The covariance biplot and stock market data: An alternative relative strength chart

G.D.I. Barr*

Department of Mathematical Statistics and Economics, University of Cape Town, Private Bag, Rondebosch 7700
Republic of South Africa

J.F. Affleck-Graves

Graduate School of Business, University of Cape Town, Private Bag, Rondebosch 7700, Republic of South Africa

Relative strength analysis is an important technique for many investment analysts. In this paper an alternative to the traditional relative strength chart is proposed, based on the covariance biplot. It is argued that this method provides more information than the traditional relative strength analysis in that it provides a visual picture of both the relative riskiness of the individual securities and the degree of co-movement between the various securities, in addition to the change in relative strength over a number of time periods. A methodology is proposed for transforming stock-market data into a form suitable for the covariance biplot routine. This transformation involves both a smoothing and scaling of the original price series data. The methodology is illustrated by application to six of the JSE Actuarial Sector Indices and a step-by-step guide to the interpretation of the covariance biplot is provided.

Relatiewesterkte-analise is 'n belangrike tegniek van baie investeringsanaliste. In hierdie artikel word 'n alternatief tot die tradisionele relatiewesterktegrafiek voorgestel wat gebaseer is op die kovariansie-tweepuntgrafiek. Daar word geredeneer dat hierdie metode meer inligting as die tradisionele analise verskaf, deurdat dit 'n visuele voorstelling van beide die relatiewe risiko van die individuele sekuriteite aan die een kant en die graad van medebeweging tussen die verskillende sekuriteite aan die ander kant, verskaf, benewens die verandering in die relatiewe sterkte van 'n aantal tydsperiodes. 'n Metodologie vir die transformasie van aandelemarkdata na 'n geskikte vorm vir die kovariansie-tweepuntroetine word voorgestel. Hierdie transformasie behels beide gladmaking en skalering van die oorspronklike prysreeksdata. Die metodologie word ge'illustreer deur toepassing by ses van die Johannesburgse Effektebeurs se 'Actuarial Sector Indices' en 'n stap-vir-stap interpretasiegids vir die kovariansie-tweepuntgrafiek word voorsien.

*To whom correspondence should be addressed

Introduction

The covariance biplot is a useful graphical tool for understanding and demonstrating the relationships in multivariate time series data. The biplot was devised by Gabriel (1971; 1972; 1981) and the extensive literature on the subject has been reviewed by Greenacre and Underhill (1982) and Greenacre (1984).

In this article we demonstrate how this graphical display technique can be used to interpret and track the relative movement of a number of Johannesburg Stock Exchange (JSE) sector indices over the period 1973 1985. It will be shown how, from a single two-dimensional plot, one can infer information about relative riskiness, covariation, and relative strength of a number of securities (or sector indices). Moreover the interpretation of the plot is straightforward and does not necessitate an understanding of the theoretical procedures underlying it.

In the second section the mathematical theory underlying the covariance biplot is briefly explained. A methodology for applying the covariance biplot to stock-market data ia proposed in the third section and illustrated on six sector indices chosen from the set of Johannesburg Stock Exchange Actuarial Indices. The covariance biplot for these six sectors for the period 1973–1985 is presented in the final section, which also provides a detailed step-by-step explanation of how the biplot should be interpreted. The article closes with a brief section of conclusions.

Mathematical procedure

The covariance biplot treats a $n \times m$ matrix as two clouds of points, n points representing the rows (or time points) and m points representing the columns (sector indices, in this case). The algorithm projects the clouds onto a low (usually two)-dimensional subspace, in such a way that the maximum amount of the variability in the data matrix is captured.

Let X be the $n \times m$ matrix to be biplotted. X is centred by subtracting column means to give:

$$Y = X - i/n\,(ii^{\,t}X)$$

where $i =$ a unit vector of dimension n; and $i^{\,t} =$ the transpose of i.

