Shape Constrained Regression -Connections the Frontier Analysis

Andrew Johnson
Texas A&M University
Osaka University

Comparison of Methods

Nonparametric methods are asymptotically consistent, so why would you want to impose axioms and risk misspecification?

Comparison of Methods

Nonparametric methods are asymptotically consistent, so why would you want to impose axioms and risk misspecification?

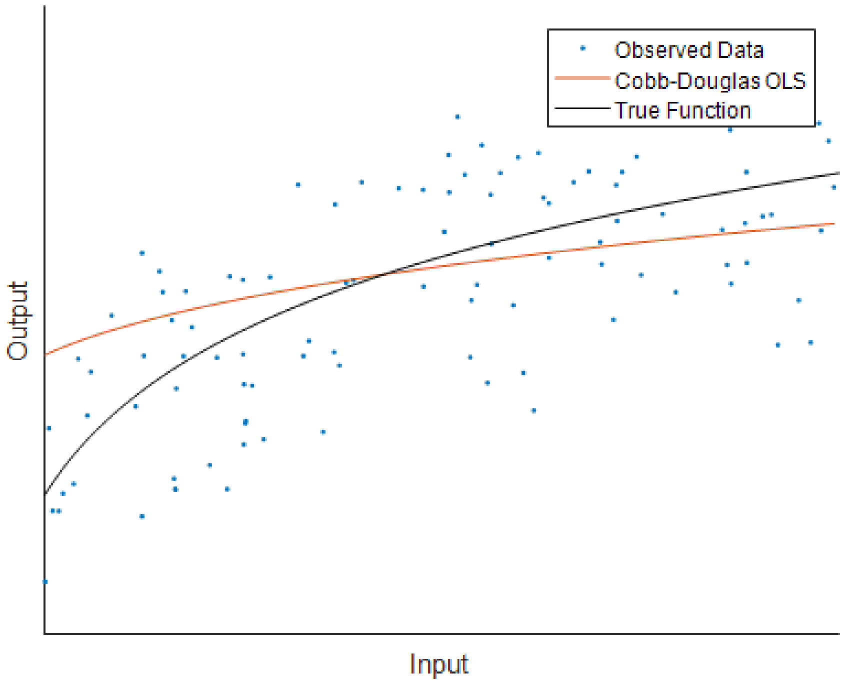
- Improved finite sample performance

  - Unrestricted nonparametric estimators converge slowly and are typically hard to interpret

  - Parametric assumptions are for statistical/computational convenience, but often contradict production theory

Comparison of Methods

Parametric



Common Models:

- Cobb-Douglas
- Translog

Advantages:

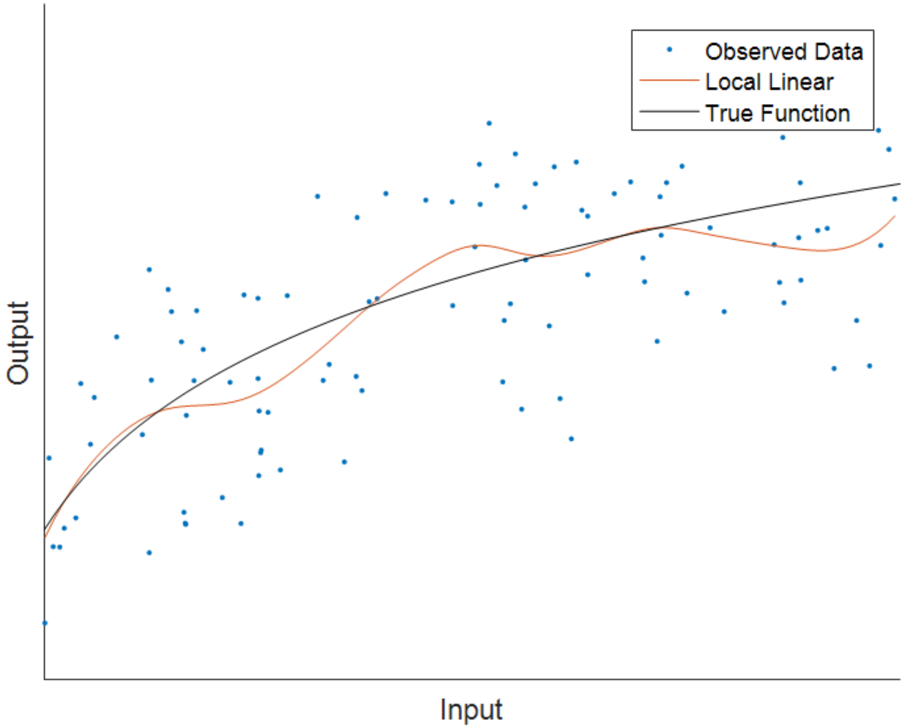
- Computational Speed

Drawbacks:

- Functional Misspecification
- Satisfies Economic Theory Axioms
- Not very flexible

Comparison of Methods

Non-Parametric



Common Models:

- Kernel regression
- Local Maximum Likelihood

Advantages:

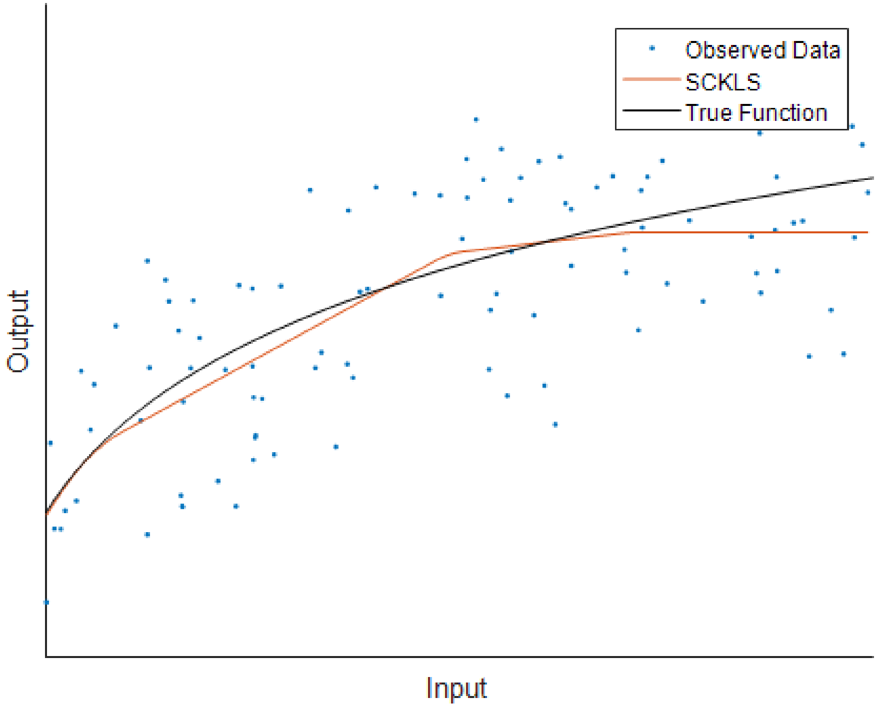
- Very Flexible

Drawbacks:

- Violation of Economic Theory Axioms
- Problem of Interpretability

Comparison of Methods

Shape constrained Nonparametric



Common Models:

- CNLS
- SCKLS

Advantages:

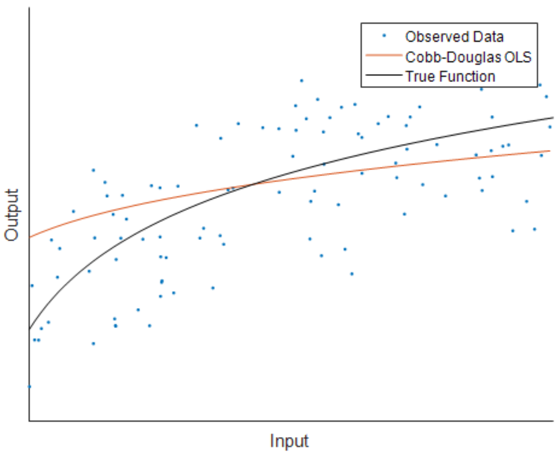
- Finite Sample Performance Improvement
- Satisfies Economic Theory Axioms
- No need of Functional Specification

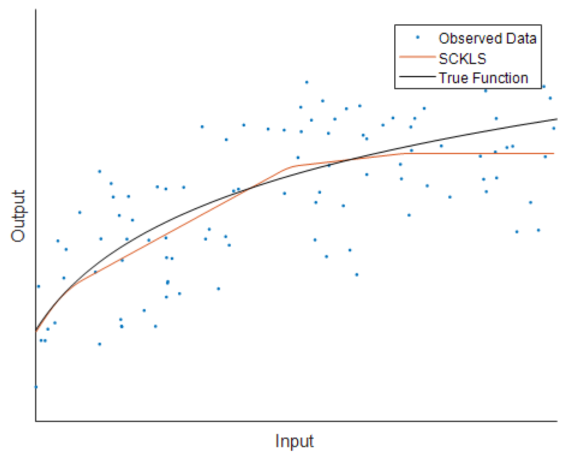
Drawbacks:

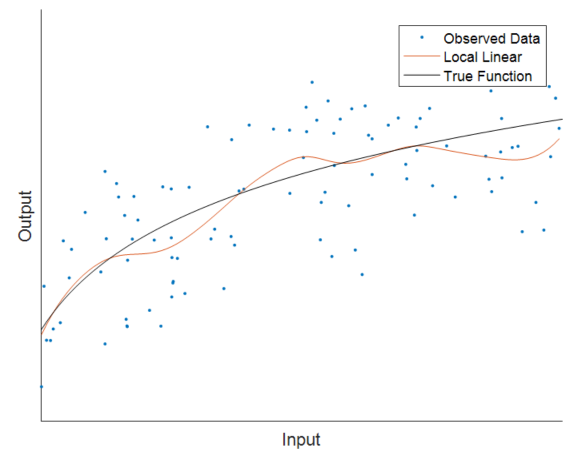
- Computational Time

Comparison of Methods

Parametric (too restrictive)	**Shape constrained Nonparametric** (balanced)	**Nonparametric** (too flexible)

