A Regularized Difference-of-coefficient Approach in Mediation Analysis

Yu-Bo Wang, Ph.D.

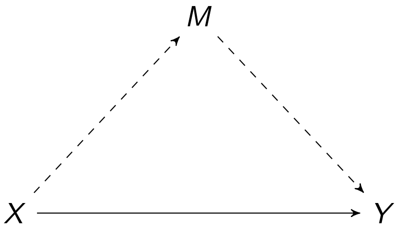
Biostatistics and Bioinformatics Branch
Division of Intramural Population Health Research
NIH/NICHD

May 12, 2018

Motivating Example

- In the Longitudinal Investigation of Fertility and the Environment (LIFE) Study, a mixture of endocrine-disrupting chemicals were assessed in relation to a spectrum of reproductive outcomes. (Louis et al. (2011))
- A scientific question: whether the effects of chemicals on pregnancy outcome are through and around mediators.

Naive Approach



- Only reveal the marginal direct and indirect effect
- Inflated type I error

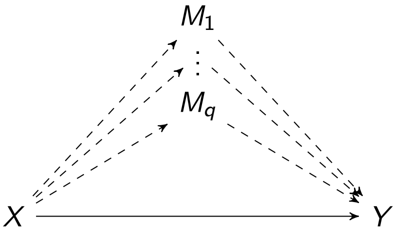
Mediation Analysis with Regularization

- Goal: Provide a method that can simultaneous handle all exposures and mediators.
- Problem: collinearity and violate model parsimony.
- Solution: Regularization
  - Simultaneous variable selection and estimation
  - Model Parsimony

- Baron-Kenny model
- Bayesian Least Absolute Shrinkage and Selection Operator (Park and Casella, 2008)

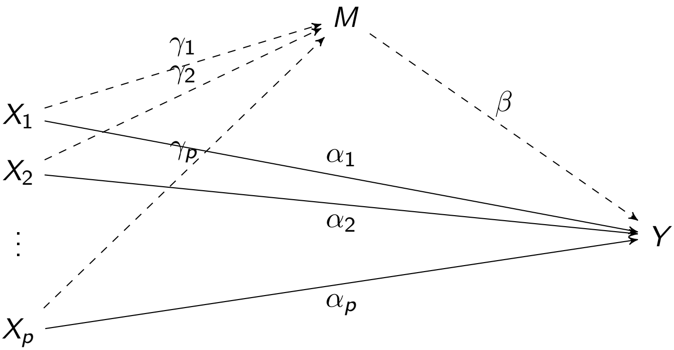
Current Regularized Approach

- An exposure and q mediators



- This structure is common in a clinical trial, but not for LIFE study, which is involved in multiple exposures and mediators.

Structure in An Environmental Study



Baron and Kenny Model

- $\boldsymbol{y} = (y_1, y_2, \ldots, y_n)'$ is a continuous outcome vector of n subjects.
- $\boldsymbol{X} = (x_{ij}) = (\boldsymbol{x}_1, \boldsymbol{x}_2, \ldots, \boldsymbol{x}_n)' = (\boldsymbol{x}_{(1)}, \boldsymbol{x}_{(2)}, \ldots, \boldsymbol{x}_{(p)})$ is a $n \times p$ predictor/exposure matrix.
- $\boldsymbol{m} = (m_1, m_2, \ldots, m_n)'$ is a mediator vector.
- Baron and Kenny model:

$$\begin{aligned} y_i &= \mu_1 + \boldsymbol{x}_i'\boldsymbol{\alpha} + \beta m_i + \varepsilon_{1i}, \\ m_i &= \mu_2 + \boldsymbol{x}_i'\gamma + \varepsilon_{2i}, \end{aligned} \tag{1}$$

  where $\varepsilon_{1i} \overset{iid}{\sim} N(0, \sigma_1^2)$, $\varepsilon_{2i} \overset{iid}{\sim} N(0, \sigma_2^2)$, and $\varepsilon_{1i} \perp \varepsilon_{2i}$.
- $\boldsymbol{\alpha}$: direct effects
- $\beta\boldsymbol{\gamma}$: indirect effects.

