Sabbatical Report
For Spring 2012

Laxman Hegde, Department of Mathematics
Frostburg State University
lhegde2@frostburg.edu

July 27, 2012

1   My Proposal

Title: A Revisit to Canonical Correspondence Analysis (CCA)

Time Frame: Spring 2012

Pupose: Enhance academic research, computational and teaching skills

A main reason for approving my sabbatical proposal was the research paper (jointly with Dr. Naik of ODU) that we published in 1999, entitled Canonical Correspondence Analysis in SAS Software. This paper has received (and continues to receive) world-wide attention of the researchers in ecology. It needed to be updated to enhance its value to the future users. There were three objectives in my proposal as listed here.

Research: Undertake research activities on the topic of CCA.

SAS: Write descriptive interpretations to complex mathematical quantities generated by our SAS program.

Workshop: Organize a workshop (jointly with Biology Department of FSU and Appalachian Laboratory) for the researchers who use CCA.

2   Talks

I presented **four** talks in the spring 2012 that are direct outcomes of my work related to the sabbatical objectives. Also I submitted an abstract in May 2012 for my **fifth** talk labelled NESUG in the following table.

At the very outset, broadly speaking, I believe that all these talks meet my **Research and Workshop** goals of the proposal. Preparing these talks required a fair amount of readings in the field of CCA, theory and applications.

Label	Date	Location	Title
SIAM	3/24	ODU,Norfolk, VA	Canonical Correspondence Analysis: Some new insights and interpretations
MAA	4/14	Stevenson, Owings Mills, MD	Correspondence Analysis
FSU	4/27	FSU, Frostburg, MD	Interpreting Canonical Correspondence Analysis (CCA)
TASTE	5/24	FSU, Frostburg, MD	Preparing a Slide Show of Scientific Material Using LaTex
NESUG	11/11	Baltimore, MD	Implementing and Interpreting Canonical CorresPondence Analysis in SAS

Please note the expansion for some of the abbreviations (Talk Labels).

SIAM: Society for Industrial and Applied Mathematics

MAA: Mathematical Association of America

NESUG: North East SAS Users Group

Please refer to the conference schedules and emails for more details. These documents are also posted in my website addressed in Documents Section 6 of this report.

2.1 More on Talks

- SIAM:

  This talk was well attended. In this talk, I presented the main mathematical ideas of CCA. There were several graduate students who told me that they were not familiar with this topic.

  Interestingly, the NSF Director of the division of Mathematical Sciences was present in my talk. I got an opportunity to know him personally.

- MAA:

  The main participants in this talk consisted of mathematics teachers and students in undergraduate colleges in DC/MD/VA section of MAA. In this

talk, I presented how to interpret Correspondence Analysis results in an application.

As a side note, I met the President of DC/MD/VA section of MAA and offered my willingness to coordinate a future meeting at FSU campus.

- FSU:

  This talk was arranged with the help of the Biology Department of FSU. Besides a few faculty members, there were more than 10 graduate students attended my talk. In this talk, I presented the details of conducting CCA in applications and interpreting the SAS output. This was one of my main goals in the proposal.

  Furthermore, I offered my help to work with all students who are interested to use CCA in their graduate work.

- TASTE:

  Mrs. Cindy Hay (FSU) requested me to present this talk for TASTE since many people on our campus are not aware of the software Latex/Beamer. During my SIAM talk, I learnt a lot about the software Beamer to make slides involving complex mathematical and scientific material. In this talk, I went over a talk presented by another researcher at MIT on Latex/Beamer.

  Although this talk is not directly related to my proposal, this is an indirect benefit to FSU community as a result of my sabbatical work.

- NESUG

  Although this talk is scheduled to be held in November 2012, the preparatory work (making contacts, early correspondences, writing proposals and abstract) began in late April and early May 2012. In this talk, I will demonstrate how to implement and interpret our new SAS program.

  As part of this presentation, I have already submitted a paper (draft version) to be published in the NESUG Conference Proceedings. I am hoping that this publication will bring a wider publicity to our modified SAS program.

3   About Objectives

As I mentioned earlier in this report, I believe that my talks and publication (NESUG) satisfy the intended objectives stated in my proposal. Here in this section I will attempt to go over some extra details about my goals.

3.1 Research Activities

In the month of March 2012, I spent two weeks with Dr. Dayanand Naik of Old Dominion University. During this time, I had fruitful discussions with him about certain inferential aspects of CCA. This motivated my SIAM talk. Also we studied several research papers related to my proposal objectives.

In addition to internet searches, I visited FSU library several times researching advanced level statistics books. **As a result, I wrote a short article on linear algebra of CCA**. I intend to submit this article in a journal during the next academic year. This article has greatly enhanced my mathematical knowledge of CCA and in turn, it will help my teaching skills for linear algebra course which I have been teaching for the last 3 years.