The singular value decomposition of Y is then computed to yield:

$$Y = UD_a V^{\,t}$$

where $D_a = \text{diag}\,(a_1, a_2, \dots a_m)$ with $a_1 \leqslant a_2 \leqslant a_3 \cdots \leqslant a_m$.

The rows of the $n \times m$ matrix, $F = U$, and the $m \times m$ matrix, $G = VD_a$ yield the coordinates for the graphical displays of the row and column points, respectively. The display in p dimensions (usually two) is obtained by considering the first p columns of F and G. The interpretation of the covariance biplot is in terms of the distances of the points from the origin (or

lengths) and the cosines of the angles that pairs of points subtend at the origin (angles between two points). The features exhibited by the covariance biplot can then be listed as follows:

a) The quality of the biplot display in p dimension is given by

$$q_p = \sum_{i=1}^{p} a_i^2 \Big/ \sum_{i=1}^{m} a_i^2$$

b) The length of the i th row point multiplied by the length of the j th column point multiplied by the cosine of the angle between them is an approximation to the element in row i and column j of the matrix Y. Therefore row points which plot in the same half-plane as a particular column point have values which are above the mean of that column point; row points which plot in the opposite half plane have values which are below the mean of that column point; row points which lie along a line at right angles to a line from the origin to the column point have values approximately equal to the mean of that column point.

Consider a hypothetical matrix with three rows and only one column with the biplot shown in Figure 1.

From Figure 1 it can be seen that the first element (R_1) of the column (C) is above the mean of the column because it plots in the same half plane as C. The second element (R_2) of the column approximates to the mean because it lies on the line drawn perpendicular to the line from the origin to the column point (C). Finally, the biplot indicates that the third element (R_3) of the column is below the mean. A corollary of this feature is that if two row points are similar they will plot very close together.
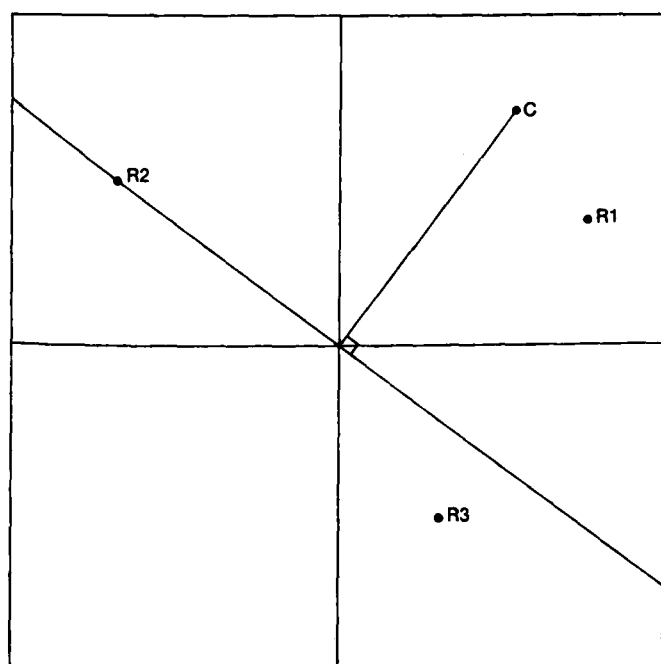

Figure 1 Biplot for hypothetical matrix with 3 rows and only 1 column

c) The length of the column points are approximations to the standard deviation of the column and the cosine of the angle between column points is an approximation to the correlation between the respective column variables. Therefore two highly correlated column variables will lie in the same direction from the origin, negatively correlated column variables will lie on opposite sides of the origin, and if the correlation between two column variables is close to 0. they tend to lie at right angles to each other.

Methodology

To illustrate the application of the covariance biplot in analysing economic data, data were collected on the following six JSE Actuarial Sector Indices:
1. All gold
2. Metals and minerals
3. Banks
4. Industrial holding
5. Motors
6. Stores

These sectors were chosen to give a reasonably wide representation of the sectors of the Johannesburg Stock Exchange. It is not implied, however, that these are either the most important sectors or that they together entirely represent the JSE. They were merely chosen for illustrative purposes.

