ivd: An R Package for Individual Variance Detection using Spike-and-Slab Priors

Summary

Research in fields such as psychology and education typically focus on average performance (e.g., students scores or patients' self-reported symptoms of a psychopathology), but the *consistency* or variability of that performance can also carry important information about the units of study (Leckie et al. 2023, rast_modeling_2012). Standard statistical approaches often treat residual variability as nuisance or assume it is constant across groups. The `ivd` package implements the Spike-and-Slab Mixed-Effects Location Scale Model (SS-MELSM) as a Bayesian framework to explicitly model and detect heterogeneous residual variances.

`ivd` allows users to simultaneously estimate a model for the means (location) and a model for the residual variability (scale). It employs a spike-and-slab prior on the random effects of the scale component as a probabilistic selection mechanism, distinguishing between clusters that have "usual" variability (shrunk to the population mean, or the spike component) and those with "unusual" variability (belonging to the slab) (Mitchell and Beauchamp 1988). This allows the identification of individual units that deviate significantly from the norm; for example, schools with students unusually consistent or inconsistent test scores or patients with unstable (or very stable) symptoms (Leckie et al. 2014).

Statement of Need

Standard mixed-effects models partition variance into between-group and within-group components, typically assuming the within-group residual variance (σ^2) is homogeneous across all groups (Raudenbush and Bryk 2002). While Mixed-Effects Location Scale Models (MELSM) relax this assumption by allowing σ^2 to vary as a function of covariates and random effects (Hedeker, Mermelstein, and Demirtas 2008), they do not inherently provide a decision rule for identifying *which* specific units deviate significantly from the average consistency.

In the `ivd` package MELSMs are expanded with the implementation of the SS-MELSM. This method computes a Posterior Inclusion Probability (PIP) for each cluster's scale random effect. The model simultaneously estimates fixed and random effects for both the location and the scale. For the scale model, the residual standard deviation σ_{ij} is modeled as a log-linear function of covariates. The random effects associated with the scale intercept are subject to the spike-and-slab prior, where a binary indicator δ_{jk} determines inclusion. A high PIP for δ_j indicates strong evidence that the unit belongs to the "slab" distribution, as it translates into a larger Bayes factor for including the kth random effect for the jth school Williams, Martin, and Rast (2021). The PIPs are determined by the proportion of MCMC samples where $\delta_{jk} = 1$.

The package serves as a user-friendly frontend for NIMBLE (Valpine et al. 2017), allowing SS-MELSMs to be fitted using standard `R` formula syntax without needing to write custom BUGS or NIMBLE code. Popular probabilistic programming languages like Stan do not support discrete parameters directly, making the implementation of such models difficult for applied researchers. The core function, `ivd()`, accepts two formulas: a `location_formula` for the mean structure and a `scale_formula` for the within-cluster variance structure. `ivd` leverages the future package (Bengtsson 2021), enabling parallel processing of Markov Chain Monte Carlo (MCMC) chains.

Analysis and Visualization

The `summary()` method returns the fixed effects for both location and scale, along with the PIPs for the random scale effects. Convergence of each estimate is summarised with computation of $\hat{R}$, and estimation efficiency by the effective sample size (Vehtari et al. 2021).

To facilitate the detection of atypical units, the package includes several visualization methods, specified by the `type` argument in the `plot()` function: `pip` is the default argument and is used to show units exceeding a certain probability threshold (e.g., 0.75 as default). `funnel` visualizes the relationship between the PIP and the estimated within-cluster standard deviation. Finally, `outcome` displays the interaction between average performance and consistency. For MCMC convergence diagnostics, the `codaplot()` function offers access to trace and density plots for specific parameters.

Model Comparison

Comparison of competing models is supported through the Widely Applicable Information Criterion (WAIC), which can be computed during estimation (`WAIC = TRUE` is set by default). Furthermore, `ivd` stores the pointwise log-likelihood matrix, making it compatible with the loo package (Vehtari et al. 2024) for computing predictive accuracy with the Pareto smoothed importance sampling Leave-one-out cross-validation (PSIS-LOO) (Vehtari, Gelman, and Gabry 2017).