Least Absolute Shrinkage and Selection Operator

- Lasso of Tibshirani (1996): $L1$ penalty on $\boldsymbol{\alpha}$ so that

$$\min_{\boldsymbol{\alpha}}(\boldsymbol{y} - \mu_1 \boldsymbol{J}_n - \boldsymbol{X}\boldsymbol{\alpha})'(\boldsymbol{y} - \mu_1 \boldsymbol{J}_n - \boldsymbol{X}\boldsymbol{\alpha}) + \lambda \sum_{j=1}^{p} |\alpha_j| \qquad (2)$$

- Bayesian Lasso of Park and Casella (2008)
  - Conditional Laplace prior on $\boldsymbol{\alpha}$.
  - Posterior distribution of $\boldsymbol{\alpha}$ is proportional to

$$\exp\left[-\frac{1}{2\sigma_1^2}\left(\boldsymbol{y} - \mu_1 \boldsymbol{J}_n - \boldsymbol{X}\boldsymbol{\alpha}\right)'\left(\boldsymbol{y} - \mu_1 \boldsymbol{J}_n - \boldsymbol{X}\boldsymbol{\alpha}\right) - \frac{\lambda}{\sigma_1}\sum_{j=1}^{p}|\alpha_j|\right] \qquad (3)$$

  - The solution to (2) is the posterior mode of (3).

Regularized Two-Staged (regTS) Approach

- Conditional Laplace priors on $\boldsymbol{\alpha}$ and $\boldsymbol{\gamma}$.
- Unimodality assured for the posterior distribution of $\boldsymbol{\alpha}$ and $\boldsymbol{\gamma}$
- Problems:
  - Unimodality is not guaranteed for $\beta\boldsymbol{\gamma}$.

		β			
		1	2	5	
γ	1	0.09	0.18	0.03	0.30
	2	0.18	0.36	0.06	0.60
	3	0.03	0.06	0.01	0.10
		0.30	0.60	0.10	

  - Lack of unimodality: slow convergence of the Gibbs sampling, and less meaningful estimates.
  - γ itself is less meaningful in mediation analysis.

Regularized Product-of-coefficient Approach

- Regularization on $\boldsymbol{\alpha}$ and $\beta\boldsymbol{\gamma}$ instead: unimodality assured for the posterior distribution of $\boldsymbol{\alpha}$ and $\boldsymbol{\gamma}$, and for $\boldsymbol{\alpha}$ and $\beta\boldsymbol{\gamma}$.
- Over penalization on β:
  - Suppose a conditional normal prior $N(\mathbf{0}, \sigma^2 I_p)$ is assigned to $\beta\boldsymbol{\gamma}$.
  - It can be viewed as a conditional prior $N(0, \sigma^2/\boldsymbol{\gamma}'\boldsymbol{\gamma})$ proposed for β.
  - When p increases or γ become large, the prior for β degenerates to zero so that the posterior distribution of β.

Regularized Difference-of-coefficient (regDOC) Approach

- Alternative version of (1) (MacKinnon and Dwyer (1993)):

$$y_i = \mu_1 + \boldsymbol{x}_i'\boldsymbol{\alpha} + \beta m_i + \varepsilon_{1i},$$
$$y_i = \mu_3 + \boldsymbol{x}_i'\boldsymbol{\alpha}^* + \varepsilon_{3i},$$

  where $\varepsilon_{3i} \overset{iid}{\sim} N(0, \sigma_3^2)$.
- $\boldsymbol{\alpha}^* - \boldsymbol{\alpha} = \beta\boldsymbol{\gamma}$ and $\sigma_1^2 < \sigma_3^2$.
- Let $\boldsymbol{\delta} = \beta\boldsymbol{\gamma}$ and $\sigma_3^2 = \eta\sigma_1^2$ where $\eta \geq 1$.
- Regularization on $\boldsymbol{\alpha}$ and $\boldsymbol{\delta}$.