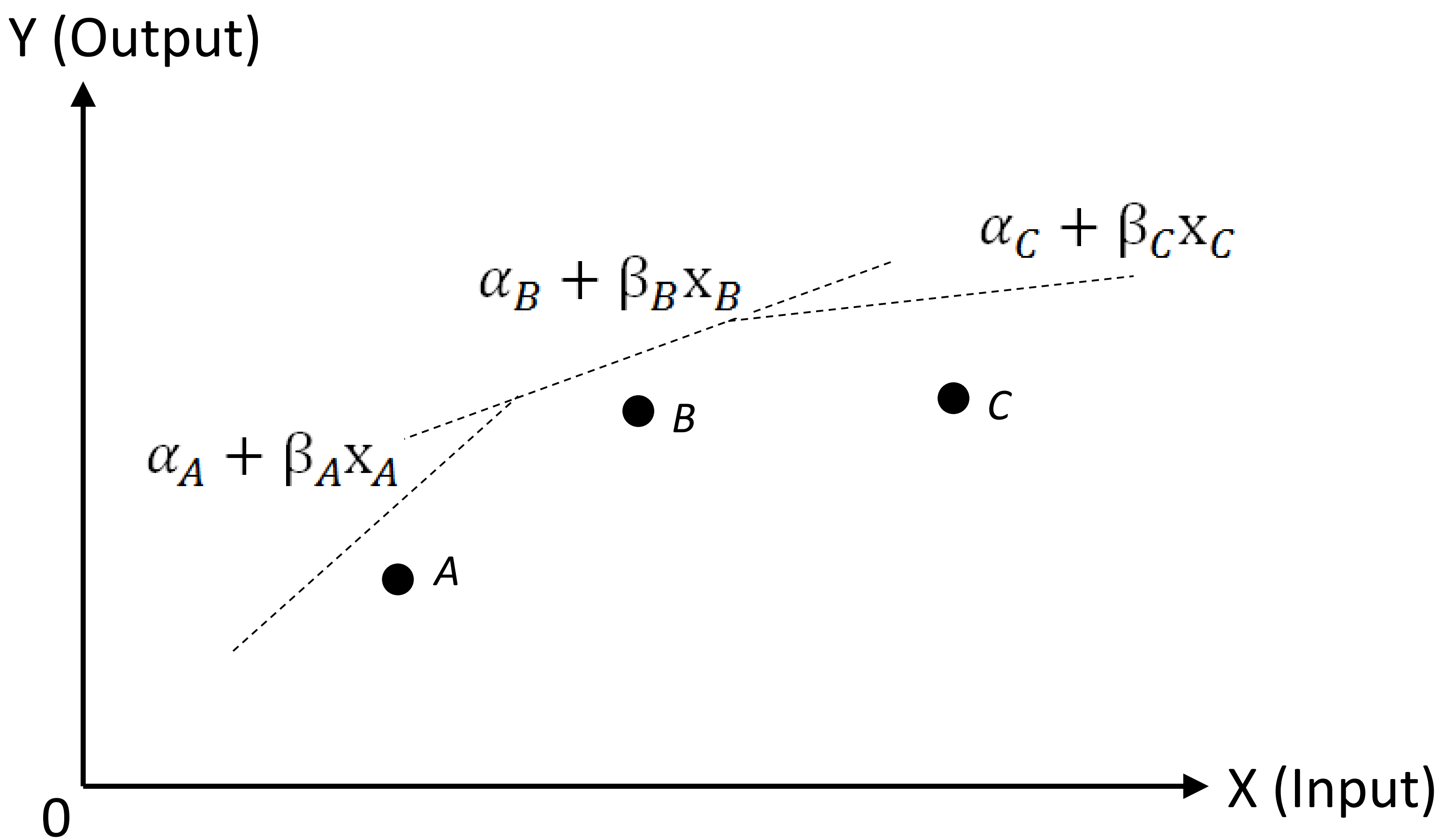






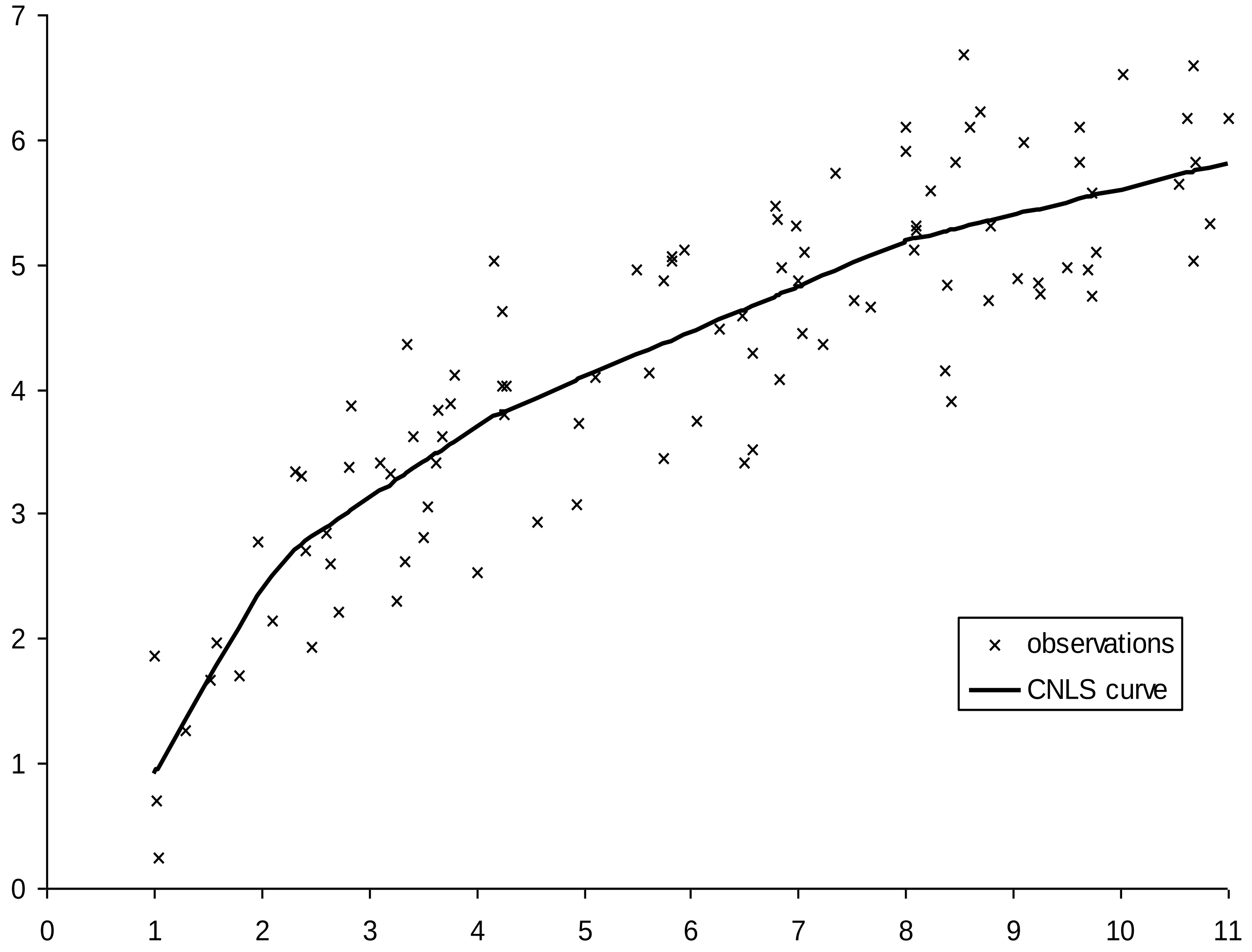
Convex Nonparametric Least Squares

Convex nonparametric least squares
Single-Input Single-Output



Convex nonparametric least squares

- Simulated example



Convex nonparametric least squares (CNLS)

- Univariate case: Hildreth (1954), Hanson & Pledger (1976); Groeneboom et al. (2001)
- Multivariate case (Kuosmanen 2008, *Ectr. J.*):

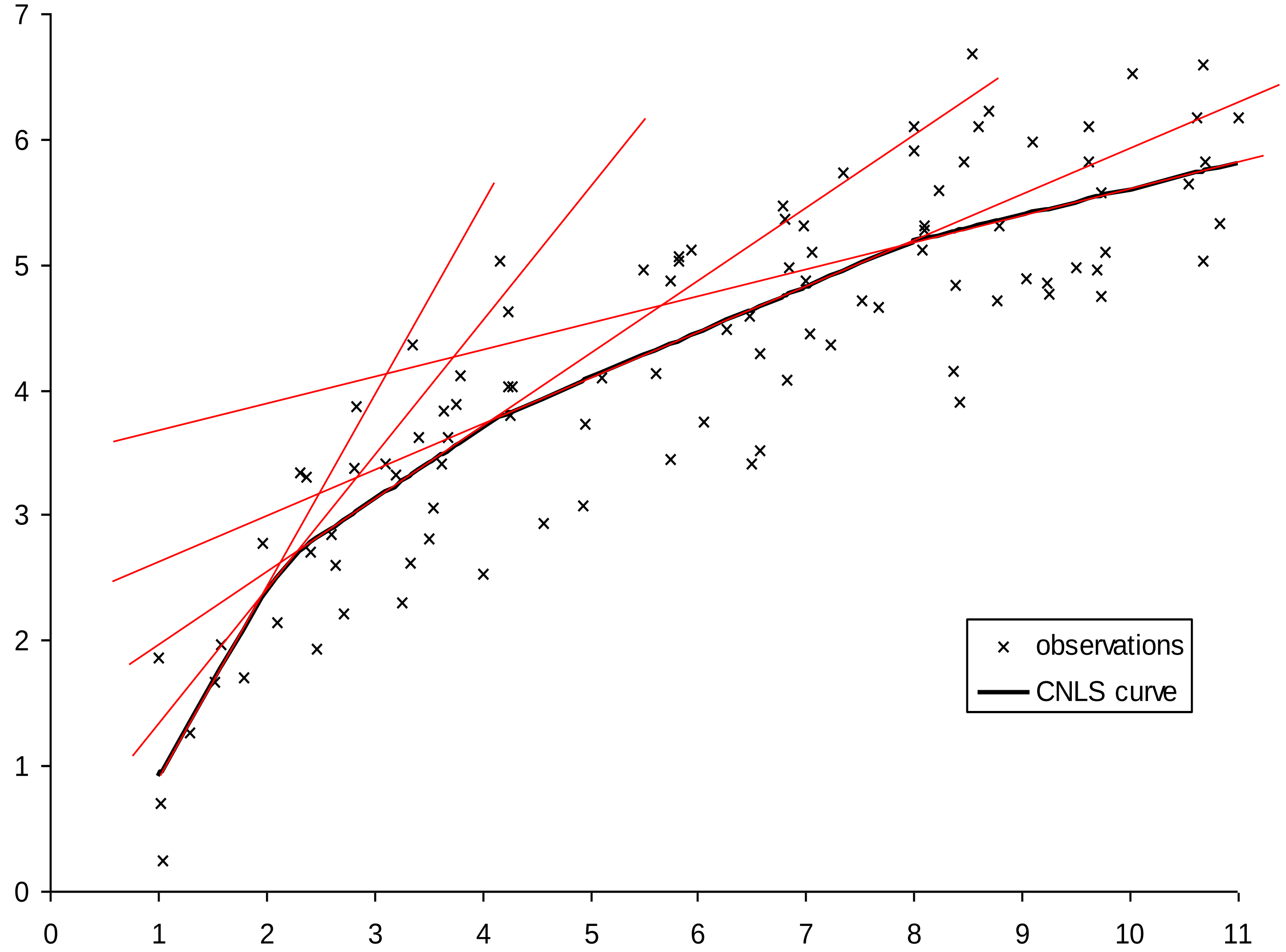
$$\min_{\alpha,\boldsymbol{\beta},\varepsilon} \sum_{i=1}^{n} \varepsilon_i^2$$
$$y_i = f(\mathbf{x}_i) + \varepsilon_i \quad \forall i$$
f monotonic increasing and concave

$$\min_{\alpha,\boldsymbol{\beta},\varepsilon} \sum_{i=1}^{n} \varepsilon_i^2$$
$$y_i = \alpha_i + \boldsymbol{\beta}_i'\mathbf{x}_i + \varepsilon_i \quad \forall i$$
$$\alpha_i + \boldsymbol{\beta}_i'\mathbf{x}_i \le \alpha_h + \boldsymbol{\beta}_h'\mathbf{x}_i \quad \forall h, i$$
$$\boldsymbol{\beta}_i \ge \mathbf{0} \quad \forall i$$

- Representation theorem: Optimal solution to the finite problem on the right is always one of the optimal solutions to the problem on the left
- Unbiasedness and consistency: Seijo & Sen (2011; *An. Stat.*)

