Nonparametrics and Local Methods: Semiparametrics

C.Conlon

February 27, 2021

Applied Econometrics

The Seminonparametric Approach

- If we are "pretty sure" that f is almost $f_{m,\sigma}$ for some family of densities indexed by (m, σ), then we can choose a family of positive functions of increasing complexity $P^1_\theta, P^2_\theta, \ldots$
- Choose some M that goes to infinity as n does (more slowly), and maximize over (m, σ, θ) the loglikelihood

$$\sum_{i=1}^{n} \log f_{m,\sigma}(y_i) P^M_\theta(y_i).$$

It works. . . but it is hard to constrain it to be a density for large M.

Mixtures of Normals

A special case of seminonparametrics, and usually a very good approach: Let $y|x$ be drawn from

$N(m_1(x,\theta), \sigma_1^2(x,\theta))$ with probability $q_1(x,\theta)$;

$\ldots$

$N(m_K(x,\theta), \sigma_K^2(x,\theta))$ with probability $q_K(x,\theta)$.

where you choose some parameterizations, and the q_k's are positive and sum to 1.

Can be estimated by maximum-likelihood:

$$\max_\theta \sum_{i=1}^n \log\left(\sum_{k=1}^K \frac{q_k(x_i,\theta)}{\sigma_k(x,\theta)}\phi\left(\frac{y_i - m_k(x_i,\theta)}{\sigma_k(x_i,\theta)}\right)\right).$$

Usually works very well with $K \leq 3$ (perhaps after transforming y to $\log y$, e.g).

Seminonparametric (=Flexible) Regression

Idea: we add regressors when we have more data

$\rightarrow$ **series or sieve estimators**: choose a basis of functions $P_k(x_i)$ (x_i^k, or orthogonal polynomials, or sines...)

$\rightarrow$ run *linear regression* $y_i = \sum_{k=1}^{M} P_k(x_i)\theta_k + \epsilon_i$

a reasonable compromise (again, M must go to infinity, more slowly than n).

Still curse of dimensionality, and nonparametric asymptotics.

Splines: trading off fit and smoothness

Choose some $0 < \lambda < \infty$ and

$$\min_{m(.)} \sum_i (y_i - m(x_i))^2 + \lambda J(m),$$

Then we "obtain" the natural cubic spline with knots=$(x_1, \ldots, x_n)$:

- m is a cubic polynomial between consecutive x_i's
- it is linear out-of-sample
- it is C^2 everywhere.

"Consecutive" implies one-dimensional... harder to generalize to $p_x > 1$.

Orthogonal polynomials: check out Chebyshev, $1, x, 2x^2 - 1, 4x^3 - 3x \ldots$ (on $[-1, 1]$ here.)

Additive models

Additive model: $y = \alpha + \sum_{j=1}^{p} + f_j(X_j) + +\epsilon$
Backfitting algorithm: start with $\hat{a} = \overline{y}_n$, and some zero–mean guesses $\hat{f}_j \equiv 0$. Then for $j = 1, \ldots, p, \ldots, 1, 2, \ldots, p, \ldots,$

1. Define

$$
\begin{aligned}
f_j &\leftarrow S_j[\{y_i - \hat{\alpha} - \sum_{k \neq j} \hat{f}_k(x_{ik})\}_1^N] \\
f_j &\leftarrow \hat{f}_j - \frac{1}{N} \sum_{i=1}^{N} \hat{f}_j(x_{ij}).
\end{aligned}
$$

2. Regress $\hat{y}$ on x_j to get R_j; then replace $\hat{r}_j$ with $R_j - \frac{1}{n}\sum_i \hat{r}_j(x_{ji})$ (where S_j is some cubic smoothing spline).
3. Iterate until $\hat{f}_j$ doesn't change.