Methods for Estimating Hidden Structure and Network Transitions in Genomics
Daniel Schlauch, PhD Candidate
Department of Biostatistics, Harvard University
April 27, 2017

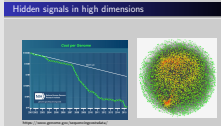
Methods for Estimating Hidden Structure and Network Transitions in Genomics

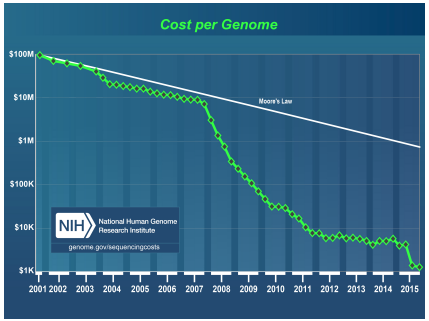
Daniel Schlauch, PhD Candidate

Department of Biostatistics, Harvard University

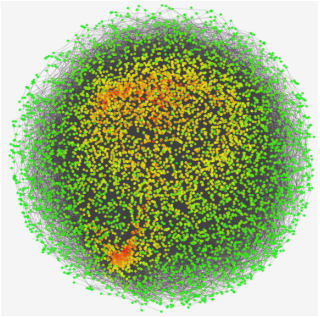
April 27, 2017

Hidden signals in high dimensions



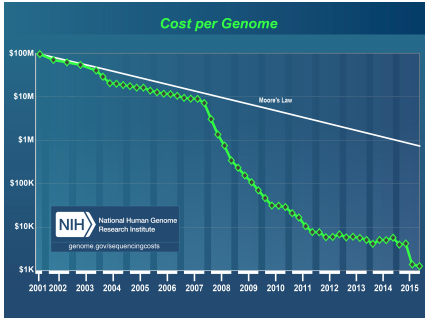


https://www.genome.gov/sequencingcostsdata/

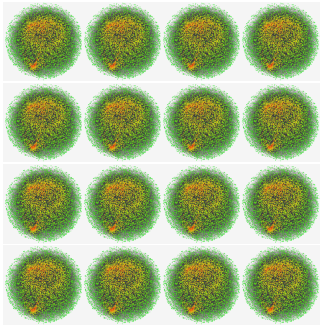


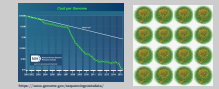
- Economics: New technologies have allowed us access to unprecendented data
- Microarrays, then seq, allowed us access to numerous fields of 'omics.
- Biology isn't simple
- But... snapshots of hairballs, which are noisy and convoluted

Hidden signals in high dimensions



https://www.genome.gov/sequencingcostsdata/

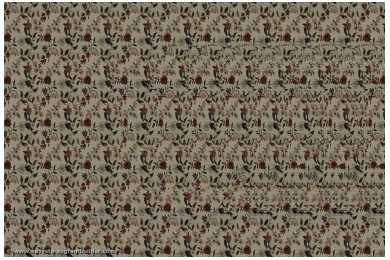




- Economics: New technologies have allowed us access to unprecendented data
- Microarrays, then seq, allowed us access to numerous fields of 'omics.
- Biology isn't simple
- But... snapshots of hairballs, which are noisy and convoluted

Hidden signals in high dimensions

- Lots of samples across LOTS of features
- Single snapshots of complex, fluid, heterogeneous landscapes
- Noisy, structured data



How do we separate the structure to reveal higher-level components?

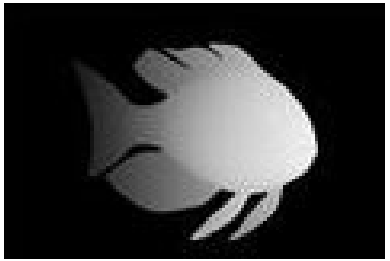
Hidden signals in high dimensions
- Lots of samples across LOTS of features
- Single snapshots of complex, fluid, heterogeneous landscapes
- Noisy, structured data

How do we separate the structure to reveal higher-level components?

- High Dimensional data
- This data has high amounts of technical and biological noise
- Single snapshots don't tell the whole picture.
- Collection of often heterogenous cells
- We often want to understand how features interact and lead to higher level phenotypes
- Challenging because structure comes in many forms.

Hidden signals in high dimensions

- Lots of samples across LOTS of features
- Single snapshots of complex, fluid, heterogeneous landscapes
- Noisy, structured data



How do we separate the structure to reveal higher-level components?

Hidden signals in high dimensions
- Lots of samples across LOTS of features
- Single snapshots of complex, fluid, heterogeneous landscapes
- Noisy, structured data

How do we separate the structure to reveal higher-level components?

- What's important is understanding that structure can be many things... pathways to cancer, or
- Microarrays, then seq, allowed us access to numerous fields of 'omics.
- But... snapshots of hairballs

Table of Contents

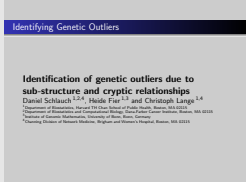
Table of Contents
Introduction
Identification of Genetic Outliers
Batch Effect on Covariance Structure
State Transitions Using Gene Regulatory Network Models
Future Work
Appendix

Identifying Genetic Outliers

Identification of genetic outliers due to sub-structure and cryptic relationships

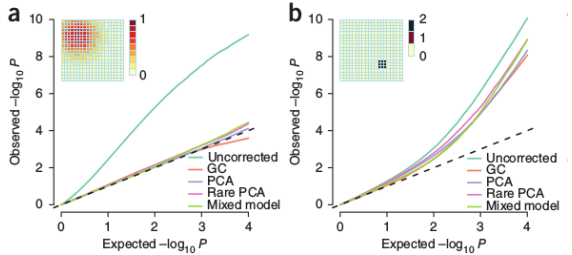
Daniel Schlauch [1,2,4], Heide Fier [1,3] and Christoph Lange [1,4]

[1] Department of Biostatistics, Harvard TH Chan School of Public Health, Boston, MA 02115
[2] Department of Biostatistics and Computational Biology, Dana-Farber Cancer Institute, Boston, MA 02115
[3] Institute of Genomic Mathematics, University of Bonn, Bonn, Germany
[4] Channing Division of Network Medicine, Brigham and Women's Hospital, Boston, MA 02115

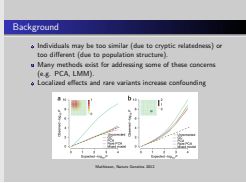


Background

- Individuals may be too similar (due to cryptic relatedness) or too different (due to population structure).
- Many methods exist for addressing some of these concerns (e.g. PCA, LMM).
- Localized effects and rare variants increase confounding



Mathieson, Nature Genetics 2012



- Inflation of type I error rate

Motivation

We want to create a similarity measure that...

