Network interference estimators

Vanessa McNealis

2024-01-18

Load dependencies and source script

```
library(boot)
library(tidyverse)
library(igraph)
library(data.table)
library(pbapply)
library(lme4)
library(statnet)
library(intergraph)
library(statmod)
library(fastGHQuad)
library(geex)
library(here)
source(url("https://github.com/vanessamcnealis/Network-Interference-estimators/blob/main/Estimating_fun
```

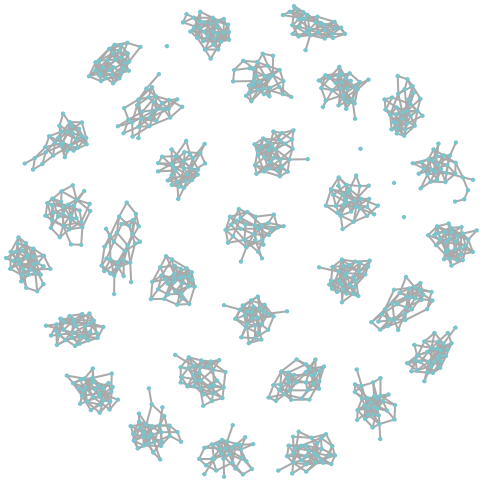
Data preparation

The data at hand `dat` consists of one simulated dataset used for the simulation studies presented in the paper. The data generating scheme for the assigned treatment Z, the outcome Y, and the covariates X_1, X_2 and X_3 is described in the manuscript.

```
dat  <- read.csv("https://github.com/vanessamcnealis/Network-Interference-estimators/blob/main/dat.csv?
load(url("https://github.com/vanessamcnealis/Network-Interference-estimators/blob/main/network.RData?ra
```

Network visualization

```
plot(network, vertex.size=1, vertex.label=NA, vertex.color="cadetblue3",
     vertex.frame.color="cadetblue3")
```



Prepare the data for modeling

```
# Total number of nodes
N <- nrow(dat)

#Number of components
m <- length(unique(dat$component.id))

# Adjacency matrix associated with the graph
A <- get.adjacency(network)

# Create a variable indicating the degree of a node
dat$degree <- as.numeric(A %*% rep(1, N))

# Create a variable indicating the number of treated neighbours for each node
dat$k_treated <-
    as.numeric(A %*% matrix(dat$Z, ncol = 1, nrow = nrow(dat)))

# Create a variable indicated the proportion of treated neighbours for each node
dat$prop_treated <- ifelse(dat$degree == 0, 0, dat$k_treated/dat$degree)

# Change this vector for different alpha values
alpha <- c(0.2, 0.5, 0.8)
```

Modeling

In this section, we will apply the IPW, REG, DR-BC, and DR-WLS estimators for the average potential outcomes and associated causal contrasts. The average potential outcome for a unit depends on an individual exposure/treatment z and treatment coverage $\alpha \in (0, 1)$, where the treatment coverage represents the counterfactual probability that first-order neighbours receive the exposure/treatment. For each estimation method, the R programs will output the **point estimate and the estimated variance for average potential outcomes** $\mu_{0\alpha}, \mu_{1\alpha}, \mu_{\alpha}$ under three different treatment coverages (i.e., $\alpha \in \{0.2, 0.5, 0.8\}$), respectively. Also, the R program will output the point estimation and estimated variance of four causal effects: Direct (DE), Indirect (IE), Total (TE), and Overall effect (OE).

Equations for calculating the four causal effect estimates are:

- $\hat{DE}(\alpha) = \hat{\mu}_{1\alpha} - \hat{\mu}_{0\alpha}$
- $\hat{IE}(\alpha_0, \alpha_1) = \hat{\mu}_{0\alpha_1} - \hat{\mu}_{0\alpha_0}$
- $\hat{TE}(\alpha_0, \alpha_1) = \hat{\mu}_{1\alpha_1} - \hat{\mu}_{0\alpha_0}$
- $\hat{OE}(\alpha_0, \alpha_1) = \hat{\mu}_{1\alpha_1} - \hat{\mu}_{0\alpha_0}$,

where α_0 and α_1 represent distinct treatment coverages with $\alpha_0 \neq \alpha_1$. If the vector supplied is $\alpha = (0.2, 0.5, 0.8)$, the program will treat values in sequential order and compare

- $\alpha_1 = 0.5$ vs. $\alpha_0 = 0.2$
- $\alpha_1 = 0.8$ vs. $\alpha_0 = 0.2$
- $\alpha_1 = 0.8$ vs. $\alpha_0 = 0.5$.

