}$ and $Pt_{(2)}$ are orthogonal, and therefore that the ellipses are surrounded by tangential rectangles.

Figure 2. Solution for P with $t_i't_i=1$. Scale and labels as in Figure 1. The essential property is that $Pt_{(1)}$ and $Pt_{(2)}$ are not orthogonal; ellipses are circumscribed by tangential parallelograms.

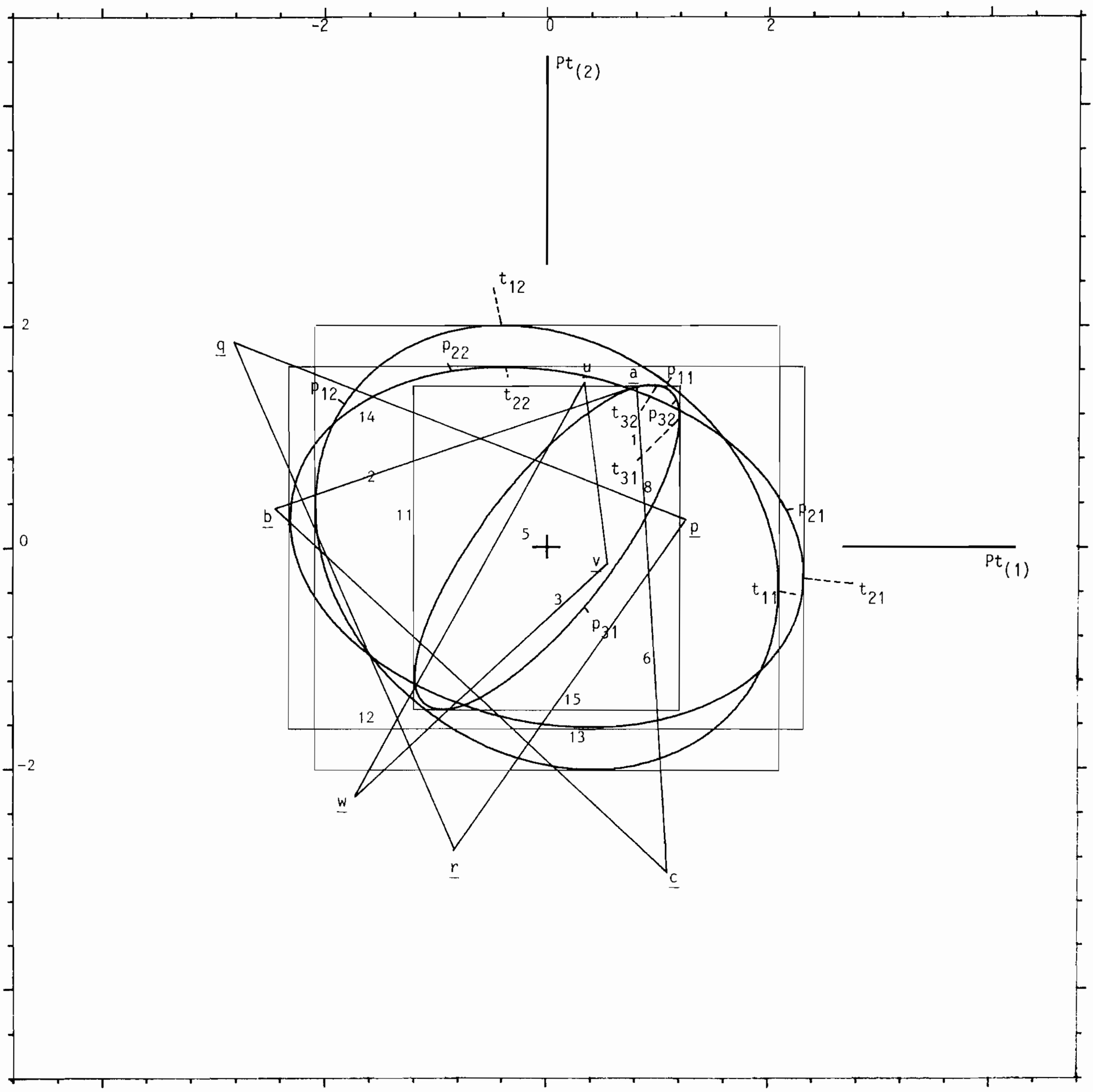
Figure 3. Solution for P with $T_i'T_i=I$. Scale and labels as in Figure 1. Vectors $Pt_{(1)}$ and $Pt_{(2)}$ are orthogonal (as in Figure 1), but tangential parallelograms no longer are orthogonal to $Pt_{(1)}$ or $Pt_{(2)}$. Instead, the parallelogram for the P_i-ellipse has sides parallel to $P_i t_{i(1)}$ or $P_i t_{i(2)}$.

