Inference for Large-Scale Systems of Linear Inequalities

Zheng Fang	Andres Santos	Azeem Shaikh	Alex Torgovitsky
TAMU	UCLA	U. Chicago	U. Chicago

March 30, 2022

The Question

Let i.i.d. sample $\{Z_i\}_{i=1}^n$ with $Z \sim P \in \mathbf{P}$ and suppose there is a parameter

$\beta(P) \in \mathbf{R}^p$ that is unknown but estimable

The Question

Let i.i.d. sample $\{Z_i\}_{i=1}^n$ with $Z \sim P \in \mathbf{P}$ and suppose there is a parameter

$\beta(P) \in \mathbf{R}^p$ that is unknown but estimable

We aim to test whether distribution P satisfies the following null hypothesis

$$H_0 : P \in \mathbf{P}_0 \qquad H_1 : P \in \mathbf{P} \setminus \mathbf{P}_0$$

where

$$\mathbf{P}_0 \equiv \{P \in \mathbf{P} : \beta(P) = Ax \text{ for some } x \geq 0\}$$

Key Structure

- The $p \times d$ matrix A is known.
- $x \geq 0$ with $x \in \mathbf{R}^d$ denotes all coordinates of x are non-negative.

Example: Nevo et al. (2016)

Type $h \in \{1, \ldots, H\}$ consumer, data plans $k \in \{1, \ldots, K\}$, time t utility

$$u_h(c_t, y_t, v_t; k) = v_t(\frac{c_t^{1-\zeta_h}}{1-\zeta_h}) - c_t(\kappa_{1h} + \frac{\kappa_{2h}}{\log(s_k)}) + y_t$$

for i.i.d. shock v_t, data usage c_t, data speed s_k, numeraire good y_t.

Example: Nevo et al. (2016)

Type $h \in \{1, \ldots, H\}$ consumer, data plans $k \in \{1, \ldots, K\}$, time t utility

$$u_h(c_t, y_t, v_t; k) = v_t(\frac{c_t^{1-\zeta_h}}{1-\zeta_h}) - c_t(\kappa_{1h} + \frac{\kappa_{2h}}{\log(s_k)}) + y_t$$

for i.i.d. shock v_t, data usage c_t, data speed s_k, numeraire good y_t.

For overage price p_k, fee F_k, data allowance $\bar{C}_k$, type h utility from plan k is

$$\max_{c_1,\ldots,c_T} \sum_{t=1}^{T} E_h[u_h(c_t, y_t, v_t; k)]$$

s.t. $F_k + p_k \max\{C_T - \bar{C}_k, 0\} + Y_T \leq I,\ C_T = \sum_{t=1}^{T} c_t,\ Y_T = \sum_{t=1}^{T} y_t$

Example: Nevo et al. (2016)

For Z observed plan choice and data usage, and m known moment function

$$E_P[m(Z)] = \sum_{h=1}^{H} E_h[m(Z)]x_h$$

where $x = (x_1, \ldots, x_H)$ are unknown proportions of each type in population.

Example: Nevo et al. (2016)

For Z observed plan choice and data usage, and m known moment function

$$E_P[m(Z)] = \sum_{h=1}^{H} E_h[m(Z)]x_h$$

where $x = (x_1, \ldots, x_H)$ are unknown proportions of each type in population.

Goal: Inference on counterfactual demand, which for known a_h equals

$$\sum_{h=1}^{H} a_h x_h$$

Example: Nevo et al. (2016)

For Z observed plan choice and data usage, and m known moment function

$$E_P[m(Z)] = \sum_{h=1}^{H} E_h[m(Z)]x_h$$

where $x = (x_1, \ldots, x_H)$ are unknown proportions of each type in population.

Goal: Inference on counterfactual demand, which for known a_h equals

$$\sum_{h=1}^{H} a_h x_h$$

Current Approach

- Build large grid of types, solve $E_h[m(Z)]$ for each type.
- Estimate proportions $x = (x_1, \ldots, x_h)$ by constrained GMM.
- Inference via bootstrap ... but bootstrap can fail.

Example: Nevo et al. (2016)

Instead, test if counterfactual demand equals hypothesized λ by testing if

$$\beta(P) = Ax \text{ for some } x \geq 0$$

with

$$\beta(P) \equiv \begin{pmatrix} E_P[m(Z)] \\ 1 \\ \lambda \end{pmatrix} \qquad A \equiv \begin{pmatrix} E_1[m(Z)] & \cdots & E_H[m(Z)] \\ 1 & \cdots & 1 \\ a_1 & \cdots & a_H \end{pmatrix}$$

Comments

- Confidence region through test inversion (in λ).
- We do not require proportion of types to be identified.
- In Nevo et al. (2016) $p \approx 120000$ and $d \approx 16800$.

Example: Honore and Lleras-Muney (2006)

Impact of War on Cancer

- (S_1, S_2) competing risks (e.g. cardio vascular disease and cancer).
- D an indicator for whether war on cancer policy in effect.
- Unspecified distribution for (S_1, S_2), and for unknown α and β assume

$$(T^*, I) = \begin{cases} (\min\{S_1, S_2\}, \arg\min\{S_1, S_2\}) & \text{if } D = 0 \\ (\min\{\alpha S_1, \beta S_2\}, \arg\min\{\alpha S_1, \beta S_2\}) & \text{if } D = 1 \end{cases}$$

Example: Honore and Lleras-Muney (2006)

Impact of War on Cancer

- (S_1, S_2) competing risks (e.g. cardio vascular disease and cancer).
- D an indicator for whether war on cancer policy in effect.
- Unspecified distribution for (S_1, S_2), and for unknown α and β assume

$$(T^*, I) = \begin{cases} (\min\{S_1, S_2\}, \arg\min\{S_1, S_2\}) & \text{if } D = 0 \\ (\min\{\alpha S_1, \beta S_2\}, \arg\min\{\alpha S_1, \beta S_2\}) & \text{if } D = 1 \end{cases}$$

Partial Identification

- We see (T, D, I) where T is interval censored version of T^*.
- Parameter (α, β) partially identified (even without interval censoring).

Goal: Construct confidence region for identified set for (α, β).

Example: Honore and Lleras-Muney (2006)

Key: (α, β) in the identified set iff there is some distribution p on $\mathcal{S}(\alpha, \beta)$ with

$$\sum_{(s_1, s_2) \in \mathcal{S}_{k,i,d}(\alpha,\beta)} p(s_1, s_2) = P(T = t_k,\ I = i | D = d)$$

where $\mathcal{S}(\alpha, \beta)$, $\mathcal{S}_{k,i,d}(\alpha, \beta) \subseteq \mathcal{S}(\alpha, \beta)$ are finite sets depending on (α, β).

Example: Honore and Lleras-Muney (2006)

Key: (α, β) in the identified set iff there is some distribution p on $\mathcal{S}(\alpha, \beta)$ with

$$\sum_{(s_1,s_2)\in\mathcal{S}_{k,i,d}(\alpha,\beta)} p(s_1, s_2) = P(T = t_k,\ I = i|D = d)$$

where $\mathcal{S}(\alpha, \beta)$, $\mathcal{S}_{k,i,d}(\alpha, \beta) \subseteq \mathcal{S}(\alpha, \beta)$ are finite sets depending on (α, β).

For Confidence Region

- Map $\beta(P)$ into conditional probabilities (and adding up restriction).
- Map x into unknown distribution p satisfying restriction.
- For each candidate (α, β) sets $\mathcal{S}_{k,i,d}(\alpha, \beta)$ map into matrix A.
- Test null hypothesis that (α, β) is in identified set by testing whether

$$\beta(P) = Ax \text{ for some } x \geq 0$$

Additional Applications

Treatment Effects

Balke & Pearl (1994, 1997), Angrist & Imbens (1995), Kline & Walters (2016), Laffers (2019), Machado, Shaikh & Vytlacil (2019), Kamat (2019)

Feasibility of Linear Program

Honore & Lleras-Muney (2006), Honore & Tamer (2006), Torgotvitsky (2019), Tebaldi, Torgovisky & Yang (2019).

Revealed Preferences

Manski (2014), Deb, Kitamura, Quah & Stoye (2017), Kitamura & Stoye (2018), Lazzati, Quah & Shirai (2018).

Key Challenge: "Large" p and $d \Rightarrow$ Computational scalability important

Related Literature

Moment Inequalities

- $P \in \mathbf{P}_0$ if and only if $\beta(P)$ is in set defined by inequalities (in $\mathbf{R}^p$).
- Challenge: For large p, d, computing inequalities is prohibitive.