We first specify the models for the outcome and treatment propensity (note that these correspond to the oracle models). In the model for the treatment propensity, we include a random effect term for the component/subgraph. Note that for DR-WLS, the outcome formula supplied should not include terms for the treatment assignment as this method stratifies the sample based on the treatment indicator.

```
propensity_formula <- Z ~ abs(X1) + I(X2*abs(X1)) +X3+ (1|component.id)
outcome_formula <- Y ~ Z + prop_treated + prop_treated*Z + abs(X1) + X2 +
  abs(X1)*X2
outcome_formula_drwls <- Y ~ prop_treated + abs(X1) + X2 + abs(X1)*X2
```

Inverse probability-of-treatment weighting (IPW) estimates

```
ipw_estimates <- ipw_estimation(data=dat,
                                propensity_formula = propensity_formula,
                                alpha = alpha)
ipw_estimates
```

```
## [[1]]
##                              estimand alpha  estimate    variance     CI_lb
## 1      Average potential outcome z=0   0.2 0.7890423 0.012590455 0.5691202
## 2      Average potential outcome z=1   0.2 3.6660878 0.080109420 3.1113473
## 3 Marginal average potential outcome   0.2 1.3644514 0.011342743 1.1557107
## 4      Average potential outcome z=0   0.5 1.0958977 0.008673339 0.9133647
## 5      Average potential outcome z=1   0.5 3.7855197 0.032398978 3.4327317
## 6 Marginal average potential outcome   0.5 2.4407087 0.010583058 2.2390794
## 7      Average potential outcome z=0   0.8 1.3498347 0.050227980 0.9105754
## 8      Average potential outcome z=1   0.8 4.8279392 0.090306441 4.2389498
## 9 Marginal average potential outcome   0.8 4.1323183 0.058075838 3.6599880
##      CI_ub
## 1 1.008964
## 2 4.220828
## 3 1.573192
## 4 1.278431
## 5 4.138308
## 6 2.642338
```

```
## 7 1.789094
## 8 5.416929
## 9 4.604649
##
## [[2]]
##           estimand alpha0 alpha1  estimate    variance       CI_lb     CI_ub
## 1    Direct effect    0.2    0.2 2.8770455 0.092196904  2.28192316 3.4721679
## 2    Direct effect    0.5    0.5 2.6896220 0.039812403  2.29854947 3.0806945
## 3    Direct effect    0.8    0.8 3.4781045 0.151343193  2.71562263 4.2405864
## 4  Indirect effect    0.2    0.5 0.3068554 0.009683359  0.11398696 0.4997238
## 5  Indirect effect    0.2    0.8 0.5607924 0.062497779  0.07081007 1.0507746
## 6  Indirect effect    0.5    0.8 0.2539370 0.041766839 -0.14661962 0.6544936
## 7     Total effect    0.2    0.5 2.9964774 0.056778515  2.52945238 3.4635023
## 8     Total effect    0.2    0.8 4.0388969 0.098598268  3.42346112 4.6543327
## 9     Total effect    0.5    0.8 3.7320415 0.100688673  3.11011597 4.3539671
## 10  Overall effect    0.2    0.5 1.0762573 0.016628373  0.82351786 1.3289967
## 11  Overall effect    0.2    0.8 2.7678669 0.070519255  2.24738940 3.2883444
## 12  Overall effect    0.5    0.8 1.6916096 0.083045085  1.12679613 2.2564231
```

In the output above,

- **Section** [[**1**]] displays the point estimates, estimated variances, and 95% Wald-type confidence intervals for the average potential outcomes under three different treatment coverages (i.e., $\alpha = 0.2, 0.5, 0.8$) and each individual exposure (i.e. $z = 0, 1$, and marginal). The marginal average potential outcome is the average potential outcomes for a particular treatment coverage, regardless of individual exposure status.
- **Section** [[**2**]] displays the point estimates (estimand = "Direct effect", "Indirect effect", "Total effect", "Overall effect" in the output), estimated variances, and 95% Wald-type confidence intervals for four causal contrasts: direct, indirect, total and overall effects, corresponding to two distinct treatment coverages α_0 and α_1.