Figure 4. Solution for $P\emptyset$ with $t't=k$. Ellipses are drawn on scale 4/3. Vectors $P_i\emptyset_i t_{i(j)}$ (or their corresponding point on the ellipse) have label t_{ij}. Principal components $p_{i1}\emptyset_{i1}$ and $p_{i2}\emptyset_{i2}$ have label p_{i1} and p_{i2}. Geometrical properties are the same as in Figure 1. However, there is much less emphasis on the distance between infrequent categories <u>b</u> and <u>c</u>, or <u>q</u> and <u>r</u>, or <u>u</u> and <u>w</u>.

Figure 5. Solution for $P\emptyset$ with $t_i't_i=1$. Scale and labels as in Figure 4. The tangential parallelograms of $P_1\emptyset_1$ and $P_2\emptyset_2$ have been cut off arbitrarily at top and bottom. Vectors $p_{21}\emptyset_{21}$, $P_2\emptyset_2 t_{2(1)}$, and $P_2\emptyset_2 t_{2(2)}$ are so close together that only one label (t_2) has been used for all three.

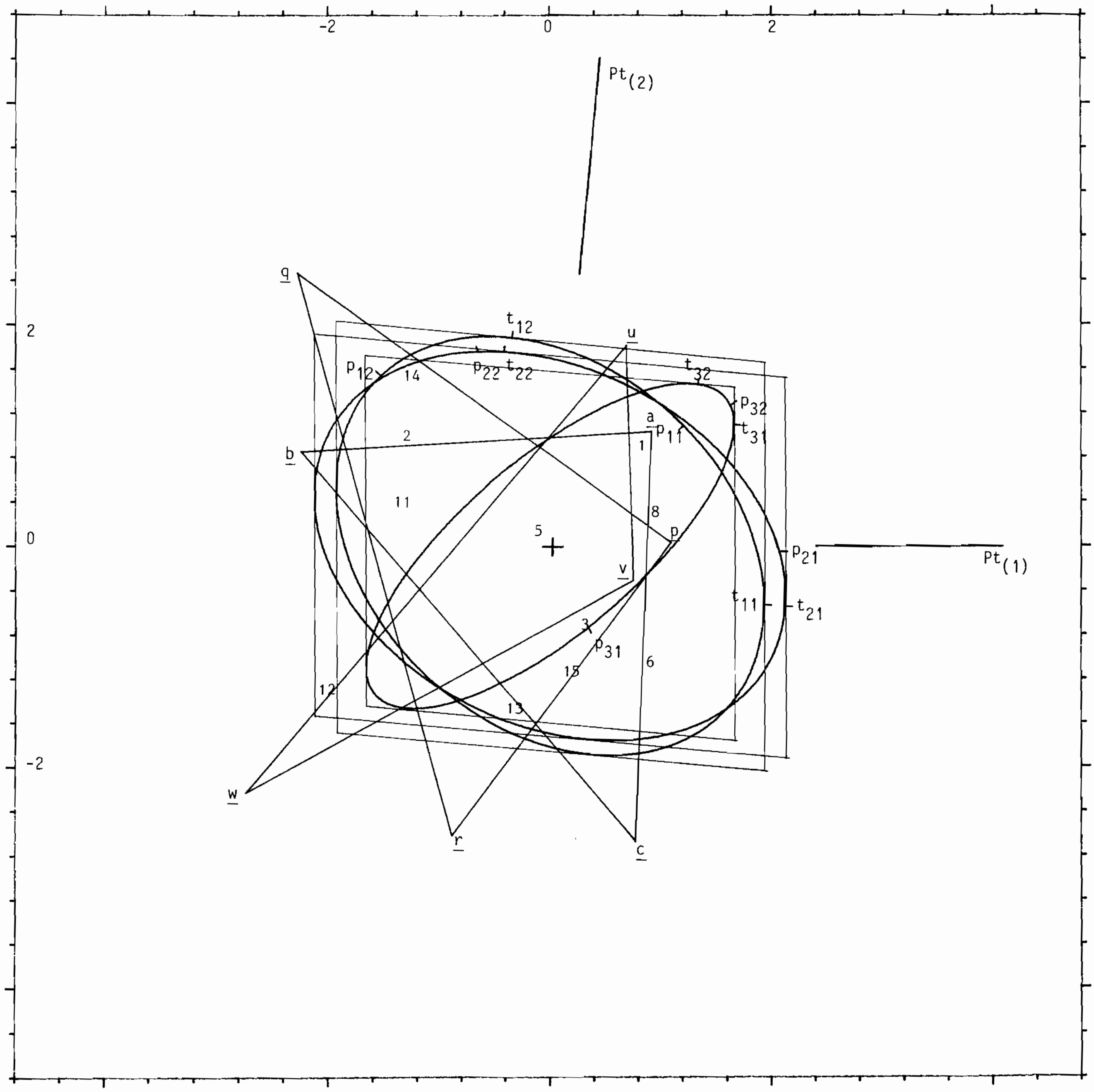
Figure 6. Solution for $P\emptyset$ with $T_i'T_i=I$. Scale and labels as in Figure 4.