3.2 SAS Program

By comparing the two versions (Old vs New) of our SAS program, we notice some significant changes between the two versions. Quantitatively speaking, we see that the number of pages in the new version increased to 11 pages from 8 pages in the old version. Qualitatively speaking, I like to point out the following important changes in the new version.

- Easy readability of SAS codes with comments.

- Several new SAS modules for more efficiency.

- More flexible for producing Biplot, an important output for end users.

3.3 Workshop

My FSU talk (April 27th) was presented in a workshop style. In this talk, I demonstrated interpreting SAS output using a very well known (widely used) example in the field of ecology.

I had contacted Prof. Robert Hilderbrand (AL) to conduct a regional workshop to the researchers in the field of ecology. Unfortunately, Prof. Hilderbrand mentioned that most researchers at AL switched to a new free software, R, instead of SAS which is a very expensive software.

More importantly, I have emailed to several researchers (who had sought our old SAS program) to see if they are interested in attending my workshop at FSU. Instead, they preferred to receive our new SAS program along with the file InterpretSas.pdf (my FSU talk) which contains interpreting SAS output. **Accordingly, I have been sending these documents as and when I receive such requests. I have saved all of the email correspondences as supporting documents.**

4 Conclusions

I gained increased competence in providing statistical consultancy and computational services to graduate students in Biology Department and AL. It is quite likely that some of these students will be analyzing their data using CCA and SAS program under my guidance in the coming years. Future users of our modified SAS program will appreciate it more, and in turn, FSU will earn more reputation.

5 Acknowledgments

I like to thank FSU administration (Math Chair $\rightarrow$ the Dean $\rightarrow$ the Provost $\rightarrow$ the President) for supporting my sabbatical proposal. Also I like to thank Dr. Dayanand Naik of ODU for being my mentor in doing this work.

6 Documents

There are several supporting documents accompanying this report. To reduce the size of this report, I have decided to put them in my website. **If I were to print all these documents, they will add to about 100 pages!**

- My website: http://faculty.frostburg.edu/math/lhegde

- If you are reading this report electronically with an internet service, you may click the following link to read my documents. My Site.

Here is a list of all the supporting documents that I have uploaded to my website.

- Sabbatical Proposal (Proposal.pdf) (7 pages)

- Old SAS Program (OldSas.pdf) (8 pages)

- New SAS Program (NewSas.pdf) (11 pages)

- Interpreting SAS program (InterpretSas.pdf) (41 slides, about 24 pages)

- MAA Talk (MaaTalk.pdf) (49 slides, about 10 pages)

- Taste Talk (TasteTalk.pdf) (58 slides, 58 pages)

- Draft version of research paper to be published (NesugPaper.pdf) (10 pages)

- Introduction to Linear Algebra for CCA (IntroLA.pdf) (8 pages)

- Siam Conference Schedule (SiamSch.pdf)(1 page)

- Maa Conference Schedule (MaaSch.pdf) (4 pages)

- Taste Conference Schedule (TasteSch.pdf) (11 pages)

- FSU Biology Department Schedule (FsuSch.pdf) (3 pages)

- NESUG Coference Acceptance Email (NesugSch.pdf) (3 pages)

7 References

For my talks and articles, I needed to read several research articles and books related to CCA. I have listed some important references which I used for my work.

Ter Braak, C. J. F. (1986). Canonical correspondence analysis: A new eigenvector technique for multivariate direct gradient analysis. *Ecology*, **67**, 1167-1179.

Ter Braak, C. J. F. (1988). A Fortran program for canonical community CANOCO, ordination by partial detrended canonical correlation analysis, principal components analysis and redundancy analysis (version 2.1), Microcomputer Power, USA.

Greenacre, M. J. (1984). *Theory and Applications of Correspondence Analysis.* Academic Press, London, England.

Hegde, L. M. and Naik, D. N. (1999). Canonical correspondence analysis in SAS software. *Proceedings of the Twenty-Fourth Annual SAS Users Group International (SUGI) Conference,* paper **278**, 1607-1613.

Khattree, R. and Naik, D. N. (2000). *Multivariate Data Reduction and Discrimination with SAS Software.* SAS institute Inc., Cary, North Carolina, and J. Wiley and Sons, New York, USA.

Hegde, L. M. and Naik, D. N. (2008). Canonical Correspondence Analysis : Some New Interpretations And Computations Using SAS, *Journal of Statistics and Applications*,**Volume 3**,Issue 2,277-302

Jan de Leeuw and Patrick Mair, Simple and Canonical Correspondence Analysis Using the R Package anacor