Complete Model Specification

$$\alpha_j|\sigma_1^2 \overset{iid}{\sim} \text{Laplace}\left(0, \frac{\sqrt{\sigma_1^2}}{\lambda_1}\right)$$

$$\delta_j|\sigma_1^2, \eta \overset{iid}{\sim} \text{Laplace}\left(0, \frac{\sqrt{\eta\sigma_1^2}}{\lambda_2}\right)$$

$$\sigma_1^2, \eta \sim \frac{1}{\sigma_1^2} I\left\{\eta \geq 1\right\}$$

$$\lambda_1^2 \sim \text{Gamma}(\theta_1, \nu_1)$$

$$\lambda_2^2 \sim \text{Gamma}(\theta_2, \nu_2)$$

$$\mu_1 \sim N(0, \sigma_0^2)$$

$$\mu_3 \sim N(0, \sigma_0^2)$$

$$\beta|\sigma_1^2 \sim N(0, \sigma_0^2\sigma_1^2)$$

Unimodality

- Let $\boldsymbol{\alpha}^\star = \boldsymbol{\alpha}/\sqrt{\sigma_1^2}$, $\boldsymbol{\beta}^\star = \boldsymbol{\beta}/\sqrt{\sigma_1^2}$, $\boldsymbol{\delta}^\star = \boldsymbol{\delta}/\sqrt{\sigma_1^2}$, $\rho = 1/\sqrt{\sigma_1^2}$, and $\phi = 1/\sqrt{\eta}$.
- $\pi(\boldsymbol{\alpha}, \beta, \boldsymbol{\delta}, \sigma_1^2, \eta | D)$ is proportional to

$$\begin{aligned}
\rho^{2n+2p+3}\phi^{n+p} I\{\phi \leq 1\} \exp\Big[&-\frac{1}{2}(\rho\tilde{\boldsymbol{y}}_1 - \boldsymbol{X}\boldsymbol{\alpha}^\star - \beta^\star \boldsymbol{m})'(\rho\tilde{\boldsymbol{y}}_1 - \boldsymbol{X}\boldsymbol{\alpha}^\star - \beta^\star \boldsymbol{m}) \\
&- \frac{\phi^2}{2}(\rho\tilde{\boldsymbol{y}}_2 - \boldsymbol{X}(\boldsymbol{\alpha}^\star + \boldsymbol{\delta}^\star))'(\rho\tilde{\boldsymbol{y}}_2 - \boldsymbol{X}(\boldsymbol{\alpha}^\star + \boldsymbol{\delta}^\star)) \\
&- \lambda_1 \sum_{j=1}^{p} |\alpha_j^\star| - \frac{1}{2\sigma_0^2}(\beta^\star)^2 - \lambda_2\phi^2 \sum_{j=1}^{p} |\delta_j^\star| \Big],
\end{aligned}$$

where $\tilde{\boldsymbol{y}}_1 = \boldsymbol{y} - \mu_1 \boldsymbol{J}_n$ and $\tilde{\boldsymbol{y}}_2 = \boldsymbol{y} - \mu_3 \boldsymbol{J}_n$.

Unimodality

The log posterior is

$$
\begin{aligned}
&\ln c + (2n + 2p + 3) \ln \rho + (n + p) \ln \phi + \ln(I\{\phi \leq 1\}) \\
&- \frac{1}{2}(\rho \tilde{\boldsymbol{y}}_1 - \boldsymbol{X}\boldsymbol{\alpha}^\star - \beta^\star \boldsymbol{m})'(\rho \tilde{\boldsymbol{y}}_1 - \boldsymbol{X}\boldsymbol{\alpha}^\star - \beta^\star \boldsymbol{m}) \\
&- \frac{\phi^2}{2}(\rho \tilde{\boldsymbol{y}}_2 - \boldsymbol{X}(\boldsymbol{\alpha}^\star + \boldsymbol{\delta}^\star))'(\rho \tilde{\boldsymbol{y}}_2 - \boldsymbol{X}(\boldsymbol{\alpha}^\star + \boldsymbol{\delta}^\star)) - \lambda_1 \sum_{j=1}^{p} |\alpha_j^\star| - \frac{1}{2\sigma_0^2}(\beta^\star)^2 - \lambda_2 \phi^2 \sum_{j=1}^{p} |\delta_j^\star|.
\end{aligned}
$$

