Comparing Matrix Decomposition Methods for Meta-analysis and Reconstruction of Cognitive Neuroscience Results

Kevin Gold
Rochester Institute of Technology
kgold@mail.rit.edu

Catherine Havasi
MIT Media Lab
havasi@media.mit.edu

Michael Anderson
Franklin and Marshall
michael.anderson@fandm.edu

Kenneth Arnold
MIT Media Lab
kcarnold@media.mit.edu

Abstract

The results of 2,256 neuroimaging experiments were analyzed using singular value decomposition (SVD) and non-negative matrix factorization (NMF) to extract patterns in the data. To evaluate the techniques' efficacy at capturing regularities in the data, one positive and one negative result from each of 100 random experiments were treated as missing, and the values were iteratively reconstructed using each technique for dimensionality reduction. Under the best conditions, precision and recall of roughly 78% was achieved for each method. Weighting the domain matrix and area matrix to have equal first eigenvalues before combining them, a technique known as blending, significantly improved results for both methods. While using unnormalized data appeared to produce a peak in results for 10-15 dimensions, normalizing to take into account variation in the popularity of experiment types removed the effect. The basis vectors produced by each method do not support the idea that current cognitive ontologies map well to individual brain areas.

Introduction

One of the most tantalizing promises of machine learning is its potential to inform other areas of science where the theories are still developing. Neuroscience is an excellent example of a discipline in which there is a wealth of detail, but a paucity of data-grounded methods for producing and verifying broad theories. A typical functional magnetic resonance imaging (fMRI) experiment tests only a small experimental manipulation for its effect on overall brain activity, but because the experimental paradigm for how to conduct one of these studies is clear, there is a wealth of these highly specific fMRI results. Our goal in this paper is to show how to take these results and, using dimensionality reduction techniques such as singular value decomposition (SVD) and non-negative matrix factorization (NMF), extract the higher-level patterns of brain activity that span many experiments. We introduce in this paper a methodology for evaluating how well these low-dimensional approximations reflect reality – namely, by examining how well values omitted from the data can be reconstructed using the approximations. We also show that two techniques for balancing the importance of all the data – namely, blending and normalization – are likely necessary for distinguishing effects such as experiment popularity from actual results.

Although the prevailing view of brain function assumes brain areas are highly selective and specialized, with each area responding to a restricted class of inputs and contributing primarily to a single cognitive domain such as language or motor control, the actual degee of selectivity of individual brain regions has increasingly come into question (Poldrack 2006; Anderson 2010). Though looking at a single brain-imaging experiment can appear to confirm the hypothesis of a very specific local function when the experimental manipulation produces the desired result, looking at many studies often reveals that any given brain area is more versatile than the single experiment would suggest. A statistical analysis of 1,469 brain imaging experiments shows that most brain areas appear to be reused across a variety of domains (Anderson 2010). Moreover, it appears to be not the individual area that is selective to a given cognitive domain, but networks of different areas. Nevertheless, the actual degree of selectivity of brain networks is still largely unknown.

The current study uses a database of 2,603 brain-imaging experiments to further investigate the selectivity of brain networks. The experiments were extracted from 824 journal articles published between 1996 and the present, including, for instance, all qualifying articles from the *Journal of Cognitive Neuroscience* (Anderson, Brumbaugh, and Suben 2010). All the experiments in the database use functional magnetic resonance imaging (fMRI), a technique for determining which areas of the brain are active when the subject is engaged in particular tasks. The experiments in the database produce images of brain activity by finding areas[1] of the brain that are significantly more activated [2] under some experimental condition than in a control condition. For each experiment, the data also includes a list of the cognitive domains that the experiment was addressing – for example, "Perception: Vision" or "Cognition: Language" – using the BrainMap (Fox and Lancaster 2002;

[1] Based on standard Freesurfer regions; http://surfer.nmr.mgh.harvard.edu/

[2] Here, "activated" means an increase in the Blood-Oxygenation Level Dependent (BOLD) signal that fMRI detects. The BOLD signal is an indirect indicator of neural activity; see (Logothetis 2007; Logothetis et al. 2001) for an overview.

Laird, Lancaster, and Fox 2005) classification system.