plot 1

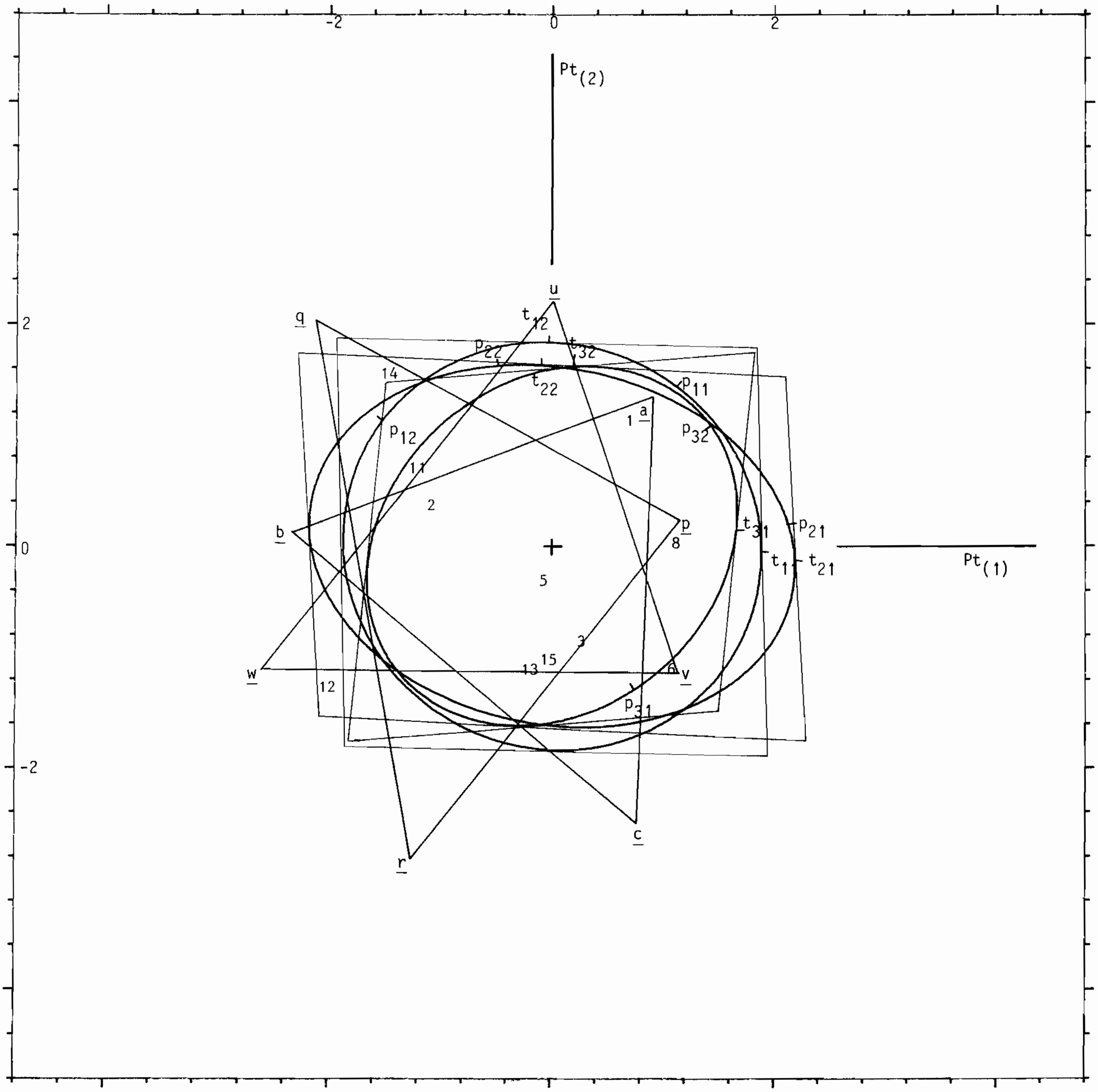


-2
0
2
2
0
-2
Pt(2)
Pt(1)
t12
p22
t22
p12
p11
t32
p32
t31
p21
t11
t21
p31
q
b
p
u
a
v
w
r
c
1
2
3
5
6
8
11
12
13
14
15

