Relative Sparsity for Medical Decision Problems

Sam Weisenthal
Medical Scientist Research Symposium, 2025
Thesis advised by Dr. Sally Thurston and Dr. Ashkan Ertefaie
(Some of my own, updated thoughts here)

University of Rochester
Rochester, New York

Framework

- Initial state, random vector
  - Example: (initial heart rate, initial blood pressure)
  - Draw from distribution
- Treatment, binary random variable
  - Example: vasopressors (yes or no) for hypotension
  - Call distribution of treatment given initial state a "policy"
    - Example: assign treatment, conditional on initial state, with probability 0.3
- Final state, random vector
  - Example: (final heart rate, final blood pressure)
  - Draw final state conditional on treatment and initial state
- Sampling distribution
  - Draw an initial state, a treatment (via policy), and then a final state

Reward maximization

- Reward is a function of the initial state, treatment, and final state
  - Example: final blood pressure in setting of hypotension
- The expected reward depends on the sampling distribution
  - It is thus a function of the treatment policy
  - One can thus obtain the expected-reward-maximizing policy
    - Policy search (Williams, 1992; Bagnell et al., 2003; Ng and Jordan, 2013; Sutton and Barto, 2018)
    - Antecedents: de Montmort (1713); Bernoulli (1738); Savage (1951); Bellman (1957); Lindley (1975)

Estimation

- The standard of care (policy) generates observational clinical data
- A density transform gives the expected reward of any policy as a function of said standard of care
  - If we assume: consistency, no interference, **positivity, no unmeasured confounding**
    - We can find the reward-maximizing policy using observational data
  - Importance sampling (Precup, 2000; Owen, 2013)
  - Inverse probability weighting (Horvitz and Thompson, 1952; Robins et al., 1994, 1993)

Trust region policy optimization and our methodology

- Trust region policy optimization (Schulman et al., 2015)
  - From robotics
  - Objective

$$\pi^{\text{new}} = \arg\max_{\pi} V_0(\pi) - \gamma KL_0(\pi, \pi^{\text{old}})$$

    - Stabilizes optimization
- Methodology
  - Choose an interpretable, identifiable policy
  - Assume that the standard of care is similar
  - Sets the stage for **relative sparsity,**

$$\beta_{0,\lambda} = \arg\max_{\beta} V_0(\beta) - \lambda||\beta - b_0||_1$$

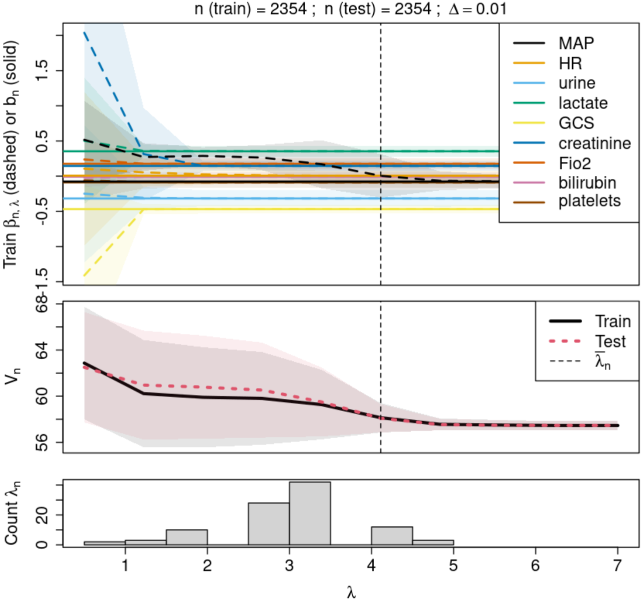
    - A window

Real data analysis

- We illustrate the proposed method on the MIMIC IV critical care database (Johnson et al. (2016))
  - Consider a decision in the medical intensive care unit
    - Management of hypotension (during initial 30 minutes)
    - Treatment is vasopressors (yes or no)
    - Reward is final blood pressure

Real data analysis

- Note, $\beta_{n,\lambda} = \arg\max_\beta V_n(\beta, b_n) - \lambda||\beta - b_n||_1$



Discussion

- Relative sparsity is a window into policy search
  - Describes what change is recommended
  - Could provide an easily justifiable policy
  - Provides another layer of assessment
- Future work
  - A more flexible model for the standard of care