Weekly closing prices were collected on each of these sectors from 1 July 1973 to 30 June 1985. These data were then converted into half-yearly data by computing the average level of the index in each succeeding half year. Thus the index data were reduced to a sequence of 24 values per index as follows:

$$I_{it.} = \sum_{k=i}^{26} (I_{itk}/26)$$

where $I_{it.}$ = the average level of index i in half year t $(i = 1, 2,...6,; t = 1, 2, 3,...24)$; and I_{itk} = the level of index i in week k in half year t.

These sequences of the I_{it} were then normalized so that each index had a value of 100 for the first data point (i.e. the second half of 1973). That is,

$$IS_{it} = 100 \times I_{it.}/I_{i1.}$$

where IS_{it} = the average of index i in half year t after normalizing to 100 in the second half of 1973.

As the object of the covariance biplot is to show the relative position of the individual indices, both with respect to one another and with respect to time, the data for each half year and each index were scaled by dividing the value of the index in a particular half year by the average value of all six indices in that half year. That is,

$$X_{it} = IS_{it}\Big/ \sum_{j=1}^{6} (IS_{jt}/6)$$

This scaling procedure merely converts the average level of the index in a particular half year to a value which

reflects the movement of the sector in that half year relative to the movement of all other sectors (in the study) in that half year.

This is of course nothing more than the familiar relative strength concept widely used by technical analysts in analysing movements in share prices and economic time series. The sequence of 24 X_{it} values for each sector provided the input data for the covariance biplot routine.

Results

The output of the covariance biplot routine (using computer programs written by L.G. Underhill of the Department of Mathematical Statistics) is shown, in two dimensions, in Figure 2. In this figure, the six sectors are represented by the numerals 1 to 6 (in the order presented in the Methodology section) while the time points are represented by the letters A to X, where A represents the second half of 1973, B represents the first half of 1974, C the second half of 1974, etc.

As Figure 1 contains a great deal of information , it may appear somewhat confusing to the reader at first glance. Therefore, we have broken this sketch into a number of simpler sketches in order to illustrate the concept of the biplot and to demonstrate the various kinds of analyses that can be made using it. Once the reader has learned how to examine the individual components of the biplot, he will be able to interpret Figure 2 more easily which is, of course, the only output from the biplot routine.

Let us begin by examining Figure 3 which is the covariance biplot showing only the six indices (i.e. with the time points removed).