Convex nonparametric least squares

supporting hyperplanes



Computational Issues

CNLS Computation

– Convex Nonparametric Least Squares (CNLS)

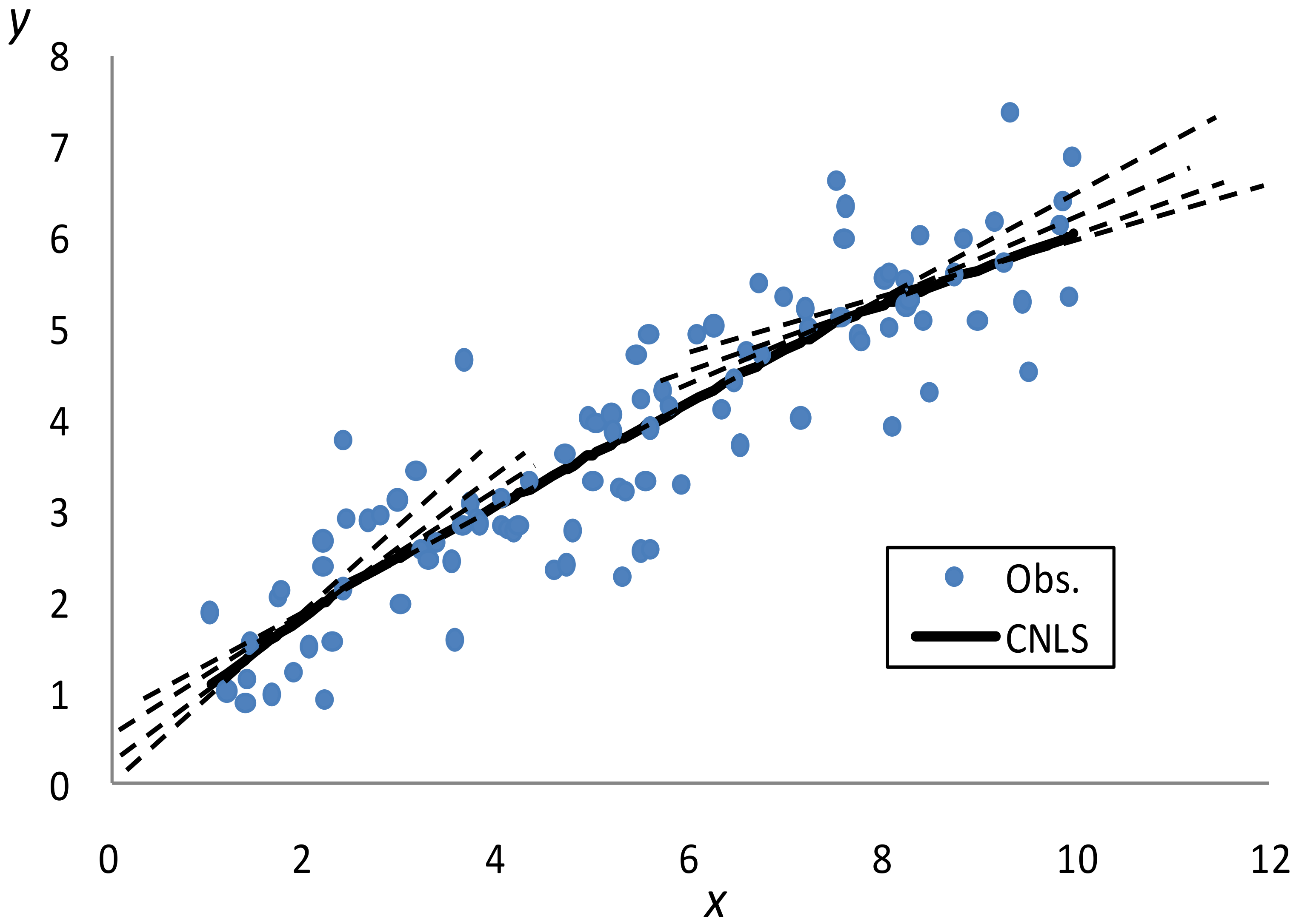
$$\min_{\alpha,\beta,\varepsilon}\left\{\sum_{i=1}^{n}\varepsilon_i^2 \,\middle|\, \begin{array}{ll} y_i = \alpha_i + \boldsymbol{\beta}_i^{\mathrm{T}}\mathbf{x}_i + \varepsilon_i \quad \text{for } i = 1,\dots,n & (1a) \\ \alpha_i + \boldsymbol{\beta}_i^{\mathrm{T}}\mathbf{x}_i \le \alpha_h + \boldsymbol{\beta}_h^{\mathrm{T}}\mathbf{x}_i \quad \text{for } i,h = 1,\dots,n \text{ and } i \ne h & (1b) \\ \boldsymbol{\beta}_i \ge 0 \quad \text{for } i = 1,\dots,n, & (1c) \end{array}\right\}$$

- 1st constraint: linear regression
- 2nd constraint: convexity using Afriat inequalities
- 3rd constraint: monotonicity

– Computation burden

- 2nd constraints will generate n(n-1) constraints, where n is number of observations

CNLS Computation



CNLS Computation

- A Generic Algorithm for CNLS Model Reduction
  - Dantzig *et al.* (1954, 1959) proposed the approach of solving large-scale traveling-salesman problems
    - Solve a relaxed model
    - Iteratively add the violated "complicating" constraints
    - Stop when the optimal solution to the relaxed model is feasible for the original problem
  - Relaxed CNLS problem (RCNLS)

$$\alpha_i + \boldsymbol{\beta}_i'\mathbf{x}_i \leq \alpha_h + \boldsymbol{\beta}_h'\mathbf{x}_i \quad \forall (i,h) \in V$$

    - Identify an initial solution (2 approaches)
    - Add violated constraints iteratively (3 strategies)

CNLS Computation

<u>Generic Algorithm</u>

1. Let $t = 0$ and let V be a subset of the observation pairs.
2. Solve RCNLS to find an initial solution, $(\alpha_i^{(0)}, \beta_i^{(0)})$.
3. Do until $(\alpha_i^{(t)}, \beta_i^{(t)})$ satisfies all concavity constraints (equations (1b)):

   3.1 Select a subset of the concavity constraints that $(\alpha_i^{(t)}, \beta_i^{(t)})$ violates and let $V^{(t)}$ be the corresponding observation pairs.

