MIRT DIF

11/1/2020

1. Adaptive Lasso was adopted in high DIF proportion studies. To check whether the adaptive Lasso implementation is correct, we run adaptive Lasso in low DIF proportion condition (simulation 5).

Results of 50 Replications

Type I error and Power of Adaptive LASSO method

Group	Omnibus DIF	Group with DIF=0.5	Group with DIF=1
Power	0.675	0.090	0.675
Type I	0.00286	0.00286	0

Type I error and Power of LASSO regularization method

Group	Omnibus DIF	Group with DIF=0.5	Group with DIF=1
Power	0.625	0.14	0.625
Type I	0.0314	0.0171	0.0185

Type I error and Power of mirt LRT

Group	Omnibus DIF	Group with DIF=0.5	Group with DIF=1
Power	0.515	0.375	0.646
Type I	0.0185	0.0238	0.0149

BIC was used for parameter selection in both Lasso and adaptive Lasso. The results of adaptive Lasso were comparable to that of Lasso. So the adaptive Lasso implementation should be fine.

2. Defining the incidence matrix $\Lambda = (\lambda_{jky})$ where $\lambda_{jky} = I(\gamma_{jky} \neq 0)$. The true Λ is on the solution path of regularization methods. But BIC and GIC would select more sparse Λ. We try to use AIC to reduce penalty in adaptive Lasso (simulation 6).

Results of 50 Replications

Type I error and Power of adaptive Lasso method

Adaptive LASSO (parameter selected by GIC)

Group	Omnibus DIF	Group with DIF=0.5	Group with DIF=1
Power	0.383	0.0086	0.3836
Type I	0	0	0

Adaptive LASSO (parameter selected by AIC)

Group	Omnibus DIF	Group with DIF=0.5	Group with DIF=1
Power	0.852	0.258	0.847
Type I	0.040	0.030	0.023

Type I error and Power of Lasso method

Group	Omnibus DIF	Group with DIF=0.5	Group with DIF=1
Power	0.45	0.06	0.443
Type I	0.0067	0.0033	0.0033

Type I error and Power of mirt LRT

Group	Omnibus DIF	Group with DIF=0.5	Group with DIF=1
Power	0.656	0.185	0.746
Type I	0.0067	0.01	0.0233

The power of results selected by AIC was much better than those of BIC and GIC, and the type I error is also acceptable.

3. Simulate traits with lower correlation (Cor(θ_1,θ_2) = 0.85 -> Cor(θ_1,θ_2)=0.25)

Regular Newton's method (fisher scoring) optimization in the re-estimation step fails to converge from the regularized parameter estimates in most replications. In those cases, the first derivatives of the log likelihood were relatively large at the first iteration. Then the Newton direction $d^{(t)}$ solving from $\bigtriangledown^2 logMd^{(t)} = -\bigtriangledown logM$ would fluctuate in a big magnitude in the following iterations. Using Newton's method with step size selection by the Bisection algorithm should solve the problem.