Related Literature

Moment Inequalities

- $P \in \mathbf{P}_0$ if and only if $\beta(P)$ is in set defined by inequalities (in $\mathbf{R}^p$).
- Challenge: For large p, d, computing inequalities is prohibitive.

Shape Restrictions

- $P \in \mathbf{P}_0$ if and only if $\beta(P)$ is in convex set.
- We employ specific structure in computation and assumptions.

Related Literature

Moment Inequalities

- $P \in \mathbf{P}_0$ if and only if $\beta(P)$ is in set defined by inequalities (in $\mathbf{R}^p$).
- Challenge: For large p, d, computing inequalities is prohibitive.

Shape Restrictions

- $P \in \mathbf{P}_0$ if and only if $\beta(P)$ is in convex set.
- We employ specific structure in computation and assumptions.

Other Related Work

- Kitamura and Stoye (2018) test imposes restrictions on A (satisfied in the revealed preferences problem that motivates them).
- Andrews, Pakes $\&$ Roth (2019) find least favorable for subvector inference in a class of (conditional) moment inequalities models.
- Cox $\&$ Shi (2021) derive tuning parameter free method for inference.

1. The Geometry
2. The Test
3. Simulations

Some Notation

Question: For any $\beta \in \mathbf{R}^p$, when is $\beta = Ax$ for some $x \geq 0$?

Some Notation

Question: For any $\beta \in \mathbf{R}^p$, when is $\beta = Ax$ for some $x \geq 0$?

Three Subspaces

$$R \equiv \{b \in \mathbf{R}^p : b = Ax \text{ for some } x \in \mathbf{R}^d\}$$
$$N \equiv \{x \in \mathbf{R}^d : Ax = 0\}$$
$$N^\perp \equiv \{y \in \mathbf{R}^d : \langle y, x \rangle = 0 \text{ for all } x \in N\}$$

Some Notation

Question: For any $\beta \in \mathbf{R}^p$, when is $\beta = Ax$ for some $x \geq 0$?

Three Subspaces

$$R \equiv \{b \in \mathbf{R}^p : b = Ax \text{ for some } x \in \mathbf{R}^d\}$$
$$N \equiv \{x \in \mathbf{R}^d : Ax = 0\}$$
$$N^\perp \equiv \{y \in \mathbf{R}^d : \langle y, x \rangle = 0 \text{ for all } x \in N\}$$

Some Intuition

- If $\beta = Ax$ text for some $x \geq 0$, then in particular we must have that ...

$$\beta \in R$$

Some Notation

Question: For any $\beta \in \mathbf{R}^p$, when is $\beta = Ax$ for some $x \geq 0$?

Three Subspaces

$$R \equiv \{b \in \mathbf{R}^p : b = Ax \text{ for some } x \in \mathbf{R}^d\}$$
$$N \equiv \{x \in \mathbf{R}^d : Ax = 0\}$$
$$N^\perp \equiv \{y \in \mathbf{R}^d : \langle y, x \rangle = 0 \text{ for all } x \in N\}$$

Some Intuition

- If $\beta = Ax$ text for some $x \geq 0$, then in particular we must have that ...

$$\beta \in R$$

- If $\beta = Ax_1$ for some x_1 and $x_2 \in N$ then $\beta = A(x_1 + x_2)$ so ...

$\Rightarrow$ Intuitively, if $x_1 \ngeq 0$, then maybe can fix it by moving along N

Simple Lemma

Lemma: If $\beta \in R$, then there is unique $x^\star \in N^\perp$ satisfying the equality

$$\beta = Ax^\star$$

Simple Lemma

Lemma: If $\beta \in R$, then there is unique $x^\star \in N^\perp$ satisfying the equality

$$\beta = Ax^\star$$

Key Implications

- $\beta = Ax$ with $x \geq 0$ requires $\beta \in R$.

Simple Lemma

Lemma: If $\beta \in R$, then there is unique $x^\star \in N^\perp$ satisfying the equality

$$\beta = Ax^\star$$

Key Implications

- $\beta = Ax$ with $x \geq 0$ requires $\beta \in R$.
- Moreover, the above lemma implies set of solutions to $\beta = Ax$ equals

$$\{x \in \mathbf{R}^d : Ax = \beta\} = x^\star + N$$

Simple Lemma

Lemma: If $\beta \in R$, then there is unique $x^\star \in N^\perp$ satisfying the equality

$$\beta = Ax^\star$$

Key Implications

- $\beta = Ax$ with $x \geq 0$ requires $\beta \in R$.
- Moreover, the above lemma implies set of solutions to $\beta = Ax$ equals

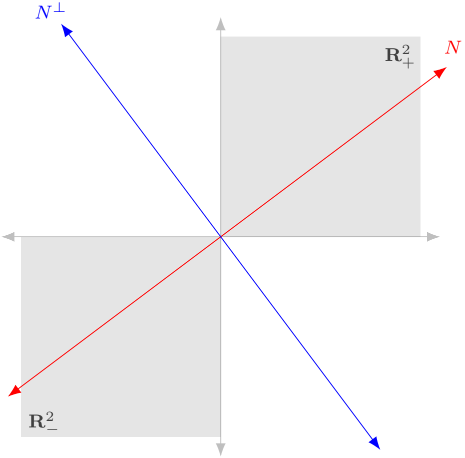
$$\{x \in \mathbf{R}^d : Ax = \beta\} = x^\star + N$$

- Whether $\beta = Ax$ for some $x \geq 0$ characterized by $\beta \in \mathbf{R}^p$, $x^\star \in \mathbf{R}^d$ via

(i) $\beta \in R$ (ii) $\{x^\star + N\} \cap \mathbf{R}^d_+ \neq \emptyset$

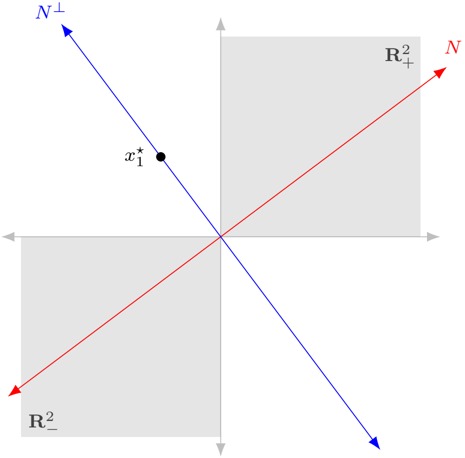
Key Challenge: Obtaining tractable characterization for (ii).

Geometric Intuition



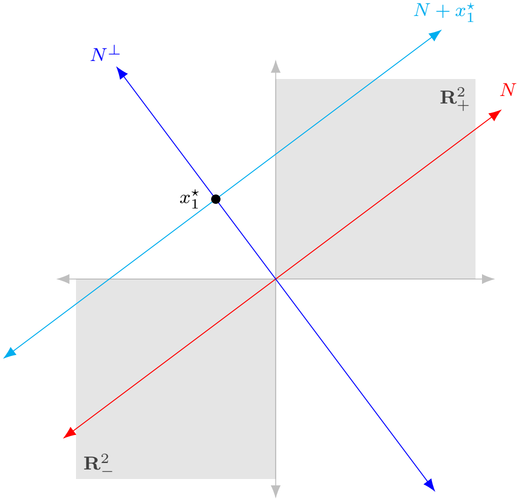
Geometric Intuition

Example: Suppose $\beta = Ax_1^\star$ with $x_1^\star \in N^\perp$ ... is there positive solution?



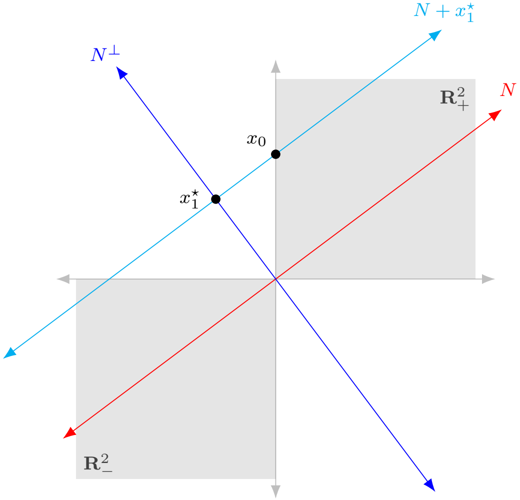
Geometric Intuition

Example: Suppose $\beta = Ax_1^\star$ with $x_1^\star \in N^\perp$ ... is there positive solution?