   3.2 Set $V = V \cup V^{(t)}$.

   3.3 Solve RCNLS to obtain solution $(\alpha_i^{(t+1)}, \beta_i^{(t+1)})$.

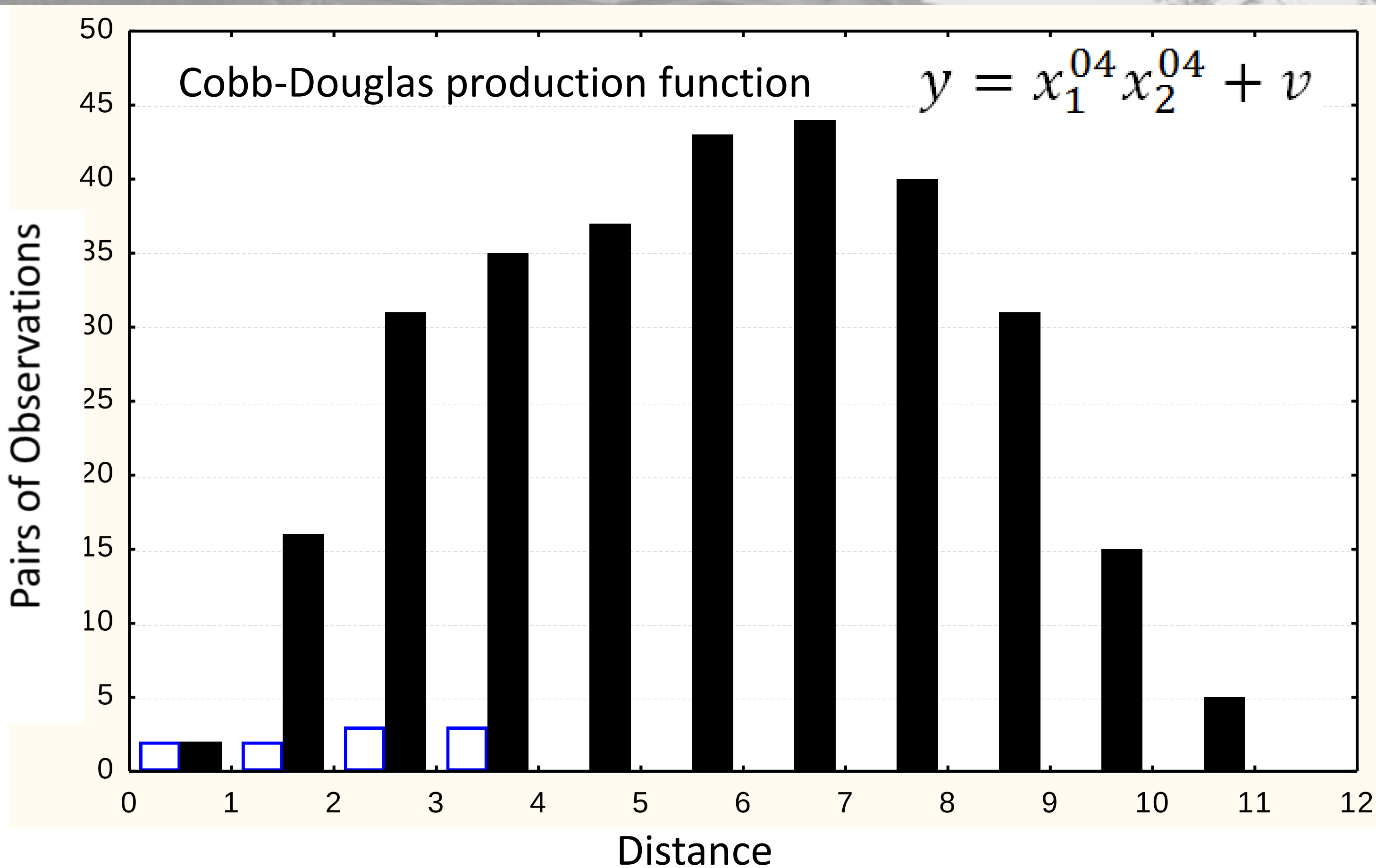
   3.4 Set $t = t + 1$

CNLS Computation

- Initial solution identification
  - the number of hyperplanes to construct the function is generally much smaller than n.
  - predict the relevant concavity constraints

2) Sweet Spot Approach (Distance measure)
  - The range between the 0 percentile and the δ_i th percentile is defined as the Sweet Spot.
  - Include the concavity constraints corresponding to the observations whose distance to observation i is less than a pre-specified threshold value δ_i.

CNLS Computation



Constraints corresponding to nearby observations are significantly more likely to be relevant.

CNLS Computation

- Add Violated Concavity Constraints
  - Generate an initial solution quickly,
  - Plug solution into the CNLS model and identify which convexity constraint violated.
  - Iteratively add some set of the violated constraints (*V*)

$$\min_{\alpha,\beta,\varepsilon}\left\{\sum_{i=1}^{n}\varepsilon_i^2 \;\middle|\; \begin{array}{c} y_i = \alpha_i + \boldsymbol{\beta}_i'\mathbf{x}_i + \varepsilon_i \quad \forall i = 1,\dots,n; \\ \alpha_i + \boldsymbol{\beta}_i'\mathbf{x}_i \le \alpha_h + \boldsymbol{\beta}_h'\mathbf{x}_i \quad \forall (i,h) \in V; \\ \boldsymbol{\beta}_i \ge 0 \quad \forall i = 1,\dots,n; \end{array}\right\}$$

Taxonomy of Methods

			Parametric	Nonparametric	
				Local averaging	Axiomatic
Conditional Mean			OLS Gauss (1795), Legendre (1805)	Kernel regression Nadaraya (1964), Watson (1964)	Convex regression Hildreth (1954), Hanson and Pledger (1976)
Frontier	Deterministic	Sign constrained	Parametric Programming Aigner and Chu (1968)	Nonparametric Programming Post et al. (2002)	Data Envelopment Analysis Farrell (1957), Afriat (1972), Charnes et al. (1978)
		2-stage	Corrected OLS Winsten (1957), Greene (1980)	Corrected Kernel Regression Kneip and Simar (1996)	Corrected CNLS Kuosmanen and Johnson (2010)
	Stochastic	Maximum likelihood	Stochastic Frontier Analysis Aigner et al. (1977), Meeusen and van den Broeck (1977)	Local-Likelihood Kumbhakar et al. (2007)	Banker and Maindiratta (1992)
		2-stage	Stochastic Frontier Analysis Aigner et al. (1977), Meeusen and van den Broeck (1977)	Semi-nonparametric SFA Fan et al. (1996)	Stochastic Nonparametric Envelopment of Data Kuosmanen and Kortelainen (2012)
				Shape Constrained Kernel Weighted Least Squares Du et al. (2013), Yagi et al. (2017)	

Data envelopment analysis (DEA)

Charnes, Cooper & Rhodes (1978), EJOR

Dual problem:

$$
\begin{aligned}
Eff_i^{-1} &= \min \phi_i \\
s.t. \qquad & \phi_i \mathbf{x}_i \geq \boldsymbol{\lambda}_i' \mathbf{X} \\
& \mathbf{y}_i \leq \boldsymbol{\lambda}_i' \mathbf{Y} \\
& \boldsymbol{\lambda} \geq \mathbf{0}
\end{aligned}
$$

X is *n*x*m* matrix of inputs

Y is *n*x*s* matrix of outputs

Stochastic Frontier

- Stochastic frontier uses regression to estimate an average function in the first stage and then shift the function to estimate a frontier production function

- SF has four common efficiency score distributions
  - Exponential
  - Half-normal
  - Truncated Normal
  - Gamma

Stochastic Frontier Step 1

- Given the model $\ln y = h(\ln x, \beta) + \varepsilon$
- Perform Ordinary Least Squares (OLS) regression
  - Results β (coefficients of each independent variable) and ε (error term)
  - If ε is not distributed normally then it may contain efficiency information, however, if ε is skewed then β_0 is inconsistent

Stochastic Frontier Step 2

Decomposition under the maintained assumptions of half-normal inefficiency and normal noise

$$M_2 = \left[\frac{\pi - 2}{\pi}\right]\sigma_u^2 + \sigma_v^2$$

$$M_3 = \left(\sqrt{\frac{2}{\pi}}\right)\left[\frac{4}{\pi} - 1\right]\sigma_u^3$$

$$\hat{M}_2 = \sum_{i=1}^{n}(\hat{\varepsilon}_i - \overline{\varepsilon})^2 / n$$

$$\hat{M}_3 = \sum_{i=1}^{n}(\hat{\varepsilon}_i - \overline{\varepsilon})^3 / n$$

Stochastic Frontier – Distribution Options

- Run log-likelihood regression to determine the values of the parameters
  - If ε is skewed then we assume $\varepsilon = v - u$ where v is a noise term distributed $N(0, \sigma_v^2)$ and we need an assumption about the distribution of u
  - $u \sim N^+ (0, \sigma_u^2)$ half-normal assumption
  - $u \sim E (\lambda)$ exponential assumption
  - $u \sim N^+ (\mu, \sigma_u^2)$ truncated normal assumption
  - $u \sim G (\alpha, \lambda)$ gamma assumption

