# Money Moves the Pen

## Link Prediction in Congress Bill Co-Sponsorship Networks Using Political Donor Network Information

*Yi Zhong, Eddie Chen —— CS224W Fall 2018 II Class Project*

## Introduction

Political collaboration is an important part of legislative life, and congress bill cosponsorships provide a rich source of information about the social network between legislators, and serving as a proxy to understand legislators' "connectedness" and collaboration graph. Moreover, according to Mark Twain, "we have the best government that money can buy" - money and politics have already been intertwined. In this project, we applied social network analysis tools on political donation networks and congress bill cosponsorship networks, and framed our research problem as a link prediction task on congress bill cosponsorship networks using political campaign donation records for the US (Congress and Presidential Campaigns) with its network characteristics. We modeled and presented graph characteristics of the two political networks, and showed investigation results of link prediction using various supervised learning techniques for this project. We then compared models' performance to a naive baseline to come up with evaluations.

## Group lasso

One approach is the *group lasso*:

minimize $f(x) + \lambda \sum_{i=1}^{N} \|x_i\|_2$

*i.e.*, like lasso, but require groups of variables to be zero or not

- also called $\ell_{1,2}$ mixed norm regularization

## Structured group lasso

Another approach is the *structured group lasso*:

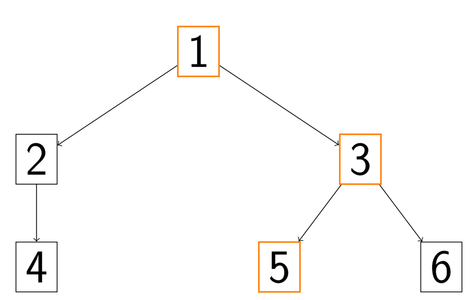
minimize $f(x) + \sum_{i=1}^{N} \lambda_i \|x_{g_i}\|_2$

where $g_i \subseteq [n]$ and $\mathcal{G} = \{g_1, \dots, g_N\}$

- like group lasso, but the groups can overlap arbitrarily
- particular choices of groups can impose 'structured' sparsity
- *e.g.*, topic models, selecting interaction terms for (graphical) models, tree structure of gene networks, fMRI data
- generalizes to the **composite absolute penalties family**:

  $r(x) = \|(\|x_{g_1}\|_{p_1}, \dots, \|x_{g_N}\|_{p_N})\|_{p_0}$

## Hierarchical selection




- $\mathcal{G} = \{\{4\}, \{5\}, \{6\}, \{2,4\}, \{3,5,6\}, \{1,2,3,4,5,6\}\}$
- nonzero variables form a rooted and connected subtree
  - if node is selected, so are its ancestors
  - if node is not selected, neither are its descendants

## Algorithm

We solve this problem using an ADMM lasso implementation:

```
prox_f = @(v,rho) (rho/(1 + rho))*(v - b) + b;
prox_g = @(v,rho) (max(0, v - 1/rho) - max(0, -v - 1/rho));

AA = A*A';
L  = chol(eye(m) + AA);

for iter = 1:MAX_ITER
    xx = prox_g(xz - xt, rho);
    yx = prox_f(yz - yt, rho);

    yz = L \ (L' \ (A*(xx + xt) + AA*(yx + yt)));
    xz = xx + xt + A'*(yx + yt - yz);

    xt = xt + xx - xz;
    yt = yt + yx - yz;
end
```

## Line search

If $L$ is not known (usually the case), can use the following line search:

---

**given** $x^k$, $\lambda^{k-1}$, and parameter $\beta \in (0,1)$.
Let $\lambda := \lambda^{k-1}$.
**repeat**
1. Let $z := \mathbf{prox}_{\lambda g}(x^k - \lambda \nabla f(x^k))$.
2. **break if** $f(z) \le \hat{f}_\lambda(z, x^k)$.
3. Update $\lambda := \beta\lambda$.

