A graph feature autoencoder for predicting perturbations to steady state gene expression

C.W. Seitz

July 2, 2022

1 Revealing causal relations and predicting gene expression

Suppose that we wanted to predict the response of gene expression to some external perturbation, like an osmotic shock and treatment with a cytokine. In principle, it is *not* possible to infer that response unless a detailed list of biochemical reactions is known apriori - a list that is very rare and likely doesn't exist at all. Luckily, there is another pharmacological problem which can potentially be addressed using this formalism. If the drug of interest is known, and can potentially induce gene expression through unknown mechanisms, those mechanisms can be identified given the right dataset. We do not need to measure detailed interactions between variables, for example, the precise effect of histone modifications and chromatin structure on the rate of transcription. Instead, these interactions can be inferred from the data as long as those variables can be measured with reasonable precision. Ultimately, this framework cannot be used for screening a large panel of potential drugs but it is appropriate for learning mechanisms of say adaptive resistance to drug treatment. We can begin to answer questions like: what are the responsible mechanisms for drug resistance in single cells? In any case, this will require some effort for the learning methods to become interpretable in a biologically significant way. The mechanisms will have more of a qualitative description as well, unless a deep network were to be supplemented by more biophysical characterization. This could be a valuable tool in trying to understand which genetic mechanisms are responsible for the development of resistant states. Deep learning is a key method for integrating information provided by sequencing tools that are becoming readily available.

Another potential application (which is arguably less valuable/feasible) is *prediction* - the ability to predict the effects of the fold-change of one gene on its neighbors in the gene regulatory network. This requires some kind of domain knowledge for selecting the gene of interest and genes affected downstream. It is less feasible because drugs often don't act on genes directly but instead target certain proteins, presenting many unknowns between disruption

of protein-mediated signaling and transcription.

Attributed graph convolutional autoencoders have a well-developed theoretical basis. The major problem in their application here is that few, if any, frameworks are capable of end to end learning of the graph structure and the data distribution. For example, the seminal work by Kipf and Welling on GCNs requires that the adjacency matrix for the underlying graph is known apriori. One possible solution, is to make use of some of the information theoretic techniques for constructing a graph representation of high dimensional data using either pairwise mutual information or partial information decomposition.

Another point that I would like to make is that most graph oriented frameworks in this context consider interactions between coding genes only. Few have published work that incorporates information on chromatin structure in order to identify non-coding regions acting as distal enhancers to a particular coding sequence. The direct binding of the protein product of a coding gene at the promoter is not terribly likely, and our graphical models should therefore incorporate additional nodes for non-coding sequences as well. Edges drawn between noncoding and coding nodes would represent the involvement of a transcription factor bound at the noncoding sequence in the transcription of the coding sequence. This is slightly problematic, because information theoretic techniques dont make sense to infer noncoding to coding edges. These edges would need to be inferred from structural data alone. Also the graph convolution doesn't really make sense here either, and it would need to incorporate some boolean logic based on proximity to noncoding regions. So you would need to determine if the graph convolution is really the best representation here or an alternative transcriptional logic is more appropriate.