Least Squares Interpretation of DEA

Frontier (Sign Constrained)

	parametric	*non-parametric*
central tendency	**Ordinary Least Squares** $\min_{\alpha,\beta,\varepsilon}\left\{\sum_{l=1}^{n}\varepsilon_i^2 \middle\vert y_i=\alpha+\boldsymbol{\beta}'\mathbf{x}_i+\varepsilon_i\ \forall i=1,\ldots,n\right\}$	**Convex Nonparametric Least Squares** $\min_{\alpha,\beta,\varepsilon}\left\{\sum_{l=1}^{n}\varepsilon_i^2 \middle\vert \begin{array}{l} y_i=\alpha_i+\boldsymbol{\beta}_i'\mathbf{x}_i+\varepsilon_i\ \forall i=1,\ldots,n; \\ \alpha_i+\boldsymbol{\beta}_i'\mathbf{x}_i\le\alpha_h+\boldsymbol{\beta}_h'\mathbf{x}_i\ \forall h,i=1,\ldots,n; \\ \boldsymbol{\beta}_i\ge\mathbf{0}\ \forall i=1,\ldots,n \end{array}\right\}$
frontier; sign-constraints	**Parametric Programming** $\min_{\alpha,\beta,\varepsilon}\left\{\sum_{l=1}^{n}\varepsilon_i^2 \middle\vert \begin{array}{l} y_i=\alpha+\boldsymbol{\beta}'\mathbf{x}_i+\varepsilon_i\ \forall i=1,\ldots,n \\ \varepsilon_i\le\mathbf{0}\ \forall i=1,\ldots,n \end{array}\right\}$	**Sign Constrained Convex Nonparametric Least Squares** $\min_{\alpha,\beta,\varepsilon}\left\{\sum_{l=1}^{n}\varepsilon_i^2 \middle\vert \begin{array}{l} y_i=\alpha_i+\boldsymbol{\beta}_i'\mathbf{x}_i+\varepsilon_i\ \forall i=1,\ldots,n; \\ \alpha_i+\boldsymbol{\beta}_i'\mathbf{x}_i\le\alpha_h+\boldsymbol{\beta}_h'\mathbf{x}_i\ \forall h,i=1,\ldots,n; \\ \boldsymbol{\beta}_i\ge\mathbf{0}\ \forall i=1,\ldots,n \\ \varepsilon_i\le\mathbf{0}\ \forall i=1,\ldots,n \end{array}\right\}$

Least squares interpretation

- Standard DEA problem
  - Single output, multiple inputs
  - Output orientation
  - Variable returns to scale

$$\max_{\boldsymbol{\lambda} \geq \mathbf{0}} \theta$$

$$\theta y_i \leq \sum_{h=1}^{n} \lambda_h y_h$$

$$\mathbf{x}_i \geq \sum_{h=1}^{n} \lambda_h \mathbf{x}_h$$

$$\sum_{h=1}^{n} \lambda_h = 1$$

Least squares interpretation

- Transform to additive output efficiency as

$$\max_{\boldsymbol{\lambda} \geq \mathbf{0}} \phi$$

$$y_i + \phi \leq \sum_{h=1}^{n} \lambda_h y_h$$

$$\mathbf{x}_i \geq \sum_{h=1}^{n} \lambda_h \mathbf{x}_h$$

$$\sum_{h=1}^{n} \lambda_h = 1$$

$$\theta^* = (\phi^* + y_i) / y_i$$

$$\phi^* = (\theta^* - 1) y_i$$

Least squares interpretation

- Dual problem

$$\min_{\boldsymbol{\alpha},\boldsymbol{\beta},\boldsymbol{\varepsilon}} (-\varepsilon_i)$$

$$s.t.$$

$$y_i = \alpha_i + \boldsymbol{\beta}_i'\mathbf{x}_i + \varepsilon_i$$

$$y_h \le \alpha_i + \boldsymbol{\beta}_i'\mathbf{x}_h \;\; \forall h = 1,\ldots,n$$

$$\boldsymbol{\beta}_i \ge \mathbf{0}$$

$$\varepsilon_i \le 0$$

Least squares interpretation

- Combine n LP problems to a single large LP problem

$$\min_{\boldsymbol{\alpha},\boldsymbol{\beta},\boldsymbol{\varepsilon}} \sum_{i=1}^{n} -\varepsilon_i$$

$$s.t.$$

$$y_i = \alpha_i + \boldsymbol{\beta}_i'\mathbf{x}_i + \varepsilon_i \quad \forall i = 1,\ldots,n$$

$$y_h \leq \alpha_i + \boldsymbol{\beta}_i'\mathbf{x}_h \quad \forall i,h = 1,\ldots,n$$

$$\boldsymbol{\beta}_i \geq \mathbf{0} \quad \forall i = 1,\ldots,n$$

$$\varepsilon_i \leq 0 \quad \forall i = 1,\ldots,n$$

Least squares interpretation

- Insert y_h into the concavity constraint and substitute indices *h,i*

$$\min_{\boldsymbol{\alpha},\boldsymbol{\beta},\boldsymbol{\varepsilon}} \sum_{i=1}^{n} -\varepsilon_i$$

$$s.t.$$

$$y_i = \alpha_i + \boldsymbol{\beta}_i'\mathbf{x}_i + \varepsilon_i$$

$$\alpha_i + \boldsymbol{\beta}_i'\mathbf{x}_i \le \alpha_h + \boldsymbol{\beta}_h'\mathbf{x}_i \;\; \forall h,i = 1,\ldots,n$$

$$\boldsymbol{\beta}_i \ge \mathbf{0} \;\; \forall i = 1,\ldots,n$$

$$\varepsilon_i \le 0 \;\; \forall i = 1,\ldots,n$$

Least squares interpretation

- Apply quadratic transformation to the epsilon

$$\min_{\boldsymbol{\alpha},\boldsymbol{\beta},\boldsymbol{\varepsilon}} \sum_{i=1}^{n} \varepsilon_i^2$$

$$s.t.$$

$$y_i = \alpha_i + \boldsymbol{\beta}_i'\mathbf{x}_i + \varepsilon_i$$

$$\alpha_i + \boldsymbol{\beta}_i'\mathbf{x}_i \le \alpha_h + \boldsymbol{\beta}_h'\mathbf{x}_i \;\; \forall h,i = 1,\ldots,n$$

$$\boldsymbol{\beta}_i \ge \mathbf{0}$$

$$\varepsilon_i \le 0 \;\; \forall i = 1,\ldots,n$$

Least squares interpretation

- DEA as a sign-constrained least-squares problem

$$
\begin{aligned}
&\min_{\boldsymbol{\alpha},\boldsymbol{\beta},\boldsymbol{\varepsilon}} \sum_{i=1}^{n} \varepsilon_i^2 \\
&s.t. \\
&y_i = \alpha_i + \boldsymbol{\beta}_i'\mathbf{x}_i + \varepsilon_i \\
&\alpha_i + \boldsymbol{\beta}_i'\mathbf{x}_i \le \alpha_h + \boldsymbol{\beta}_h'\mathbf{x}_i \;\; \forall h,i = 1,\ldots,n \\
&\boldsymbol{\beta}_i \ge \mathbf{0} \;\; \forall i = 1,\ldots,n \\
&\varepsilon_i \le 0 \;\; \forall i = 1,\ldots,n
\end{aligned}
$$

Least squares interpretation

- Relaxing the sign-constraint

$$\min_{\boldsymbol{\alpha},\boldsymbol{\beta},\boldsymbol{\varepsilon}} \sum_{i=1}^{n} \varepsilon_i^2$$

$$s.t.$$

$$y_i = \alpha_i + \boldsymbol{\beta}_i'\mathbf{x}_i + \varepsilon_i$$

$$\alpha_i + \boldsymbol{\beta}_i'\mathbf{x}_i \le \alpha_h + \boldsymbol{\beta}_h'\mathbf{x}_i \;\; \forall h,i = 1,\ldots,n$$

$$\boldsymbol{\beta}_i \ge \mathbf{0} \;\; \forall i = 1,\ldots,n$$

- The CNLS problem of Hildreth (1954) and Hanson and Pledger (1976) in the multivariate case: Kuosmanen (2008), *Econometric Journal*.