**return** $\lambda^k := \lambda$, $x^{k+1} := z$.

---

typical value of $\beta$ is $1/2$, and

$$\hat{f}_\lambda(x,y) = f(y) + \nabla f(y)^T (x-y) + (1/2\lambda)\|x-y\|_2^2$$

## Convergence proof

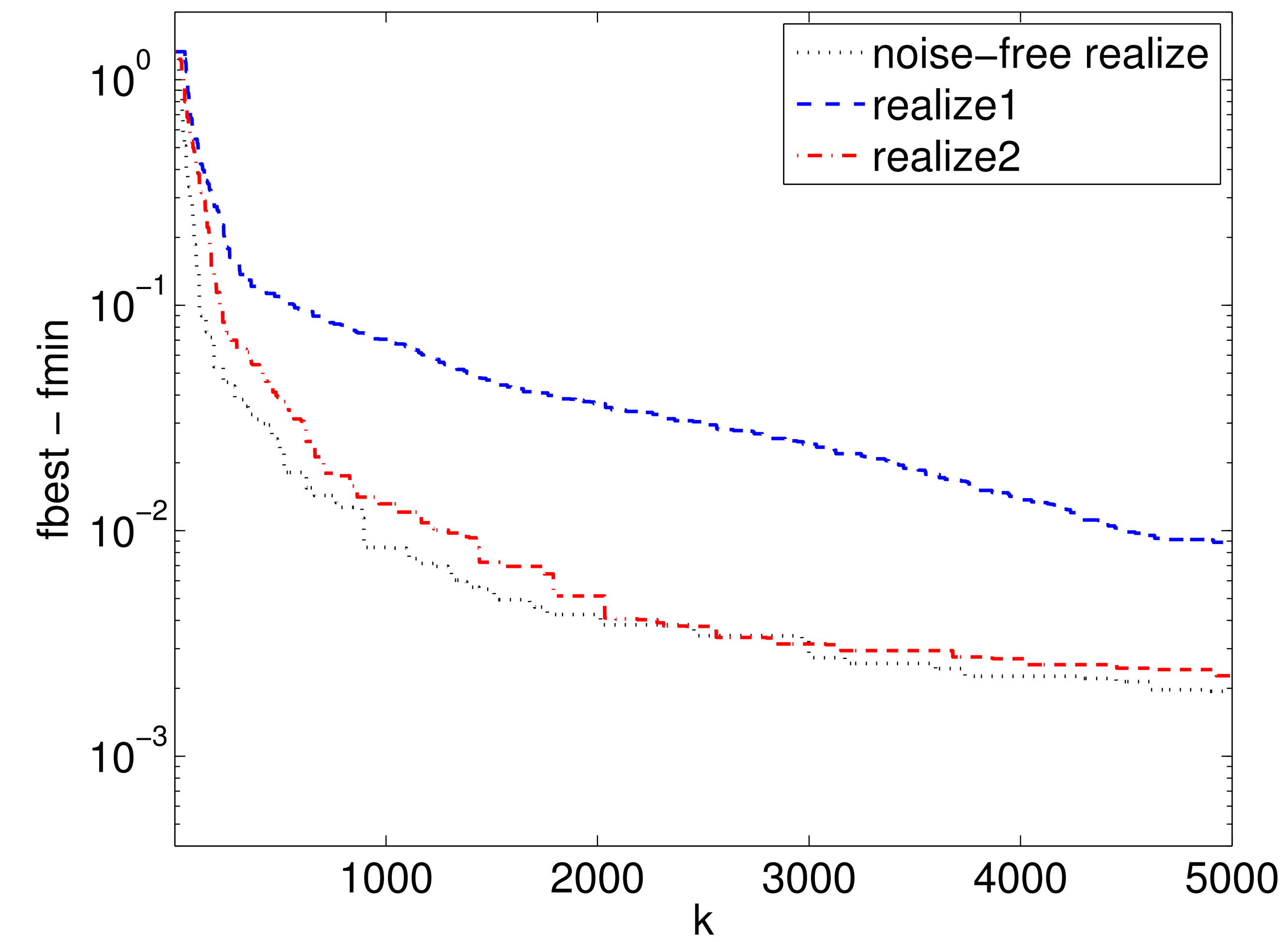
Lorem ipsum dolor sit amet, consectetur adipisicing elit, sed do eiusmod tempor incididunt ut labore et dolore magna aliqua. Ut enim ad minim veniam, quis nostrud exercitation ullamco laboris nisi ut aliquip ex ea commodo consequat. Duis aute irure dolor in reprehenderit in voluptate velit esse cillum dolore eu fugiat nulla pariatur. Excepteur sint occaecat cupidatat non proident, sunt in culpa qui officia deserunt mollit anim id est laborum.

## Numerical example

Consider a numerical example with $f(x) = \|Ax - b\|_2^2$ with $A \in \mathbf{R}^{10 \times 100}$ and $b \in \mathbf{R}^{10}$. Entries of $A$ and $b$ are generated as independent samples from a standard normal distribution. Here, we have chosen $\lambda$ using cross validation.

## Results

On this numerical example, the ADMM method converges quickly. We give two realizations corresponding to different parameters $A$ and $b$.




## Conclusion

Lorem ipsum dolor sit amet, consectetur adipisicing elit, sed do eiusmod tempor incididunt ut labore et dolore magna aliqua. Ut enim ad minim veniam, quis nostrud exercitation ullamco laboris nisi ut aliquip ex ea commodo consequat. Duis aute irure dolor in reprehenderit in voluptate velit esse cillum dolore eu fugiat nulla pariatur. Excepteur sint occaecat cupidatat non proident, sunt in culpa qui officia deserunt mollit anim id est laborum.

## Acknowledgements

This material is based upon work supported by the X Fellowship.