Gibbs Sampling

- The identity in Andrews and Mallows (1974):

$$\pi(\alpha_j|\sigma_1^2) = \int_0^\infty \frac{1}{\sqrt{2\pi\sigma_1^2\tau_{j\mathrm{A}}}} \exp\left[-\frac{\alpha_j^2}{2\sigma_1^2\tau_{j\mathrm{A}}}\right] \frac{\lambda_1^2}{2} \exp\left[-\frac{\lambda_1^2\tau_{j\mathrm{A}}}{2}\right] d\tau_{j\mathrm{A}}$$

and

$$\pi(\delta_j|\sigma_1^2,\eta) = \int_0^\infty \frac{1}{\sqrt{2\pi\eta\sigma_1^2\tau_{j\mathrm{D}}}} \exp\left[-\frac{\delta_j^2}{2\eta\sigma_1^2\tau_{j\mathrm{D}}}\right] \frac{\lambda_2^2}{2} \exp\left[-\frac{\lambda_2^2\tau_{j\mathrm{D}}}{2}\right] d\tau_{j\mathrm{D}}.$$

where $(\tau_{1\mathrm{A}}, \tau_{2\mathrm{A}}, \ldots, \tau_{p\mathrm{A}})$ and $(\tau_{1\mathrm{D}}, \tau_{2\mathrm{D}}, \ldots, \tau_{p\mathrm{D}})$ are non-negative latent variables.

- All full conditional distributions are analytically tractable: Gibbs sampling.

Extended Approach to the Case with Multiple Mediators

- $\boldsymbol{m}_i = (m_{i1}, m_{i2}, \ldots, m_{iq})'$ is the mediator vector of i^{th} subject.
- $$y_i = \mu_1 + \boldsymbol{x}_i'\boldsymbol{\alpha} + \boldsymbol{m}_i'\boldsymbol{\beta} + \varepsilon_{1i},$$
  $$y_i = \mu_3 + \boldsymbol{x}_i'\boldsymbol{\alpha}^* + \varepsilon_{3i},$$
- $\boldsymbol{\alpha}^* = \boldsymbol{\alpha} + \boldsymbol{\delta}$, and $\boldsymbol{\delta} = (\delta_1, \delta_2, \ldots, \delta_p)'$ denotes the vector of total indirect effect of each exposure through all mediators.
- Regularization on each total indirect effect.
- All full conditional distributions are analytically tractable: Gibbs sampling.

Application to the LIFE Study

- Outcome: time to pregnancy
- Exposures: a male partner's serum concentrations of 36 PCB congeners
- Mediator: paternal total sperm count
- Models considered: B-K, regTS, and regDOC.
- An MCMC sample of 15,000 iterations with a 5,000 burn-ins

