A Graphical Method for Assessing the Results of Fuzzy Clustering

Patrick Bobbitt[1]
Bureau of Labor Statistics, 2 Massachusetts Ave. NE, Room 3655, Washington, DC 20212
bobbitt_p@bls.gov

KEY WORDS: clustering, K-Means clustering, fuzzy clustering

1. Introduction

In one sense, statistics is practice of summarizing data by reducing the dimensionality of a data set in a sensible way. One way in which the dimensionality of a data set can be reduced is to classify the n items in a dataset into k clusters. One such method of classification is known as the k-means clustering algorithm. In this algorithm the statistician uses a measure of distance between the items to define items in a dataset as similar, and then classifies that item as having membership to one of k clusters.

In some instances, the k-means algorithm may reduce the dimensionality a bit too far since it does not give any insight into how close specific elements of a cluster were to being classified differently. To aid in this conundrum, a method of 'fuzzy clustering' was developed that assigns a probability[1] to each element of being included in each of the k cluster. While inserting this extra information certainly addresses the issue at hand, it also increases the dimensionality of the summary from a $(1 \text{ x } \mathrm{n})$ matrix (i.e. a single column with a number to classify the cluster), to a $(k \text{ x } n)$ matrix (i.e. a probability of inclusion to each of the k clusters for each element in the data set).

Graphics have long been a staple for statisticians to reduce the dimension of data. The most readily used example of this is the use of the scatter-plot to display the association between two variables. This form of summary works because there is a 1:1 mapping between the association one sees in the points and the mathematical model used to describe the line. In this paper, we propose to use color to similarly reduce the dimensionality of the $(k \text{ x } n)$ fuzzy clustering summary, enabling the user to gain inference from the fuzzy clustering techniques without having to examine a dramatically increased set of numbers.

Section 2 of this paper gives a review of how color is reproduced, and defines the Red, Green, Blue (RGB) color space. Section 3 defines the space of the $(k \text{ x } n)$ summary produced by fuzzy clustering techniques (fuzzy space), suggests a possible extension to the method, and demonstrates a useful map between the fuzzy space and the RGB color space. Section 4 covers two specific examples of applying this map, and Section 5 gives overall conclusions as well as points for further study.

The technique presented in the paper is the culmination of an effort to present the results of several measures computed to summarize the 'quality' of each of the 120 price indexes published monthly at the International Price Program (IPP).

2. Color

Overview

Color exists as a response of our eyes observing specific wavelengths of light. White light comes from a source that contains all wavelengths of visible light, and our usual notions of color are responses to light that has been 'filtered' by reflecting back from a substance that 'filters out' other specific wavelengths of light within the white light. This form of color production is called a 'subtractive' system of color reproduction, and examples would be the color you see outside in everyday objects. Another method of observing color is to observe frequencies of light being produced at the source. This form of color production is referred to as an 'additive' system of color reproduction and examples would be the light found in lasers, given off from burning wood or natural gas.

Depending upon the method of production of color, there exist four primary colors that cannot be obtained by mixing any other. For the subtractive method these colors are Cyan, Magenta and Yellow, with Black being obtained by the presence of all, and white being obtained by the presence of none. For the additive method these colors are Red, Green and Blue, with Black being obtained by the presence of none, and white being obtained by the presence of all.

Because all other colors are merely 'blendings' of these primary colors depending upon the method of production, they form a hyper-plane in four dimensions, and can be described with a 'color cube' (see Figure 1). In practice we can't get every color since there are problems finding 'pure' primary colors with which to blend, this leads print media to use a four color process (Cyan, Magenta, Yellow, and Black commonly referred to as a CMYK process) to reproduce color, and Television/Computer/Film media to use a RGB (Red, Green, Blue) process to reproduce color.

As this paper is being reproduced electronically, we will focus on the RGB color model

RGB Color model

The standard method of displaying color on a RGB color monitor is to divide the amount of Red, Green and Blue into factional levels of intensity and then 'light' the three

[1] Opinions expressed in this paper are those of the author and do not constitute policy of the Bureau of Labor Statistics.
[2] These values aren't really probabilities, rather 'fractional classifications.' We chose to use the term since they have similar properties, and we felt their use would ease the reader's transition for notation and vocabulary, and broaden the appeal.

levels on a given point. For a Television screen, these fractional levels are continuous, for a computer screen, they are discrete, and for the 'richest' implementation each of the primary colors is broken up into 256 different levels, yielding 256^3 distinct colors.

Because primary colors can't be reconstructed by any of the other colors in the visible spectrum, it should be clear that they form a basis for the color cube, and the colors displayed on a computer screen are being modeled mathematically by a subset of $\Re^3$ (a three dimensional Cartesian set of real numbers), specifically a unit cube. Table 1 gives a listing of the colors and the vertices.