plot 2

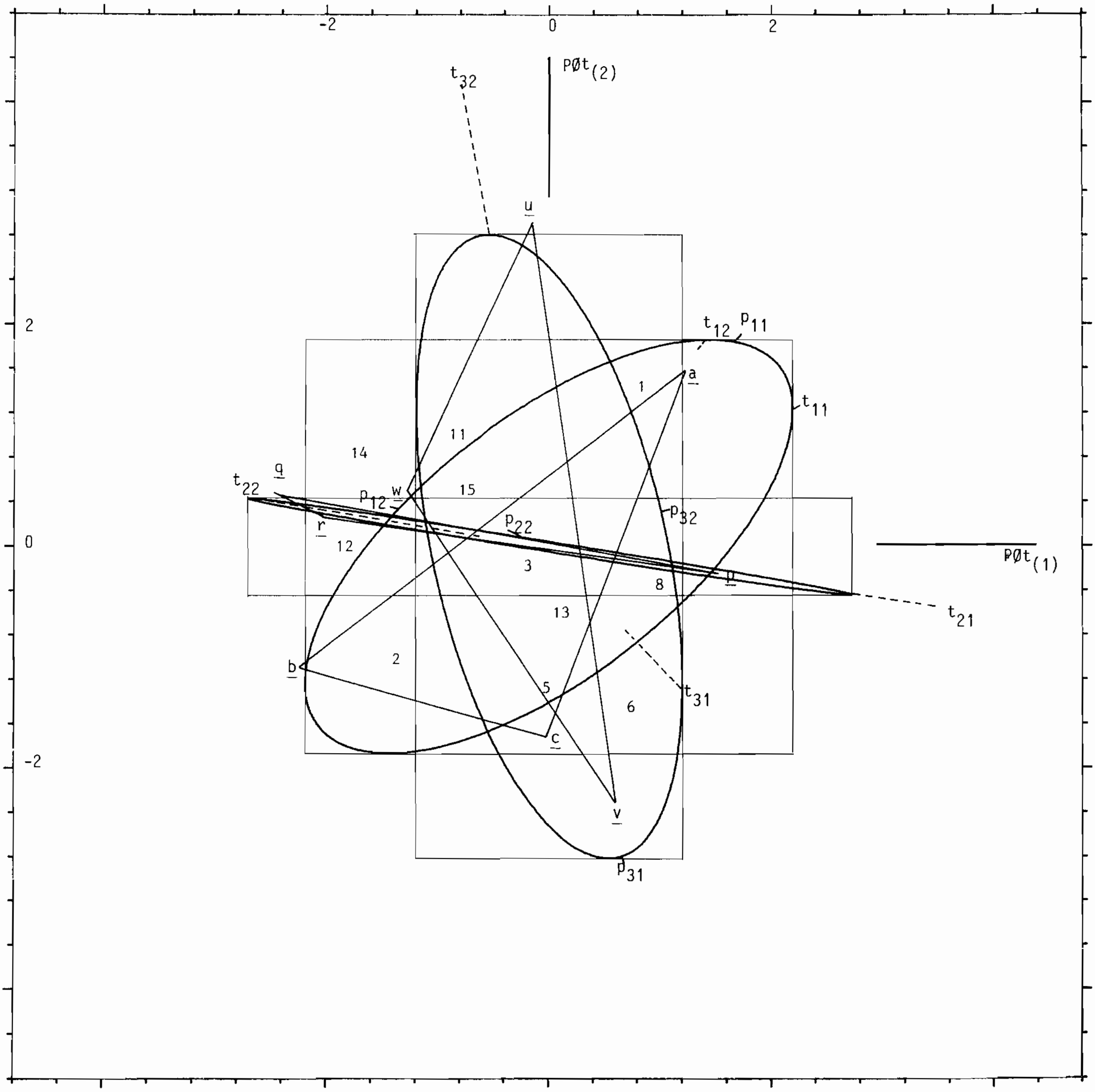

-2
0
2
2
0
-2
Pt(2)
Pt(1)
q
u
b
a
p
v
w
r
c
t12
t22
t32
t31
t11
t21
p12
p22
p32
p11
p21
p31
1
2
3
5
6
8
11
12
13
14
15

