Exploring Structures in the Private Foundation Ecosystem

Thomas Dudley[1], Alex Fetisova[2], Darren Hau[3]

[1]GSB, [2]MS&E, [3]EE

Motivation & Data Sources

Motivation:

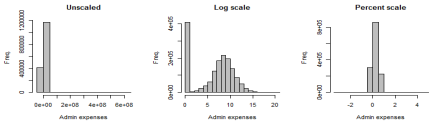
Private foundations play a key role in funding the social sector in the United States. All told, there are nearly 90,000 private foundations that manage approximately $700 billion in assets, give nearly $50 billion in grants each year, and benefit from large tax breaks[1].

Despite their importance in the social sector (and their favorable tax status), private foundations have very little external accountability.

To better understand how these organizations behave, we apply unsupervised learning methods to analyze the landscape of private foundations in the United States. Are there previously unused traits we can use to cluster foundations?

Data Sources:

By law, private foundations are required to report key financial information to the IRS via an annual filling of form 990-PF. The National Center for Charitable Statistics has collected the full set of form 990-PFs for the years 1989 – 2013.



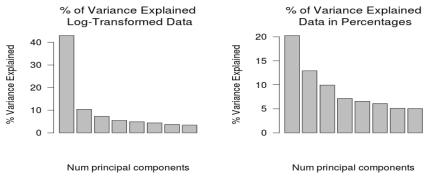
Above: Log- and percent-scaled data to account for significant size skew.

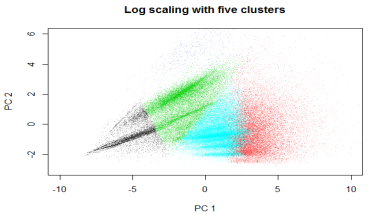
Feature Subset	PC1	PC2
Expenses (log-scaled)	Size of expenses	Officer Compensation – Contributions Paid
Assets (log-scaled)	Size of portfolio	Government Bonds – Corporate Stock/Bonds

Analysis Techniques

PCA:

Using R, we applied PCA to find low-dimension representations of our data. We applied it to both the log and the percent transformations. The following figure shows the percent of variance explained by each principal component.



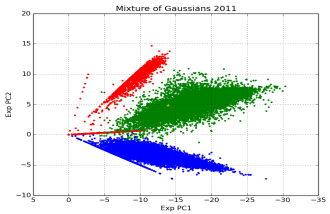


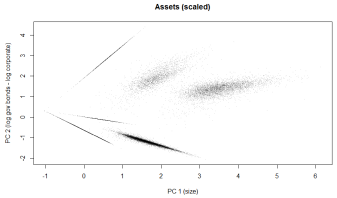
K-Means:

Our original intent was to run k-means. The ratio of between cluster sum of squares and total sum of squares provides a good approximation of the variance explained by the model, and we achieved decent values (between 0.6 – 0.7 for 10-20 clusters). However, when plotting the first two PCs, we saw that k-means was not picking up on the Gaussian-looking distribution, and there appeared to be far fewer than 10 clusters.

Mixture of Gaussians:

To capture the Gaussian distributions, we moved to a Gaussian mixture model, using Python's scikit-learn module. To make the analysis more interpretable, we selected a subset of our features to interpret. The GMM gave us easily-interpretable clusters for expenses and assets.



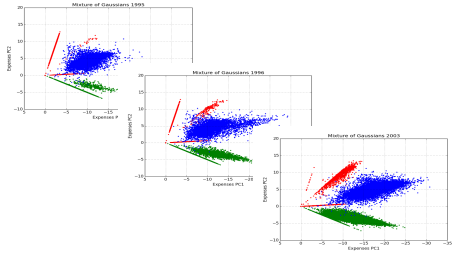


Above: GMM run on asset variables. We identify 3 main clusters which correlate with the riskiness of the organizations' portfolios.

Left: GMM run on expense variables. We identify 3 main clusters – those who become more efficient with size, those who stay roughly as efficient, and those whose efficiency diminishes with size. There is a surprisingly clear distinction between the three groups. The yellow star indicates a chapter of the American Red Cross.

Results & Interpretation

The biggest finding was that we could clearly group foundations into 3 groups by expenses or assets. This trend mostly held from 1995-2011, though we see a divergence from 2 clusters to 3 clusters in 1995 and 1996. The new third cluster represents foundations which have much higher officer compensation than direct grant contributions.



This poses some interesting questions for policymakers. Was there a change in regulations which facilitated this rise in high-expense, low-impact organizations? Or do these 3rd-cluster organizations represent more targeted but effective grant-giving?

Conclusions & Future Directions

Using a mixture of Gaussians model shows promise in clustering private foundations. While we did not have time to run a thorough analysis, we believe it is possible to extend the GMM approach to larger subsets of the data.

Additionally, this approach can be used on other organizations as well. An analysis on for-profit companies and politically-active organizations would be highly relevant in both fields.

References & Acknowledgements

We would like to thank Professor Ng for teaching an intellectually challenging class, and our project mentor Bryan McCann for supervising our work.

1. https://www.irs.gov/uac/SOI-Tax-Stats-Domestic-Private-Foundation-and-Charitable-Trust-Statistics#3