Importance of the LS interpretation

- Enhances the statistical foundation of DEA
  - DEA is not so different from regression analysis
- Integration of tools and techniques from the regression analysis to DEA
  - Goodness of fit statistics (R^2)
  - Stochastic noise term
  - Contextual variables
  - Panel data modeling
  - Etc.

Implications of Regression Formulation

Kuosmanen and Johnson (2010) Proposition 3.1 reveals new possibilities for adapting tools and concepts of regression analysis to the DEA framework. For example, DEA lacks a meaningful goodness-of-fit statistic. Given the least-squares formulation derived in this paper, we could apply the standard coefficient of determination

$$R^2 = \frac{\sum_{i=1}^{n}(\hat{y}_i - \bar{y})^2}{\sum_{i=1}^{n}(y_i - \bar{y})^2} = 1 - \frac{\sum_{i=1}^{n}(\varepsilon_i)^2}{\sum_{i=1}^{n}(y_i - \bar{y})^2} = 1 - \frac{\sum_{i=1}^{n}((1-\theta_i)y_i)^2}{\sum_{i=1}^{n}(y_i - \bar{y})^2}$$

Least-squares formulation of DEA

Kuosmanen & Johnson (2010) *Oper. Res.*

- DEA can be formulated as the sign-constrained nonparametric least squares problem

$$\min_{\alpha,\boldsymbol{\beta},\varepsilon} \sum_{i=1}^{n} \varepsilon_i^2$$

$$y_i = \alpha_i + \boldsymbol{\beta}_i'\mathbf{x}_i + \varepsilon_i \quad \forall i \qquad \text{(regression equation)}$$

$$\alpha_i + \boldsymbol{\beta}_i'\mathbf{x}_i \le \alpha_h + \boldsymbol{\beta}_h'\mathbf{x}_i \quad \forall h,i \qquad \text{(concavity)}$$

$$\boldsymbol{\beta}_i \ge \mathbf{0} \quad \forall i \qquad \text{(monotonicity)}$$

$$\varepsilon_i \le 0 \quad \forall i$$

- Least squares problem solved simultaneusly for all firms.
- DEA is a nonparametric, axiomatic counterpart to Aigner & Chu (1968) parametric programming

Stochastic non-parametric envelopment of data (StoNED)

Kuosmanen & Kortelainen (2012) *JPA*

Encompassing frontier model:

$$y = f(\mathbf{x}) - u + v$$

- *f* is a monotonic increasing and concave frontier production function (possibly CRS)
- *u* is an asymmetric inefficiency term (half-normal)
- *v* is the random noise term (normal)

Note:

- DEA model (Banker, 1993) obtained by setting $\sigma_v = 0$.
- SFA model (ALS ’77) obtained by setting $f(\mathbf{x}) = \mathbf{x}'\boldsymbol{\beta}$

Stochastic non-parametric envelopment of data (StoNED)

Kuosmanen & Kortelainen (2012) *JPA*

Stepwise approach (analogous to MOLS):

1) CNLS estimation:

$$\min_{\alpha,\boldsymbol{\beta},\varepsilon} \sum_{i=1}^{n} \varepsilon_i^2$$

$$y_i = \alpha_i + \boldsymbol{\beta}_i'\mathbf{x}_i + \varepsilon_i \quad \forall i \qquad \text{(regression equation)}$$

$$\alpha_i + \boldsymbol{\beta}_i'\mathbf{x}_i \le \alpha_h + \boldsymbol{\beta}_h'\mathbf{x}_i \quad \forall h,i \qquad \text{(concavity)}$$

$$\boldsymbol{\beta}_i \ge \mathbf{0} \quad \forall i \qquad \text{(monotonicity)}$$

2) Method of moments (ALS ’77) or pseudolikelihood estimation (Fan et al., 1996) of standard deviations σ_u , σ_v

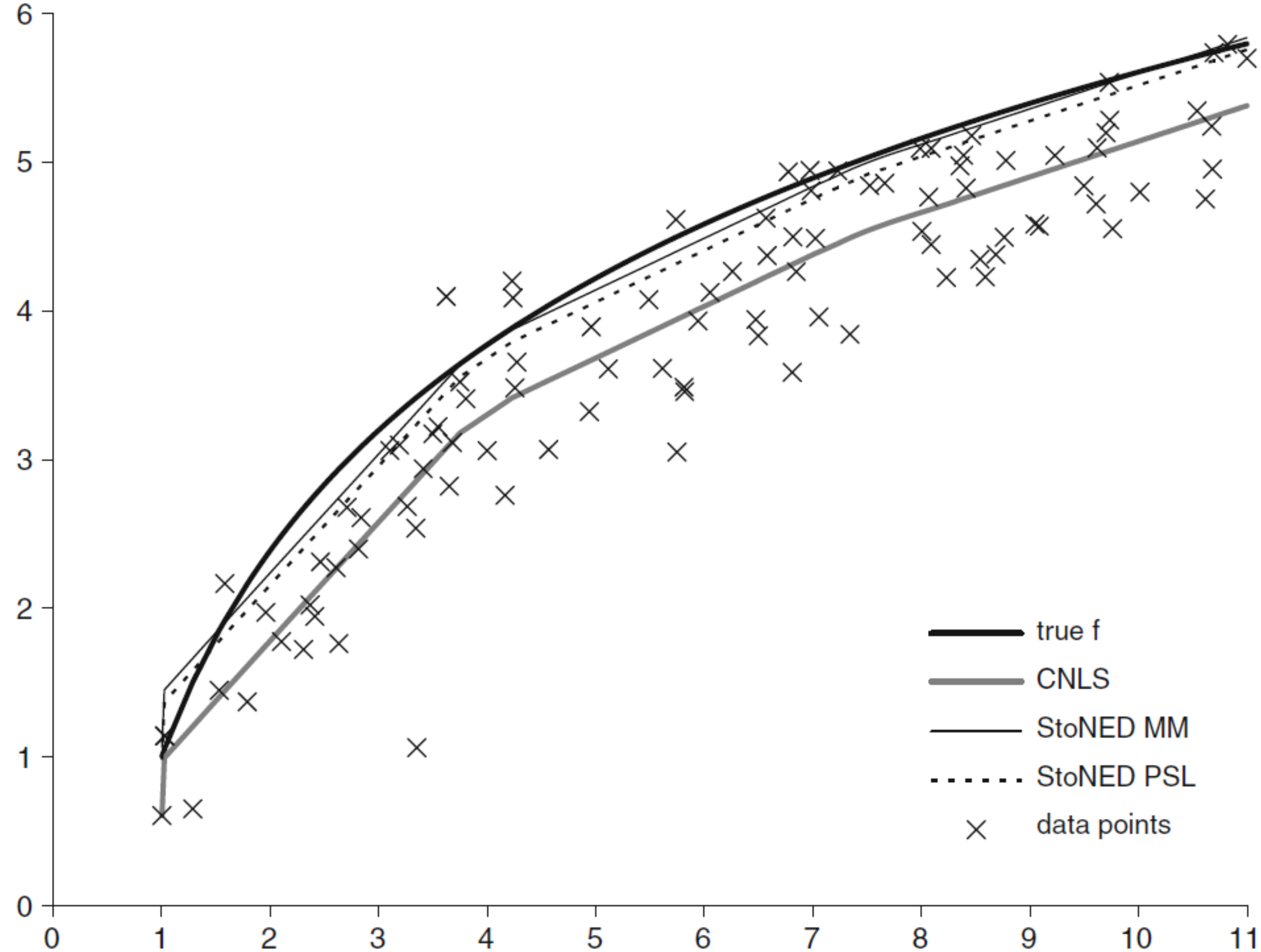
3) Shift the estimated curve upward by expected inefficiency.

