Econometrics 2

Lecture 9: LAD, Linear Programming & Algorithmic Solution

AUEB | Spring Semester 2026

Taught by: Prof. S. Arvanitis

Notes transcribed & digitized by: Ch. Dourgoutis & T. Kourtalis

These notes cover the handwritten material from the lecture session. They are supplementary to the official slides of Prof. S. Arvanitis, available at:

Official LAD Slides (eClass)

---

1. Why the Conditional Median? (Comments)

(Comments made by the professor during the lecture.)

The question is: why should we care about finding $\text{Med}(Y_{(i)} \mid X_n)$? Two reasons were given:

1. It may interest us **independently**, for example in questions that concern income distributions.

2. **Statistical difficulties** related to the properties of the distribution of $\varepsilon_{(i)} \mid X_n$. For example, if the conditional distribution behaves "as towards extreme values", in the sense that as $x \to +\infty$:

$$P\big(|\varepsilon_{(i)}| > x \mid X_n\big) \sim \frac{c}{x^{\alpha}}, \qquad c > 0, \quad \alpha \in (0, \bar{\alpha})$$

   When $\alpha \leq 2$: the $\text{Var}(\varepsilon_{(i)} \mid X_n)$ does not exist.
   When $\alpha \leq 1$: neither does $\mathbb{E}(\varepsilon_{(i)} \mid X_n)$.

   Consequently, problems can arise in the interpretation of $X_n\beta_0$ (when $0 \leq \alpha \leq 1$, it will not be the conditional mean of Y_n given X_n), and difficulties arise with the asymptotic properties of the OLSE:

   (a) When $\alpha \leq 1 \Rightarrow$ the OLSE is **not consistent**.

(b) When $1 < \alpha \leq 2 \Rightarrow$ the OLSE will have a **slower rate of convergence** to β_0 than the classical $\sqrt{n}$, and will **not** be asymptotically normal. (The relevant limiting theorems are the Central Limit Theorems and the Limit Theorems for α-stable distributions, specifically the Central & Boundary Theorems.)

(Note: Under some assumptions, even when $\alpha < 1$, the LADE can still be consistent, with rate $\sqrt{n}$ and asymptotically normal.)

2. Property of Absolute Value

We recall the decomposition of the absolute value, which will be needed for the LP reformulation below.

$$|x| = \begin{cases} -x, & x < 0 \\ x, & x \geq 0 \end{cases} \;=\; -x \cdot \mathbf{1}_{x<0} + x \cdot \mathbf{1}_{x\geq 0}$$

Define the indicator functions:

$$\mathbf{1}_{x<0} = \begin{cases} 1, & x < 0 \\ 0, & x \geq 0 \end{cases} \qquad \mathbf{1}_{x\geq 0} = \begin{cases} 0, & x < 0 \\ 1, & x \geq 0 \end{cases}$$

Therefore any real number x can be written as:

$$x = x \cdot \mathbf{1}_{x\geq 0} + x \cdot \mathbf{1}_{x<0}$$

i.e., decomposed into its positive part $x^+ := x \cdot \mathbf{1}_{x\geq 0} \geq 0$ and negative part $x^- := -x \cdot \mathbf{1}_{x<0} \geq 0$, so that $x = x^+ - x^-$ and $|x| = x^+ + x^-$.

3. From LAD to Linear Programming (Part I)

(Continuation from the slides, cf. the official slides for the full formal statement.)

We showed in the slides that:

$$|Y_{(i)} - X_{(i)}\beta| = (Y_{(i)} - X_{(i)}\beta) \cdot \mathbf{1}_{Y_{(i)} \geq X_{(i)}\beta} + (X_{(i)}\beta - Y_{(i)}) \cdot \mathbf{1}_{Y_{(i)} < X_{(i)}\beta}$$

Now introduce auxiliary variables:

$$u^+_{(i)} := (Y_{(i)} - X_{(i)}\beta)\, \mathbf{1}_{(Y_{(i)} \geq X_{(i)}\beta)}, \qquad u^-_{(i)} := (Y_{(i)} - X_{(i)}\beta)\, \mathbf{1}_{(Y_{(i)} < X_{(i)}\beta)}$$

(Note: using the notation from the slides, $u^-_{(i)} := (X_{(i)}\beta - Y_{(i)})\, \mathbf{1}_{(Y_{(i)} < X_{(i)}\beta)}$.)

Then we have $Y_{(i)} - X_{(i)}\beta = u^+_{(i)} - u^-_{(i)}$ with $u^+_{(i)}, u^-_{(i)} \geq 0$, and therefore:

$$|Y_{(i)} - X_{(i)}\beta| = u^+_{(i)} + u^-_{(i)}$$

By defining $\beta = \beta^+ - \beta^-$ where $\beta^+, \beta^- \geq 0$, this gives us:

$$\hat{\beta}_{LAD} \;\in\; \arg\min_{\beta \in \Theta} \sum_{i=1}^{n} |Y_{(i)} - X_{(i)}\beta| \;=\; \arg \min_{\beta^+, \beta^- \geq 0} \sum_{i=1}^{n} \big(u^+_{(i)} + u^-_{(i)}\big)$$

subject to:

$$Y_{(i)} - X_{(i)}\beta^+ + X_{(i)}\beta^- = u^+_{(i)} - u^-_{(i)}, \qquad u^+_{(i)},\, u^-_{(i)} \geq 0$$

This is a **Linear Programming problem** (Γ Π).

4. Canonical Form of a Linear Program

A linear program in canonical form is:

$$\min_{x \in \mathbb{R}^m} c'x \qquad \text{subject to} \quad Ax = b,\ x \geq 0$$

where the inequalities between vectors are interpreted componentwise.

In the LAD case with $\Theta = \mathbb{R}^p$, we set $m = 2p + 2n$ and:

$$x = \begin{pmatrix} \beta^+ \\ \beta^- \\ u^+ \\ u^- \end{pmatrix}, \quad c = \begin{pmatrix} 0_{p\times 1} \\ 0_{p\times 1} \\ \mathbf{1}_{n\times 1} \\ \mathbf{1}_{n\times 1} \end{pmatrix}, \quad A = \begin{pmatrix} X_n & -X_n & I_{n\times n} & -I_{n\times n} \end{pmatrix}, \quad b = Y_n$$

5. Monte Carlo Setup

(As described in the handwritten notes and elaborated in the slides.)

We consider the model:

$$Y_{(i)} = X_{(i)}\beta + \varepsilon_{(i)} = (1,\, z_{(i)})(\beta_0,\, \beta_1)' + \varepsilon_{(i)}$$

Single-valued (generic) design:

$$z_{(i)} \sim N(0,1), \qquad \varepsilon_{(i)} \sim \text{Laplace}(0,\sigma), \qquad \text{iid}$$

Multi-valued (degenerate) design:

$$z_{(i)} \in \{-1,\, 0,\, 1\}, \quad \#\{i : z_{(i)} = k\} \approx \tfrac{n}{3}, \quad \varepsilon_{(i)} \in \{-\sigma,\, +\sigma\}, \quad P(\varepsilon_{(i)} = \sigma) = P(\varepsilon_{(i)} = -\sigma) = \tfrac{1}{2}$$

*(Note: In the degenerate design the median of $\varepsilon_{(i)}$ is not unique, see slides for a full discussion of the consequences for identification and for why the two solvers, **`highs-ds`** and **`highs-ipm`**, may return different solutions.)*