Targeted learning under shape constraints

Anders Munch and Thomas Gerds

September 11, 2023

Motivation

The road map of causal learning tells us to incorporate all the knowledge that we have into the statistical model for the distribution of the data

In many real applications, subject matter knowledge is available regarding the shape of the underlying conditional density and regression functions

Examples of biologically motivated shapes are

- risk of disease is not decreasing with age (given other covariates)
- The risk of disease should be a monotone function of age (given other covariates)
- The number of comorbidities increases the risk of disease
- the effect of a biomarker on the risk of disease is an unimodal function (given other covariates)

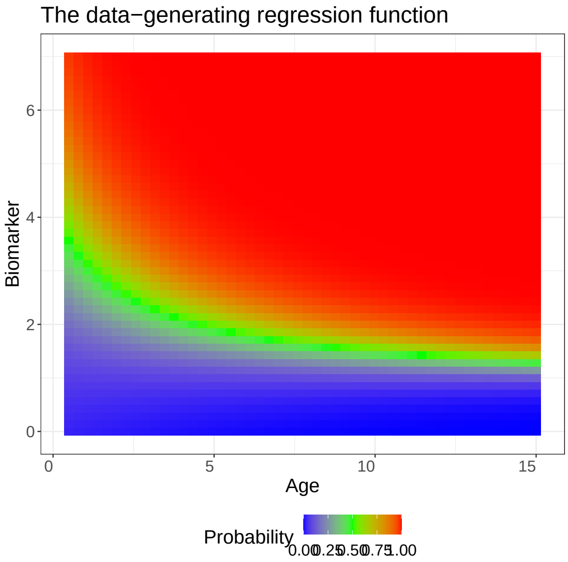
Working hypotheses

- Shape constraints can be incorporated into machine learning for nuisance parameters
- Biologically motivated shape constraints may lead to improved estimators

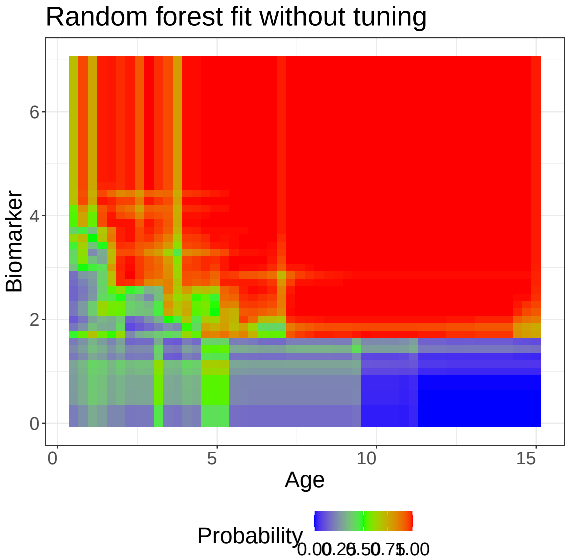
Goal for the workshop

- Discuss some initial hypotheses and ideas
- Help us move in the right research direction

Multivariate shape constraints



Machine learning (from the shelf)



Shape constraints

Examples

- Monotonicity
- Unimodality
- Convexity
- Log-concave densities

Constraints imposed on target or nuisance parameter

Shape constraint on a function-valued target parameter has been considered [e.g., Groeneboom and Jongbloed, 2014, Westling and Carone, 2020, Wu and Westling, 2022]. We will mostly discuss imposing shape-constraints on nuisance parameters.

Information bounds

Claim 1

Most shape constraints will not restrict the tangent space, and hence imposing shape constraints does not change the information bound for a statistical estimation problem.

- Which shape constraints (if any) is this true for?
- Can we still expect to improve a TMLE by imposing shape constraints on the nuisance parameters?

Information bounds

Claim 1

Most shape constraints will not restrict the tangent space, and hence imposing shape constraints does not change the information bound for a statistical estimation problem.

- Which shape constraints (if any) is this true for?
- Can we still expect to improve a TMLE by imposing shape constraints on the nuisance parameters?