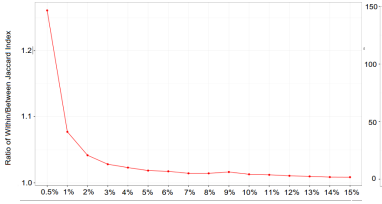
- is more sensitive to fine scale population stratification
- estimates relatedness
- can be used as a formal test for cryptic relatedness
- can be used as a formal test for population structure

Motivation
We want to create a similarity measure that...
- is more sensitive to fine scale population stratification
- estimates relatedness
- can be used as a formal test for cryptic relatedness
- can be used as a formal test for population structure

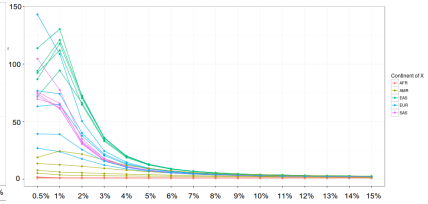
Rare variants

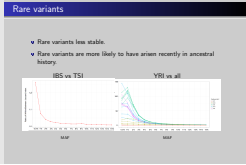
- Rare variants less stable.
- Rare variants are more likely to have arisen recently in ancestral history.

IBS vs TSI



YRI vs all





- RVs become fixed at 0% with high probability over a relatively short timeframe.

Test Statistic

STEGO: Similarity Test for Estimating Genetic Outliers

$$s_{i,j} = \frac{\sum_{k=1}^{N} w_k \mathbf{G}_{i,k} \mathbf{G}_{j,k}}{\sum_{k=1}^{N} I\left[\sum_{l=1}^{2n} \mathbf{G}_{l,k} > 1\right]}$$

where

$$w_k = \begin{cases} \frac{\binom{2n}{2}}{\binom{\sum_{l=1}^{2n} \mathbf{G}_{l,k}}{2}} & \sum_{l=1}^{2n} \mathbf{G}_{l,k} > 1 \\ 0 & \sum_{l=1}^{2n} \mathbf{G}_{l,k} \leq 1 \end{cases}$$

Under H_0, homogeneity,

$$s_{i,j} \sim N\left(1, \sigma_{i,j}^2\right)$$

- Distribution under the null has expectation 1, and variance is defined ONLY by the allele frequencies

Tests of Heterogeneity

$$H_0 : E\left[s_{i,j}\right] = 1 \forall i, j \in 1 \ldots n$$

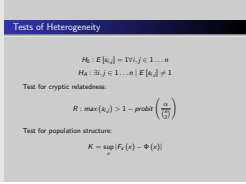
$$H_A : \exists i, j \in 1 \ldots n \mid E\left[s_{i,j}\right] \neq 1$$

Test for cryptic relatedness:

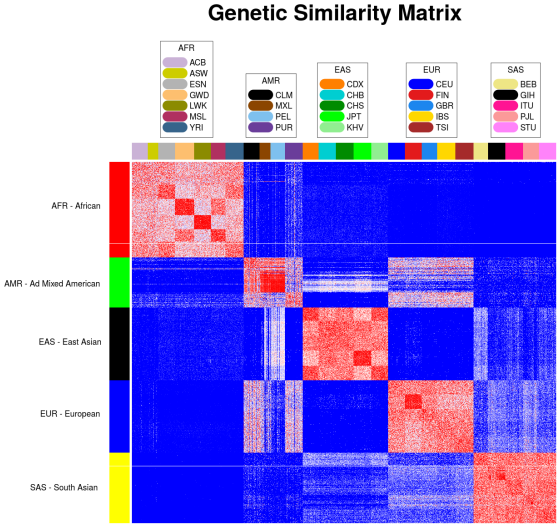
$$R : max\left(s_{i,j}\right) > 1 - probit\left(\frac{\alpha}{\binom{n}{2}}\right)$$

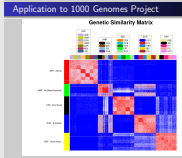
Test for population structure:

$$K = \sup_x \left|F_s\left(x\right) - \Phi\left(x\right)\right|$$

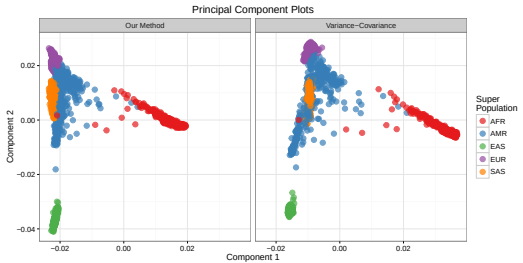


Application to 1000 Genomes Project

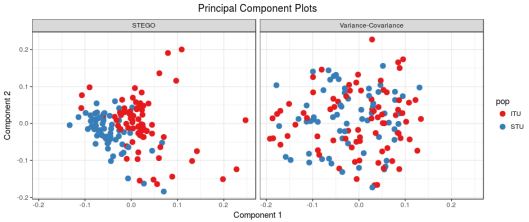




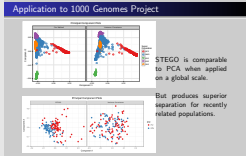
Application to 1000 Genomes Project



STEGO is comparable to PCA when applied on a global scale.

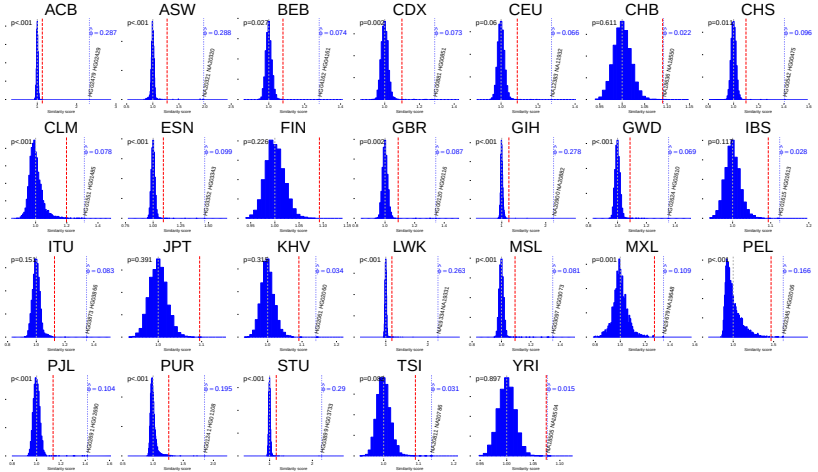


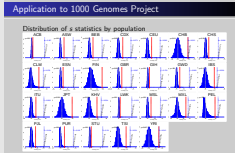
But produces superior separation for recently related populations.