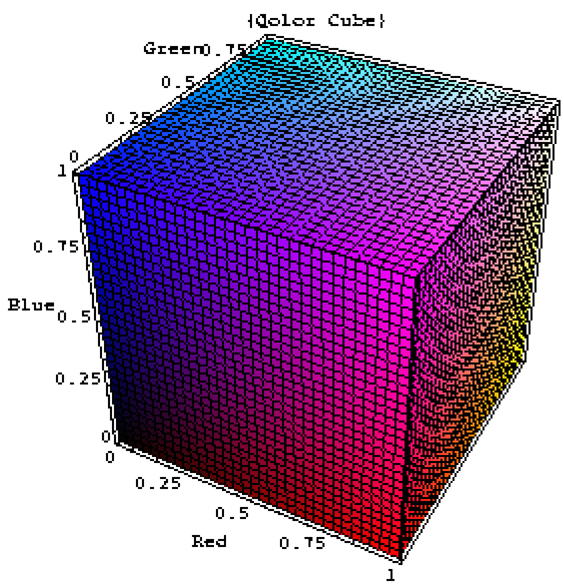

Figure 1: color cube

Table 1: Colors and their RGB colors space coordinates

Vertices	Corresponding color
(0,0,0)	Black
(1,0,0)	Red
(0,1,0)	Green
(1,1,0)	Yellow
(0,0,1)	Blue
(1,0,1)	Magenta
(0,1,1)	Cyan
(1,1,1)	White

3. Spaces induced by fuzzy clustering

Overview

In this section, we will outline the assumptions and goals of the known methods of fuzzy clustering ('general' fuzzy clustering), and describe the space of fuzzy clustering results. We then give a method for selecting an acceptable color scheme for each of the fuzzy spaces (induced by k=2, k=3, k=4) that can be possibly mapped into the RGB color-space. We will then propose an extension to the goals of known fuzzy clustering methods, constrained fuzzy clustering; and then describe the space of the results that one would encounter *if* such an algorithm was used that optimized those goals, as well as a method for selecting acceptable color schemes.

General Fuzzy Clustering

The goal of 'general' fuzzy clustering is to assign a probability of inclusion, $\pi_{i,j}$ for observation i into cluster j for each of the k known clusters. No assumption is made about possible constraints about membership to these clusters, and it is presumed that every point belongs to at least one of the clusters. The fuzzy clustering algorithm we chose to use was the 'fanny' algorithm used by S plus. This seeks to minimize:

$$\text{Objective Function} = \sum_{j=1}^{k} \frac{\sum_{i,v=1}^{n} \pi_{ij}^2 \pi_{vj}^2 d(i,j)}{2\sum_{v=1}^{n} \pi_{vj}^2}$$

where the dissimilarities $d(i,j)$ are known and the memberships $\pi_{i,j}$ are unknown. The assumptions of this algorithm are:

1. $\pi_{i,j} \geq 0$ for all $i = 1,\ldots,n$ and $j = 1,\ldots,k$
2. $\sum_{j=1}^{k} \pi_{i,j} = 1$ for all $i = 1,\ldots,n$

To visualize the space of possible values, consider first the simplest case when $k = 2$. In this case, we can easily see that we could describe this space as the interval from 0 to 1 (this would be the front edge on the triangle formed by the plane defined by the fuzzy space shown in Figure 2). Clearly any particular data point's 'fuzzy' assignment could be described by this segment (since $\pi_{i,1} = a \Rightarrow \pi_{i,2} = 1 - a, \forall a \in (0,1)$). If we had k=3, and we did not constrain inclusion in any way (i.e. membership into any cluster i, does not exclude membership into any other cluster), the triangular plane shown in figure 2 will describe the space of all possible $\pi_{i,j}$.

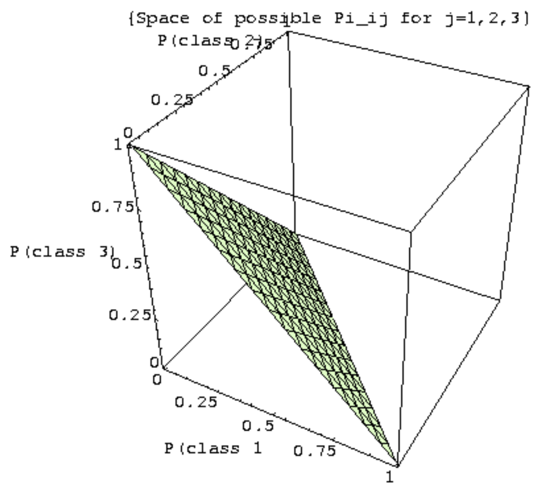

Figure 2: Space of possible values for π_{ij}

In general we could define the entire set of values for π_{ij} (the probability for observation i being in cluster j) as being the set of points defined by the following:

$$\left\{ \Pi = \begin{pmatrix} \pi_1 \\ \pi_2 \\ \vdots \\ \pi_k \end{pmatrix} : \sum_{i=1}^{k} \pi_i = 1, \pi_i \in [0,1] \right\}$$