Acknowledgements

- Advisors
  - Ashkan Ertefaie, Sally Thurston
  - Lab / seminar / reading groups
- Committee members, defense chair, collaborators
  - Brent Johnson, Mike McDermott, Kah Poh Loh. Hae-Ryung Park, Marissa Sobolewski, Joseph Futoma (MIMIC code)
- University of Rochester Biostatistics Department
  - Mike McDermott, Matt McCall, Karin Gasaway, Mike Conklin, Keith Bourgeois, Louis, Brian Griebner, Debra Perno, Kristine Pittinaro
  - Faculty who taught courses, supervised consulting, funded (Sally Thurston, Brent Johnson, Mechelle Sanders, Kevin Fiscella), BST students
- University of Rochester Medical Scientist Training Program, School of Medicine and Dentistry
  - Kerry O'Banion, Laurie Steiner, Alysha Taggart, Cathy Senecal-Rice, Jason Norris, Nicole Cupp, MSTP students
  - Cheryl Kodjo, David Lambert, Christopher Mooney, URSMD staff, Medical students
- ART year / RCHI / CTSI / TBS
  - Martin Zand, Henry Kautz, Vincent Silenzio, PJ Simpson-Haidaris
- University of Rochester Center for Integrated Research Computing (CIRC)
  - Brendan Mort, Harry Stern, Carl Schmidtmann, Tod Romo
- Research reported here was supported by the National Institutes of Health (NIH) under award numbers
  - T32ES007271 (National Institute of Environmental Health Sciences (NIEHS))
  - T32GM007356 (National Institute of General Medical Sciences (NIGMS))

Thank you!

References

James Bagnell, Sham M Kakade, Jeff Schneider, and Andrew Ng. Policy search by dynamic programming. *Advances in neural information processing systems*, 16, 2003.

Richard Bellman. A markovian decision process. *Journal of mathematics and mechanics*, pages 679–684, 1957.

Daniel Bernoulli. Exposition of a new theory on the measurement of risk. 1738.

Pierre Rémond de Montmort. *Essay d'analyse sur les jeux de hazard*. C. Jombert, 1713.

Daniel G Horvitz and Donovan J Thompson. A generalization of sampling without replacement from a finite universe. *Journal of the American statistical Association*, 47(260):663–685, 1952.

Alistair EW Johnson, Tom J Pollard, Lu Shen, H Lehman Li-Wei, Mengling Feng, Mohammad Ghassemi, Benjamin Moody, Peter Szolovits, Leo Anthony Celi, and Roger G Mark. Mimic-iii, a freely accessible critical care database. *Scientific data*, 3(1):1–9, 2016.

Dennis V Lindley. The role of utility in decision-making. *Journal of the royal college of physicians of London*, 9(3): 225, 1975.

Andrew Y Ng and Michael I Jordan. Pegasus: A policy search method for large mdps and pomdps. *arXiv preprint arXiv:1301.3878*, 2013.

Art B. Owen. *Monte Carlo theory, methods and examples*. 2013.

Doina Precup. Eligibility traces for off-policy policy evaluation. *Computer Science Department Faculty Publication Series*, page 80, 2000.

James M Robins, Andrea Rotnitzky, and Lue Ping Zhao. Estimation of regression coefficients when some regressors are not always observed. *Journal of the American statistical Association*, 89(427):846–866, 1994.

James M Robins et al. Information recovery and bias adjustment in proportional hazards regression analysis of randomized trials using surrogate markers. In *Proceedings of the Biopharmaceutical Section, American Statistical Association*, volume 24, page 3. San Francisco CA, 1993.

Leonard J Savage. The theory of statistical decision. *Journal of the American Statistical association*, 46(253): 55–67, 1951.

John Schulman, Sergey Levine, Pieter Abbeel, Michael Jordan, and Philipp Moritz. Trust region policy optimization. In *International conference on machine learning*, pages 1889–1897. PMLR, 2015.

Richard S Sutton and Andrew G Barto. *Reinforcement learning: An introduction*. 2018.

Ronald J Williams. Simple statistical gradient-following algorithms for connectionist reinforcement learning. *Reinforcement learning*, pages 5–32, 1992.