Application to 1000 Genomes Project

Distribution of s statistics by population

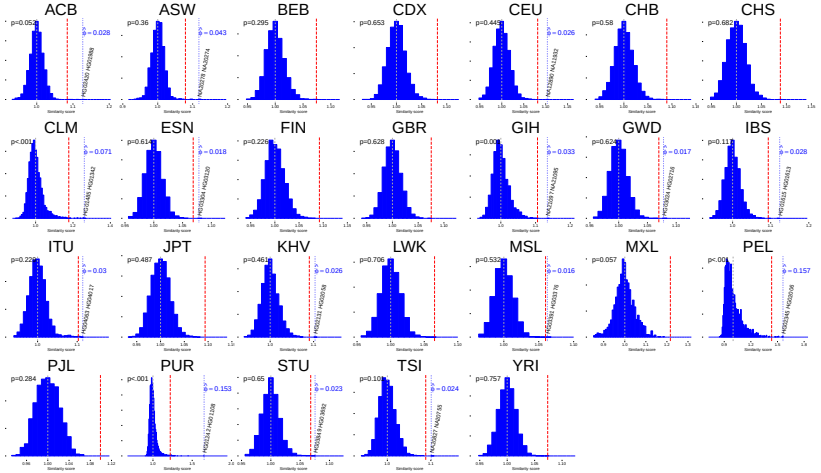


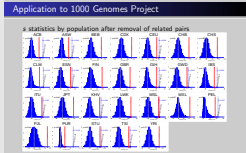


- Describe red-line, black line, and labeled pair.

Application to 1000 Genomes Project

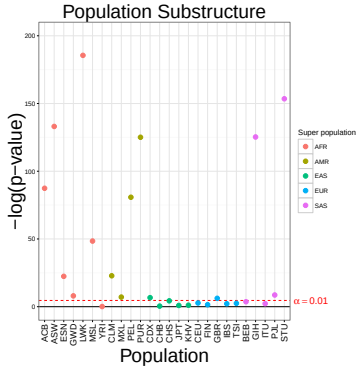
s statistics by population after removal of related pairs



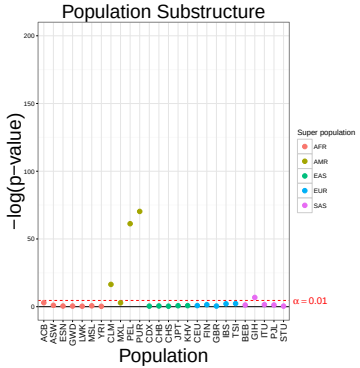


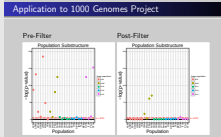
Application to 1000 Genomes Project

Pre-Filter



Post-Filter



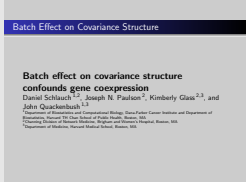


Batch Effect on Covariance Structure

Batch effect on covariance structure confounds gene coexpression

Daniel Schlauch [1,2], Joseph N. Paulson [2], Kimberly Glass [2,3], and John Quackenbush [1,3]

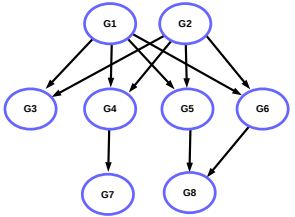
[1]Department of Biostatistics and Computational Biology, Dana-Farber Cancer Institute and Department of Biostatistics, Harvard TH Chan School of Public Health, Boston, MA
[2]Channing Division of Network Medicine, Brigham and Women's Hospital, Boston, MA
[3]Department of Medicine, Harvard Medical School, Boston, MA

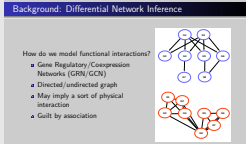


Background: Differential Network Inference

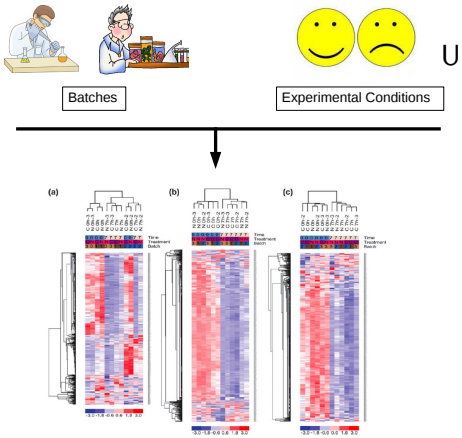
How do we model functional interactions?

- Gene Regulatory/Coexpression Networks (GRN/GCN)
- Directed/undirected graph
- May imply a sort of physical interaction
- Guilt by association





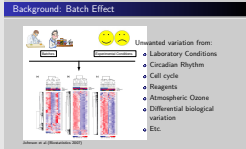
Background: Batch Effect



Johnson et al.(Biostatistics 2007)

Unwanted variation from:

- Laboratory Conditions
- Circadian Rhythm
- Cell cycle
- Reagents
- Atmospheric Ozone
- Differential biological variation
- Etc.



Methods for Controlling Batch Effect

Location scale model:

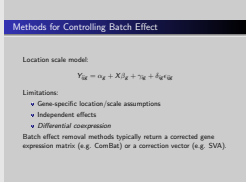
$$Y_{ijg} = \alpha_g + X\beta_g + \gamma_{ig} + \delta_{ig}\epsilon_{ijg}$$

Limitations:

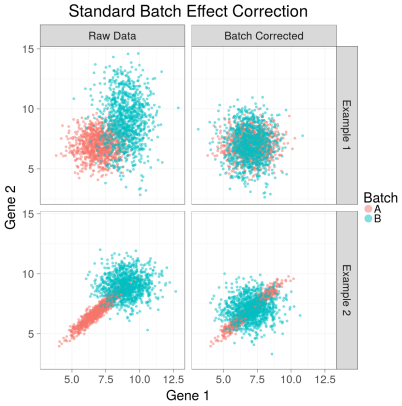
- Gene-specific location/scale assumptions
- Independent effects
- *Differential coexpression*

Batch effect removal methods typically return a corrected gene expression matrix (e.g. ComBat) or a correction vector (e.g. SVA).