Specifically, this would be a hyper-plane in k-dimensions. It should also be clear that the 'span' of the space[2] is k-1. Because of this, the largest k for which of the space can be visualized in three dimensions is the case of k=4. The shape of that space is the wedge shown in Figure 3. Specifically, this figure is the projection of Π onto the RGB color-space (shown in Figure 1). This can easily be seen by holding the fourth component of the vector constant. When this is done it then becomes the 'location' term for the triangular plane that is created by the other three free variables. If we vary this fourth component over all possible values, you will obtain the solid wedge shown in Figure 3. This 'wedge' also represents the greatest variation of color we can display the results of 'general' fuzzy clustering methods using the RGB color-space model of color reproduction.

Picking Colors for k=4

It should be evident that the wedge shown in Figure 3 provides a map from the set of all possible fuzzy values to particular palette of colors. Since this map is based upon a projection from a four dimensional space (the hyperplane) to a three dimensional space (the wedge), it isn't a 1:1 mapping[3]. As will be discussed at the end of this section, the mapping is 'close enough' to 1:1 for our purposes to still be useful. This map was obtained by setting the level of Red to the possible values of π_1, the level of green to the possible values of π_2, and the level of Blue to the possible values of π_3. This pallet is then described as the set of all colors within the wedge of the color cube in the direction of the red (1,0,0), green (0,1,0), or blue (0,0,1) vertices. Other pallets could be obtained by rotating the projection inside the RGB color-space. An example of one such rotation would be to replace the values of π_1 with π_1' where $\pi_1' = 1 - \pi_1$, moving the vertices of the wedge that lies directly across from its hypotenuse[4] (the face of the wedge with vertices (1,0,0), (0,1,0), (0,0,1)) from (0,0,0) to (1,0,0). This process of rotation will give the eight possible (there are eight vertices on a cube) color schemes for k=4 shown in Table 2.

Table 2: Possible color schemes for k=4

Vertices Opposite Hypotenuse (coordinates)	Corresponding Colors (coordinates)
Black (0,0,0)	Black, Red, Green, Blue
Red (1,0,0)	Red, Magenta, Yellow, Black
Green (0,1,0)	Green, Cyan, Black, Yellow
Yellow (1,1,0)	Yellow, White, Red, Green
Blue (0,0,1)	Blue, Black, Cyan, Magenta
Magenta (1,0,1)	Magenta, Red, White, Blue
Cyan (0,1,1)	Cyan, Green, Blue, White
White (1,1,1)	White, Yellow, Magenta, Cyan

Picking Colors for k=3

When considering only three clusters in which to assign membership of the data points, there are two methods for picking the colors to use:

The first simply picks the colors of the vertices of the hypotenuse of given wedge. Figure 3 demonstrates the picking of colors Red, Blue and Green for the case of k=3. There are another seven possible schemes available using this method, one for each of the rotations outlined picking colors for k=4.

The second is to apply a linear projection from a given hypotenuse selected via method 1 to one of the six faces on the color cube. A particular example of this would be to choose the color scheme {Black, Blue, Red} obtained by projecting the hypotenuse shown in Figure 3 onto the face of the RGB color-space with vertices Black, Blue, Magenta, Red. The effect of this second method is to provide another forty-eight valid color schemes. While we run into the same

[2] We say 'span' here since space defined here is not a linear subspace, and as such the notions of span introduced in Linear Algebra wouldn't generally be true. Specifically, it is the minimum number of dimensions needed to reproduce the space.

[3] To see this, consider this simpler case where k=3. As shown in Figure 2 the fuzzy space is an equilateral triangle with vertices (1,0,0), (0,1,0), (0,0,1). To project this triangle onto the right triangle formed by its shadow in the axis plane, we would have to distort it slightly.

[4] We use hypotenuse to describe the surface of the wedge with the largest surface area

distance distortion problem touched upon for the k=4 case, it is of little consequence to our application for this case as well.