Direct Effects - Part I

PCB Congener	regDOC			B-K			regTS		
	Mean	95% HPD Int.		Mean	95% HPD Int.		Mean	95% HPD Int.	
		Lower	Upper		Lower	Upper		Lower	Upper
#28	-0.005	-0.306	0.287	-0.845	-3.653	1.844	-0.001	-0.311	0.298
#44	0.015	-0.251	0.335	-0.157	-2.001	1.710	0.007	-0.304	0.304
#49	0.028	-0.241	0.331	0.766	-1.257	2.744	0.012	-0.279	0.325
#52	0.004	-0.294	0.291	0.314	-1.595	2.211	-0.004	-0.294	0.320
#66	-0.008	-0.307	0.263	0.262	-2.047	2.472	-0.010	-0.327	0.284
#74	-0.073	-0.365	0.181	-0.359	-1.286	0.541	-0.049	-0.352	0.215
#87	0.016	-0.136	0.179	0.058	-0.245	0.366	0.008	-0.168	0.195
#99	-0.105	-0.364	0.112	-0.350	-0.914	0.268	-0.055	-0.307	0.175
#101	-0.034	-0.233	0.145	-0.058	-0.531	0.427	-0.028	-0.253	0.177
#105	0.039	-0.164	0.257	0.306	-0.448	0.981	0.025	-0.195	0.265
#110	-0.016	-0.152	0.123	-0.117	-0.415	0.193	-0.014	-0.168	0.151
#114	-0.035	-0.197	0.111	-0.041	-0.327	0.255	-0.033	-0.224	0.138
#118	0.024	-0.190	0.270	0.003	-0.760	0.816	0.013	-0.222	0.259
#128	-0.061	-0.235	0.084	-0.221	-0.525	0.066	-0.052	-0.244	0.124
#138	0.205	-0.061	0.538	0.929	0.269	1.627	0.117	-0.141	0.440
#146	-0.028	-0.234	0.168	-0.023	-0.491	0.428	-0.023	-0.267	0.202
#149	0.027	-0.209	0.265	0.272	-0.828	1.327	0.008	-0.242	0.267
#151	0.002	-0.249	0.230	0.210	-0.896	1.330	-0.007	-0.260	0.249

Direct Effects - Part II

PCB Congener	regDOC			B-K			regTS		
	Mean	95% HPD Int.		Mean	95% HPD Int.		Mean	95% HPD Int.	
		Lower	Upper		Lower	Upper		Lower	Upper
#153	-0.040	-0.336	0.221	-0.638	-1.549	0.266	-0.015	-0.327	0.264
#156	-0.018	-0.214	0.182	-0.110	-0.577	0.383	-0.010	-0.242	0.200
#157	-0.031	-0.182	0.110	0.009	-0.248	0.291	-0.032	-0.201	0.137
#167	-0.006	-0.128	0.107	-0.023	-0.206	0.162	-0.004	-0.143	0.126
#170	0.016	-0.218	0.270	-0.032	-0.888	0.911	0.018	-0.252	0.278
#172	-0.028	-0.207	0.123	-0.051	-0.378	0.276	-0.015	-0.213	0.171
#177	-0.114	-0.437	0.134	-0.733	-1.560	0.083	-0.068	-0.397	0.192
#178	-0.028	-0.215	0.165	-0.059	-0.473	0.332	-0.025	-0.250	0.183
#180	0.071	-0.183	0.378	0.589	-0.582	1.832	0.041	-0.238	0.340
#183	-0.031	-0.271	0.177	-0.171	-0.733	0.389	-0.012	-0.268	0.235
#187	0.083	-0.145	0.377	0.491	-0.348	1.307	0.054	-0.206	0.362
#189	-0.090	-0.216	0.038	-0.121	-0.295	0.043	-0.085	-0.232	0.051
#194	-0.027	-0.221	0.177	-0.222	-0.743	0.294	-0.022	-0.262	0.195
#195	0.046	-0.101	0.207	0.040	-0.249	0.324	0.048	-0.122	0.235
#196	-0.024	-0.280	0.204	-0.195	-0.947	0.631	-0.015	-0.281	0.252
#201	-0.005	-0.223	0.205	-0.006	-0.712	0.753	-0.004	-0.238	0.230
#206	0.005	-0.178	0.182	0.097	-0.452	0.623	0.004	-0.197	0.208
#209	0.077	-0.053	0.227	0.103	-0.129	0.330	0.069	-0.090	0.229

