A Bayesian Optimization Approach to Estimating Expected Match Time and Organ Quality in Kidney Exchange

Naveen Durvasula – Greater DC

Motivation

End stage renal disease (ESRD) affects over 750,000 annually in the United States alone, and this number is increasing by 5% each year. However, there is only one sustainable treatment option for ESRD patients: kidney transplantation.

However, *not all kidney transplants are equal!* Depending on comparative biological information between the donor and patient, a patient's post-transplant quality of life and prognosis may vary dramatically. Kidney transplantation is time-sensitive as well, as patients may be in critical condition.

Objective

Develop an accurate, scalable method to predict match time and organ quality in kidney exchange.

Kidney Paired Donation

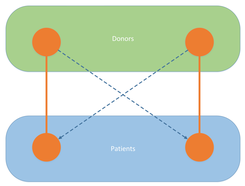


Figure 1: A KPD Cycle

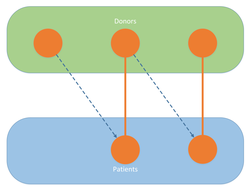


Figure 2: A KPD Chain

Suppose that an ESRD patient knows someone who is willing to donate to them, but is biologically *incompatible*. KPD allows such donor-patient pairs to enter a trading process. We formalize this by representing the pairs as *nodes* in a *compatibility graph*. Edges denote compatibility between the donor of one node and the patient of another.

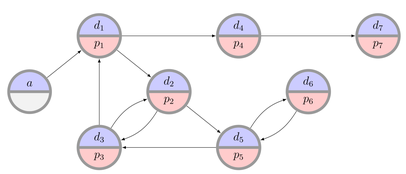


Figure 3: A compatibility graph

A *kidney exchange* finds the socially optimal set of cycles/chains that maximize a weighted edge sum. In practice, this is computationally expensive.

Abstract

Kidney exchanges allow patients with end-stage renal disease to find a lifesaving living donor by way of an organized market. However, not all patients are equally easy to match, nor are all donor organs of equal quality – some patients are matched within weeks, while others may wait for years with no match offers at all. Knowledge of expected waiting time and organ quality affects medical and insurance decisions. This work presents a principled method to estimate the expected quality of the kidney that a specific patient who enters an exchange will receive, as well as how long it will take to find that match. Estimation is performed via a novel Bayesian-optimization-based approach that learns a model of a computationally complex underlying Monte Carlo simulator. With a limited number of expensive simulation trajectories, the model produces practically-applicable results. Such fast and accurate sampling could provide medical professionals near-instantaneous access to valuable insight regarding a patient's expected outcome in a kidney exchange system.

Methodology

To accurately estimate match time and organ quality in stochastic KPD systems, I developed an open source, realistic simulator (the inner loop of Figure 4). The simulator draws from public demographic data to model the entry and exit of patients in the system. We can accurately estimate expected match time and organ quality by simulating a donor-patient pair many times and averaging the results.

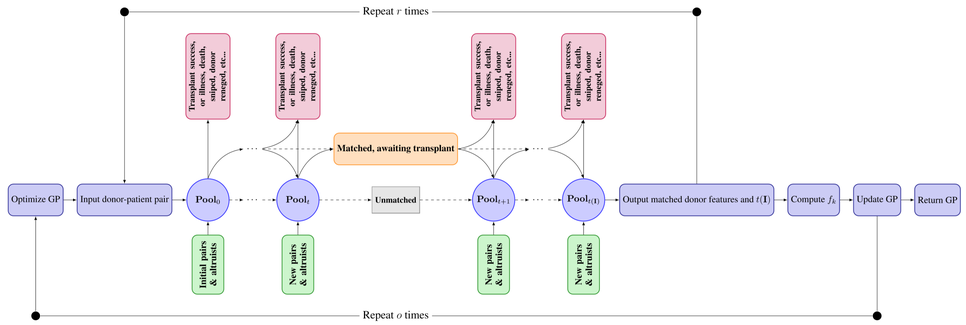
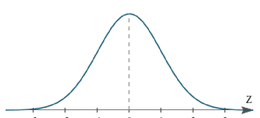