Since our database can essentially be represented as a large matrix, there are two methods that appear most promising for extracting high-level patterns of activity: singular value decomposition (SVD) and non-negative matrix factorization (NMF). We can plot each experiment in a high-dimensional space that represents the brain areas and cognitive domains involved, then use the SVD to collapse the space into a basis that more readily captures the regularities in the data. In the best case, the dimensions of the new space can represent meaningful variables – for example, an SVD on the statements in the Open Mind Common Sense database can reveal axes such as liked versus hated objects, and possible versus impossible actions (Speer, Havasi, and Lieberman 2008). Ideally, the dimensions of a space of brain experiments might represent networks of regions that meaningfully co-activate. NMF is similar, but instead makes the assumption that we are looking for a strictly additive combination of non-negative components. NMF can be used to decompose a data set of photographs of faces into parts such as eyes, noses, eyebrows, and mouths (Lee and Seung 1999). One might hope that if brain activity can be characterized by the sum of distinct networks responsible for each cognitive domain, that NMF would better characterize this activity as the sum of somewhat independent parts.

Two previous studies are most similar to the current work. One used NMF to find text associated with particular areas of the brain, using a matrix of abstract text versus brain area to find terms commonly associated with particular areas (Nielsen, Hansen, and Balslev 2004). Another used SVD and independent component analysis (ICA), a method similar to NMF without the non-negativity constraint, to find networks of co-activation in the resting brain, comparing them against networks found using the same techniques on reported experimental results (Smith et al. 2009). In both papers, a critical question was left unanswered: how would it be possible to evaluate the networks and meanings that were extracted?

In this paper, we introduce a technique for examining the goodness of fit of these methods for extracting patterns in brain activity – namely, the ability of the matrices to reliably reconstruct results when they are omitted from the data, using iterative reconstruction (Kurucz, Benczúr, and Csalogány 2007; Zhang et al. 2006). The idea of relating the quality of a model to its predictive capability is natural within machine learning, but it has apparently not been used before for neuroimaging results. By using this technique, we show that the quality of the approximation can dramatically decrease when using too many dimensions. We also show that when the data is normalized to take into account the frequency with which areas are activated and domains are tested, there is no particular difference in the number of dimensions used, nor whether SVD or NMF is used – demonstrating that large artifacts in the data can appear when metastudies do not take the frequency of experiments into account.

Methods

We have chosen two models to test in representing our data, non-negative matrix factorization and singular value decomposition. SVD is characterized by decomposing a matrix-based data representation into a set of orthonormal basis vectors which represent latent patterns in the source data set. NMF decomposes its source matrix into a set of non-negative basis vectors. Unlike SVD, an NMF decomposition is not necessarily unique or the optimal decomposition of the space. Since NMF prohibits negative values, it cannot directly represent inhibition. However, the non-negativity constraint causes the NMF to not have tendency to use these negative values to overfit a space.

To create our source space, we used data on fMRI activations from the NICAM database collected by Anderson et al (Anderson, Brumbaugh, and Suben 2010). We created two source matrices — A for brain areas which showed activation in the fMRI data and D for domain tags (such as "Cognition, Language, Phonology") used by the data set's curators. Each row of these matrices represented an experiment and each column either a brain activation or a domain tag. These two data sets were combined together to form a single source matrix. In some experiments, we then normalized this matrix to account for the fact that certain experiment areas are more popular than others.

This matrix is used as the source matrix for either SVD or NMF analysis. These decomposition techniques find a set of basis vectors which compress the space by representing it in terms of a basis of patterns or clusters which represent the variance in the data. We can see an example of this by examining how one of the SVD's basis vectors represents language-related tags an area in Figure 4. The spaces created by these methods can be evaluated to understand how various factors such as normalization and the number of basis vectors affect the model. The optimal space can then be used in future work to continue to analyze brain function.

Creating the Matrix

Our data consisted of two matrices, the brain area matrix A and the experiment domain matrix D, created from a database of 2,603 subtraction-based within-subject fMRI experiments from the NICAM database (Anderson, Brumbaugh, and Suben 2010). The experiments were extracted from 824 journal articles published between 1996 and the present, including, for instance, all qualifying articles from the *Journal of Cognitive Neuroscience* (Anderson, Brumbaugh, and Suben 2010). Of these experiments, 155 were excluded because they did not activate any of the Freesurfer regions, and 192 were excluded because they contained only the coarsest level of domain description (e.g., "Cognition"), presumably because they fit no more specific category. The remaining experiments were described by the 2,256 rows of each matrix. A was a 2256×80 matrix in which entry a_{ij} was 1 iff the Freesurfer brain area j was reported to have significant activation in experiment i, and 0 otherwise. D was a 2256×73 matrix in which entry d_{ij} was 1 iff the database reported experiment i as testing cognitive domain j. Each experiment could belong to more than one domain,

including many cases in which the domains were in different branches of the BrainMap hierarchy. Experiments classified in a sub-subdomain also resulted in a 1 in the parent classification's column.