This figure is similar to the familiar preference map (or multidimensional scaling) widely used in marketing
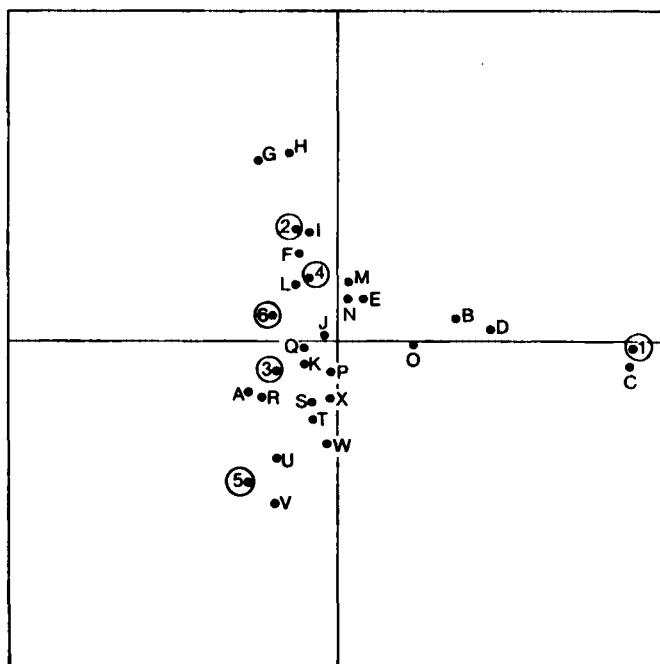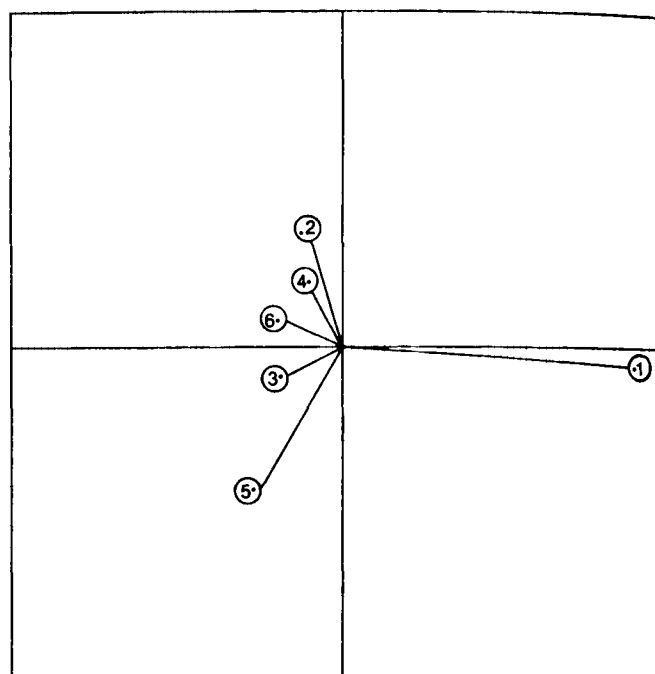

Figure 3 Covariance biplot showing only the six indices with the time points removed

research. In analysing economic data, this section of the biplot provides two different and important pieces of information.

Firstly, variability is indicated by the distance of the point from the origin. The greater the distance from the origin, the greater the variability of the data. In the case of stock-market data, variability provides a measure of risk (Markowitz, 1952). Therefore, in the above example we see that over the period examined (1973–1985), on average, the All gold sector (Sector 1) was by far the most risky of the six sectors. Indeed, Figure 3 indicates that the All gold sector was almost twice as risky as the second most risky sector, the Motor sector (Sector 5), almost three times as risky as the Metals and minerals sector (Sector 2), and four times as risky as the Banking sector (Sector 3), the Industrial holding sector (Sector 4) and the Stores sector (Sector 6).

The Motor sector was the second most risky sector followed by the Metals and minerals sector. The least risky sectors were the Banking, Industrial holding, and Stores sectors, which all have approximately the same degree of risk.

The second piece of information provided by this aspect of the biplot is the relationship which exists between the various sectors. If one draws a line from each sector's point to the origin, then the cosine of the angle between any two lines gives the correlation between the two indices concerned over the period analysed. Therefore sectors for which the angle is small tend to move in unison while sectors for which the angle is close to 180 degrees tend to move in opposite directions. Finally, sectors for which the angle is close to 90 degrees tend to move independently of one another.

Examination of Figure 3 therefore indicates that the All gold sector tended on average over this period to

Figure 2 Covariance biplot in two dimensions

move in the opposite direction to the other five sectors. This follows since the angles between the line segment from the All gold sector and the line segments from all other sectors are greater than 90 degrees. As regards the other sectors, the figure indicates a fairly high degree of co-movement between the Motor sector (Sector 5) and the Banking sector (Sector 3) and a fairly high degree of co-movement between the Industrial holding sector (4) and the Stores sector (6).