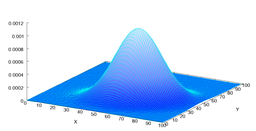
Figure 4: Actively Sampling the Black-Box Monte Carlo Simulator with Gaussian Processes

Unfortunately, as simulation is computationally expensive, this naive method is not scalable. I developed a novel active learning method to instantaneously predict the output of the simulator.

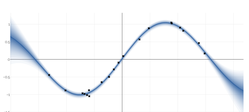
The simulator is modeled by a *Gaussian Process* (GP). The pair with highest uncertainty is sampled. Images from [1].



$\mathcal{N}(\mu, \sigma)$



$\mathcal{N}(\boldsymbol{\mu}, \boldsymbol{\Sigma})$



$\mathcal{GP}(\mu(\mathbf{x}), \kappa(\mathbf{x}, \mathbf{x}'))$

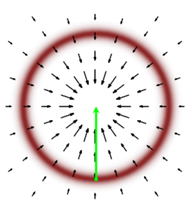


Figure 5: The gradient points to the distribution's mode

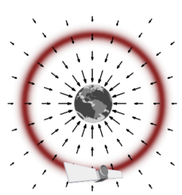


Figure 6: The typical set is like an "orbit" around them mode

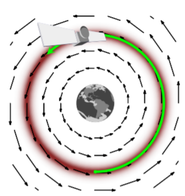


Figure 7: Hamilton's equations allow us to explore the typical set

Given the simulator output, the GP mean is updated by Bayes' Law. The predicted variance is updated by Hamiltonian Monte Carlo. Images from [2].

The GP model is compared against several standard regression methods under a simplified matching policy. Models are compared by the mean absolute residual between their predictions and the simulator output.

Experiments

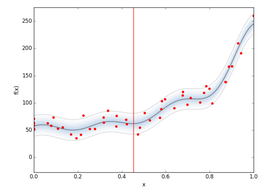


Figure 8: Prediction of expected waiting time (in weeks) by the GP model

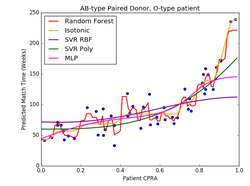


Figure 9: Prediction of expected waiting time (in weeks) by passively learned models

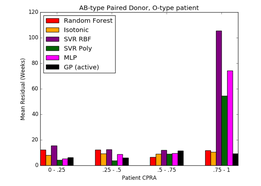


Figure 10: Mean absolute residual (weeks), O-type patient and AB-type donor

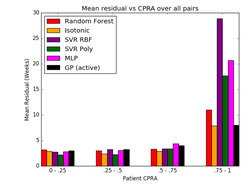


Figure 11: Mean absolute residual (weeks), all blood type pairs

Conclusions

With a limited number of expensive simulation trajectories, the model produced reliably accurate results in a proof-of-concept setting, outperforming standard regression models. The model and proof-of-concept experiments support more-intense computational experiments with a more advanced matching policy. Aside from providing instantaneous access to valuable medical insight, this work has further applications to kidney exchange, such as in the design of more socially beneficial matching mechanisms.

References

[1] GPyOpt.
GPyOpt: A Bayesian optimization framework in Python, 2016.

[2] Michael Betancourt.
A conceptual introduction to Hamiltonian Monte Carlo. *CoRR*, abs/1701.02434, 2017.

[3] Allan B Massie, Joseph Leanza, LM Fahmy, EKH Chow, Niraj M Desai, X Luo, EA King, MG Bowring, and DL Segev.
A risk index for living donor kidney transplantation. *American Journal of Transplantation*, 16(7):2077–2084, 2016.