Usage Example

The following example illustrates the workflow using the `saeb` dataset included in the package. We model mathematics proficiency for students nested within schools, predicting both the mean achievement and the residual variability as a function of student and school socioeconomic status (SES).

```r
library(ivd)
library(data.table)

## Prepare data: separate within- and between-school SES effects
data(saeb)
school_ses <- saeb[, .(school_ses = mean(student_ses, na.rm = TRUE)), by = school_id]
saeb <- saeb[school_ses, on = "school_id"]
saeb$student_ses <- saeb$student_ses - saeb$school_ses
saeb$school_ses <- c(scale(saeb$school_ses, scale = FALSE))

## Fit the Spike-and-Slab MELSM
model <- ivd(
  location_formula = math_proficiency ~ student_ses * school_ses + (1|school_id),
  scale_formula    = ~ student_ses * school_ses + (1|school_id),
  data             = saeb,
  niter            = 2000,
  nburnin          = 1000,
  WAIC             = TRUE,
  workers          = 2
)

## Summarize results
summary(model)

## Visualizations
## Identify schools with unusual variability (PIP > 0.75)
plot(model, type = "pip")
```

```r
## Inspect the relationship between PIP and within-cluster SD
plot(model, type = "funnel")

## Check convergence for the intercept
codaplot(model, parameters = "Intc")
```

Acknowledgements

Bengtsson, Henrik. 2021. "A Unifying Framework for Parallel and Distributed Processing in r Using Futures." *The R Journal* 13 (2): 208–27. https://doi.org/10.32614/RJ-2021-048.

Hedeker, Donald, Robin J. Mermelstein, and Hakan Demirtas. 2008. "An Application of a Mixed-Effects Location Scale Model for Analysis of Ecological Momentary Assessment (EMA) Data." *Biometrics* 64 (2): 627–34. https://doi.org/10.1111/j.1541-0420.2007.00924.x.

Leckie, George, Robert French, Chris Charlton, and William Browne. 2014. "Modeling Heterogeneous Variance-Covariance Components in Two-Level Models." *Journal of Educational and Behavioral Statistics* 39 (5): 307–32. https://doi.org/10.3102/1076998614546494.

Leckie, George, Richard Parker, Harvey Goldstein, and Kate Tilling. 2023. "Mixed-Effects Location Scale Models for Joint Modeling School Value-Added Effects on the Mean and Variance of Student Achievement." *Journal of Educational and Behavioral Statistics*, November, 10769986231210808. https://doi.org/10.3102/10769986231210808.

Mitchell, T. J., and J. J. Beauchamp. 1988. "Bayesian Variable Selection in Linear Regression." *Journal of the American Statistical Association* 83 (404): 1023–32. https://doi.org/10.1080/01621459.1988.10478694.

Raudenbush, S. W., and A. S. Bryk. 2002. *Hierarchical Linear Models: Applications and Data Analysis Methods.* Vol. 1. Thousand Oaks, CA: Sage Publications, Inc.

Rodriguez, Josue E., Donald R. Williams, and Philippe Rast. 2024. "Who Is and Is Not 'Average'? Random Effects Selection with Spike-and-Slab Priors." *Psychological Methods.* https://doi.org/10.1037/met0000535.

Valpine, Perry de, Daniel Turek, Christopher J. Paciorek, Clifford Anderson-Bergman, Duncan Temple Lang, and Rastislav Bodik. 2017. "Programming With Models: Writing Statistical Algorithms for General Model Structures With NIMBLE." *Journal of Computational and Graphical Statistics* 26 (2): 403–13. https://doi.org/10.1080/10618600.2016.1172487.

Vehtari, Aki, Jonah Gabry, Måns Magnusson, Yuling Yao, Paul-Christian Bürkner, Topi Paananen, and Andrew Gelman. 2024. "Loo: Efficient Leave-One-Out Cross-Validation and WAIC for Bayesian Models." https://mc-stan.org/loo/.

Vehtari, Aki, Andrew Gelman, and Jonah Gabry. 2017. "Practical Bayesian Model Evaluation Using Leave-One-Out Cross-Validation and WAIC." *Statistics and Computing* 27: 1413–32. https://doi.org/10.1007/s11222-016-9696-4.

Vehtari, Aki, Andrew Gelman, Daniel Simpson, Bob Carpenter, and Paul-Christian Bürkner. 2021. "Rank-Normalization, Folding, and Localization: An Improved R^ for Assessing Convergence of MCMC (with Discussion)." *Bayesian Analysis* 16 (2). https://doi.org/10.1214/20-BA1221.

Williams, Donald R., Stephen R. Martin, and Philippe Rast. 2021. "Putting the Individual into Reliability: Bayesian Testing of Homogeneous Within-Person Variance in Hierarchical Models." *Behavior Research Methods*, November. https://doi.org/10.3758/s13428-021-01646-x.