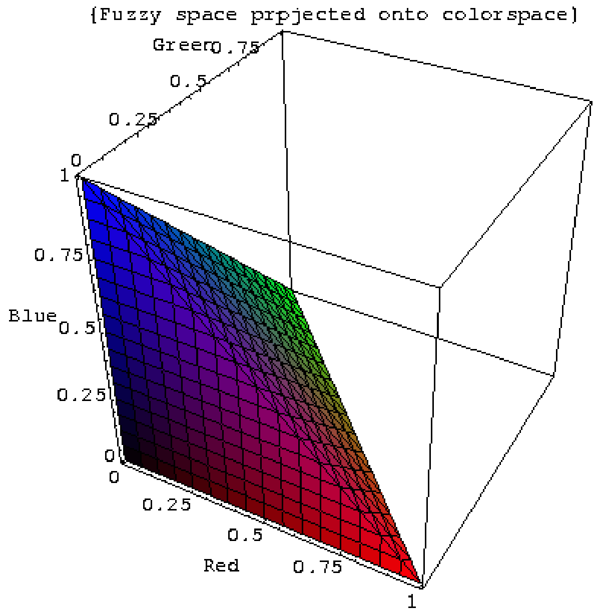

Figure 3: Constrained color cube

Picking Colors for k=2

In the case of two clusters, we can pick any path joining any vertices in the color space. Specific examples of this would be Red and Black (set level of Red equal to the value π_1), Blue and Yellow (set level of Blue equal to the value of π_1 and the levels of Red and Green equal to value of π_2).

Comments

The problem of the distortion in distance that is encountered by projecting a space that is described by k variables onto a space that is described by k-1 can be seen by observing Figure 3. The first hint of trouble comes by observing that the area of the hypotenuse of the wedge is certainly greater than the area of the triangle that forms the 'floor' of the wedge. The problem is compounded a little further by noticing that our notion of distance is really one of how many colorshifts do we have between two points (as opposed to our more natural Euclidean measure of distance). In the appendix we show that this measure of distance is quite likely outside the bounds of most people's ability to perceive the distance between colors. In short: projecting the spaces from their native k description to the projected k-1 description has no identifiable impact on the map.

Constrained fuzzy clustering

For the simplest case, consider k=3. Suppose that we <u>know</u> that if data point i is a member of cluster 1, then it could not possibly be a member of cluster 3; however, this data point could also be a member of cluster 2. More generally, suppose we can have 'mingling' between clusters 1 and 2 **or** clusters 2 and 3, but not 1 and 3. We could think of this space graphically as the union of two cases general fuzzy clustering when k = 2, specifically two (0,1) line segments joined at a central point. A specific example of this would be the classification of items being 'good', 'bad', or 'fair'. If you know that the method you are using to suggest classification is reasonable, it isn't really possible for an item to be fractionally a member of both 'good' and 'bad' without being a member of 'fair'. More to the point, an item can't be both 'good' and 'bad', but could be 'good' and 'fair' *or* 'fair' and 'bad'. We can observe this graphically by considering Figure 4.

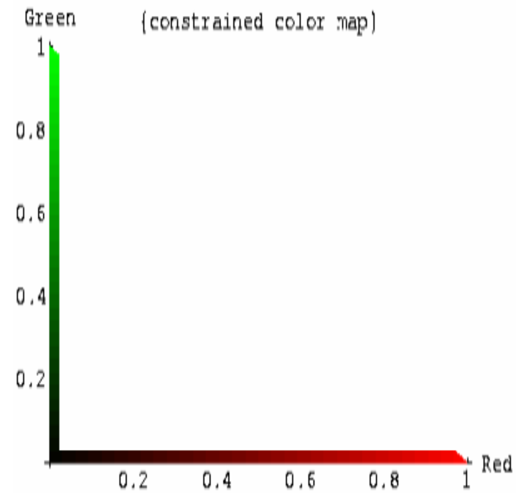

Figure 4: Color map for fuzzy space constrained Red and Blk or Blk and Grn

Suppose now we have k = 4. If we <u>know</u> that we can have mingling between clusters 1, 2, and 3; **or** mingling between clusters 2, 3, and 4, we could represent this space graphically as the union of the two projected triangles given for the general fuzzy clustering when k = 3 that share an adjacent base (see Figure 5).