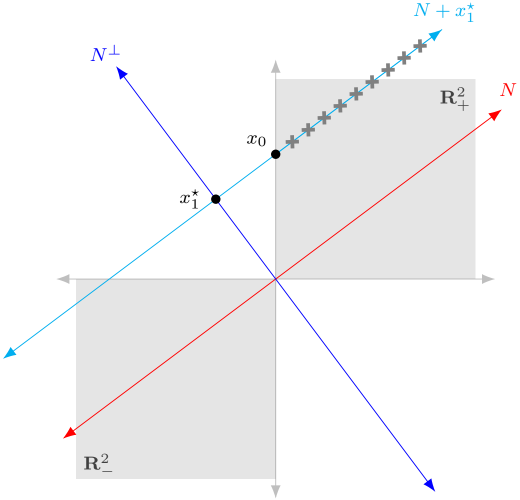
Geometric Intuition

Example: Suppose $\beta = Ax_1^\star$ with $x_1^\star \in N^\perp$ ... is there positive solution?



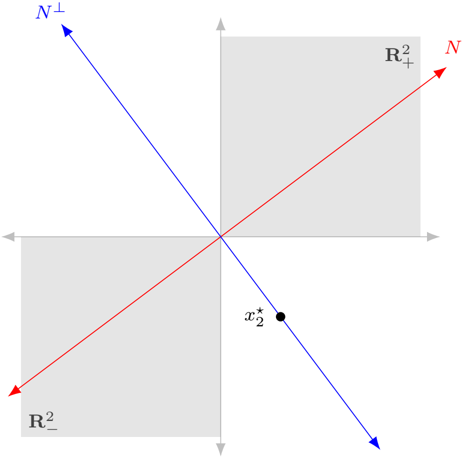
Geometric Intuition

Example: Suppose $\beta = Ax_1^\star$ with $x_1^\star \in N^\perp$ ... is there positive solution?



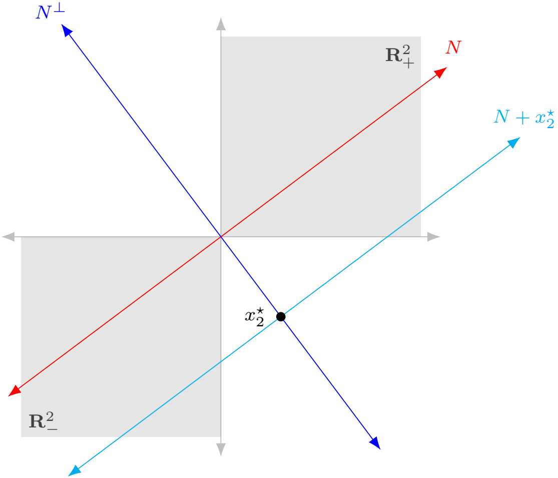
Geometric Intuition

Question: What if instead $\beta = Ax_2^\star$ with $x_2^\star \in N^\perp$?



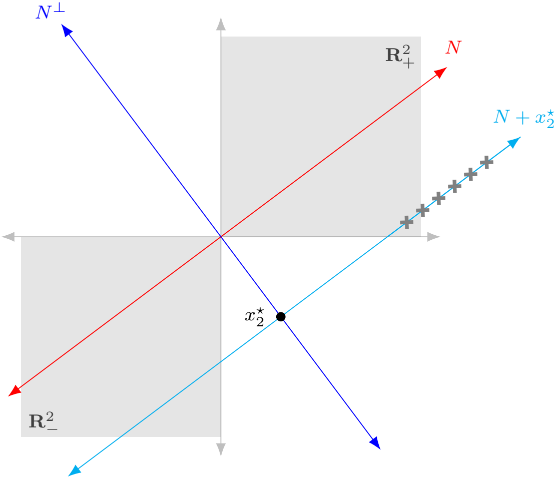
Geometric Intuition

Question: What if instead $\beta = Ax_2^\star$ with $x_2^\star \in N^\perp$?



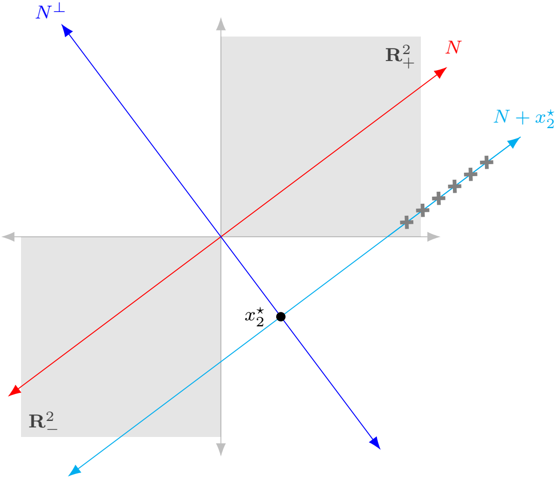
Geometric Intuition

Question: What if instead $\beta = Ax_2^\star$ with $x_2^\star \in N^\perp$?

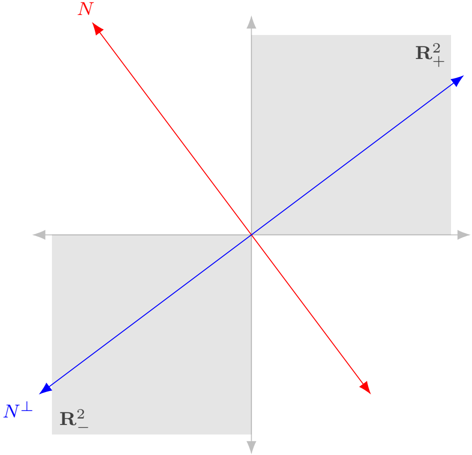


Geometric Intuition

Note: In this example positive solution always exists (provided $\beta \in R$)

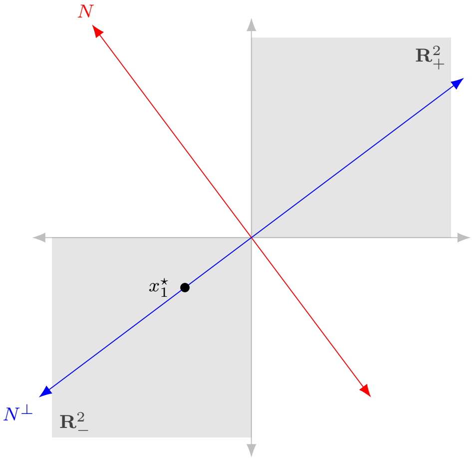


Geometric Intuition



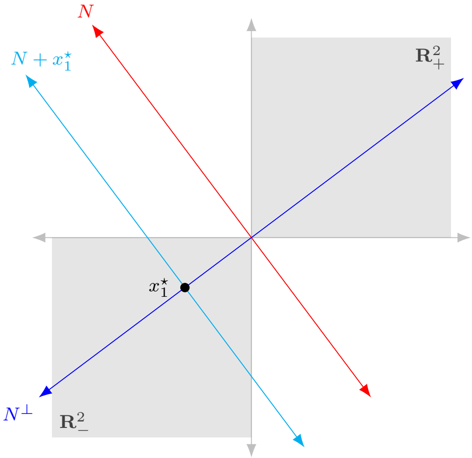
Geometric Intuition

Example: Suppose $\beta = Ax_1^\star$ with $x_1^\star \in N^\perp$ ... is there a positive solution?



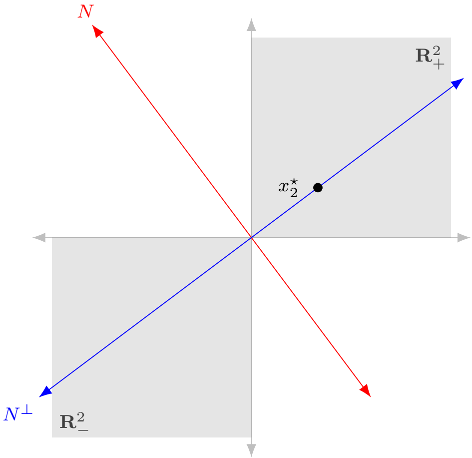
Geometric Intuition

Example: Suppose $\beta = Ax_1^\star$ with $x_1^\star \in N^\perp$ ... is there a positive solution?



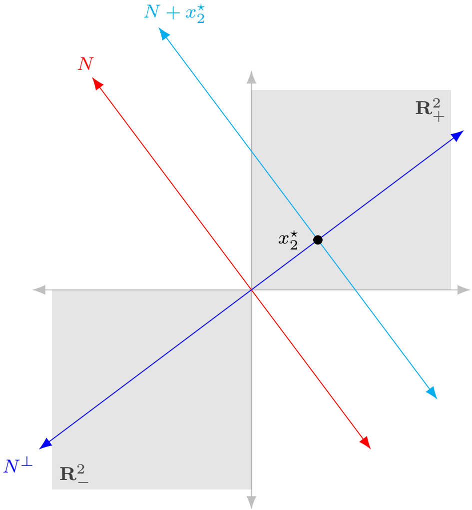
Geometric Intuition

Question: What if instead $\beta = Ax_2^\star$ with $x_2^\star \in N^\perp$?