Blending

We must combine A and D together so we can create a single space to reason over that is influenced by both the domain tags and brain activations. Thus, our final target matrix has rows that represent experiments and columns that represent either a brain area or a domain tag. We explored simply concatenating the two matrices, as compared to using the blending methodology, which aims to balance the influence that each data set has on the final results.

We combined A and D into a single 2256×153 matrix $B = [f_1 A \quad f_2 D]$. The factorization of B gives a vector space that represents correlations within and across both matrices. In concatenation, $f_1 = f_2 = 1$; in blending (Havasi et al. 2009), the weighting factors f_1, f_2 are set so that one matrix does not overpower the other: $f_i = 1/\sigma_{1i}$, the first singular value of the matrix.

Normalization

Normalization takes into account the fact that certain types of experiments are more popular in the literature than others, and attempts to correct for that fact in the source matrix. We would like the columns of B to sum to one, so that each experiment or tag is given equal weight in the decomposition.

In the case of blending, both an unnormalized and normalized combined matrix were examined; in the normalized case, each entry was divided by the Euclidean norm of its column to balance the importance of each domain and area.

Evaluation Methodology

To test the SVD and NMF's ability to reconstruct missing values, 100 experiments were selected at random, and a random positive entry was set to 0 for that experiment, to create a matrix B^0. Another zero entry was selected from each experiment for testing precision later. The positive and negative entries selected resulted in a mask M where $m_{ij} = 1$ if the entry was being tested, 0 otherwise. In the case of SVD, the entries were iteratively reconstructed by factoring B^t into U^t,S^t, and V^t, truncating to the first k singular values, reconstituting the matrix $R^t = U_k \Sigma_k V_k^T$, and setting $B_{ij}^{t+1} = R_i^t j$ if $M_{ij} = 1$ and B_{ij} otherwise. That is, the test values were set to the reconstituted values from dimensionality reduction, and all other values were left alone. This process was repeated until the process converged, as determined by the root mean square of the difference between B^i and R^i falling below a constant ($\epsilon = 0.0001$).

The reconstitution for NMF was similar; the reconstituted entries was replaced at each step with entries from $W_k * H_k$, where these were the low-dimensional approximations returned by the NMF. The constant ϵ for convergence was kept the same (ϵ = 0.0001), though because NMF has random starting points and local minima, the reconstruction did not necessarily converge; it was also terminated when a reconstruction resulted in a larger difference from the previous matrix than at the last step. After convergence, or this termination, the NMF was run 10 times with random initialization on the reconstructed matrix, and the lowest error approximation was taken.

In both cases, to decide whether a reconstructed entry should be considered to return "true" for the purpose of evaluation, the matrix entry was compared to a threshold of $0.05\lambda_i$, where λ_i was the submatrix's blending factor (1 if no blending was performed). This threshold was chosen after some experimentation to produce a good tradeoff between precision and recall for both SVD and NMF; we tried values between 0.01 and 0.5 at $k = 10$, 20, and 30. In the case of normalized entries, this threshold was also scaled by the column's normalization factor.

We repeated the steps above for $k = 5, 10, 15, 20, 25, 30, 35, 40, 45$, and 50 dimensions, with the same 100 experiments run for each condition. The whole run was then repeated 10 times, each with a different 100 experiments as test examples. We calculated precision as the proportion of reconstructed values that were nonzero in the original matrix, and recall as the proportion of 100 omitted positive values that were correctly reconstructed. The F measure weights both of these factors equally.