plot 3

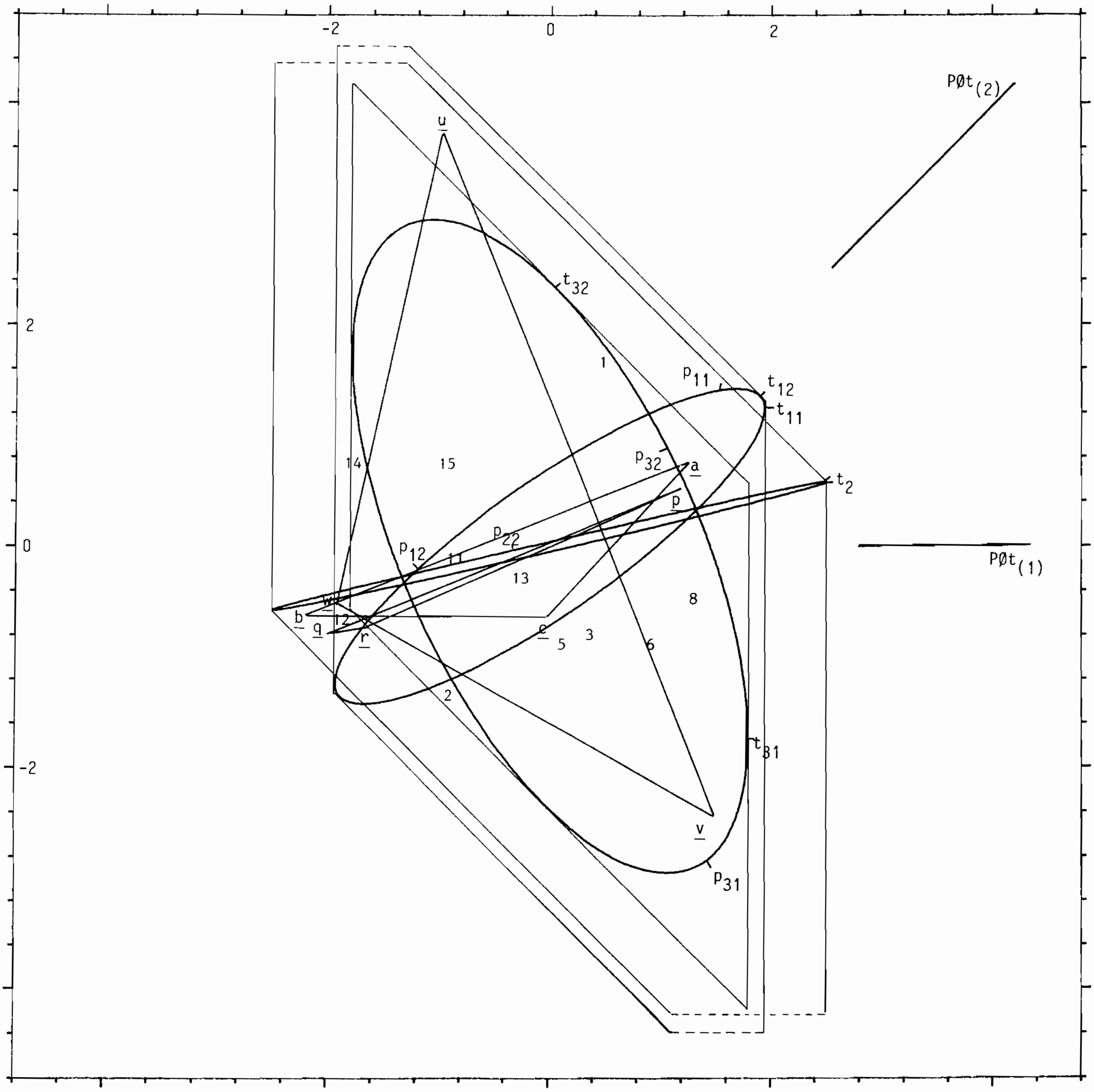


-2
0
2
Pt(2)
Pt(1)
u
q
t12
p22
t32
t22
14
p11
a
1
p12
p32
11
2
b
p
t31
p21
8
t1
t21
5
3
15
13
w
12
v
p31
c
r