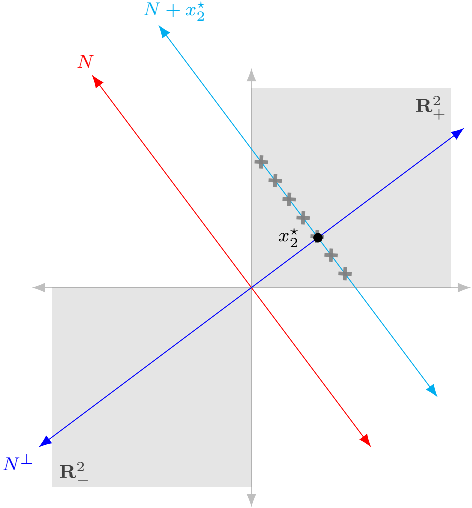
Geometric Intuition

Question: What if instead $\beta = Ax_2^\star$ with $x_2^\star \in N^\perp$?



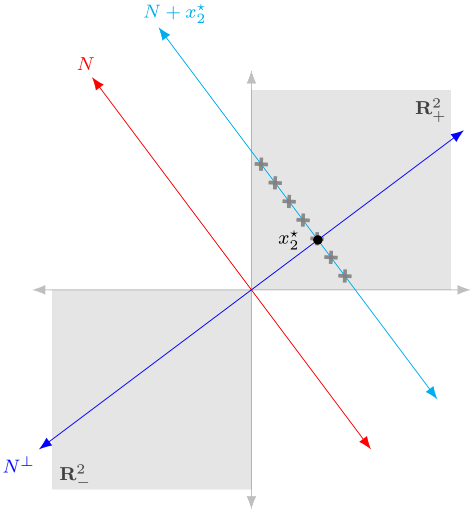
Geometric Intuition

Question: What if instead $\beta = Ax_2^\star$ with $x_2^\star \in N^\perp$?



Geometric Intuition

Note: Positive solution exists if and only if $x^\star \in \mathbf{R}^2_+$ (provided $\beta \in R$)



Geometric Characterization

(i) $\beta \in R$ (ii) $\{x^\star + N\} \cap \mathbf{R}^d_+ \neq \emptyset$

Goal: Obtain alternative characterization that suggests natural test statistic.

Geometric Characterization

(i) $\beta \in R$ (ii) $\{x^\star + N\} \cap \mathbf{R}^d_+ \neq \emptyset$

Goal: Obtain alternative characterization that suggests natural test statistic.

Theorem: There is an $x_0 \in \mathbf{R}^d_+$ satisfying $Ax_0 = \beta$ if and only if

(i) $\beta \in R$ (ii) $\langle s, x^\star \rangle \leq 0$ for all $s \in N^\perp \cap \mathbf{R}^d_-$

Comments

- Condition (i) yields "equalities" and (ii) yields "inequalities."
- (ii) equivalent to angles between $x^\star$ and $N^\perp \cap \mathbf{R}^d_-$ are obtuse.
- Reflects dependence on $x^\star$ and "orientation" of $N^\perp$ in $\mathbf{R}^d$.

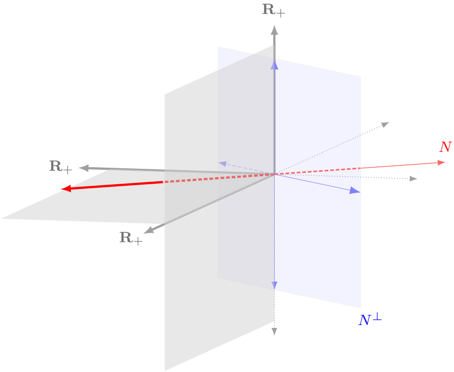
Geometric Intuition

$\langle s, x^\star \rangle \leq 0$ for all $s \in N^\perp \cap \mathbf{R}^d_-$

Geometric Intuition

$$\langle s, x^\star \rangle \leq 0 \text{ for all } s \in N^\perp \cap \mathbf{R}^d_-$$

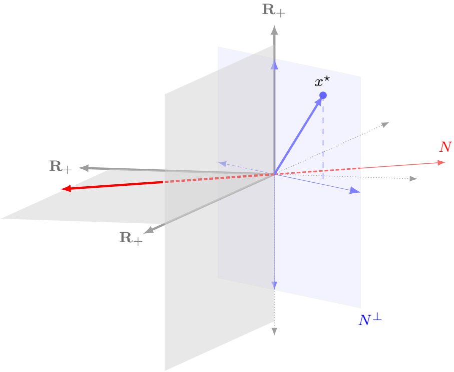
Equivalent angle between $x^\star$ and $N^\perp \cap \mathbf{R}^d_- = \{(0,0,\lambda) : \lambda \leq 0\}$ obtuse.



Geometric Intuition

$$\langle s, x^\star \rangle \leq 0 \text{ for all } s \in N^\perp \cap \mathbf{R}^d_-$$

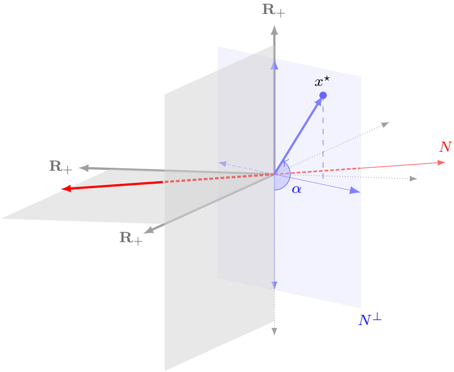
Equivalent angle between $x^\star$ and $N^\perp \cap \mathbf{R}^d_- = \{(0,0,\lambda) : \lambda \leq 0\}$ obtuse.



Geometric Intuition

$$\langle s, x^{\star} \rangle \leq 0 \text{ for all } s \in N^{\perp} \cap \mathbf{R}^d_-$$

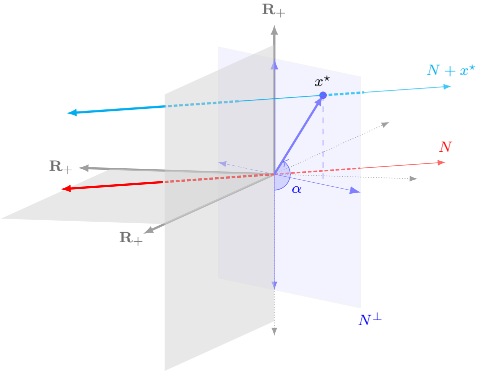
Equivalent angle between $x^{\star}$ and $N^{\perp} \cap \mathbf{R}^d_- = \{(0,0,\lambda) : \lambda \leq 0\}$ obtuse.



Geometric Intuition

$$\langle s, x^\star \rangle \leq 0 \text{ for all } s \in N^\perp \cap \mathbf{R}^d_-$$

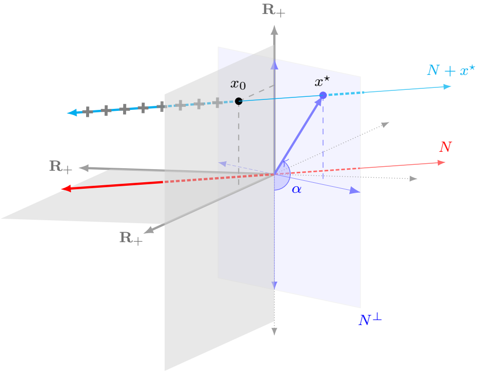
Equivalent angle between $x^\star$ and $N^\perp \cap \mathbf{R}^d_- = \{(0,0,\lambda) : \lambda \leq 0\}$ obtuse.



Geometric Intuition

$$\langle s, x^\star \rangle \leq 0 \text{ for all } s \in N^\perp \cap \mathbf{R}^d_-$$

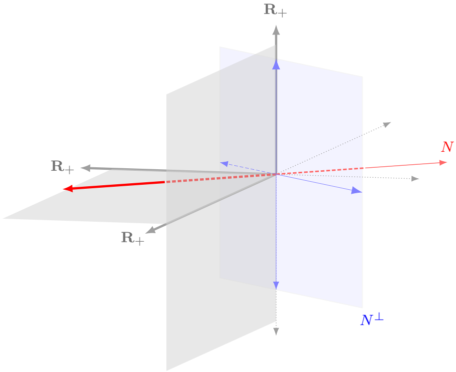
Equivalent angle between $x^\star$ and $N^\perp \cap \mathbf{R}^d_- = \{(0,0,\lambda) : \lambda \leq 0\}$ obtuse.