- point 1



Limitations to common batch effect correction methods

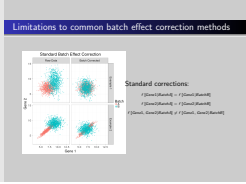


Standard corrections:

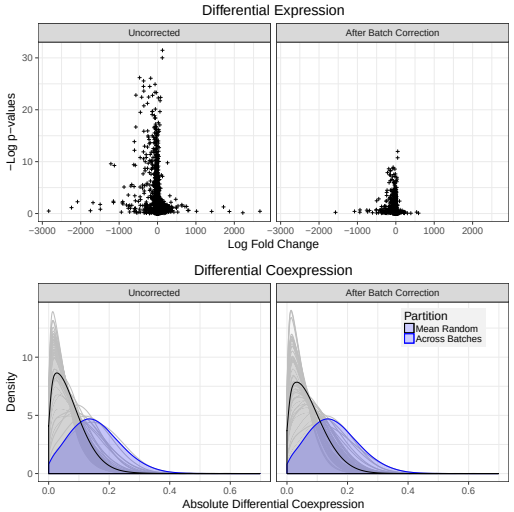
$$f\left[Gene1|BatchA\right] = f\left[Gene1|BatchB\right]$$

$$f\left[Gene2|BatchA\right] = f\left[Gene2|BatchB\right]$$

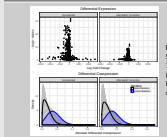
$$f\left[Gene1, Gene2|BatchA\right] \neq f\left[Gene1, Gene2|BatchB\right]$$



Limitations to existing batch effect correction methods



ENCODE Project:
50k genes
126 samples
(63 patients each RNA-seq'ed at 2 centers)



Estimating the conditional coexpression matrix

Proposed method: **CMA**, Coexpression model adjustment
Motivating concepts:

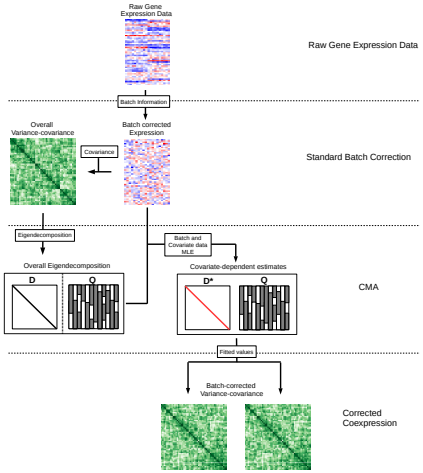
- Provide a regression framework for the coexpression matrix.
- Estimate a reduced number of parameters.

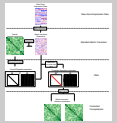
Our approach:

- Exploit modular nature of gene expression patterns.
- Define our parameters as functions of components of variation.
- Estimate the eigenvalue contribution of each eigenvector.

Estimating the conditional coexpression matrix
Proposed method: CMA, Coexpression model adjustment
Motivating concepts:
Provide a regression framework for the coexpression matrix.
Estimate a reduced number of parameters.
Our approach:
Exploit modular nature of gene expression patterns.
Define our parameters as functions of components of variation.
Estimate the eigenvalue contribution of each eigenvector.

Example Workflow

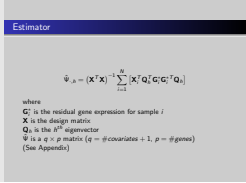




Estimator

$$\hat{\Psi}_{\cdot,h} = \left(\mathbf{X}^T\mathbf{X}\right)^{-1} \sum_{i=1}^{N} \left[\mathbf{X}_i^T \mathbf{Q}_h^T \mathbf{G}_i^* \mathbf{G}_i^{*T} \mathbf{Q}_h\right]$$

where
$\mathbf{G}_i^*$ is the residual gene expression for sample i
$\mathbf{X}$ is the design matrix
$\mathbf{Q}_h$ is the h^{th} eigenvector
$\hat{\Psi}$ is a $q \times p$ matrix ($q = \#covariates + 1$, $p = \#genes$)
(See Appendix)



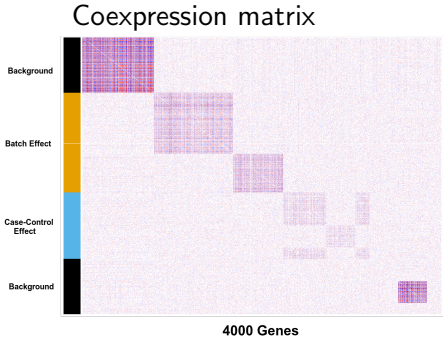
- We can use Ψ in the standard regression way... differential coexpression, fitted coexpression...

Simulations

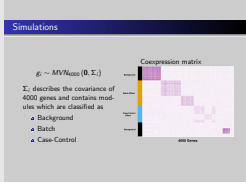
$$g_i \sim MVN_{4000}(\mathbf{0}, \Sigma_i)$$

Σ_i describes the covariance of 4000 genes and contains modules which are classified as

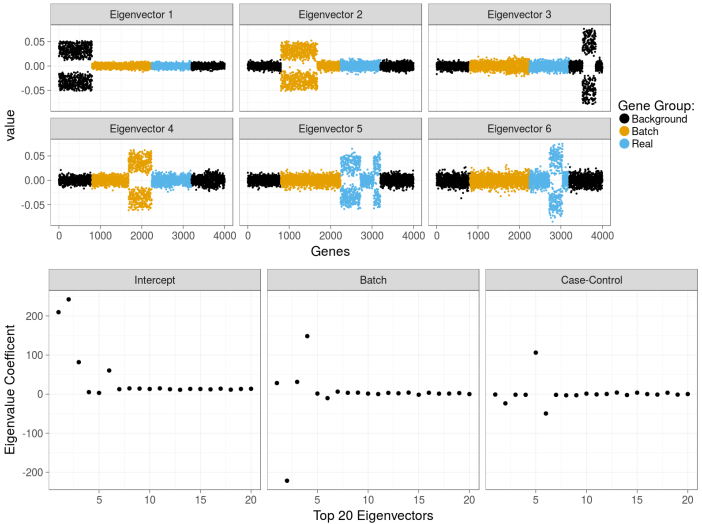
- Background
- Batch
- Case-Control

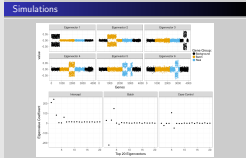


- Describe naive approach

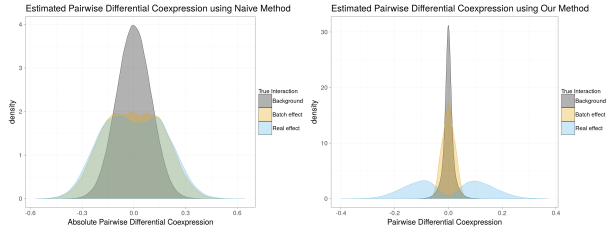


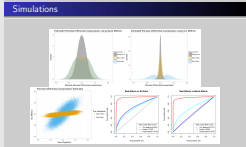
Simulations





Simulations





Application to data from COPDGene Study

The COPDGene Study (GSE42057):