Results

In the unnormalized, unblended case, SVD consistently outperformed NMF for all numbers of dimensions, though the difference was negligible for their best performance (Figure 1). SVD's best average F measure was at $k = 10$ ($F = 0.740$), with 78% average precision and 72% recall. NMF's best F measure also occurred at $k = 10$ ($F = 0.736$), with 76% average precision and 71% average recall. The difference in F measures between the two methods across the ten runs was not significant for $k = 10$, but was significant for $k \geq 35$ via a nonparametric Wilcoxon ranksum test ($p < 0.05$).

Blending (weighting the area and domain matrices to have equal first eigenvalues) appeared to improve performance across the board for both algorithms, and was of particular benefit to the NMF algorithm (Figure 2). Blending factors of $\lambda_A = 0.0187, \lambda_D = 0.362$ were found to balance the eigenvalues of A and D, thus weighting the domain matrix three times as heavily as the area matrix. The best performance for SVD was now at $k = 15$ ($F = 0.784$), with 83% precision and 75% recall. Best performance for NMF was still at $k = 10$ ($F = 0.783$), with 80% precision and 77% recall. Again, the difference between the algorithms was not significant when comparing their best performance, but now, NMF performed significantly better ($p < 0.05$) than SVD when compared at 30, 40, and 50 dimensions. Blending significantly improved both algorithms' performance across all dimensions ($p < 10^{-9}$ (SVD) and $p < 10^{-24}$ (NMF) combining all F measures), and in particular significantly improved the algorithms' performance when compared at only their best k ($p < 0.02$).

Normalization of the columns of the matrix, which served to remove bias toward particular areas, had little effect on the algorithms' best performance, but dramatically improved

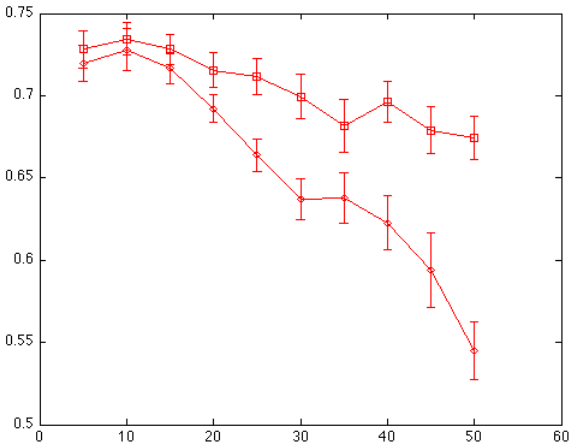

Figure 1: F measures for the unblended, unnormalized matrix for dimensionality reduction to varying numbers of dimensions k (x-axis) for SVD (squares) and NMF (diamonds). Between 10 and 15 dimensions appear to be the ideal number, with performance dropping as dimensions increase. SVD produces the best performance in this naive case. (Bars are standard error across 10 runs.)

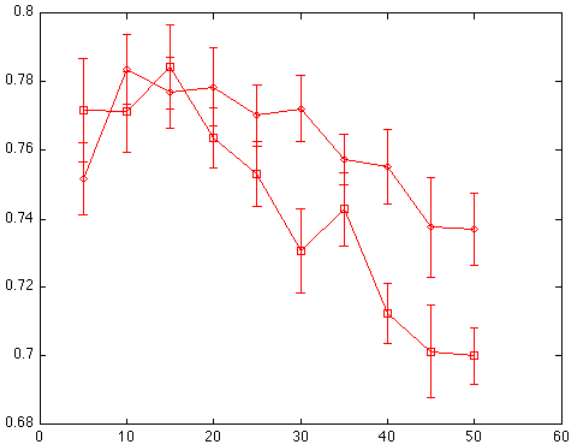

Figure 2: F measures for the blended (i.e., equally weighting areas and domains) unnormalized matrix for varying k (x-axis) for SVD (squares) and NMF (diamonds). Overall performance is better once the two data sources are weighted to contribute equally, and predictive performance with more dimensions does not suffer as much. NMF benefits most from the weighting, now outperforming SVD with more dimensions.