In demonstrating the use of the biplot in analysing the time plot of points. we have found it convenient to examine the sequence in five distinct phases. This is done in Figures 4 to 8.
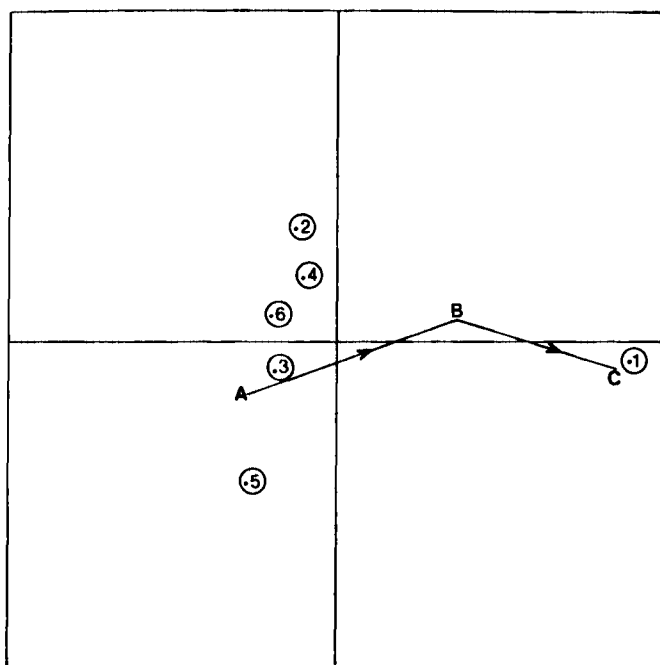

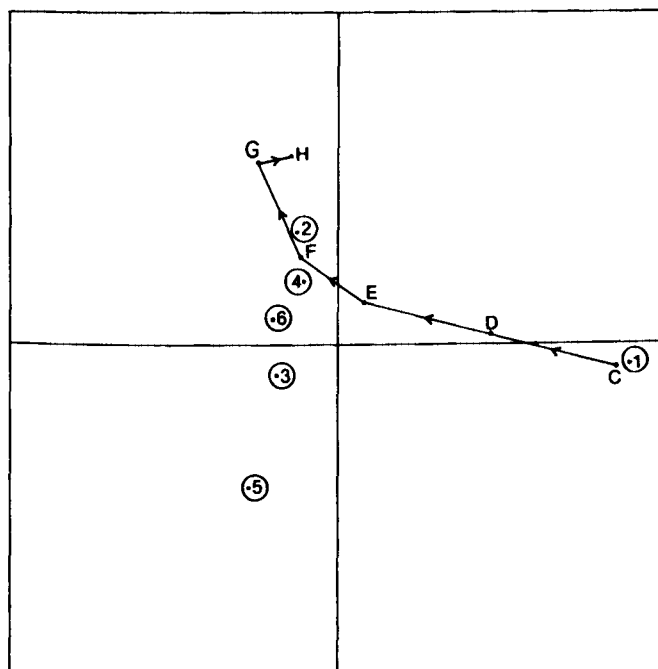**Figure 4** Movement during the first three half years

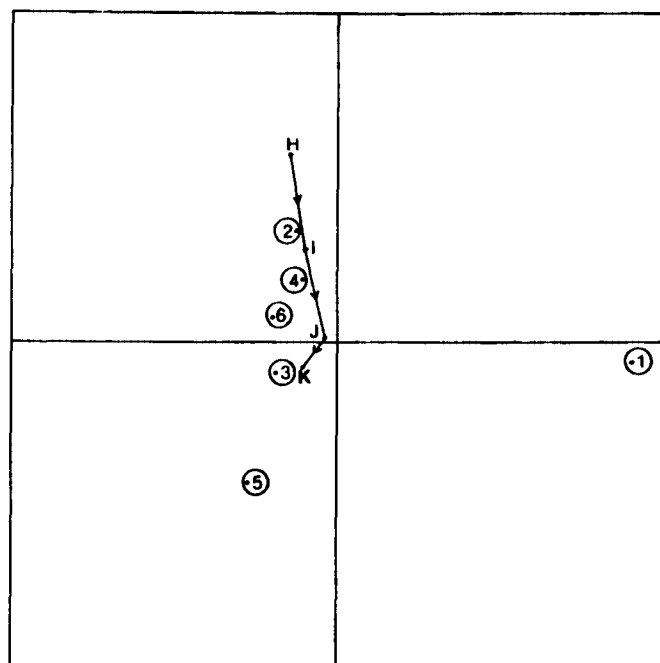**Figure 5** Relative movements of six sectors