Geometric Intuition

$$\langle s, x^\star \rangle \leq 0 \text{ for all } s \in N^\perp \cap \mathbf{R}^d_-$$

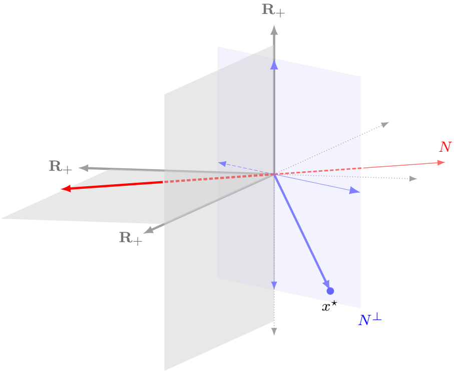
Equivalent angle between $x^\star$ and $N^\perp \cap \mathbf{R}^d_- = \{(0, 0, \lambda) : \lambda \leq 0\}$ obtuse.



Geometric Intuition

$$\langle s, x^\star \rangle \leq 0 \text{ for all } s \in N^\perp \cap \mathbf{R}^d_-$$

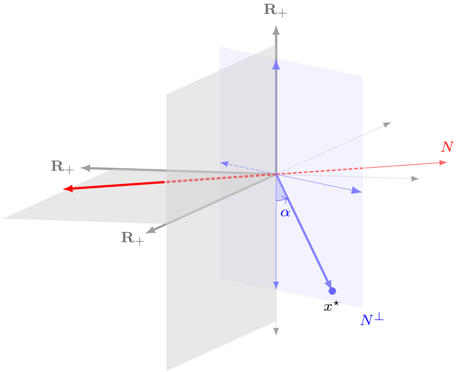
Equivalent angle between $x^\star$ and $N^\perp \cap \mathbf{R}^d_- = \{(0,0,\lambda) : \lambda \leq 0\}$ obtuse.



Geometric Intuition

$$\langle s, x^\star \rangle \leq 0 \text{ for all } s \in N^\perp \cap \mathbf{R}^d_-$$

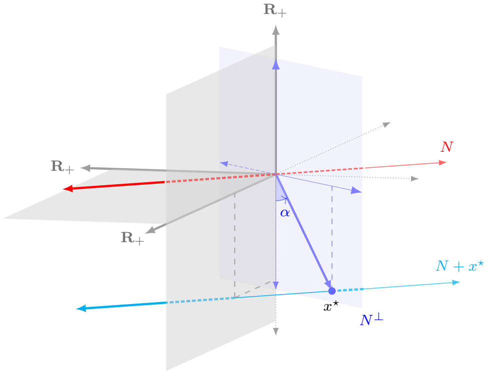
Equivalent angle between $x^\star$ and $N^\perp \cap \mathbf{R}^d_- = \{(0,0,\lambda) : \lambda \leq 0\}$ obtuse.



Geometric Intuition

$$\langle s, x^{\star} \rangle \leq 0 \text{ for all } s \in N^{\perp} \cap \mathbf{R}^d_{-}$$

Equivalent angle between $x^{\star}$ and $N^{\perp} \cap \mathbf{R}^d_{-} = \{(0, 0, \lambda) : \lambda \leq 0\}$ obtuse.



Test Statistic

Key: For $x^\star(P) \in N^\perp$ solving $\beta(P) = Ax^\star(P)$, $P \in \mathbf{P}_0$ if and only if

(i) $\beta(P) \in R$ (ii) $\langle s, x^\star(P)\rangle \leq 0$ for all $s \in N^\perp \cap \mathbf{R}^d_-$

Test Statistic

Key: For $x^\star(P) \in N^\perp$ solving $\beta(P) = Ax^\star(P)$, $P \in \mathbf{P}_0$ if and only if

(i) $\beta(P) \in R$ (ii) $\langle s, x^\star(P) \rangle \leq 0$ for all $s \in N^\perp \cap \mathbf{R}^d_-$

For talk only: Assume $R = \mathbf{R}^p$ so condition (i) is automatically satisfied.

Test Statistic

Key: For $x^\star(P) \in N^\perp$ solving $\beta(P) = Ax^\star(P)$, $P \in \mathbf{P}_0$ if and only if

(i) $\beta(P) \in R$ (ii) $\langle s, x^\star(P) \rangle \leq 0$ for all $s \in N^\perp \cap \mathbf{R}^d_-$

For talk only: Assume $R = \mathbf{R}^p$ so condition (i) is automatically satisfied.

The Pseudoinverse

- Under $R = \mathbf{R}^p$, for any $b \in \mathbf{R}^p$ there is unique $x(b) \in N^\perp$ solving

$$b = Ax(b)$$

Test Statistic

Key: For $x^\star(P) \in N^\perp$ solving $\beta(P) = Ax^\star(P)$, $P \in \mathbf{P}_0$ if and only if

(i) $\beta(P) \in R$ (ii) $\langle s, x^\star(P) \rangle \leq 0$ for all $s \in N^\perp \cap \mathbf{R}^d_-$

For talk only: Assume $R = \mathbf{R}^p$ so condition (i) is automatically satisfied.

The Pseudoinverse

- Under $R = \mathbf{R}^p$, for any $b \in \mathbf{R}^p$ there is unique $x(b) \in N^\perp$ solving

$$b = Ax(b)$$

- Under $R = \mathbf{R}^p$, the (MP) pseudoinverse $A^\dagger$ of A is $d \times p$ matrix solving

$$x(b) = A^\dagger b$$

Test Statistic

$$\langle s, x^\star(P)\rangle \leq 0 \text{ for all } s \in N^\perp \cap \mathbf{R}^d_-$$

Test Statistic

$$\langle s, x^\star(P)\rangle \leq 0 \text{ for all } s \in N^\perp \cap \mathbf{R}^d_-$$

... or equivalently, since $A^\dagger \beta(P) = x^\star(P)$, we may re-write condition as

$$\langle s, A^\dagger \beta(P)\rangle \leq 0 \text{ for all } s \in N^\perp \cap \mathbf{R}^d_-$$

Test Statistic

$$\langle s, x^\star(P)\rangle \leq 0 \text{ for all } s \in N^\perp \cap \mathbf{R}^d_-$$

... or equivalently, since $A^\dagger \beta(P) = x^\star(P)$, we may re-write condition as

$$\langle s, A^\dagger \beta(P)\rangle \leq 0 \text{ for all } s \in N^\perp \cap \mathbf{R}^d_-$$

... or equivalently, since range$\{A^\dagger\} = N^\perp$, we may re-write condition as

$$\langle A^\dagger s, A^\dagger \beta(P)\rangle \leq 0 \text{ for all } s \in \mathbf{R}^p \text{ s.t. } A^\dagger s \leq 0 \text{ (in } \mathbf{R}^d\text{)}$$

Test Statistic

$\langle A^\dagger s, A^\dagger \beta(P) \rangle \leq 0$ for all $s \in \mathbf{R}^p$ s.t. $A^\dagger s \leq 0$ (in $\mathbf{R}^d$)

Test Statistic

$$\langle A^\dagger s, A^\dagger \beta(P)\rangle \leq 0 \text{ for all } s \in \mathbf{R}^p \text{ s.t. } A^\dagger s \leq 0 \text{ (in } \mathbf{R}^d\text{)}$$

Test Statistic

$$T_n = \sup_{s \in \hat{\mathcal{V}}_n} \langle A^\dagger s, A^\dagger \hat{\beta}_n\rangle$$

Test Statistic

$$\langle A^\dagger s, A^\dagger \beta(P)\rangle \leq 0 \text{ for all } s \in \mathbf{R}^p \text{ s.t. } A^\dagger s \leq 0 \text{ (in } \mathbf{R}^d)$$

Test Statistic

$$T_n = \sup_{s \in \hat{\mathcal{V}}_n} \langle A^\dagger s, A^\dagger \hat{\beta}_n \rangle$$

$$\hat{\mathcal{V}}_n = \{s \in \mathbf{R}^p : A^\dagger s \leq 0 \text{ and } \|\hat{\Omega}_n (AA')^\dagger s\|_1 \leq 1\}$$

Comments

- Weighting matrix $\hat{\Omega}_n$ can be used to obtain scale invariance.
- Norm constraint ensures $T_n \neq +\infty$ with positive probability.
- Test statistic can be computed by linear programming.
- The norm $\|\cdot\|_1$ yields better coupling rates than, e.g., $\|\cdot\|_2$.