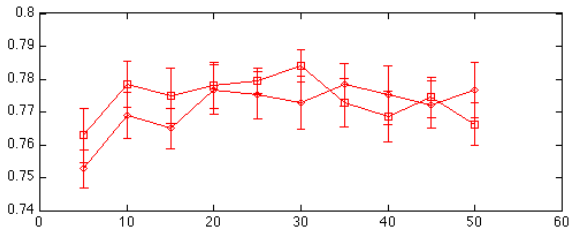

Figure 3: F measures for the blended (i.e., equally weighting areas and domains) and normalized matrix for varying k (x-axis) for SVD (squares) and NMF (diamonds). Remarkably, with normalization, the performance does not drop off much at all with more dimensions, suggesting that previous results about the true number of subcomponents of brain activity may simply be a result of oversampling particular kinds of experiments. Normalization also appears to eradicate any difference in performance between the two methods. (Graph scaled to match previous figure.)

performance for dimensionality reduction at higher dimensions (Figure 3). Best performance for SVD was again at $k = 10$ ($F = 0.778$), with a mean precision of 78% and a mean recall of 78%. This performance was not significantly different from the best unnormalized performance. Nor was the NMF performance significantly different at its best k ($k = 10$, $F = 0.769$, precision = 77%, recall = 77%, insignificant difference from SVD or unnormalized case). However, as the figure shows, the effect is striking for higher dimensions; there is no longer a significant difference for either algorithm between performance at 10 and at 50 dimensions! Normalization also removes any significant difference between the algorithms at $k = 50$. What may have seemed a fundamental insight about the true number of dimensions to posit in the brain, or whether a parts-based or holistic representation of the brain is more accurate, may actually be an artifact of a bias in how common it is to run particular kinds of experiments.

We can further examine qualitatively the results of each experiment by plotting the basis vectors created for each domain and area (Figures 5 and 6). We used SVDview (Speer et al. 2010) to visualize the basis vectors, which can help give an intuition for which domains and areas would be clustered together when a clustering algorithm is run on the reduced space. In general, areas with known function tended to occur near their respective domains – for example, "pars opercularis," a part of the language area known as Broca's area, was clustered along the same axis as the domains of phonology, semantics, and general language (Figure 4). In addition, matching areas on opposite sides of the brain – for example, the left and right superior temporal cortex – tended to land near each other in the space (Figure 5). However, there were many domains that did not fall near any particular area; and in some cases, particularly for areas typically associated with high-level cognitive function and language, left and right areas did not land near each other. In other words, although the results sometimes fit with our canonical understanding of local function, there is in general no clear

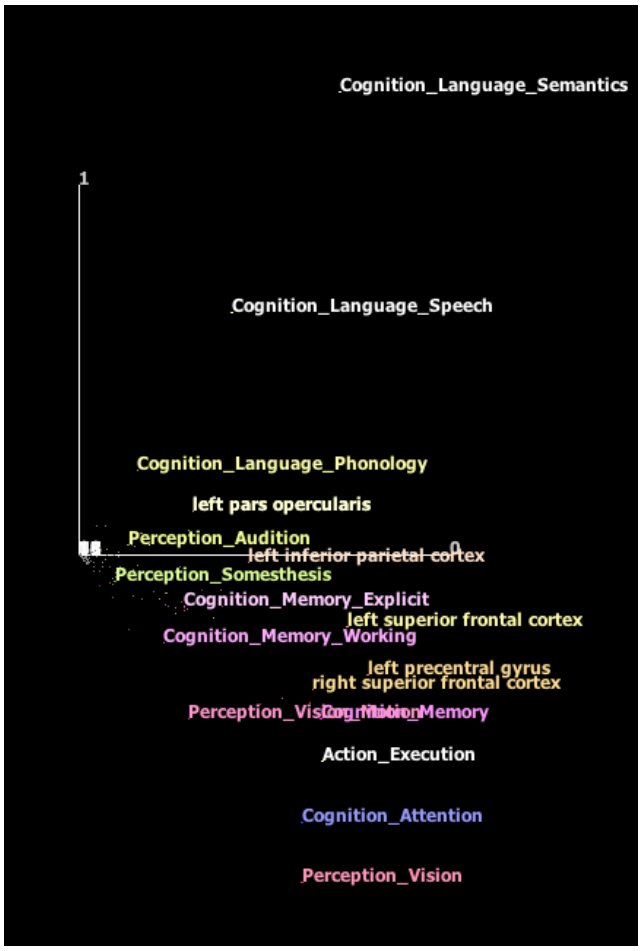

Figure 4: Plotting a projection of the first two dimensions of the 20-dimensional SVD. Related areas and domains tend to fall along an off-axis line from the origin, including semantics, language, and the left pars opercularis (Broca's area) extending to the upper right, and vision, attention, and action execution extending to the lower right.

mapping between domain and area.