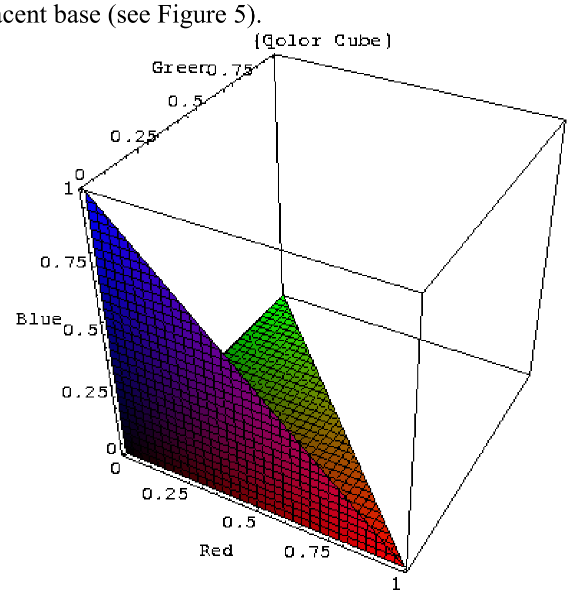

Figure 5: Color map for fuzzy space constrained by (Red and Blu and Blk) or (Red and Grn and Blk)

It should be intuitively clear that by disallowing 'mingling' you can *reduce* the dimension necessary in the RGB target space for your map. Suppose again that we *know* we have a mingling between sets that can have minglings

(12), (23), (34), (45), (56), we could map these probabilities by assigning each pair two edges of a face of the RGB colors space such that no face has more than two edges used. From the above discussion of general fuzzy clustering, mapping six clusters simply isn't possible when the assignments are obtained using 'general' fuzzy clustering.

Currently, we have not found any stock fuzzy clustering packages that allow for such limited mingling. We also acknowledge that it some cases it would be a stretch to even consider such a constraint, but in the second example given below we will outline below, there are strong reasons to believe the simplest constrained fuzzy space was reasonable.

4. Examples

Overview

All of these results are clearly dependent upon the accuracy of reproduction of color. It is our opinion that the way to view these results is to use a relatively new computer monitor that has been reasonably calibrated for 24 bit color. The simplest way to achieve this is to make sure your monitor is fairly new, and that your computer has been configured to display specifically for that monitor. Hard copies can be used; bearing in mind that accurate color reproduction can only really be obtained for 24 bit color with photo-capable printers.

Fisher Iris Data

The Fisher Iris data is a classic data set for the classification problem. To summarize, four measurements (sepal length and width, and petal length and width) were made on 150 different Irises. It was known that there were three different types of Irises measured - Setosa, Virginica, and Versicolor. The goal was to see if classification of the Iris types could be obtained by using the four measurements.

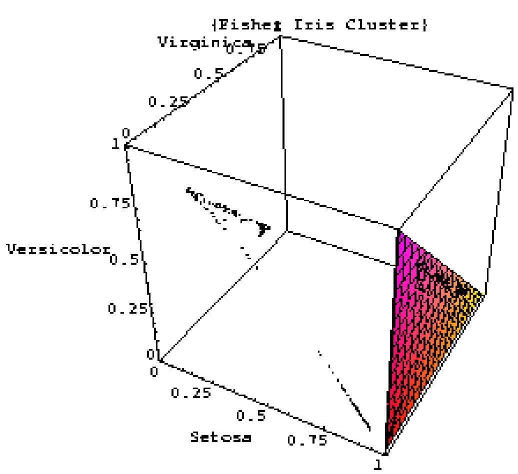

Figure 6: Fisher Iris Cluster Projection

S-plus was used to obtain probability estimates for each item's inclusion into given cluster. The first step in 'coloring' these results is to pick a color scheme. For this example, we (arbitrarily) chose to use the vertices Red (1,0,0), Magenta (1,0,1), and Yellow (1,1,0). If we plot the results of the fuzzy clustering in three dimensions (its natural space), it is clear from Figure 6 and the discussion above that a projection onto the Red, Yellow, White, Magenta surface of the RGB color space will yield a mapping that is very close to 1:1. To give added information as to possible misclassifications, we computed K-means clustering using S-plus. The idea behind this came from our observation that if the two methods (K-means and fuzzy clustering) do not provide the same results this would be a likely indicator of a boundary point. The two sets of results were summarized by using the K-means clustering as the word of the color placed over the color created by the RGB color space map. The sample table of this summary is seen in Table 3 (the entire table can be seen in the Appendix).

Table 3: Summary table for Fisher Iris Data

	fisher iris data			
obs	**P(Setosa)**	**P(Versicolor)**	**P(Virginica)**	**Cluster**
1	91.61%	4.88%	3.51%	Red (Setosa)
2	5.42%	22.68%	71.89%	Yellow (Virginica)
3	6.96%	56.38%	36.67%	Magenta (Versicolor)
4	5.73%	20.63%	73.64%	Yellow (Virginica)
5	6.11%	43.81%	50.08%	Magenta (Versicolor)
6	86.06%	8.08%	5.86%	Red (Setosa)
7	6.76%	26.34%	66.90%	Yellow (Virginica)
8	8.88%	60.13%	30.98%	Magenta (Versicolor)
9	7.39%	53.27%	39.34%	Magenta (Versicolor)
10	80.15%	11.33%	8.52%	Red (Setosa)
11	6.13%	67.33%	26.54%	Magenta (Versicolor)
12	6.53%	49.17%	44.30%	Magenta (Versicolor)
13	4.73%	24.71%	70.56%	Yellow (Virginica)
14	12.98%	66.42%	20.60%	Magenta (Versicolor)
15	4.71%	22.26%	73.03%	Yellow (Virginica)
16	7.34%	47.16%	45.50%	Magenta (Versicolor)
17	6.51%	21.32%	72.17%	Yellow (Virginica)
…	…	…	…	…

