
Formatting instructions for NIPS 2016

Elijah Ben-Michael
Department of Statistics
UC Berkeley

Runjing Liu
Department of Statistics
UC Berkeley

Jake Soloff
Department of Statistics
UC Berkeley

Abstract

The abstract paragraph should be indented $\frac{1}{2}$ inch (3 picas) on both the left- and right-hand margins. Use 10 point type, with a vertical spacing (leading) of 11 points. The word **Abstract** must be centered, bold, and in point size 12. Two line spaces precede the abstract. The abstract must be limited to one paragraph.

1 Introduction

1.1 Motivation

Voting records of legislators are commonly used by political scientists to examine relationships between legislator policy preferences, institutional structures, and legislative outcomes (Clinton et al 2004). In fact, even simple dimensionality reduction techniques on roll call data are able to uncover the political characteristics of individual representatives. For example, in figure 1a, we factored the 448×1707 matrix representing the 448 representatives and their votes on 1707 bills into two nonnegative matrices of dimensions 448×2 and 2×1707 . Plotting the 448×2 matrix where each row places a representative in a two dimensional space, we are able to clearly identify party affiliation.

Another dimensionality technique we applied to visualize voting data was principle component analysis (figure 1b). We formed the principle components by computing the two largest eigenvalues and their respective eigenvectors on the 448×448 covariance matrix of vote data, and each representative's voting profile was projected onto the space spanned by these principle components. Again, we clearly see differentiation along party lines.

Another common analysis of roll call votes that may potentially yield more subtle structures in legislative preferences is to conduct *ideal point modeling*. **maybe the next sentences goes in the introduction** { Here, a congressman and a bill is presumed to lie in a latent “ideological space,” where the probability of a “yay” or “nay” response is a function of the bill's position and the congressman's position. The congressman's position is known as an “ideal point” because his or her utility decreases as a bill's position deviates from this point. }

In Gerrish and Blei 2011, ideal points of each representative was drawn from a zero mean Gaussian prior. In this paper, we aim to obtain better estimates of the representatives' ideal points, and to do so, we incorporate data from caucus memberships using a stochastic block model (see models below) because we hypothesize that sharing caucuses with other representatives influences a representative's voting behavior. Figure 2 plots the number of shared caucuses between two representatives against the proportion of bills on which they voted the same way, and we see that the more caucuses two members share, the more likely they are to vote the same way.

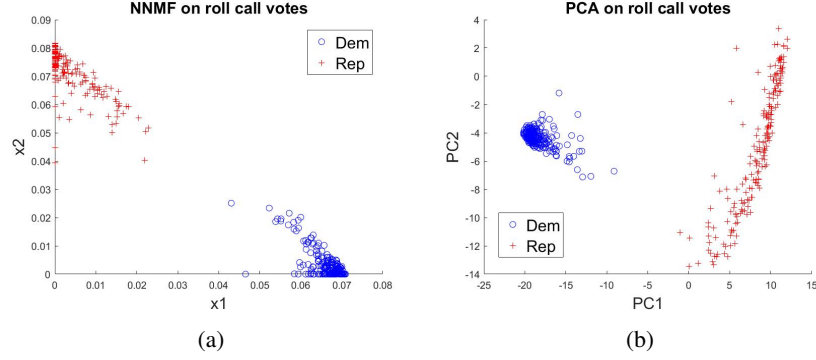


Figure 1: (a) Nonnegative matrix factorization on the 448×1707 matrix (448 representatives, 1707 bills) of roll call votes into two matrices of dimensions 448×2 and 2×1707 . The rows of the 448×2 matrices were plotted to visualize the distribution of representatives. (b) Principle component analysis on the roll call vote data. The eigenvalues and eigenvectors of the 448×448 covariance matrix of representative voting data were computed, and each representative's voting profile was projected onto the space of the two eigenvectors with the two largest eigenvalues.

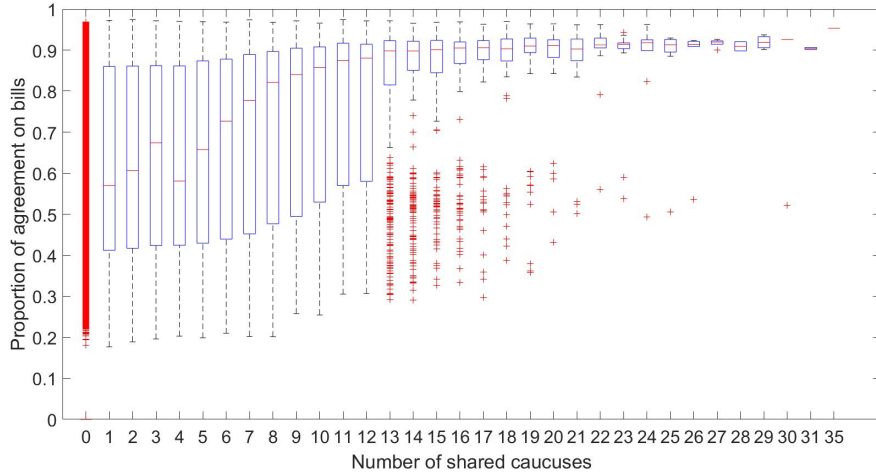


Figure 2: The distribution of agreement on bills as a function of the number of caucuses two representatives share. We see that the more caucuses people share, the more likely they are to agree on a bill.