Outcome regression (REG) estimates

```
reg_estimates <- reg_estimation(data=dat,
                                outcome_formula = outcome_formula,
                                alpha = alpha)
reg_estimates
```

```
## [[1]]
##                              estimand alpha  estimate    variance     CI_lb
## 1      Average potential outcome z=0   0.2 0.9162181 0.006834660 0.7541839
## 2      Average potential outcome z=1   0.2 3.1192325 0.008519018 2.9383307
## 3 Marginal average potential outcome   0.2 1.3568210 0.005808949 1.2074394
## 4      Average potential outcome z=0   0.5 1.2364819 0.006327718 1.0805727
## 5      Average potential outcome z=1   0.5 3.6925175 0.005844780 3.5426760
## 6 Marginal average potential outcome   0.5 2.4644997 0.004830373 2.3282804
## 7      Average potential outcome z=0   0.8 1.5567457 0.011283605 1.3485498
## 8      Average potential outcome z=1   0.8 4.2658025 0.006048942 4.1133664
## 9 Marginal average potential outcome   0.8 3.7239911 0.005470936 3.5790208
##      CI_ub
## 1 1.078252
```

```
## 2 3.300134
## 3 1.506202
## 4 1.392391
## 5 3.842359
## 6 2.600719
## 7 1.764942
## 8 4.418239
## 9 3.868961
##
## [[2]]
##           estimand alpha0 alpha1  estimate    variance     CI_lb     CI_ub
## 1    Direct effect    0.2    0.2 2.2030144 0.008516141 2.0221432 2.3838856
## 2    Direct effect    0.5    0.5 2.4560356 0.005023502 2.3171199 2.5949513
## 3    Direct effect    0.8    0.8 2.7090568 0.010155867 2.5115388 2.9065748
## 4  Indirect effect    0.2    0.5 0.3202638 0.002731414 0.2178303 0.4226973
## 5  Indirect effect    0.2    0.8 0.6405276 0.010925656 0.4356607 0.8453945
## 6  Indirect effect    0.5    0.8 0.3202638 0.002731414 0.2178303 0.4226973
## 7     Total effect    0.2    0.5 2.7762994 0.005842575 2.6264861 2.9261126
## 8     Total effect    0.2    0.8 3.3495844 0.006047411 3.1971676 3.5020012
## 9     Total effect    0.5    0.8 3.0293206 0.005370225 2.8856909 3.1729503
## 10  Overall effect    0.2    0.5 1.1076787 0.001710576 1.0266164 1.1887411
## 11  Overall effect    0.2    0.8 2.3671701 0.004034297 2.2426808 2.4916595
## 12  Overall effect    0.5    0.8 1.2594914 0.001082823 1.1949963 1.3239866
```

See the section **Inverse probability-of-treatment weighting (IPW) estimates** for an interpretation of the output.

Regression estimation with residual bias correction (DR-BC) estimates

```
drbc_estimates <- drbc_estimation(data=dat,
                                  propensity_formula=propensity_formula,
                                  outcome_formula = outcome_formula,
                                  alpha = alpha)
drbc_estimates
```

```
## [[1]]
##                              estimand alpha estimate    variance     CI_lb
## 1      Average potential outcome z=0   0.2 0.903769 0.010513000 0.7028082
## 2      Average potential outcome z=1   0.2 3.067442 0.009117773 2.8802911
## 3 Marginal average potential outcome   0.2 1.336504 0.007806633 1.1633308
## 4      Average potential outcome z=0   0.5 1.225190 0.006699357 1.0647674
## 5      Average potential outcome z=1   0.5 3.723372 0.007631144 3.5521563
## 6 Marginal average potential outcome   0.5 2.474281 0.005510269 2.3287902
## 7      Average potential outcome z=0   0.8 1.641758 0.015258094 1.3996559
## 8      Average potential outcome z=1   0.8 4.268699 0.005078439 4.1290255
## 9 Marginal average potential outcome   0.8 3.743311 0.004564508 3.6108931
##      CI_ub
## 1 1.104730
## 2 3.254593
## 3 1.509677
## 4 1.385612
## 5 3.894587
```

```
## 6 2.619771
## 7 1.883860
## 8 4.408372
## 9 3.875728
##
## [[2]]
##           estimand alpha0 alpha1  estimate    variance     CI_lb     CI_ub
## 1    Direct effect    0.2    0.2 2.1636732 0.015170763 1.9222651 2.4050812
## 2    Direct effect    0.5    0.5 2.4981820 0.006619924 2.3387136 2.6576504
## 3    Direct effect    0.8    0.8 2.6269409 0.015936635 2.3795143 2.8743675
## 4  Indirect effect    0.2    0.5 0.3214206 0.005913692 0.1706984 0.4721429
## 5  Indirect effect    0.2    0.8 0.7379887 0.015410390 0.4946816 0.9812959
## 6  Indirect effect    0.5    0.8 0.4165681 0.006842221 0.2544443 0.5786919
## 7     Total effect    0.2    0.5 2.8196027 0.012191768 2.6031906 3.0360147
## 8     Total effect    0.2    0.8 3.3649297 0.009634937 3.1725441 3.5573153
## 9     Total effect    0.5    0.8 3.0435090 0.006636411 2.8838422 3.2031759
## 10  Overall effect    0.2    0.5 1.1377770 0.004450835 1.0070189 1.2685352
## 11  Overall effect    0.2    0.8 2.4068069 0.005402432 2.2627471 2.5508666
## 12  Overall effect    0.5    0.8 1.2690298 0.002765500 1.1659592 1.3721005
```