The items that lie in the border areas can be quickly identified by the colors alone (especially if one uses the color reference table shown in Table 4) as being observations 5, 12, and 16. Closer inspection of the probabilities for these observations shows the fuzzy clustering algorithm estimates that these observations are equally likely to be Virginica or Versicolor Irises. Close inspection of all the probabilities shows that these are the only elements in the selection that are in a boundary.

Table 4: Color reference for Iris Data

Fisher iris data			
%Red	**%Magenta**	**%Yellow**	**color**
1.0	0.0	0.0	Red
0.5	0.5	0.0	Red-Magenta
0.0	1.0	0.0	Magenta
0.0	0.5	0.5	Magenta-Yellow
0.0	0.0	1.0	Yellow
0.5	0.0	0.5	Yellow-Red

The tables 3 and 4 were created using Microsoft Excel to first create a HTML page and then importing that page into Microsoft Word.

Twelve Month Variance Example

The International Price Program (IPP) of the Bureau of Labor Statistics (BLS) produces a plethora of internal reports to aid economists in their assessment of the price index quality for the strata they report on. These reports are read by economists who must make decisions about the data obtained from a given company quickly and accurately. One set of reports under development focuses on the variance in annual price changes. It would be helpful to an economist to compare the coefficient of variation of a given strata's estimate of annual price change to the coefficient of variation of annual price changes across all published strata.

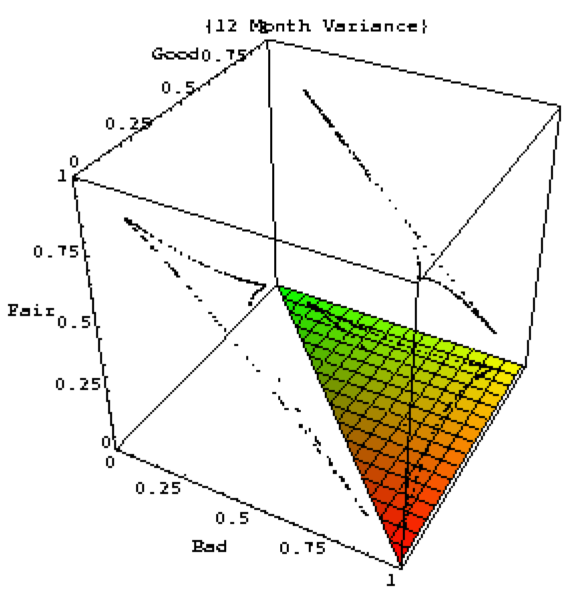

Figure 7: 12 Month Variance Rotation and Projection

Initially it was proposed to partition the set of coefficients of variation into quartiles with the coefficients in the largest quarter being 'bad', the middle half being 'fair', and the lowest quarter being 'good'. It became clear however, that this method was lacking, since there were clear clusters in the dispersion of the variances between the different strata that did not correspond with the quartiles. We opted to use this method of fuzzy clustering to give added guidance to the economist as if there were no clear clusters; the method would (on average) give no worse results than those initially proposed.

In this example (as in the last), we have three clusters; however, this case is different in that there is a clear order to the values, as well as impossibility for 'interaction' between two directions by any path other than the border of the plane. This observation suggests that the 'constrained' fuzzy clustering discussed in section 3 should be used. If it were, the theory says that all probability assignments should occur so that they lie on line segments that form the right angle of the Red, Yellow and Green triangle in Figure 7. Unfortunately, no constrained fuzzy clustering software yet exists (to our knowledge), so we again applied S-plus' fuzzy clustering algorithm to the vector of coefficients of variation, noting that the results may not be entirely accurate. Again, since we were classifying three clusters, we needed to pick a RGB color scheme, and we chose to use Red, Green, and Yellow. Inspection of Figure 7 shows that we obtained this map by moving the origin of the original coordinate system (0,0,0) to (1,1,1). We then projected this rotation onto the Black, Red, Yellow, Green face of the RGB color space. We oriented the map to put Red as 'bad' Yellow as 'fair' and Green as 'good' (see Figure 7).

