1 Feb 15 notes

Suppose that the observations are a Hermitian matrix. If they are not, then start from the Gram matrix and take a matrix sqrt of this matrix to be the observation matrix.

We would like to analyze $||\Sigma||_{\text{op}}$.

$$\Sigma = D^T M^T M D$$

Where M is the Hermitian data matrix, and D is an n-by-k matrix: $\begin{pmatrix} e_{d_1} & e_{d_2} & \dots & e_{d_k} \end{pmatrix}$, selecting k columns of M using multiplication on the right.

We introduced u, which can be initialized to be the left singular vector of MD for some well-performing D ($\sigma = |uMD|_2$ large). This is slightly useful, because inner products with u could be factored out to be computed using entries of Σ only, and discard/never synthesize an M.

Set $A = D^T M^T u u^T M D$ (large and rank-one), and $E = \Sigma - A$. Analyzing E for different permissible D should give us a smaller recursive problem, compared to analyzing Σ (shrinking the problem). However, we would like to produce a new upper-bound deterministically in polynomial time instead.

We found that E is similar to:

$$(I - \frac{1}{\sigma^2} M D u u^T D^T M) M^2$$

We hope that $||E||_{\text{op}}$ is small compared to $||\Sigma||_{\text{op}}$. Note that with a change-of-basis using the diagonalization of M, then the matrix in parentheses is still a rank-$n-1$ projection matrix. It could have a zero for the dominant eigenvalue of M^2, and that is the best case. We want a large sine of the angular comparison of the dominant eigenspaces of the projection matrix on the left and of M, and this sine can be bounded using the Davis-Kahan theorem.

Let $P = I - \frac{1}{\sigma^2} M D u u^T D^T M$. For Davis-Kahan, we should come up with our own upper-bound on $||P - \frac{1}{\sigma^2} M^2||_{\text{op}}$. Davis-Kahan bounds the dissimilarity in the largest eigenspaces of P and $\frac{1}{\sigma^2} M^2$, and if this dissimilarity obtains its maximum value, then the dominant eigenvalue would drop out of PM^2 and $||PM^2||_{\text{op}}$ would be driven by the second-largest eigenvalue of M^2 instead.

2 March 1 notes

We would like a large sine distance between the eigenvector spanned by the range of $\frac{1}{\sigma^2} M D u u^T D^T M$, and the dominant eigenvector of M. We expect:

$$\sin \angle(\frac{1}{\sigma^2} M D u u^T D^T M, M) = \cos \angle(I - \frac{1}{\sigma^2} M D u u^T D^T M, M)$$

We don't need to use $\sin^2 \theta + \cos^2 \theta = 1$, because we are actually swapping whether the vector or some orthogonal vector is used as sin or as cos.

To improve performance of Davis-Kahan, scale up the rank-one projection matrix so that it is $M D u u^T D^T M$. Apply Davis-Kahan:

$$|\sin\theta| \leq \frac{2||M(I - Duu^T D^T)M||_F}{\min(||MDuu^T D^T M||_{\text{op}}, \text{Gap}[M^2])}$$

We need a lower bound on Gap$[M^2]$. If our gap evaluation is intentionally simplistic (take $\lambda_2 = \text{Tr } M^2 - \lambda_1$), then the gap estimate is nondecreasing with λ_1. If we choose a putative eigenvector to evaluate λ_1 for Gap$[M^2]$, and later we increase our objective function, then our gap using the best sparse PCA solution so far will still hold as a lower bound. Therefore, we use Berk et. al's stochastic best-observed lower bound solution to compute the gap.