The results are even more striking when the basis vectors of the NMF are visualized (Figure 6). Here, the components of the basis vectors H are again plotted in the space, and for most of the basis vectors, *areas were not clustered with domains*. Instead, for each basis vector, one area or domain tended to dominate the vector. This is a *very* different finding from what we would expect if the current BrainMap ontology mapped well to brain areas, in which case, we would expect vectors with large area components to have large components of the corresponding domain, and vice versa.

Conclusions

The questions we begin to approach here are much larger than what we can answer in a single paper: How is the brain truly organized? How good is the current ontology for neuroscience experiments, and can we improve it through dimen-

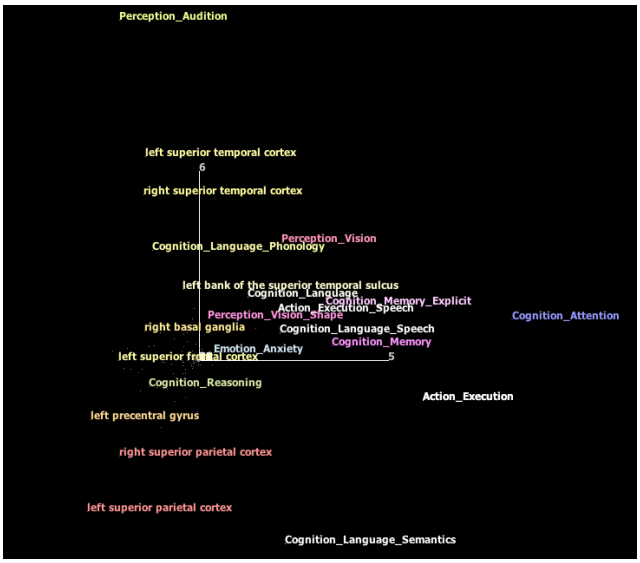

Figure 5: Plotting the domains and areas into the collapsed space produced by the SVD for 20 dimensions in SVDview. An area's mirrored counterpart on the opposite side of the brain tends to be co-activated with the area in many cases, resulting in the left and right areas plotted close to each other in the space.

sionality reduction and clustering or NMF? How specific in function are the brain's areas, and is it any easier to identify function when we look at co-activity of multiple areas? As we attempted to answer these questions, we found that we had to develop new methodology for even evaluating the quality of our dimensionality reduction – much less making any conclusions from the reduced space. We have not yet made any new neuroscientific conclusions, but we have made several contributions here in methodology.

First, our results demonstrate the importance of evaluating the quality of these reduced spaces, and give a concrete methodology for doing so: that of reconstruction of values from nearly complete information. Given our results in Figures 1 and 2, it is clear that studies that collapse to a large number of dimensions are at risk of overfitting. Our results show the importance of separating training from test data when performing metastudies. This is a useful contribution of our work.

Second, our results show the importance in scientific metastudies of normalization to take into account that certain hypotheses and experiments may be more popular than others. Had we not run our analysis on normalized data, we would likely have concluded that there are roughly 10-15 fundamental networks in the brain. Instead, we draw the conclusion that there are roughly 10-15 very commonly studied cognitive domains, and unnormalized metastudies are likely to overfit to these domains at the expense of less well-understood areas.

Third, this paper represents the first application of blending to NMF, and we have shown that blending improves results in this domain for both SVD and NMF. We have

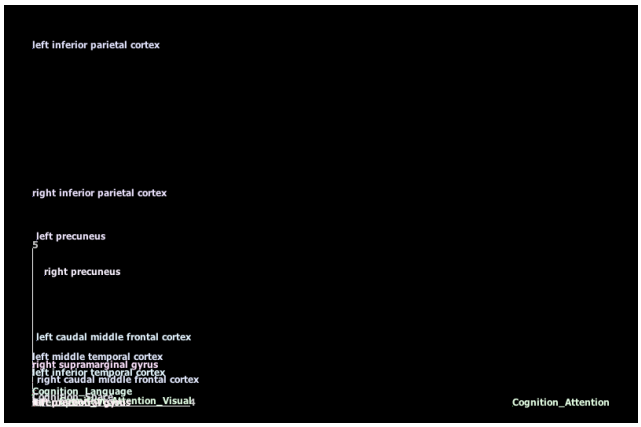

Figure 6: Visualizing two columns of the basis matrix H for the NMF factorization with 20 basis vectors in SVD-view. Though one might predict that areas would be coupled tightly with experimental domains in an NMF decomposition, this does not appear to happen at all. Instead, each basis vector is dominated by a single area or domain.

also shown that the existing reconstructive algorithms for SVD and NMF work with blended matrices as well, and have shown (but not proven) that scaling the reconstruction thresholds by the blending factor gives satisfactory performance.