In a simplified example, suppose an economist was tasked with gathering the data for published strata 23-34. As with the previous example, we chose to summarize this data by putting the K-means clustering in print, and the color derived from the fuzzy clustering as a background. Clearly the information displayed in this form is far more compact than a table that contains an indicator for the k-means clustering, and a column each for the probabilities assigned via fuzzy clustering. It is also far more informative than a picture of all the data points with their strata numbers highlighted. Yet, the economist is able to ascertain all the information from either of these figures in a far more compact form.

Table 5: Summary table for 12 month variance data

sample stm#	sampling stratum description	coeff var	cluster
23	TEXTILE AND TEXTILE ARTICLES	0.257399	green
24	COTTON	0.523422	yellow
26	SWEATERS, PULLOVERS, SWEATSHIRTS, AND SIMILAR ARTICLES, KNITTED OR CROCHETED	0.451669	yellow
28	MEN'S SUITS, BLAZERS, TROUSERS, BIB OVERALLS, SHORTS (OTHER THAN SWIMWEAR)	0.544465	yellow
29	WOMEN'S SUITS,	0.708234	red

	BLAZERS, DRESSES, SKIRTS, TROUSERS, SHORTS (OTHER THAN SWIMWEAR)		
30	MEN'S OR BOYS' SHIRTS	0.684721	red
32	OTHER MADE UP TEXTILE ARTICLES; SETS; WORN CLOTHING AND WORN ARTICLES" RAGS	0.311910	green
33	FOOTWEAR, HEADGEAR, UMBRELLAS, WHIPS, FEATHERS; ARTIFICIAL FLOWERS	0.784319	red
34	FOOTWEAR, GAITERS AND THE LIKE; PARTS OF SUCH ARTICLES	0.677354	red

Looking at this set, sample stm# 24 and 28 stand out as being potential 'border' points. However, if we compare these problem points to the reference strip found in Table 6 we may conclude there aren't any border points at all.

Table 6: Reference table for 12 month variance data

12 month variance data

%Red	%yellow	%Green	color
1.0	0.0	0.0	Red
0.5	0.5	0.0	Red-Yellow
0.0	1.0	0.0	Yellow
0.0	0.5	0.5	Yellow-Green
0.0	0.0	1.0	Green
0.5	0.0	0.5	Red-Green

Indeed, perusal of the probability assignments suggests that all the points are solidly in the cluster assigned by K-means clustering. The worst case was stratum 28 where the Red, Yellow, Green fuzzy probabilities are (29.76%, 56.12%, 14.11%). Not really a good glowing endorsement for any particular cluster, but still more yellow than any other cluster.

An economist looking at this data will see that some of their strata fall into a 'high' annual variance group, and is free to determine if these fluctuations are occurring in items that historically have had high price volatility or if the variation may be due to other, more controllable factors such as Low response or outliers. The economist can also see that there are no borderline 'high variance' strata based upon the k-means vs. fuzzy clustering results.

While this particular table was created in an identical fashion as above, in its final implementation, the table could be compiled using a Java program, and incorporated with other useful information to the Economist with a fuzzy triple for the three clusters given when the mouse is hovered over the cell in question.

5. Conclusions

The uses and benefits of clustering data have been well established over the last few years. Graphics have also long been a staple for statisticians to reduce the dimension of data, while still providing useful information. This paper combined these two methods and introduced a graphical enhancement to clustering output. We believe that this method provides a robust and reasonably accurate way to display the results of fuzzy clustering to give guidance to the user as to how 'close' a particular data point is to the boundary of its cluster by comparing the corresponding fuzzy cluster color to the 'true' color of being in each group alone.

While we are aware that certain low level distortions occur during the projections in the color space we suggest their use in the case of k = 3 as short cut for coding, and point out that they really are not necessary. We also note that mapping to the color-space when k = 4 is not possible without the projections.

6. Appendix

Gradients and distance changes

Define the fuzzy space with the following :

$$f(x,y) = 1 - x - y$$

let $\mathbf{v}$ be a vector in any direction : $\langle v_1, v_2 \rangle$

then $\mathbf{u} = \dfrac{\mathbf{v}}{\|\mathbf{v}\|}$

from this, the gradient can easily be computed :

$$\nabla f(x,y) = \langle -1, -1 \rangle$$

giving way to the familiar derivative

in the direction of $\mathbf{u}$:

$$\nabla f(x,y) \bullet \mathbf{u}$$

At this point, it is important to recognize the overarching goal: to ascertain the loss of color induced by the projection. The first step towards this end is to determine how much given color in the fuzzy space 'covers' a given color in the projected space. Since the directional derivative will give us the change in *f(x,y)* for a unit change in the direction of **u**, it's inverse will give the amount of change on the plane that contains **u** (i.e. the projected plane) for a one unit change in *f(x,y)*. If we then compute distance of a color shift in the direction of **u** in the projected plane, we can find the number of colors shifts in the fuzzy plane needed to reach the next color shift in the projected plane by division. If we cleverly pick **v** to have a length that is the length of the projected plane's colorshift, we get the following result:

$$\text{CS} = \frac{\|\mathbf{v}\|}{\dfrac{1}{\left|\nabla f(x,y) \bullet \dfrac{\mathbf{v}}{\|\mathbf{v}\|}\right|}}$$
$$= \left|\nabla f(x,y) \bullet \mathbf{v}\right|$$