Claim 2

Constructing a TMLE under a shape constrained model will typically result in a sub-model that is not contained in the shape constrained model.

- Is this a problem?

(Un)necessary restrictions on nuisance parameters?

Undersmoothing

It has been argued that undersmoothing estimators of nuisance parameters can provide better estimators of a low-dimensional target parameter [e.g., Goldstein and Khasminskii, 1996, Hjort and Walker, 2001, van der Laan et al., 2022]. Could shape constrained estimators provide unnecessary smoothing of nuisance parameter estimators, which might in fact be damaging?

(Un)necessary restrictions on nuisance parameters?

Undersmoothing

It has been argued that undersmoothing estimators of nuisance parameters can provide better estimators of a low-dimensional target parameter [e.g., Goldstein and Khasminskii, 1996, Hjort and Walker, 2001, van der Laan et al., 2022]. Could shape constrained estimators provide unnecessary smoothing of nuisance parameter estimators, which might in fact be damaging?

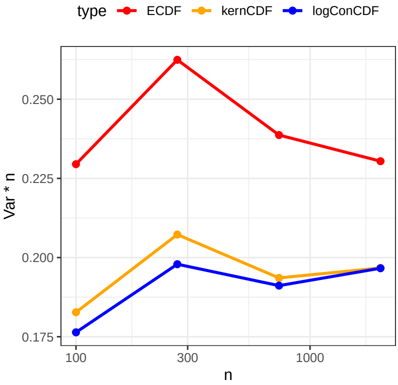
Biologically reasonable nuisance parameter estimators?

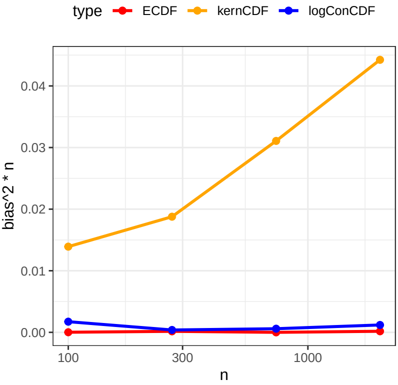
Should we pay attention to whether nuisance parameters are estimated by biologically meaningful estimators?

Should we accept a biologically unreasonable estimator of a nuisance parameter as long as it provides a good estimator of the target parameter?

Estimating a cumulative distribution function

ECDF	Empirical distribution function
kernCDF	Estimator based on smoothed kernel density estimator
logConCDF	Estimator based on log-concave density estimator [Dümbgen and Rufibach, 2009, Rufibach and Duembgen, 2023]





Challenges for future research

- Should we distinguish between learning Q vs G parts of a causal model/information loss model?
- How do we translate "marginal" smoothness constraints into constraints on a multivariate function?
- In longitudinal settings: need to discuss shape-constraints on the history (filtration)

References

L. Dümbgen and K. Rufibach. Maximum likelihood estimation of a log-concave density and its distribution function: Basic properties and uniform consistency. 2009.

L. Goldstein and R. Khasminskii. On efficient estimation of smooth functionals. *Theory of Probability & Its Applications*, 40(1):151–156, 1996.

P. Groeneboom and G. Jongbloed. *Nonparametric estimation under shape constraints*. Number 38. Cambridge University Press, 2014.

N. L. Hjort and S. G. Walker. A note on kernel density estimators with optimal bandwidths. *Statistics & Probability Letters*, 54(2):153–159, 2001.

K. Rufibach and L. Duembgen. *logcondens: Estimate a Log-Concave Probability Density from Iid Observations*, 2023. URL `https://CRAN.R-project.org/package=logcondens`. R package version 2.1.8.

M. J. van der Laan, D. Benkeser, and W. Cai. Efficient estimation of pathwise differentiable target parameters with the undersmoothed highly adaptive lasso. *The International Journal of Biostatistics*, 2022.

T. Westling and M. Carone. A unified study of nonparametric inference for monotone functions. *Annals of statistics*, 48(2):1001, 2020.

Y. Wu and T. Westling. Nonparametric inference under a monotone hazard ratio order. *arXiv preprint arXiv:2205.01745*, 2022.