Test Statistic

Assumption T

- $\hat{\beta}_n$ is function of i.i.d. sample $\{Z_i\}_{i=1}^n$ with $Z_i \sim P \in \mathbf{P}$.
- $\hat{\Omega}_n$ is consistent for Ω uniformly in $P \in \mathbf{P}$ (under $\|\cdot\|_{o,\infty}$).
- For some sequence $a_n \downarrow 0$ and influence function ψ we have

$$\|\Omega^{\dagger}\{\sqrt{n}\{\hat{\beta}_n - \beta(P)\} - \frac{1}{\sqrt{n}}\sum_{i=1}^{n}\psi(Z_i)\}\|_\infty = O_P(a_n)$$

Comments

- Weighting matrix Ω need not be invertible.
- Estimator $\hat{\beta}_n$ is asymptotically linear.
- Norm $\|\cdot\|_\infty$ leads to favorable rate conditions in p.

Asymptotic Distribution

Theorem: Under Assumption T and regularity conditions, we have

$$\begin{aligned} T_n &\equiv \sup_{s \in \hat{\mathcal{V}}_n} \sqrt{n} \langle A^\dagger s, A^\dagger \hat{\beta}_n \rangle \\ &= \sup_{s \in \mathcal{V}} \langle A^\dagger s, A^\dagger \mathbb{G}_n \rangle + \sqrt{n} \langle A^\dagger s, A^\dagger \beta(P) \rangle + O_P(r_n) \end{aligned}$$

for some centered gaussian $\mathbb{G}_n \in \mathbf{R}^p$ (uniformly in $P \in \mathbf{P}$)

Comments

- Set $\mathcal{V} \subset \mathbf{R}^p$ just population analogue to $\hat{\mathcal{V}}_n$.
- Under moment conditions, $r_n \downarrow 0$ provided $p^2/n + a_n \downarrow 0$ (up to logs).
- $\| \cdot \|_1$ constraint defining $\hat{\mathcal{V}}_n$ (and $\mathcal{V}$) facilitate coupling under $\| \cdot \|_\infty$.

Critical Value

$$T_n = \sup_{s \in \mathcal{V}} \underbrace{\langle A^\dagger s, A^\dagger \mathbb{G}_n \rangle}_{\text{can be simulated}} + \underbrace{\sqrt{n} \langle A^\dagger s, A^\dagger \beta(P) \rangle}_{\text{nuisance parameter}} + O_P(r_n)$$

Like Moment Inequalities

- From geometry section, $\langle A^\dagger s, A^\dagger \beta(P) \rangle \leq 0$ for all $s \in \mathcal{V}$, $P \in \mathbf{P}_0$.
- Multiple techniques available from moment inequalities literature.

Critical Value

$$T_n = \sup_{s \in \mathcal{V}} \underbrace{\langle A^\dagger s, A^\dagger \mathbb{G}_n \rangle}_{\text{can be simulated}} + \underbrace{\sqrt{n} \langle A^\dagger s, A^\dagger \beta(P) \rangle}_{\text{nuisance parameter}} + O_P(r_n)$$

Like Moment Inequalities

- From geometry section, $\langle A^\dagger s, A^\dagger \beta(P) \rangle \leq 0$ for all $s \in \mathcal{V}$, $P \in \mathbf{P}_0$.
- Multiple techniques available from moment inequalities literature.

... But Different

- Replace $\sqrt{n} \langle A^\dagger s, A^\dagger \beta(P) \rangle$ with zero (may not be least favorable).
- Moment selection (e.g., Andrews $\&$ Soares 2010), two step procedures (e.g., Romano, Shaikh $\&$ Wolf 2014) can suffer in power.

Key: Nuisance parameter has additional structure beyond it being negative!

Critical Value

First Step

$$\hat{\beta}_n^{\mathrm{r}} \in \arg\min_{b \in \mathbf{R}^p} \sup_{s \in \hat{\mathcal{V}}_n} |\langle A^\dagger s, A^\dagger \hat{\beta}_n - A^\dagger b \rangle| \text{ s.t. } Ax = b \text{ for some } x \geq 0$$

Critical Value

First Step

$$\hat{\beta}_n^{\mathrm{r}} \in \arg\min_{b \in \mathbf{R}^p} \sup_{s \in \hat{\mathcal{V}}_n} |\langle A^\dagger s, A^\dagger \hat{\beta}_n - A^\dagger b \rangle| \text{ s.t. } Ax = b \text{ for some } x \geq 0$$

Bootstrap Statistic

$$T_n^\star \equiv \sup_{s \in \hat{\mathcal{V}}_n} \langle A^\dagger s, A^\dagger \hat{\mathbb{G}}_n^\star \rangle + \lambda_n \sqrt{n} \langle A^\dagger s, A^\dagger \hat{\beta}_n^{\mathrm{r}} \rangle$$

where $1 \geq \lambda_n \downarrow 0$ and $\hat{\mathbb{G}}_n^\star = \sqrt{n}\{\hat{\beta}_n^\star - \hat{\beta}_n\}$ with $\hat{\beta}_n^\star$ "bootstrapped" $\hat{\beta}_n$

Critical Value

First Step

$$\hat{\beta}_n^{\mathrm{r}} \in \arg\min_{b \in \mathbf{R}^p} \sup_{s \in \hat{\mathcal{V}}_n} |\langle A^\dagger s, A^\dagger \hat{\beta}_n - A^\dagger b\rangle| \text{ s.t. } Ax = b \text{ for some } x \geq 0$$

Bootstrap Statistic

$$T_n^\star \equiv \sup_{s \in \hat{\mathcal{V}}_n} \langle A^\dagger s, A^\dagger \hat{\mathbb{G}}_n^\star\rangle + \lambda_n \sqrt{n} \langle A^\dagger s, A^\dagger \hat{\beta}_n^{\mathrm{r}}\rangle$$

where $1 \geq \lambda_n \downarrow 0$ and $\hat{\mathbb{G}}_n^\star = \sqrt{n}\{\hat{\beta}_n^\star - \hat{\beta}_n\}$ with $\hat{\beta}_n^\star$ "bootstrapped" $\hat{\beta}_n$

Critical Value

$$\hat{c}_n(1-\alpha) \equiv \inf\{u : P(T_n^\star \leq u | \{Z_i\}_{i=1}^n) \geq 1 - \alpha\}$$

Some Intuition

Question: Why does this bootstrap approximation control size?

$$T_n^\star \equiv \sup_{s \in \hat{\mathcal{V}}_n} \langle A^\dagger s, A^\dagger \hat{\mathbb{G}}_n^\star \rangle + \lambda_n \sqrt{n} \langle A^\dagger s, A^\dagger \hat{\beta}_n^{\mathrm{r}} \rangle$$

Some Intuition

Question: Why does this bootstrap approximation control size?

$$T_n^\star \equiv \sup_{s \in \hat{\mathcal{V}}_n} \langle A^\dagger s, A^\dagger \hat{\mathbb{G}}_n^\star \rangle + \lambda_n \sqrt{n} \langle A^\dagger s, A^\dagger \hat{\beta}_n^{\mathrm{r}} \rangle$$

$$\approx \sup_{s \in \hat{\mathcal{V}}_n} \langle A^\dagger s, A^\dagger \hat{\mathbb{G}}_n^\star \rangle + \lambda_n \sqrt{n} \langle A^\dagger s, A^\dagger \beta(P) \rangle \qquad \text{(if } \lambda_n \to 0\text{)}$$

Some Intuition

Question: Why does this bootstrap approximation control size?

$$\begin{aligned}
T_n^\star &\equiv \sup_{s\in\hat{\mathcal{V}}_n} \langle A^\dagger s, A^\dagger \hat{\mathbb{G}}_n^\star\rangle + \lambda_n\sqrt{n}\langle A^\dagger s, A^\dagger \hat{\beta}_n^{\mathrm{r}}\rangle \\
&\approx \sup_{s\in\hat{\mathcal{V}}_n} \langle A^\dagger s, A^\dagger \hat{\mathbb{G}}_n^\star\rangle + \lambda_n\sqrt{n}\langle A^\dagger s, A^\dagger \beta(P)\rangle && \text{(if } \lambda_n \to 0\text{)} \\
&\geq \sup_{s\in\hat{\mathcal{V}}_n} \langle A^\dagger s, A^\dagger \hat{\mathbb{G}}_n^\star\rangle + \sqrt{n}\langle A^\dagger s, A^\dagger \beta(P)\rangle && \text{(by } \langle A^\dagger s, A^\dagger \beta(P)\rangle \leq 0\text{)}
\end{aligned}$$