- 136 Individuals
- Affymetrix Human Genome U133 Plus 2.0 microarrays
- 18,960 genes
- 42 Smoker Controls, 94 COPD subjects

$$\mathbf{S} \sim COPD + Gender + Age + Packyears$$

Application to data from COPDGene Study
The COPDGene Study (GSE42057):
136 Individuals
Affymetrix Human Genome U133 Plus 2.0 microarrays
18,960 genes
42 Smoker Controls, 94 COPD subjects
$\mathbf{S} \sim COPD + Gender + Age + Packyears$

- It's difficult to evaluate results... functional enrichment...

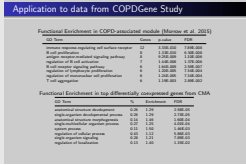
Application to data from COPDGene Study

Functional Enrichment in COPD-associated module (Morrow et al. 2015)

GO Term	Genes	p-value	FDR
immune response-regulating cell surface receptor	12	3.33E-010	7.89E-008
B cell proliferation	8	1.33E-010	6.30E-008
antigen receptor-mediated signaling pathway	8	9.25E-009	1.10E-006
regulation of B cell activation	7	1.44E-008	1.37E-006
B cell receptor signaling pathway	6	1.64E-009	2.59E-007
regulation of lymphocyte proliferation	6	1.20E-005	7.34E-004
regulation of mononuclear cell proliferation	6	1.24E-005	7.34E-004
T cell aggregation	6	1.19E-003	2.89E-002

Functional Enrichment in top differentially coexpressed genes from CMA

GO Term	%	Enrichment	FDR
anatomical structure development	0.26	1.29	2.58E-05
single-organism developmental process	0.26	1.29	2.73E-05
anatomical structure morphogenesis	0.14	1.46	1.60E-04
single-multicellular organism process	0.27	1.25	4.01E-04
system process	0.11	1.50	1.46E-03
regulation of cellular process	0.43	1.12	5.86E-03
single organism signaling	0.28	1.21	7.89E-03
regulation of localization	0.13	1.40	1.15E-02



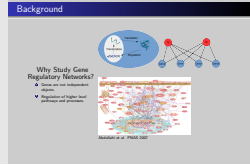
State Transitions Using Gene Regulatory Network Models

Estimating Drivers of Cell State Transitions Using Gene Regulatory Network Models

Daniel Schlauch [1,2], Kimberly Glass [2,3], Craig P. Hersh [2,3,4], Edwin K. Silverman [2,3,4] and John Quackenbush [1,2,3]

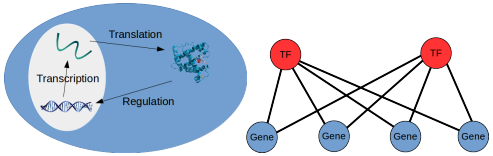
[1] Department of Biostatistics and Computational Biology, Dana-Farber Cancer Institute and Department of Biostatistics, Harvard TH Chan School of Public Health, Boston, MA
[2] Channing Division of Network Medicine, Brigham and Women's Hospital, Boston, MA
[3] Department of Medicine, Harvard Medical School, Boston, MA
[4] Pulmonary and Critical Care Division, Brigham and Women's Hospital and Harvard Medical School, Boston, MA

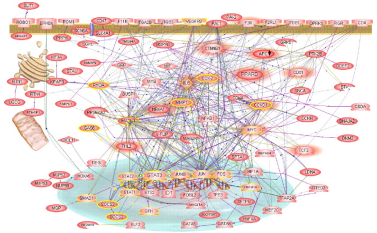
Background



Why Study Gene Regulatory Networks?

- Genes are not independent objects.
- Regulation of higher level pathways and processes.



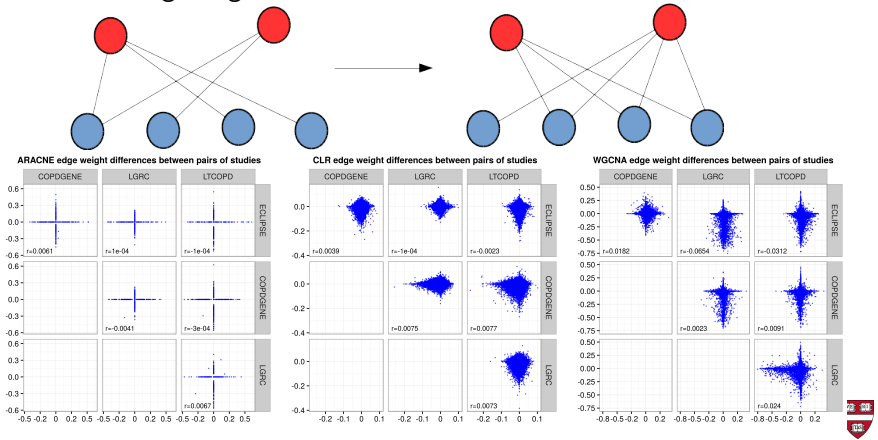


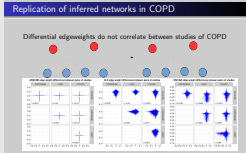
Abdollahi et al. PNAS 2007

- Measurements of gene expression are at the mRNA level.
- Measurements only consist of mRNA abundance.
- Experimental data is collected as static snapshots.
- Biological variability can be difficult to induce
-
- Gene expression measurements are noisy.
- Model complexity may require the estimate of too many model parameters.
- May be computationally intractable.
- May be statistically undetermined. “The curse of dimensionality”