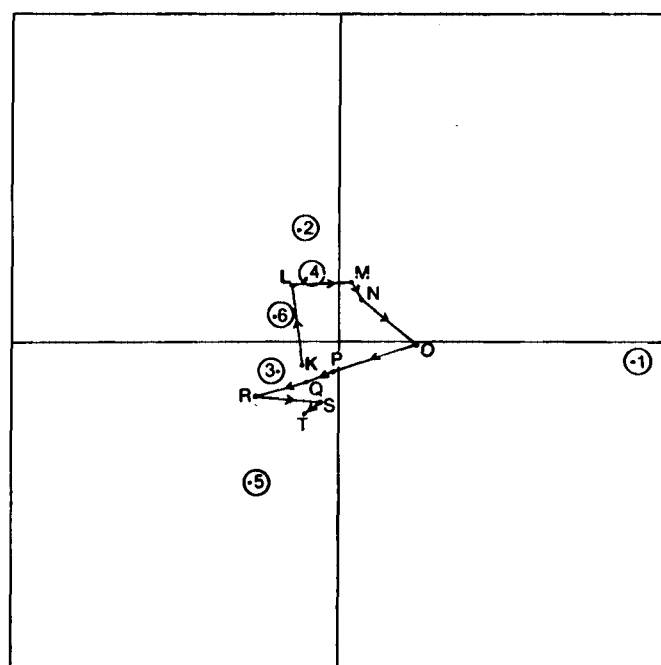**Figure 6** Part of the covariance biplot showing relative movements

Figure 7 Relative movements of six sectors

Figure 4 demonstrates the movement during the first three half years (represented by points A to C); that is, from the second half of 1973 to the second half of 1974. This period represented a major boom in the Gold sectors of the JSE and is illustrated in Figure 4 by a clear movement towards the All gold sector and away from the other sectors. This shows that during this period the All gold sector gained significantly in relative strength vis-a-vis the other sectors.

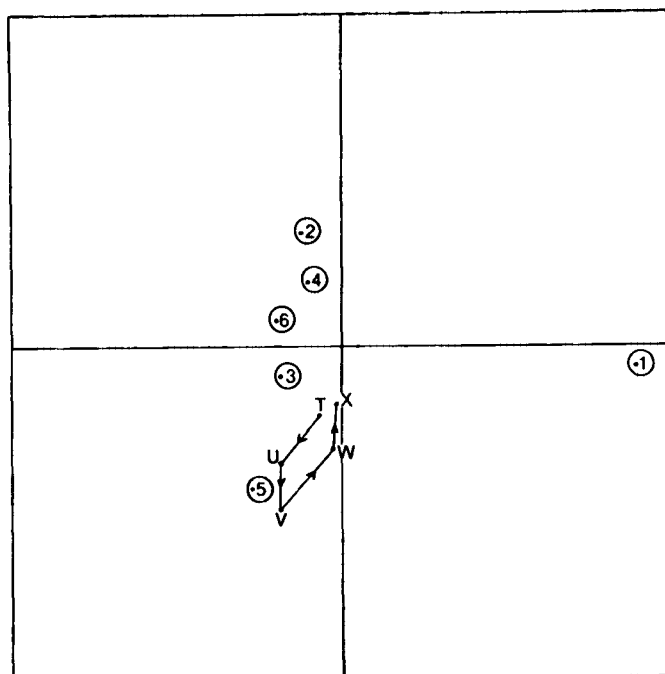In Figure 5, the relative movements of these six

