Ryan Cumings-Menon
Research Statement

My research focuses on the formulation of nonparametric quantitative methods and the applications of these methods in microeconometrics. My job market paper uses optimal transport to create a framework for the estimation and inference of densities subject to shape constraints. I also have written a paper on a nonparametric estimator of a mode conditional on covariates. Both of these papers can be found on my website at `https://sites.google.com/view/ryancumingsmenon`. My final paper, which I am working on with Professor Elizabeth Powers, estimates the mental health impacts of having a third child. These papers are described below in more detail.

Shape Constrained Density Estimation Via Optimal Transport

Estimating a density by optimizing a fidelity criterion subject to a shape constraint is capable of ensuring the parsimony of a density estimate without the requirement of penalty or smoothing parameters. Moreover, the accuracy of these estimators often compares favorably to those that do require these parameters. This feature has motivated much of the research in shape constrained density estimation. Common examples include methods that constrain the density estimate to be monotone, $\log-$concave, or $\rho-$concave.

There are also a wide variety of shape constraints in economic theory, but, in this setting, they are more often referred to as regularity conditions. For example, if $f(x)$ denotes a density of private valuations of risk neutral bidders, with a corresponding distribution function denoted by $F(x)$, Myerson (1981) shows that the virtual valuations function, defined by $J(x) := x - (1 - F(x))/f(x)$, must be strictly increasing in order for an auction that awards the item to the highest bidder to be revenue maximizing. For a partial list of more examples see (Ewarhart, 2013; Bagnoli and Bergstrom, 2005).

This paper proposes a framework to estimate a shape constrained density and to test if a population density satisfies a shape constraint. Our approach has the advantage of allowing both procedures to be implemented over a large class of constraints.[1] Thus, the proposed estimator allows for the estimation of densities that satisfy regularity conditions for structural modeling without parametric assumptions, and, if a theoretical result implies a density satisfies a particular shape constraint, it allows for the estimation of densities conditional on this additional data. The test allows for the verification that a density satisfies the necessary and sufficient conditions of theoretical results. Fundamentally, the formulation of these methods is possible over a large class of constraints because the fidelity criterion to the data has an unconstrained minimum that is itself a consistent density estimator.

The paper initially presents all of the results in the context of density estimation subject to a $\rho-$concavity constraint. In this context, Theorems 2 and 4 provide the limiting density of the estimator, and these results also imply that the estimator is consistent when the population

[1] As discussed in the introduction of the paper, the commonly used fidelity criterions in shape constrained density estimation would not provide a convincing way to achieve this goal when the constraint is sufficiently weak. This would rule out many of the constraints of interest in economics. In short, this is partly because the parsimony and the existence of the estimator cannot be ensured in these cases.

density satisfies the shape constraint. In addition, Theorem 5 introduces a test for the null hypothesis that the population density satisfies the shape constraint. This test is based on comparing the fidelity criterion at the unconstrained minimum to its corresponding value at the shape constrained density estimate. Theorem 6 shows the optimization problem defining the estimator is strictly convex and provides the gradient and Hessian of the fidelity criterion, and Theorem 7 provides the sufficient conditions for all of these results to hold under an arbitrary shape constraint.

The paper considers all density estimators that are defined as a transformation of the choice variables of the optimization problem subject to a set of constraints on this choice variable. Most would likely view the sufficient conditions for convexity of the optimization problem to be the strongest assumptions that are required for these generalizations. Specifically, these are convexity of the feasible set and either convexity or concavity of the transformation of the choice variable. Since these assumptions may not hold for all formulations of the transformation and constraints, the final part of Theorem 7 shows that the objective function is convex in a region near the unconstrained global minimum even when these assumptions do not hold. This allows for the optimization algorithm to be initialized at this global minimum, and, as long as the unconstrained minimum is sufficiently close to the feasible set, the convergence path of the optimization algorithm will often stay inside this region.

The application uses data from the California Department of Transportation to explore if their decision to award construction contracts with the use of a first price auction is cost minimizing. The density estimate of the marginal costs of construction firms defined by the unconstrained minimum of the fidelity criterion does not satisfy Myerson's (1981) regularity condition. However, the estimate, subject to this constraint, appears to match the data closely, and the proposed test also fails to reject the null hypothesis that the regularity condition is satisfied.

A Nonparametric Modal Regression

The inherently local nature of the mode makes estimators of the mode one of the more robust notions of central tendency. For example, univariate mode estimators are robust to outliers moving further away from the mode, switching from one side of the mode to the other, or being removed from the dataset entirely. Since estimators of the mode rely on a proportion of the data that converges to zero, this property also generally holds asymptotically for any datapoint that is not located in the neighborhood of the mode. Partly for this reason, these estimators have also been generalized to estimators of the mode of a dependent variable, conditional on covariates. These regressions can be viewed as estimators of the most common or typical value of the dependent variable given the covariates.

In this paper I propose an alternative starting point for nonparametric modal regressions. Specifically, a graph is constructed using the data, which is used to define several paths that balance a tradeoff between distance between subsequent points in the direction of the covariates and curvature in the direction of the dependent variable. Afterward, these paths are combined to define the final regression. This approach lends three primary advantages to the method. First, it allows for the formulation of an algorithm that computes the proposed regression with a worst case time complexity of $O(n^2)$ or an approximation of the regression with a strictly faster time complexity that depends on the desired accuracy of the approximation. Our partially unoptimized code is generally faster than the classic formulations when the sample size is greater than approximately 1,500.

Second, while the method allows for the specification of a smoothing parameter for increased accuracy, this is not a requirement as long as the conditional mode can be expressed as a function of the covariates. Third, even without specifying this smoothing parameter, the method's accuracy appears to compare favorably to methods based on kernel density estimation when the conditional mode and the value of the joint density at the conditional mode are both nonlinear. The paper explores this further through simulations.

After providing a consistency result, we move onto an application using data on high school GPA, University of Illinois at Urbana-Champaign (UIUC) GPA, and ACT score of all undergraduates at UIUC between 2000 and 2010. We estimate the most typical value of UIUC GPA conditional on high school GPA and ACT score, and find a relatively weak association between ACT score on UIUC GPA.

The Mental Health Effects of Having a Third Child

Professor Elizabeth Powers and I are also working on a paper in which we estimate the causal effect of having a third child on the mental health of parents. We use the identification strategy that was first described by Angrist and Evans (1998). Specifically, since parents in the United States are more likely to have a third child after having two children of the same gender, the gender of the first two children can be used to construct instrumental variables for having a third child. We find some evidence that having a third child negatively impacts the happiness of mothers but not fathers. We are also planning on comparing these estimates with at least one alternative identification strategy.

References

Angrist, J. D., & Evans, W. N. (1998). Children and their parents' labor supply: Evidence from exogenous variation in family size. *The American Economic Review, 88*(3), 450-477.

Bagnoli, M., & Bergstrom, T. (2005). Log-concave probability and its applications. *Economic Theory, 26*(2), 445-469.

Ewerhart, C. (2013). Regular type distributions in mechanism design and $\rho-$concavity. *Economic Theory, 53*(3), 591-603.

Myerson, R. B. (1981). Optimal auction design. *Mathematics of operations research, 6*(1), 58-73.