Replication of inferred networks in COPD

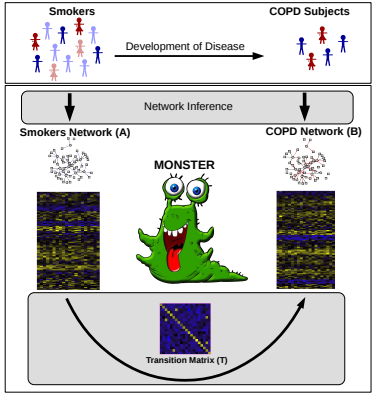
Differential edgeweights do not correlate between studies of COPD

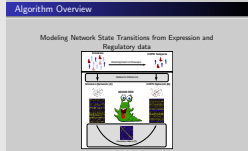




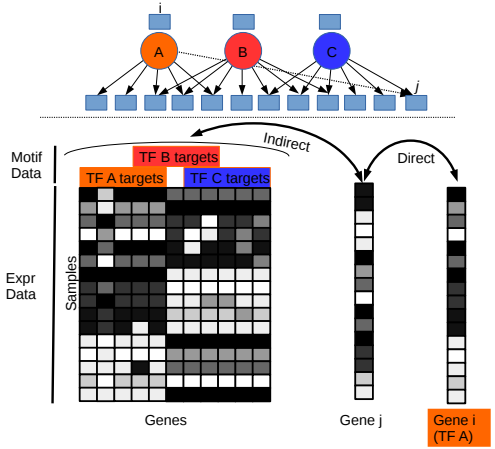
Algorithm Overview

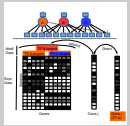
Modeling Network State Transitions from Expression and Regulatory data





Network Inference





Network Inference

Direct Evidence:

$$d_{i,j} = cor\left(g_i, g_j | \left\{g_{k,-i} : k \neq i, k \in \mathbf{TF}\right\}\right)^2$$

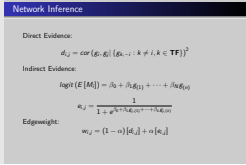
Indirect Evidence:

$$logit\left(E\left[M_i\right]\right) = \beta_0 + \beta_1 g_{(1)} + \cdots + \beta_N g_{(n)}$$

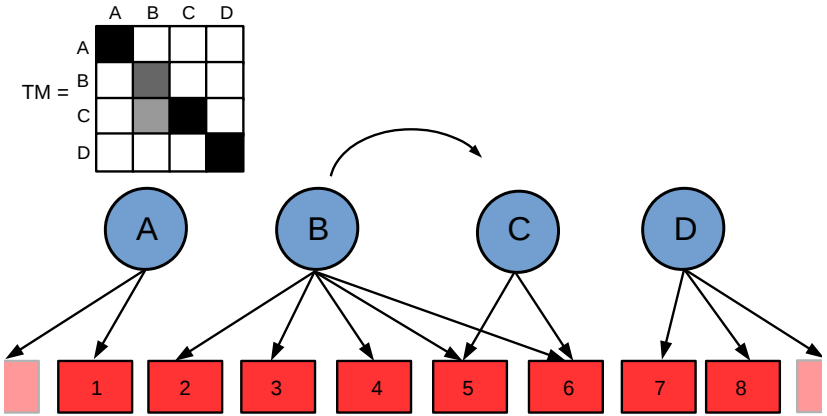
$$e_{i,j} = \frac{1}{1 + e^{\beta_0 + \beta_1 g_{j,(1)} + \cdots + \beta_k g_{j,(k)}}}$$

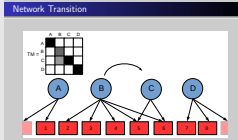
Edgeweight:

$$w_{i,j} = (1 - \alpha)\left[d_{i,j}\right] + \alpha\left[e_{i,j}\right]$$

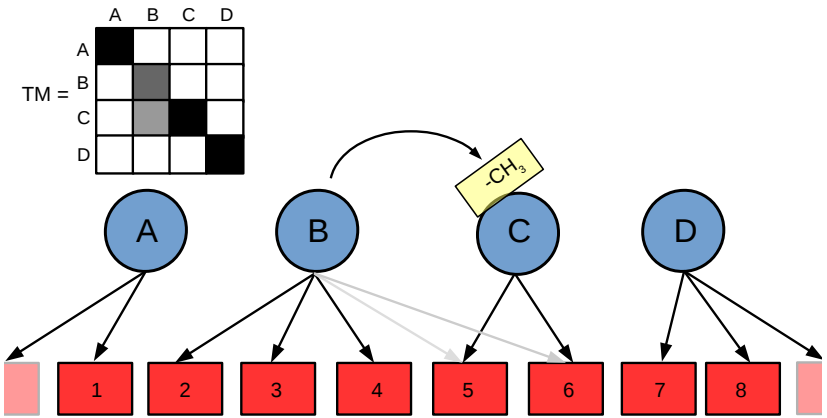


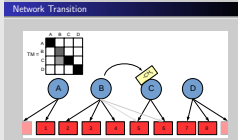
Network Transition





Network Transition





Network Transition

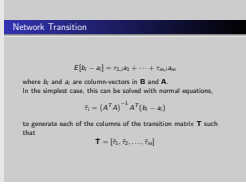
$$E[b_i - a_i] = \tau_{1,i} a_1 + \cdots + \tau_{m,i} a_m$$

where b_i and a_i are column-vectors in $\mathbf{B}$ and $\mathbf{A}$.
In the simplest case, this can be solved with normal equations,

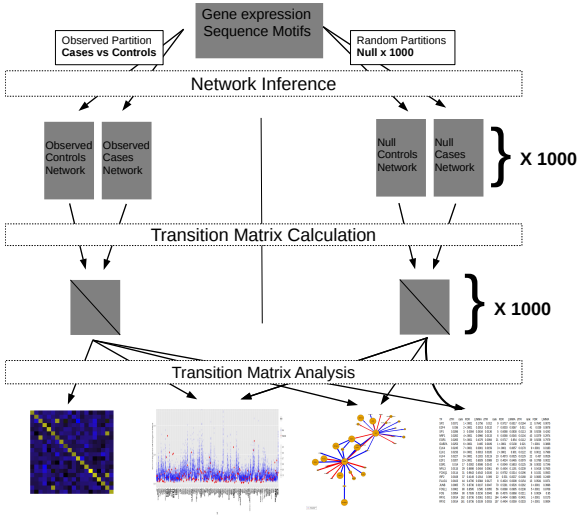
$$\hat{\tau}_i = \left(A^T A\right)^{-1} A^T (b_i - a_i)$$