Some Intuition

Question: Why does this bootstrap approximation control size?

$$\begin{aligned}
T_n^\star &\equiv \sup_{s\in\hat{\mathcal{V}}_n} \langle A^\dagger s, A^\dagger \hat{\mathbb{G}}_n^\star\rangle + \lambda_n\sqrt{n}\langle A^\dagger s, A^\dagger \hat{\beta}_n^{\mathrm{r}}\rangle \\
&\approx \sup_{s\in\hat{\mathcal{V}}_n} \langle A^\dagger s, A^\dagger \hat{\mathbb{G}}_n^\star\rangle + \lambda_n\sqrt{n}\langle A^\dagger s, A^\dagger \beta(P)\rangle && \text{(if } \lambda_n \to 0\text{)} \\
&\geq \sup_{s\in\hat{\mathcal{V}}_n} \langle A^\dagger s, A^\dagger \hat{\mathbb{G}}_n^\star\rangle + \sqrt{n}\langle A^\dagger s, A^\dagger \beta(P)\rangle && \text{(by } \langle A^\dagger s, A^\dagger \beta(P)\rangle \leq 0\text{)} \\
&\overset{d}{\approx} \sup_{s\in\mathcal{V}} \langle A^\dagger s, A^\dagger \mathbb{G}_n\rangle + \sqrt{n}\langle A^\dagger s, A^\dagger \beta(P)\rangle && \text{(bootstrap cons.)}
\end{aligned}$$

Some Intuition

Question: Why does this bootstrap approximation control size?

$$
\begin{aligned}
T_n^\star \equiv & \sup_{s \in \hat{\mathcal{V}}_n} \langle A^\dagger s, A^\dagger \hat{\mathbb{G}}_n^\star \rangle + \lambda_n \sqrt{n} \langle A^\dagger s, A^\dagger \hat{\beta}_n^{\mathrm{r}} \rangle & \\
\approx & \sup_{s \in \hat{\mathcal{V}}_n} \langle A^\dagger s, A^\dagger \hat{\mathbb{G}}_n^\star \rangle + \lambda_n \sqrt{n} \langle A^\dagger s, A^\dagger \beta(P) \rangle & \text{(if } \lambda_n \to 0\text{)} \\
\geq & \sup_{s \in \hat{\mathcal{V}}_n} \langle A^\dagger s, A^\dagger \hat{\mathbb{G}}_n^\star \rangle + \sqrt{n} \langle A^\dagger s, A^\dagger \beta(P) \rangle & \text{(by } \langle A^\dagger s, A^\dagger \beta(P) \rangle \leq 0\text{)} \\
\overset{d}{\approx} & \sup_{s \in \mathcal{V}} \langle A^\dagger s, A^\dagger \mathbb{G}_n \rangle + \sqrt{n} \langle A^\dagger s, A^\dagger \beta(P) \rangle & \text{(bootstrap cons.)} \\
\approx & \, T_n & \text{(by theorem)}
\end{aligned}
$$

Key: Bootstrap provides uniform upper bound ... but is it conservative?

Some Intuition

Suppose: P is fixed and $n \to \infty$ (i.e. pointwise, not uniform analysis)

$$T_n \approx \sup_{s \in \mathcal{V}} \langle A^\dagger s, A^\dagger \mathbb{G}_n \rangle + \sqrt{n} \langle A^\dagger s, A^\dagger \beta(P) \rangle$$

Some Intuition

Suppose: P is fixed and $n \to \infty$ (i.e. pointwise, not uniform analysis)

$$\begin{aligned} T_n &\approx \sup_{s \in \mathcal{V}} \langle A^\dagger s, A^\dagger \mathbb{G}_n \rangle + \sqrt{n} \langle A^\dagger s, A^\dagger \beta(P) \rangle \\ &= \max\{0, \sup_{s \in \mathcal{V}} \langle A^\dagger s, A^\dagger \mathbb{G}_n \rangle + \sqrt{n} \langle A^\dagger s, A^\dagger \beta(P) \rangle\} \qquad \text{(since } 0 \in \mathcal{V}\text{)} \end{aligned}$$

Some Intuition

Suppose: P is fixed and $n \to \infty$ (i.e. pointwise, not uniform analysis)

$$\begin{aligned} T_n &\approx \sup_{s \in \mathcal{V}} \langle A^\dagger s, A^\dagger \mathbb{G}_n \rangle + \sqrt{n} \langle A^\dagger s, A^\dagger \beta(P) \rangle \\ &= \max\{0, \sup_{s \in \mathcal{V}} \langle A^\dagger s, A^\dagger \mathbb{G}_n \rangle + \underbrace{\sqrt{n} \langle A^\dagger s, A^\dagger \beta(P) \rangle}_{\text{if } \ll 0 \text{ then } s \text{ drops out}} \} \qquad \text{(since } 0 \in \mathcal{V}\text{)} \end{aligned}$$

Some Intuition

Suppose: P is fixed and $n \to \infty$ (i.e. pointwise, not uniform analysis)

$$\begin{aligned} T_n &\approx \sup_{s \in \mathcal{V}} \langle A^\dagger s, A^\dagger \mathbb{G}_n \rangle + \sqrt{n} \langle A^\dagger s, A^\dagger \beta(P) \rangle \\ &= \max\{0, \sup_{s \in \mathcal{V}} \langle A^\dagger s, A^\dagger \mathbb{G}_n \rangle + \sqrt{n} \langle A^\dagger s, A^\dagger \beta(P) \rangle\} && \text{(since } 0 \in \mathcal{V}\text{)} \\ &\approx \max\{0, \sup_{s \in \mathcal{V}} \langle A^\dagger s, A^\dagger \mathbb{G}_n \rangle \text{ s.t. } \langle A^\dagger s, A^\dagger \beta(P) \rangle = 0\} && \text{(}P\text{ is fixed)} \end{aligned}$$

Some Intuition

Suppose: P is fixed and $n \to \infty$ (i.e. pointwise, not uniform analysis)

$$\begin{aligned} T_n &\approx \sup_{s \in \mathcal{V}} \langle A^\dagger s, A^\dagger \mathbb{G}_n \rangle + \sqrt{n} \langle A^\dagger s, A^\dagger \beta(P) \rangle \\ &= \max\{0, \sup_{s \in \mathcal{V}} \langle A^\dagger s, A^\dagger \mathbb{G}_n \rangle + \sqrt{n} \langle A^\dagger s, A^\dagger \beta(P) \rangle\} && \text{(since } 0 \in \mathcal{V}\text{)} \\ &\approx \max\{0, \sup_{s \in \mathcal{V}} \langle A^\dagger s, A^\dagger \mathbb{G}_n \rangle \text{ s.t. } \langle A^\dagger s, A^\dagger \beta(P) \rangle = 0\} && \text{(}P\text{ is fixed)} \end{aligned}$$

What About Bootstrap?

Some Intuition

Suppose: P is fixed and $n \to \infty$ (i.e. pointwise, not uniform analysis)

$$\begin{aligned} T_n &\approx \sup_{s \in \mathcal{V}} \langle A^\dagger s, A^\dagger \mathbb{G}_n \rangle + \sqrt{n} \langle A^\dagger s, A^\dagger \beta(P) \rangle \\ &= \max\{0, \sup_{s \in \mathcal{V}} \langle A^\dagger s, A^\dagger \mathbb{G}_n \rangle + \sqrt{n} \langle A^\dagger s, A^\dagger \beta(P) \rangle\} && \text{(since } 0 \in \mathcal{V}\text{)} \\ &\approx \max\{0, \sup_{s \in \mathcal{V}} \langle A^\dagger s, A^\dagger \mathbb{G}_n \rangle \text{ s.t. } \langle A^\dagger s, A^\dagger \beta(P) \rangle = 0\} && (P \text{ is fixed}) \end{aligned}$$

What About Bootstrap?

$$T_n^\star \approx \sup_{s \in \mathcal{V}} \langle A^\dagger s, A^\dagger \hat{\mathbb{G}}_n^\star \rangle + \lambda_n \sqrt{n} \langle A^\dagger s, A^\dagger \beta(P) \rangle \qquad \text{(shown before)}$$