JLMS estimator can be used for estimating the conditional expectation $E(u_i \mid \varepsilon_i)$

Stochastic non-parametric envelopment of data (StoNED)

Kuosmanen & Kortelainen (2012) *JPA*

Fig. 1 Graphical illustration of the CNLS regression curve and the StoNED frontiers. The data generation process is $y_i = \ln(x_i) + 2 + v_i - u_i$, where $v_i \underset{i.i.d}{\sim} N(0, 0.6^2)$ and $u_i \underset{i.i.d}{\sim} |N(0, 0.3^2)|n$



StoNED vs. DEA

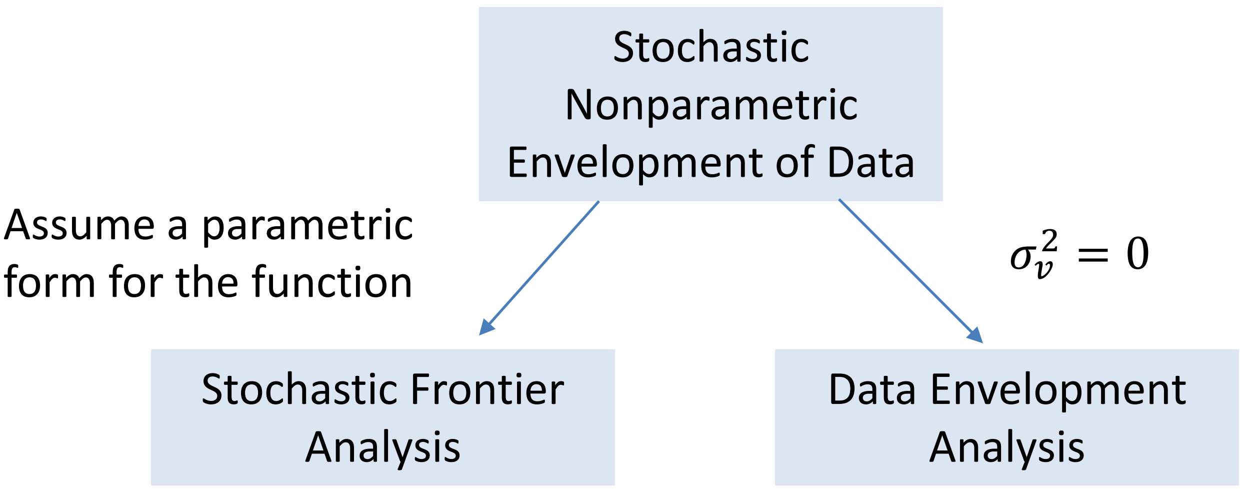
Similarities

- Nonparametric frontier
- Monotonicity and concavity
- No smoothness assumptions
- Math programming

Differences

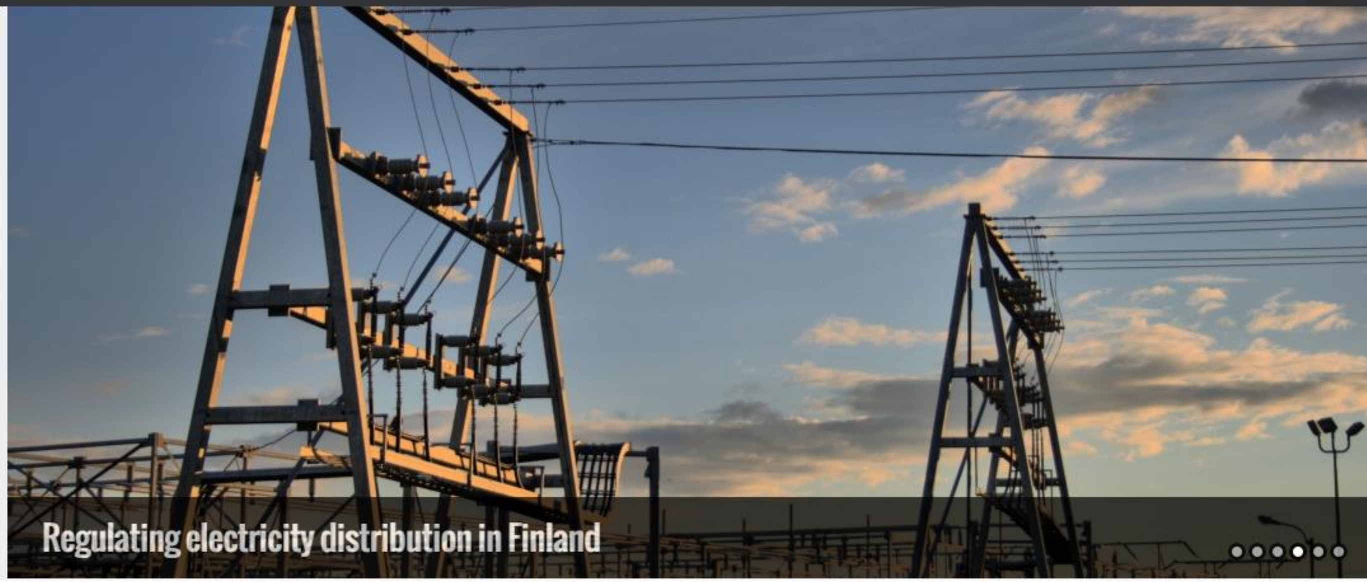
- Sign-constraint on residuals
- Inefficient observations matter
- Relax the strong assumption of ”no noise”
- Probabilistic treatment of inefficiency and noise

StoNED generalizes DEA and SFA



Andrew (Andy) L. Johnson, Ph.D.

Texas A&M University



News

[Plenary Talk at European Workshop on Efficiency and Productivity Analysis (EWEPA) XIV]
June 26th, 2015

The largest conference in the field of efficiency and productivity analysis is the European Workshop on Efficiency and Productivity Analysis (EWEPA) [...]

[Informs Annual Conference - Nov 1-4 Philadelphia PA - DEA Cluster]
June 25th, 2015

Seminars/presentation

- [November 9th – Informs Annual Conference: A Multivariate Seminonparametric Bayesian Concave Regression Method to Estimate Stochastic Frontiers]

  This presentation discusses a method that incorporates the latest advances in the Bayesian constrained regression literature offering an alternative to the current Least Squares-based and Kernel Regression-based Stochastic frontier constrained estimation methods, both in terms of runtime and of data capacity.

- [October 4 and 5: College Industry Council on

Ongoing work

- [Multi-variate Bayesian Convex Regression with Inefficiency]

  This research builds in Hannah and Dunson's Multi-variate Bayesian Convex Regression to develop a method to estimate a shape constrained production functions and potential deviations from the function representing inefficiency.

- [Shape Restricted Estimation of the Power Curve for a Wind Turbine]

  The estimation of the power curve provides an application for methods to estimate production

Thank you for your attention

- Further information available online at the

  StoNED homepage:
  http://www.nomepre.net/stoned

  Johnson lab group homepage:
  http://www.andyjohnson.guru