Moreover, figure 3 shows the relationship between representatives within several caucuses in an undirected graphical model. We first used roll call vote data to infer the graph structure among the representatives in the entire House; we assumed pairwise interactions described via an Ising model in which each node denotes a binary variable of a representative voting either yes or no. The edges were inferred using neighborhood selection, and the graphs shown in figure 3 are subsets of this full graph corresponding to members of a caucus. The connectivity ($\#edges / \#nodes(\#nodes - 1)$) of the full graph with 448 representatives is 0.064, while the connectivity within the caucus subgraphs was much higher. This suggests that a representative more likely to be influenced by a member of his caucus than another random representative in the House. Therefore, by relating these interactions among representatives to their ideal points using a stochastic block model, we hope to be able to capture more subtle patterns in estimating ideal points. In particular, we aim to extend the one dimensional ideological space in Gerrish and Blei to higher dimensions. In doing so, we hypothesize that more accurate predictions of roll call votes can be explained.

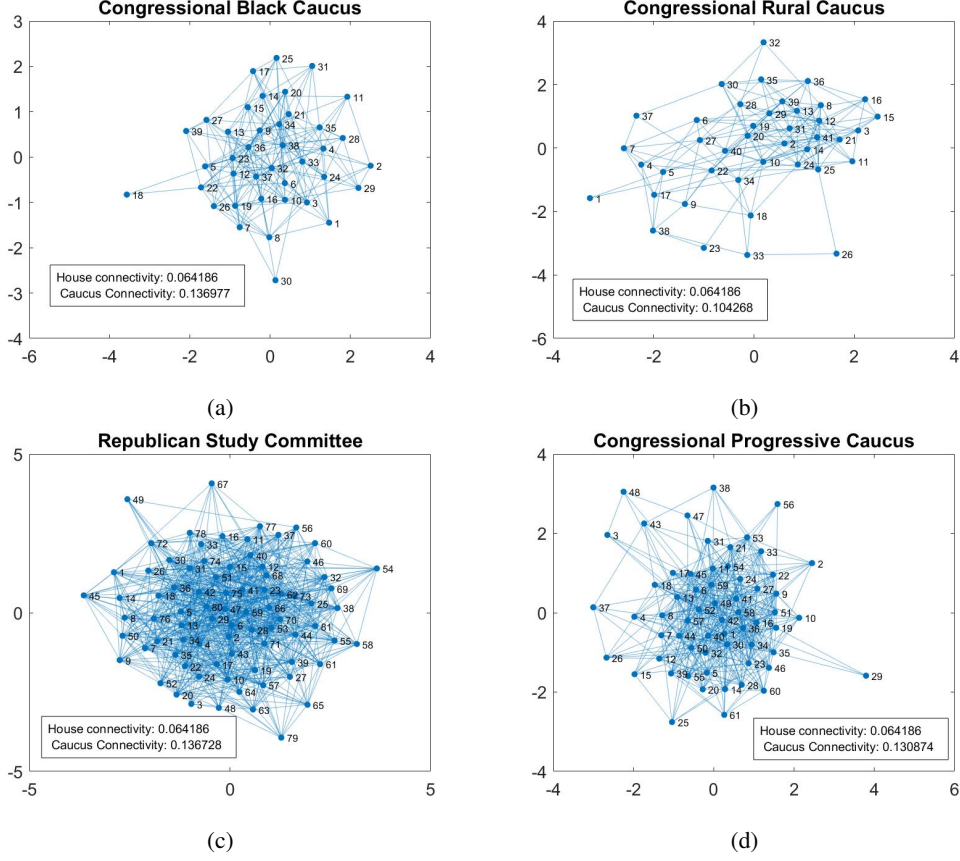


Figure 3: Graphs inferred from Neighborhood regression on House roll call vote data. Shown here are subgraphs with representatives taken from a given caucus. The caucuses are their connectivities shown here are (a) the Congressional Black Caucus, connectivity 0.137; (b) the Congressional Rural Caucus, connectivity 0.104; (c) the Republican Study Committee, connectivity 0.136; and (d) the Congressional Progressive Caucus, connectivity 0.131. In each case, the connectivity within the caucuses was higher than the connectivity of the whole graph of the House (0.064).

2 The model

We introduce models for both roll call votes and caucus memberships.

2.1 Ideal point model

2.2 Stochastic Block Model

3 Results

4 Discussion

References

[1] Gerrish, S.M. & Blei, M.B. (2011) Predicting Legislative Roll Calls from Text. *Proceedings of the 28th International Conference on Machine Learning*

A Variational updates