Fourth, we have provided a potentially useful application for neuroscientists; an 80% precision is high enough to make it potentially worthwhile to run an experiment to look for an activation in a novel area, or to check whether an unexpected domain would evoke an area. Ideally, our method could be used to both construct new hypotheses and check data that already exists in the database; one future application we intend to pursue is to check the efficacy of the algorithm in reconstructing the missing subdomains for the studies we omitted.

This is our first work in applying SVD and NMF to cognitive neuroscience data, and many questions remain unresolved: Why does NMF not pair the domains with areas more effectively? Can we produce a better experimental ontology through these methods, and if so, how? Why is there no effective difference between the two methods on the normalized data, when they make fundamentally different assumptions about the nature of the components? We are excited by the possibility of using these techniques to generate novel discoveries in cognitive neuroscience.

Acknowledgments

The authors would like to thank Rob Speer and Jason Alonso for their comments on an early draft of the paper.

References

Anderson, M. L.; Brumbaugh, J.; and Suben, A. 2010. Investigating functional cooperation in the human brain using simple graph-theoretic methods. In *Computational Neuroscience*. New York: Springer.

Anderson, M. L. 2010. Neural reuse: A fundamental organizational principle of the brain. *Behavioral and Brain Sciences* 33:245–313.

Fox, P., and Lancaster, J. L. 2002. Mapping context and content: The BrainMap model. *Nature Rev. Neuroscience* 3:319–321.

Havasi, C.; Speer, R.; Pustejovksky, J.; and Lieberman, H. 2009. Digital intuition: Applying common sense using dimensionality reduction. *IEEE Intelligent systems* 24(4):24–35.

Kurucz, M.; Benczúr, A.; and Csalogány, K. 2007. Methods for large scale SVD with missing values. In *Proceedings of KDDCup '07*. ACM.

Laird, A.; Lancaster, J.; and Fox, P. 2005. BrainMap: The social evolution of a human brain mapping database. *NeuroInformatics* 3:65–77.

Lee, D. D., and Seung, H. S. 1999. Learning the parts of objects by non-negative matrix factorization. *Nature* 401:788=791.

Logothetis, N. K.; Pauls, J.; Trinath, M. A. T.; and Oeltermann, A. 2001. Neurophysiological investigation of the fMRI signal. *Nature* 412:150–157.

Logothetis, N. K. 2007. The ins and outs of fMRI signals. *Nature Neuroscience* 10:1230–1232.

Nielsen, F.; Hansen, L.; and Balslev, D. 2004. Mining for associations between text and brain activation in a functional neuroimaging database. *Neuroinformatics* 4:369–380.

Poldrack, R. A. 2006. Can cognitive processes be inferred from neuroimaging data? *Trends in Cognitive Sciences* 10:59–63.

Smith, S. M.; Fox, P. T.; Miller, K. L.; Glahn, D. C.; Fox, P. M.; Mackay, C. E.; Filippini, N.; Watkins, K. E.; Toro, R.; Laird, A. R.; and Beckmann, C. F. 2009. Correspondence of the brain's functional architecture during activation and rest. *PNAS* 106(31):13040 — 13045.

Speer, R.; Havasi, C.; Treadway, N.; and Lieberman, H. 2010. Finding your way in a multi-dimensional semantic space with luminoso. In *Proceedings of Intelligent User Interfaces*.

Speer, R.; Havasi, C.; and Lieberman, H. 2008. AnalogySpace: Reducing the dimensionality of common sense knowledge. *Proceedings of AAAI 2008*.

Zhang, S.; Wang, W.; Ford, J.; and Makedon, F. 2006. Learning from incomplete ratings using non-negative matrix factorization. In *Proc. of the 6th SIAM conference on data mining*, 549–553.