plot 4

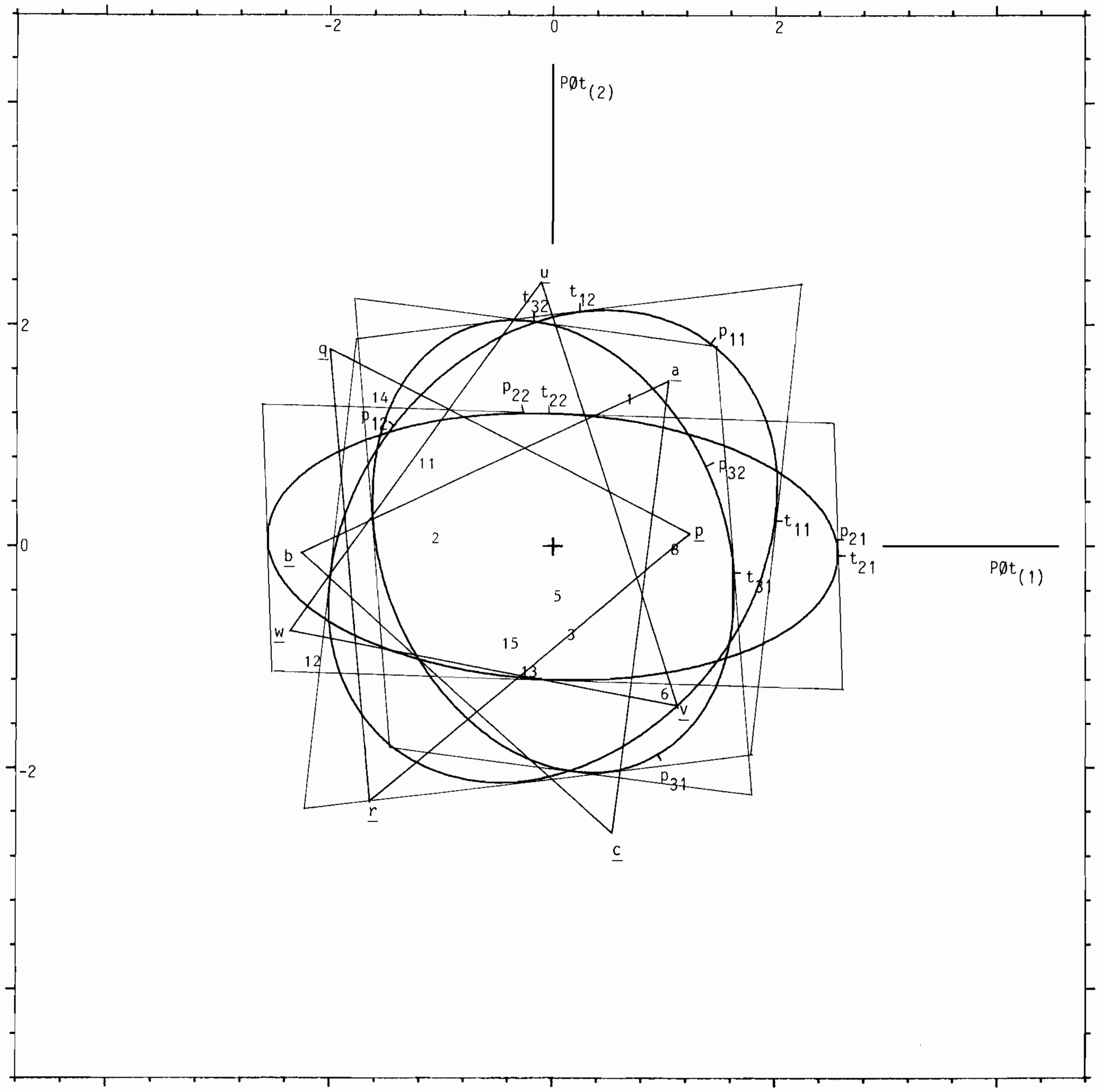

-2
0
2
2
0
-2
PØt(2)
PØt(1)
t32
u
t12
p11
a
t11
1
11
14
q
t22
p12
w
15
p32
p22
r
12
3
p
8
t21
13
2
b
5
t31
6
c
v
p31

plot 5



-2
0
2
2
0
-2
PØt(2)
PØt(1)
u
t32
1
p11
t12
t11
p32
a
t2
14
15
p
p22
p12
11
13
8
w
b
12
q
r
c
5
3
6
2
t31
v
p31

plot 6



-2
0
2
PØt(2)
PØt(1)
u
t32
t12
p11
q
a
p22 t22
p12
14
11
1
p32
t11
p21
t21
2
b
p
8
t31
5
w
15
3
12
13
6
v
p31
-2
r
c