Sparse Methods for Medical Imaging Decomposition

Ben Kandel

June 4, 2012

Abstract

This is the abstract of the document. In short, it's clear to me that sparsity is cool and worthwhile. Now what to do with it...

I have no idea what to do with it. Oh well. We'll have to figure it out.

1 Methodology

There are a lot of things.

1. One basic idea is to do some **non-linear** sparse dimensionality reduction–not CCA, not PCA, but sparse non-linear dimensionality reduction. There are about a bajillion non-linear dimensionality reduction techniques [4]. Rueckert and his gang are using locally linear embedding, a manifold learning technique, for combining multimodal (in this case, imaging + clinical data) for classifying images [5]. Possibly sparse extensions of this (in his case, he used a kind of sparse representation of the image–only looked at the MTL/hippocampus, but it may be interesting to look at extensions of this).

2. Contiguous PCA–in practice, NMF does, in theory [1] give a parts-based representation, when the population you're looking at actually comes from bases and loading vectors. How this applies to real brains is anyone's guess. I think that by actually enforcing contiguity constraints, you might be able to do a better job than just hoping for contiguity by doing the l1 norm or something like that.

3. Isomap [3] and LLE [2] are both nonlinear dimensionality reduction techniques. I wonder if they can be extended to have a sparse representation of images. The way this would work is that you'd get some sort of sparse bases of the images and then calculate graph distances and build graphs. Ideally, there would be some sort of feedback so that the classification informs the dimensionality reduction so that the reduction/sparsification is optimally tuned to produce a graph that's useful for classification.

In general, I think that it's worth thinking a bit about how interested I am in classification (supervised learning) vs unsupervised/exploratory data analysis.

2 Applications

Possible things to look at:

1. Cardiac imaging–look at which areas of stenosis/wall thinness are most correlated with some disease outcome

2. Combining imaging and non-imaging information–Rueckert's recent paper in MedIA, looking at genomics, building prediction models for areas of brain based on genomic clusters

3. Methodology–explicit contiguity constraints. I think this may be more important than it appears . . . also maybe look at fused lasso-type penalties, but in a multivariate matrix decomposition (or for that matter, totally unsupervised context)– don't think that's been done before.

4. breast imaging (despina)? may be worthwhile . . . but not sure if it makes sense . . . maybe replace her ROI method with sparse PCA, looking for different textures and comparing against disease outcomes.

Literature Cited

[1] David Donoho and Victoria Stodden. When does Non-Negative matrix factorization give a correct decomposition into parts? In Sebastian Thrun, Lawrence Saul, and Bernhard Schlkopf, editors, *Advances in Neural Information Processing Systems 16*. MIT Press, Cambridge, MA, 2004.

[2] S.T. Roweis and L.K. Saul. Nonlinear dimensionality reduction by locally linear embedding. *Science*, 290(5500):2323–2326, 2000.

[3] J.B. Tenenbaum, V. De Silva, and J.C. Langford. A global geometric framework for nonlinear dimensionality reduction. *Science*, 290(5500):2319–2323, 2000.

[4] L. J. P. Van Der Maaten, E. O. Postma, and H. J. Van Den Herik. Dimensionality reduction: A comparative review. *Published online*, 10(February):1–35, 2007.

[5] Robin Wolz, Paul Aljabar, Joseph V Hajnal, Jyrki Lotjonen, and Daniel Rueckert. Nonlinear dimensionality reduction combining MR imaging with non-imaging information. *Medical Image Analysis*, 16(4):819–830, May 2012. PMID: 22244037.