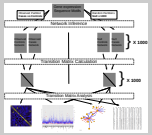
to generate each of the columns of the transition matrix $\mathbf{T}$ such that

$$\hat{\mathbf{T}} = [\hat{\tau}_1, \hat{\tau}_2, \ldots, \hat{\tau}_m]$$



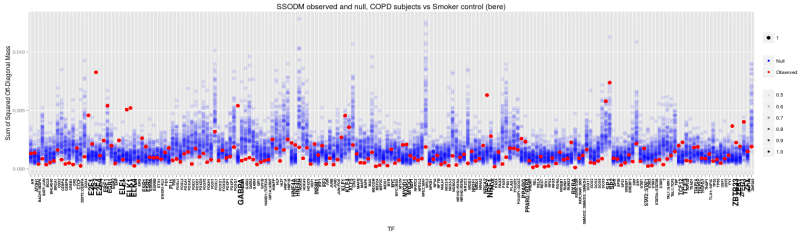
Network Transition

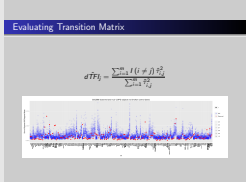




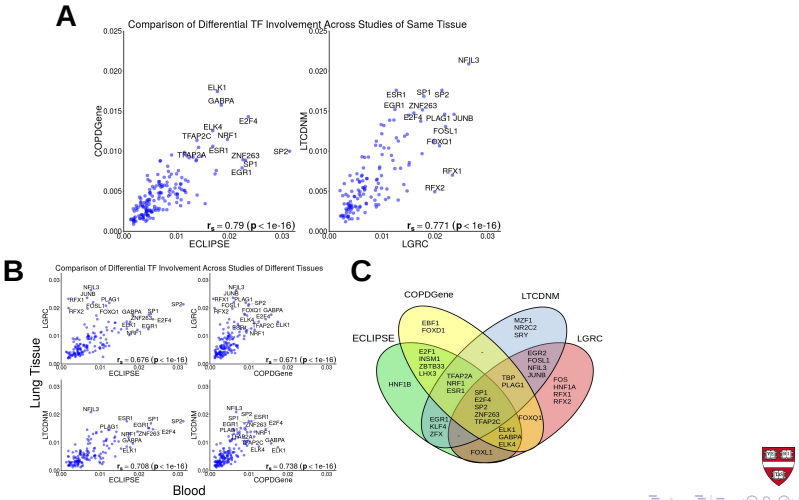
Evaluating Transition Matrix

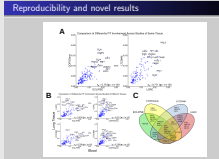
$$d\hat{T}Fl_j = \frac{\sum_{i=1}^{m} I\left(i \neq j\right) \hat{\tau}_{i,j}^2}{\sum_{i=1}^{m} \hat{\tau}_{i,j}^2}$$



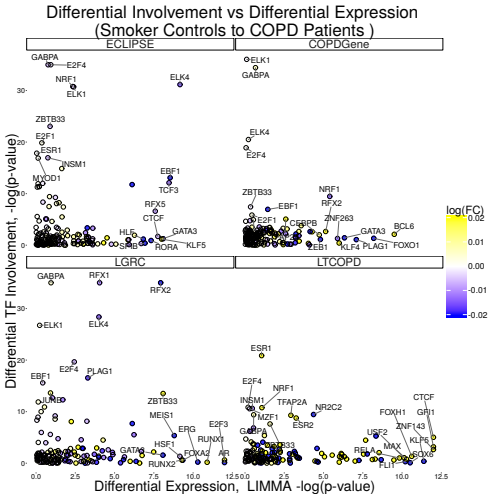


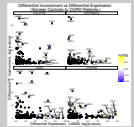
Reproducibility and novel results





Reproducibility and novel results





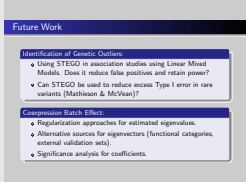
Future Work

Identification of Genetic Outliers:

- Using STEGO in association studies using Linear Mixed Models. Does it reduce false positives and retain power?
- Can STEGO be used to reduce excess Type I error in rare variants (Mathieson & McVean)?

Coexpression Batch Effect:

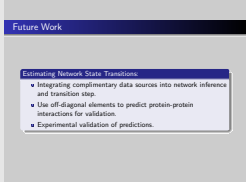
- Regularization approaches for estimated eigenvalues.
- Alternative sources for eigenvectors (functional categories, external validation sets).
- Significance analysis for coefficients.



Future Work

Estimating Network State Transitions:

- Integrating complimentary data sources into network inference and transition step.
- Use off-diagonal elements to predict protein-protein interactions for validation.
- Experimental validation of predictions.



Acknowledgements

Dissertation Committee

- John Quackenbush
- Christoph Lange
- Kimberly Glass

Channing Division of Network Medicine

- Ed Silverman
- Craig Hersh

JQ Lab

- Joe Barry
- Joey Chen
- Maud Fagny
- Marieke Kuijjer
- Camila Lopes-Ramos
- Megha Padi
- Joe Paulson
- John Platig
- Heather Selby
- Abhijeet Sonawane
- Nicole Trotman

Appendix

Batch covariance model

Consider a set of N samples with q covariates measuring gene expression across p genes. Let $\mathbf{x}_i = (x_{i1}, \ldots, x_{iq})$ denote the covariates for sample i and let $\mathbf{g}_i = (g_{i1}, \ldots, g_{ip})^T$ denote the gene expression values for sample i for the p genes.
We can express a model for the gene expression as

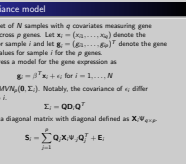
$$\mathbf{g}_i = \beta^T \mathbf{x}_i + \epsilon_i \text{ for } i = 1, \ldots, N$$

where $\epsilon_i \sim MVN_p(\mathbf{0}, \Sigma_i)$. Notably, the covariance of ϵ_i differ according to i.

$$\Sigma_i = \mathbf{Q}\mathbf{D}_i\mathbf{Q}^T$$

where $\mathbf{D}_i$ is a diagonal matrix with diagonal defined as $\mathbf{X}_i\Psi_{q\times p}$.

$$\mathbf{S}_i = \sum_{j=1}^{p} \mathbf{Q}_j\mathbf{X}_i\Psi_{.j}\mathbf{Q}_j^T + \mathbf{E}_i$$



