

# Journal

To do List 1. Find data 2. Read chapter 2 of Advances in PCA 3. Learn more about R packages. 4. Figure out which variables from songs to keep 5. Connect spyder to pip 6. Audio interface in R

Look into combinations of SPCA and KPCA. Would the transformations applied during KPCA make PCs too hard to interpret or could we find meaningful interpretations of a small number of variables under a transformation.

KPCA works best with large amount of data fairly evenly distributed along the manifold.

I was able to find a few (I think 3) papers on Sparse Kernel PCA (SKPCA). Performs KPCA but adds a penalty similar to LASSO in order to give sparser PCs.

Look into other ways to penalize model complexity and test them with SPCA/SKPCA.

I want to look further into different methods of penalizing model complexity. This will likely mean focusing on SPCA right now and later adapting it to KPCA.

Would it be possible to find a way to create principal components that tend to be sparse and are still orthogonal?

How does SPCArt help with interpretability, it seems like results would be harder to interpret due to the rotation that the data undergoes, undercutting one the main advantages of SPCA.

Topic Idea:

Look into how different methods of sparse principal component analysis effect the performance of principal component regression (sparse principal component regression) to see if or how they can improve model performance. Based on preliminary research, it looks like there is a fair amount of recent research in SPCR so combining with more recently proposed SPCA models should yield novel results. SPCR should offer improvements in model complexity while possibly being less likely to overfit training data.

This blog feed back [<https://archive.ics.uci.edu/ml/datasets/BlogFeedback>]