Indirect Effects - Part I

PCB Congener	regDOC			B-K			regTS		
	Mean	95% HPD Int.		Mean	95% HPD Int.		Mean	95% HPD Int.	
		Lower	Upper		Lower	Upper		Lower	Upper
#28	0.002	-0.308	0.321	-0.021	-0.285	0.228	0.001	-0.032	0.034
#44	0.011	-0.304	0.319	-0.098	-0.362	0.117	-0.002	-0.036	0.029
#49	0.018	-0.276	0.349	0.068	-0.120	0.302	0.002	-0.028	0.038
#52	0.008	-0.306	0.313	0.040	-0.129	0.254	0.001	-0.029	0.037
#66	-0.001	-0.336	0.290	0.027	-0.183	0.263	0.001	-0.031	0.035
#74	-0.034	-0.334	0.252	-0.010	-0.105	0.073	-0.002	-0.036	0.025
#87	0.018	-0.174	0.220	0.015	-0.018	0.058	0.009	-0.010	0.039
#99	-0.054	-0.338	0.180	-0.046	-0.157	0.042	-0.020	-0.069	0.016
#101	-0.009	-0.241	0.230	-0.001	-0.050	0.041	-0.000	-0.024	0.023
#105	0.025	-0.214	0.284	-0.006	-0.076	0.058	-0.001	-0.028	0.025
#110	-0.008	-0.188	0.166	-0.009	-0.045	0.017	-0.001	-0.021	0.016
#114	-0.007	-0.206	0.196	0.006	-0.021	0.038	0.005	-0.013	0.029
#118	0.011	-0.237	0.270	0.019	-0.056	0.103	-0.001	-0.028	0.029
#128	-0.019	-0.218	0.180	0.004	-0.022	0.034	0.003	-0.013	0.025
#138	0.072	-0.207	0.373	0.018	-0.038	0.096	0.002	-0.025	0.030
#146	-0.014	-0.266	0.220	-0.003	-0.049	0.038	-0.002	-0.029	0.020
#149	0.028	-0.230	0.298	0.034	-0.065	0.161	0.003	-0.022	0.033
#151	0.010	-0.266	0.271	-0.039	-0.172	0.069	-0.002	-0.031	0.025

Indirect Effects - Part II

PCB Congener	regDOC			B-K			regTS		
	Mean	95% HPD Int.		Mean	95% HPD Int.		Mean	95% HPD Int.	
		Lower	Upper		Lower	Upper		Lower	Upper
#153	-0.026	-0.348	0.245	-0.002	-0.090	0.084	-0.000	-0.030	0.031
#156	-0.003	-0.243	0.230	0.002	-0.044	0.050	0.002	-0.020	0.028
#157	-0.007	-0.191	0.182	0.000	-0.026	0.025	0.001	-0.016	0.018
#167	-0.000	-0.156	0.160	-0.000	-0.017	0.017	0.000	-0.015	0.015
#170	0.006	-0.270	0.294	-0.008	-0.098	0.072	-0.003	-0.034	0.025
#172	-0.013	-0.235	0.204	-0.003	-0.035	0.027	-0.005	-0.031	0.013
#177	-0.042	-0.361	0.237	0.014	-0.053	0.101	0.005	-0.022	0.040
#178	-0.013	-0.245	0.216	-0.004	-0.047	0.033	-0.003	-0.028	0.021
#180	0.025	-0.275	0.326	-0.004	-0.121	0.105	-0.001	-0.034	0.028
#183	-0.015	-0.285	0.235	-0.015	-0.077	0.035	-0.002	-0.030	0.022
#187	0.038	-0.237	0.343	0.015	-0.055	0.105	0.003	-0.026	0.034
#189	-0.019	-0.197	0.129	0.001	-0.015	0.017	0.001	-0.015	0.015
#194	-0.015	-0.256	0.219	0.018	-0.029	0.081	0.003	-0.020	0.031
#195	0.021	-0.169	0.224	0.001	-0.024	0.030	0.004	-0.013	0.026
#196	-0.018	-0.311	0.250	0.014	-0.059	0.098	0.000	-0.029	0.029
#201	-0.006	-0.258	0.240	-0.045	-0.166	0.044	-0.008	-0.041	0.016
#206	-0.004	-0.215	0.210	0.016	-0.033	0.079	0.001	-0.021	0.024
#209	0.016	-0.161	0.193	-0.009	-0.039	0.013	-0.006	-0.028	0.009

Future work

- Identify inter-generational effects.
  - Childhood Obesity
  - Maternal dietary intake
  - Child dietary intake

Happy Birthday, Professor Kuo!