Figure 8 Relative movements from first half of 1983
to first half of 1985

sectors are illustrated for the next five half years (represented by points C to H); that is, from the second half of 1974 to the first half of 1977. Examination of Figure 5 indicates that this period saw an initial movement (points C and D) away from the All gold sector towards all other sectors. Thereafter there is a continued movement away from the All gold sector (points D, E, and F) but now predominantly towards the Industrial holding sector (Sector 4) and the Metals ard minerals sector (Sector 2). There is also evidence of a degree of movement away from the Motor sector (Sector 5). Finally, this period ends with a strong movement (points F, G, and H) towards the Metals and minerals sector and away from all other sectors.

Figure 6 demonstrates that part of the covariance biplot showing relative movements from the first half of 1977 to the second half of 1978 (points H to K). The figure indicates that the period was characterized by a movement away from the Metals and minerals sector towards the Industrial holding sector, the Stores sector, the Banking sector and the Motor sector. It is interesting to note that during this period the Gold sector did not significantly gain or lose relative strength. This is evidenced in the biplot by the movement from point H through to Point K being essentially perpendicular to the line segment from the All gold sector to the origin.

Figure 7 shows the relative movements of the six sectors from the second half of 1978 to the first half of 1983 (points K to T). The points represent a fairly long period during which no clear change in the relative strengths of the various sectors is apparent. Therefore the biplot indicates that during this period there was no obvious indication of any one sector consistently gaining relative strength vis-a-vis the other sectors.

Figure 8 illustrates relative movements from the first half of 1983 to the first half of 1985 (points T, U, V, W, and X). This period is characterized by an initial period

in which the Motor sector gained relative strength vis-a-vis all other sectors (points T, U, and V). This was followed by a period in which this movement was almost exactly reversed (points V, W, and H).

Conclusion

As mentioned at the outset of the previous section, the covariance biplot results in a single sketch which is that shown in Figure 2. This single sketch provides a pictorial representation of several aspects of importance to investment analysts.

Firstly, the sketch indicates the absolute riskiness of each of the individual sectors. This is represented by the distance of the point from the origin. At the same time, the sketch indicates the relative riskiness of each sector vis-a-vis the other sectors. For each pair of securities this is indicated by the angle between the line segments drawn from each of the points to the origin.

Finally, the sketch also provides an indication of how the relative strengths to the various sectors have changed over time. This is shown by a series of points plotted successively through time. Any movement towards the point representing a sector indicates an increase in the relative strength of that sector while any movement away from the point represents a decrease in relative strength.

The use of the covariance biplot was illustrated by analysing the relative movements of six sectors of the JSE from 1973 to 1985. Of course the technique can be used on other sectors or on individual securities. Indeed, it can be applied to any set of economic time series data.

It is our contention that the covariance biplot provides a good alternative to the more traditional relative stregth analyses as it provides a single sketch for several securities simultaneously and provides additional information such as the absolute and relative riskiness of individual data points. Whether charting patterns can be established that provide a profit-making opportunity remains on open research question.

References

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

Gabriel, K.R. 1972. Analysis of meteorological data by means of cononical decomposition and biplots. *J. Appl. Meteorol.*, 11, 1071–1077.

Gabriel, K.R. 1981. Biplot display of multivariate matrices for inspection of data and diagnosis. In Barnett, V. (ed.) *Interpreting multivariate data*, 147–174. Chichester, U.K.: Wiley.

Greenacre, M.J. & Underhill, L.G. 1982. Scaling a data matrix in low-dimensional Euclidian space. In Hawkins, D.M. (ed.) *Topics in applied multivariate analysis*, 183–268. Cambridge, U.K.:Cambridge University Press.

Greenacre, M.J. 1984. *Theory and applications of correspondence analysis*. Academic Press.

Markowitz, H.M. 1952. Portfolio selection. *J. Finance*, 7, 77–91.