See the section **Inverse probability-of-treatment weighting (IPW) estimates** for an interpretation of the output.

Regression estimation with inverse-propensity weighted coefficients (DR-WLS) estimates

```
drwls_estimates <- drwls_estimation(data=dat,
                                    propensity_formula=propensity_formula,
                                    outcome_formula = outcome_formula_drwls,
                                    alpha = alpha)
drwls_estimates
```

```
## [[1]]
##                              estimand alpha  estimate    variance     CI_lb
## 1      Average potential outcome z=0   0.2 0.8945974 0.008911670 0.7095736
## 2      Average potential outcome z=1   0.2 3.1160855 0.007898553 2.9418960
## 3 Marginal average potential outcome   0.2 1.3388951 0.006880255 1.1763213
## 4      Average potential outcome z=0   0.5 1.2277460 0.006528412 1.0693837
## 5      Average potential outcome z=1   0.5 3.7213612 0.007492040 3.5517134
## 6 Marginal average potential outcome   0.5 2.4745536 0.005344578 2.3312672
## 7      Average potential outcome z=0   0.8 1.6415100 0.014890514 1.4023420
## 8      Average potential outcome z=1   0.8 4.2512543 0.005089634 4.1114272
## 9 Marginal average potential outcome   0.8 3.7293055 0.004536159 3.5972999
##      CI_ub
## 1 1.079621
## 2 3.290275
## 3 1.501469
## 4 1.386108
## 5 3.891009
## 6 2.617840
## 7 1.880678
## 8 4.391081
## 9 3.861311
```

```
##
## [[2]]
##           estimand alpha0 alpha1  estimate    variance     CI_lb     CI_ub
## 1    Direct effect    0.2    0.2 2.2214880 0.011429949 2.0119464 2.4310296
## 2    Direct effect    0.5    0.5 2.4936152 0.006662595 2.3336337 2.6535967
## 3    Direct effect    0.8    0.8 2.6097443 0.015710319 2.3640809 2.8554078
## 4  Indirect effect    0.2    0.5 0.3331485 0.004530038 0.2012321 0.4650650
## 5  Indirect effect    0.2    0.8 0.7469125 0.013410183 0.5199441 0.9738810
## 6  Indirect effect    0.5    0.8 0.4137640 0.006506461 0.2556681 0.5718599
## 7     Total effect    0.2    0.5 2.8267637 0.010539050 2.6255541 3.0279734
## 8     Total effect    0.2    0.8 3.3566569 0.008402625 3.1769951 3.5363186
## 9     Total effect    0.5    0.8 3.0235083 0.006297253 2.8679749 3.1790417
## 10  Overall effect    0.2    0.5 1.1356585 0.003651873 1.0172165 1.2541006
## 11  Overall effect    0.2    0.8 2.3904104 0.004802210 2.2545889 2.5262319
## 12  Overall effect    0.5    0.8 1.2547519 0.002359959 1.1595380 1.3499658
```

See the section **Inverse probability-of-treatment weighting (IPW) estimates** for an interpretation of the output.