That is, the number of color shifts (CS) in the fuzzy space represented by a single color in the projected space is equal to $\left|\nabla f(x,y) \bullet \mathbf{v}\right|$.

Since we are dealing with a plane, this rate of change is constant regardless of where it occurs, we can describe the set of all possible values for $\left|\nabla f(x,y) \bullet \mathbf{v}\right|$ thus:

Let

$$S_0 = \left\{\mathbf{v} : \mathbf{v} = \left\langle \frac{1}{256}, \frac{a_0}{256} \right\rangle, a_0 \in [0,1]\right\},$$

$$S_1 = \left\{\mathbf{v} : \mathbf{v} = \left\langle \frac{a_1}{256}, \frac{1}{256} \right\rangle, a_1 \in [0,1]\right\}$$

then,

$$\underset{\mathbf{v}\in S_0}{MAX}\left\|\nabla f(x,y) \bullet \mathbf{v}\right\| =$$
$$\underset{\mathbf{v}\in S_1}{MAX}\left\|\nabla f(x,y) \bullet \mathbf{v}\right\| = 2 \text{ for } \mathbf{v} = \langle 1,1 \rangle$$

Similarly,

$$\underset{\mathbf{v}\in S_0}{Min}\left\|\nabla f(x,y) \bullet \mathbf{v}\right\| =$$
$$\underset{\mathbf{v}\in S_1}{Min}\left\|\nabla f(x,y) \bullet \mathbf{v}\right\| = 1 \text{ for } \mathbf{v} = \langle 0,1 \rangle \text{ or } \mathbf{v} = \langle 1,0 \rangle$$

The projected maps outlined for the cases of k=3, k=4 aren't really 1:1 maps because there is a distortion of the distances as one moves from the kernel to the target spaces. However, given our ultimate objective, and the density of the RGB color-space we feel that the map is adequate for our goals. This assertion can be seen when k=3 by first noticing that for a plane, the gradient in the direction of **v** (denoted $\nabla f(x,y) \bullet \mathbf{v}$) can be interpreted as the number of color shifts in the fuzzy plane represented by a color shift in the projected plane.

The **v** that maximizes the gradient is $\langle 1,1 \rangle$. This maximum can be interpreted as being the direction that has the most color shifts represented by a single color shift in the projected plane. In this 'worst-case' scenario, we get the number of fuzzy plane shifts in color represented by a projected plane color-shift to be $\nabla f(x,y) \bullet \langle 1,1 \rangle = |-1-1| = 2$. Similarly, the number of fuzzy plane shifts represented by a projected plane color-shift when moving in the direction of $\langle 0,1 \rangle$ (or $\langle 1,0 \rangle$) will be the minimum number of color shifts represented a shift in the projected plane, with there being an equal number of shifts. While this seems like bad news for the worst case scenario keep in mind three facts:

1) making the colors appear 'closer' in terms of color will only give us a more 'conservative' representation of the closeness between points, and the ultimate goal of quickly identifying boundary points.
2) Both the fuzzy plane and the projected planes start and end with the same colors at their vertices
3) There are a *lot* of colors. In the worst case, we are shrinking the total number of colors from 512 (the number distinct colors across the 'slant' edge in the relevant direction) to 256. It is entirely likely that most people could not discern any color difference for the cluster of points that mingle at the borders we are hoping to classify.

All of these notions extend to the projection we have when projecting the 4D plane that results from the case of k=4 into the solid 'wedge' shown in Figure 3, with the worst case scenario now having the projected space representing three colors of the fuzzy space instead of 2 as in the above example. The same caveats that are listed above still apply as well, suggesting to us that in this case there is little cause for concern for our purposes.

7. Acknowledgements

The author would like to thank Sumer Yates, Bob Eddy, and Marvin Kasper for their invaluable editing efforts, as well as Richard Toftness for special assistance as a sounding board for initial and intermediary results.

8. References

Insightful Corporation (2001). *S-Plus 6 for Windows Guide to Statistics, Volume 2*. Insightful Corporation Seattle, Washington.