Least Squares Estimator For Batch Correction

To calculate least squares solution, $\hat{\Psi}$, we solve separately for each column of $\mathbf{Q}$.
and note that the residual matrices should be orthogonal to the hyperplane spanned by X^T.
Recall $0 = \mathbf{X^T}\left(\mathbf{Y} - \mathbf{X}\hat{\beta}\right)$

$$\mathbf{0}_q = \sum_{i=1}^{N} \mathbf{X}_i^T \left[\mathbf{Q}_h^T \left[\mathbf{G}_i^* \mathbf{G}_i^{*T} - \mathbf{Q}_h \mathbf{X}_i \hat{\Psi}_h \mathbf{Q}_h^T\right] \mathbf{Q}_h\right]$$

$$\hat{\Psi}_h = \left(\mathbf{X}^T\mathbf{X}\right)^{-1} \sum_{i=1}^{N} \left[\mathbf{X}_i^T \mathbf{Q}_h^T \mathbf{G}_i^* \mathbf{G}_i^{*T} \mathbf{Q}_h\right]$$

Least Squares Estimator For Batch Correction
To calculate least squares solution, $\hat{\Psi}$, we solve separately for each column of $\mathbf{Q}$.
and note that the residual matrices should be orthogonal to the hyperplane spanned by X^T.
Recall $0 = \mathbf{X^T}\left(\mathbf{Y} - \mathbf{X}\hat{\beta}\right)$
$\mathbf{0}_q = \sum_{i=1}^{N} \mathbf{X}_i^T \left[\mathbf{Q}_h^T \left[\mathbf{G}_i^* \mathbf{G}_i^{*T} - \mathbf{Q}_h \mathbf{X}_i \hat{\Psi}_h \mathbf{Q}_h^T\right] \mathbf{Q}_h\right]$
$\hat{\Psi}_h = \left(\mathbf{X}^T\mathbf{X}\right)^{-1} \sum_{i=1}^{N} \left[\mathbf{X}_i^T \mathbf{Q}_h^T \mathbf{G}_i^* \mathbf{G}_i^{*T} \mathbf{Q}_h\right]$

- point 1
- point 2

The Corrected Coexpression Matrix

With the estimates obtained with our method, it is straightforward to see how fitted values for the coexpression matrix for each sample or experimental condition can be obtained. Given an estimate for Ψ, $\hat{\Psi}$, we can now estimate the batch-independent coexpression structure as

$$\hat{\mathbf{S}} = \mathbf{Q} diag\left(\bar{\mathbf{X}}\hat{\Psi}\right)\mathbf{Q}^T \text{ or } \hat{\mathbf{S}} = \sum_{i=1}^{p} \bar{\mathbf{X}}\hat{\Psi}_i \mathbf{Q}_i \mathbf{Q}_i^T$$

The differential coexpression matrix between two conditions, defined in binary as column 2 of $\mathbf{X}$, is computed

$$\hat{\mathbf{W}} = \mathbf{Q} diag\left(\hat{\Psi}_{2,\cdot}\right)\mathbf{Q}^T$$

The Corrected Coexpression Matrix
With the estimates obtained with our method, it is straightforward to see how fitted values for the coexpression matrix for each sample or experimental condition can be obtained. Given an estimate for Ψ, $\hat{\Psi}$, we can now estimate the batch-independent coexpression structure as
$\hat{\mathbf{S}} = \mathbf{Q} diag\left(\bar{\mathbf{X}}\hat{\Psi}\right)\mathbf{Q}^T$ or $\hat{\mathbf{S}} = \sum_{i=1}^{p} \bar{\mathbf{X}}\hat{\Psi}_i \mathbf{Q}_i \mathbf{Q}_i^T$
The differential coexpression matrix between two conditions, defined in binary as column 2 of $\mathbf{X}$, is computed
$\hat{\mathbf{W}} = \mathbf{Q} diag\left(\hat{\Psi}_{2,\cdot}\right)\mathbf{Q}^T$

- point 1
- point 2

Test Statistic

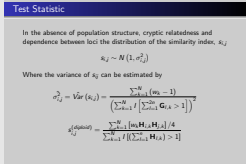
In the absence of population structure, cryptic relatedness and dependence between loci the distribution of the similarity index, $s_{i,j}$

$$s_{i,j} \sim N\left(1, \sigma^2_{i,j}\right)$$

Where the variance of s_{ij} can be estimated by

$$\hat{\sigma^2_{i,j}} = \hat{Var}\left(s_{i,j}\right) = \frac{\sum_{k=1}^{N}\left(w_k - 1\right)}{\left(\sum_{k=1}^{N} I\left[\sum_{l=1}^{2n} \mathbf{G}_{l,k} > 1\right]\right)^2}$$

$$s_{i,j}^{(diploid)} = \frac{\sum_{k=1}^{N}\left[w_k \mathbf{H}_{i,k} \mathbf{H}_{j,k}\right]/4}{\sum_{k=1}^{N} I\left[\left(\sum_{l=1}^{n} \mathbf{H}_{l,k}\right) > 1\right]}$$

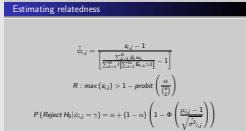


Estimating relatedness

$$\hat{\phi}_{i,j} = \frac{s_{i,j} - 1}{\left[\frac{\sum_{k=1}^{N} \hat{p}_k w_k}{\sum_{k=1}^{N} I\left[\sum_{l=1}^{2n} \mathbf{G}_{l,k} > 1\right]} - 1\right]}$$

$$R : max\left(s_{i,j}\right) > 1 - probit\left(\frac{\alpha}{\binom{n}{2}}\right)$$

$$P\left(Reject\ H_0 | \phi_{i,j} = \gamma\right) = \alpha + (1 - \alpha)\left(1 - \Phi\left(\frac{\mu_{i,j} - 1}{\sqrt{\hat{\sigma}^2{}_{i,j}}}\right)\right)$$



Network Transition Regularization

$$\mathbf{Q}_{i,j} = \begin{cases} 1 & \textit{for } i = j \neq k \\ 0 & \textit{elsewhere} \end{cases},$$

which results in the minimization of the penalized residual sum of squares

$$PRSS\left(\mathbf{T}_{\cdot,\mathbf{k}}\right) = \sum_{i=1}^{p} \left(\mathbf{B}_{i,k} - \sum_{j=1}^{m} A_{i,j} \mathbf{T}_{j,k} \right)^2 + \lambda \sqrt{\mathbf{T}'_{\cdot,k} \mathbf{Q} \mathbf{T}_{\cdot,k}}$$

An implementation of this extension is available in the R package MONSTER.