Some Intuition

Suppose: P is fixed and $n \to \infty$ (i.e. pointwise, not uniform analysis)

$$T_n \approx \sup_{s\in\mathcal{V}} \langle A^\dagger s, A^\dagger \mathbb{G}_n\rangle + \sqrt{n}\langle A^\dagger s, A^\dagger \beta(P)\rangle$$

$$= \max\{0, \sup_{s\in\mathcal{V}} \langle A^\dagger s, A^\dagger \mathbb{G}_n\rangle + \sqrt{n}\langle A^\dagger s, A^\dagger \beta(P)\rangle\} \qquad \text{(since } 0 \in \mathcal{V}\text{)}$$

$$\approx \max\{0, \sup_{s\in\mathcal{V}} \langle A^\dagger s, A^\dagger \mathbb{G}_n\rangle \text{ s.t. } \langle A^\dagger s, A^\dagger \beta(P)\rangle = 0\} \qquad (P \text{ is fixed})$$

What About Bootstrap?

$$T_n^\star \approx \sup_{s\in\mathcal{V}} \langle A^\dagger s, A^\dagger \hat{\mathbb{G}}_n^\star\rangle + \lambda_n\sqrt{n}\langle A^\dagger s, A^\dagger \beta(P)\rangle \qquad \text{(shown before)}$$

$$= \max\{0, \sup_{s\in\mathcal{V}} \langle A^\dagger s, A^\dagger \hat{\mathbb{G}}_n^\star\rangle + \lambda_n\sqrt{n}\langle A^\dagger s, A^\dagger \beta(P)\rangle\} \qquad \text{(since } 0 \in \mathcal{V}\text{)}$$

Some Intuition

Suppose: P is fixed and $n \to \infty$ (i.e. pointwise, not uniform analysis)

$$\begin{aligned} T_n &\approx \sup_{s \in \mathcal{V}} \langle A^\dagger s, A^\dagger \mathbb{G}_n \rangle + \sqrt{n} \langle A^\dagger s, A^\dagger \beta(P) \rangle \\ &= \max\{0, \sup_{s \in \mathcal{V}} \langle A^\dagger s, A^\dagger \mathbb{G}_n \rangle + \sqrt{n} \langle A^\dagger s, A^\dagger \beta(P) \rangle \} && \text{(since } 0 \in \mathcal{V}\text{)} \\ &\approx \max\{0, \sup_{s \in \mathcal{V}} \langle A^\dagger s, A^\dagger \mathbb{G}_n \rangle \text{ s.t. } \langle A^\dagger s, A^\dagger \beta(P) \rangle = 0 \} && (P \text{ is fixed}) \end{aligned}$$

What About Bootstrap?

$$\begin{aligned} T_n^\star &\approx \sup_{s \in \mathcal{V}} \langle A^\dagger s, A^\dagger \hat{\mathbb{G}}_n^\star \rangle + \lambda_n \sqrt{n} \langle A^\dagger s, A^\dagger \beta(P) \rangle && \text{(shown before)} \\ &= \max\{0, \sup_{s \in \mathcal{V}} \langle A^\dagger s, A^\dagger \hat{\mathbb{G}}_n^\star \rangle + \underbrace{\lambda_n \sqrt{n} \langle A^\dagger s, A^\dagger \beta(P) \rangle}_{\text{if } \ll 0 \text{ then } s \text{ drops out}} \} && \text{(since } 0 \in \mathcal{V}\text{)} \end{aligned}$$

Some Intuition

Suppose: P is fixed and $n \to \infty$ (i.e. pointwise, not uniform analysis)

$$\begin{aligned} T_n &\approx \sup_{s\in\mathcal{V}} \langle A^\dagger s, A^\dagger \mathbb{G}_n\rangle + \sqrt{n}\langle A^\dagger s, A^\dagger \beta(P)\rangle \\ &= \max\{0, \sup_{s\in\mathcal{V}} \langle A^\dagger s, A^\dagger \mathbb{G}_n\rangle + \sqrt{n}\langle A^\dagger s, A^\dagger \beta(P)\rangle\} && \text{(since } 0\in\mathcal{V}) \\ &\approx \max\{0, \sup_{s\in\mathcal{V}} \langle A^\dagger s, A^\dagger \mathbb{G}_n\rangle \text{ s.t. } \langle A^\dagger s, A^\dagger \beta(P)\rangle = 0\} && (P \text{ is fixed}) \end{aligned}$$

What About Bootstrap?

$$\begin{aligned} T_n^\star &\approx \sup_{s\in\mathcal{V}} \langle A^\dagger s, A^\dagger \hat{\mathbb{G}}_n^\star\rangle + \lambda_n\sqrt{n}\langle A^\dagger s, A^\dagger \beta(P)\rangle && \text{(shown before)} \\ &= \max\{0, \sup_{s\in\mathcal{V}} \langle A^\dagger s, A^\dagger \hat{\mathbb{G}}_n^\star\rangle + \lambda_n\sqrt{n}\langle A^\dagger s, A^\dagger \beta(P)\rangle\} && \text{(since } 0\in\mathcal{V}) \\ &\approx \max\{0, \sup_{s\in\mathcal{V}} \langle A^\dagger s, A^\dagger \hat{\mathbb{G}}_n^\star\rangle \text{ s.t. } \langle A^\dagger s, A^\dagger \beta(P)\rangle = 0\} && (\text{if } \lambda_n\sqrt{n}\to\infty) \end{aligned}$$

Critical Value

Assumption B

- There are random variables $\{W_{i,n}\}_{i=1}^n$ independent of $\{Z_i\}_{i=1}^n$ with

$$\|\Omega^\dagger\{\hat{\mathbb{G}}_n^\star - \frac{1}{\sqrt{n}}\sum_{i=1}^n (W_{i,n} - \bar{W}_n)\psi(Z_i)\}\|_\infty = O_P(a_n)$$

- The distribution of $\{W_{i,n}\}_{i=1}^n$ is exchangeable.

Comments

- Asymptotically linear assumption analogous to requirement on $\hat{\beta}_n$.
- Exchangeability covers multiplier, score, and nonparametric bootstrap.
- Derive coupling results for exchangeable bootstrap under $\|\cdot\|_\infty$.

Asymptotic Size

Theorem: Under Assumptions T, B, regularity conditions, and $\alpha \in (0, 0.5)$

$$\limsup_{n\to\infty} \sup_{P\in\mathbf{P}_0} P(T_n > \hat{c}_n(1-\alpha)) \leq \alpha$$

Comments

- Bootstrap coupling requires $p^2/n \downarrow 0$ (up to logs).
- Anti-concentration: Under fixed p and studentization automatic.
- Anti-concentration: Dependence on p through $(AA')^{\dagger}\mathcal{V}$.
- Conservative universal (in A) bounds on dependence on p available.
- Under same conditions, two stage critical value also valid.

1. The Geometry

2. The Test

3. Simulations

Simulation Design

$$Y = 1\{C_0 + C_1 W \geq U\}$$

$U \sim$ logistic, unobservable $V \equiv (C_0, C_1)'$ and observable W discrete.

Comments

- W, V, and U all mutually independent.
- Random coefficients logit (Fox, Kim, Ryan, Bajari, 2011).
- $C_0 \in [0.5, 1]$, $C_1 \in [-3, 0]$ with $\sqrt{d}$ points of support each.
- Support of W is evenly spaced grid on $[0, 2]$ (cardinality equals $p - 2$).
- 250 bootstrap draws, 5000 or 1000 replications.

Simulation Design

Restrictions

- For $\mathcal{V}$ support of V, $\pi(v) = P(V = v)$, and $v = (c_1, c_2)'$ we have

$$P(Y = 1|W = w) = \sum_{v \in \mathcal{V}} \pi(v) \frac{1}{1 + \exp\{-c_0 - c_1 w\}}$$

- Unknown probabilities $\{\pi(v) : v \in \mathcal{V}\}$ satisfy $\sum_{v \in \mathcal{V}} \pi(v) = 1$.

Simulation Design

Restrictions

- For $\mathcal{V}$ support of V, $\pi(v) = P(V = v)$, and $v = (c_1, c_2)'$ we have

$$P(Y = 1|W = w) = \sum_{v \in \mathcal{V}} \pi(v) \frac{1}{1 + \exp\{-c_0 - c_1 w\}}$$

- Unknown probabilities $\{\pi(v) : v \in \mathcal{V}\}$ satisfy $\sum_{v \in \mathcal{V}} \pi(v) = 1$.

Parameter of Interest

- Consumer type $v = (c_0, c_1)$ with price $\bar{w}$ has purchase prob. elasticity

$$\epsilon(v, \bar{w}) \equiv c_0 \bar{w}(1 - \frac{1}{1 + \exp\{-c_0 - c_1 \bar{w}\}})$$

- Inference on $F(t|\bar{w}) \equiv P(\epsilon(V, \bar{w}) \leq t) = \sum_{v \in \mathcal{V}} \pi(v) 1\{\epsilon(v, \bar{w}) \leq t\}$

Design Partially Identified

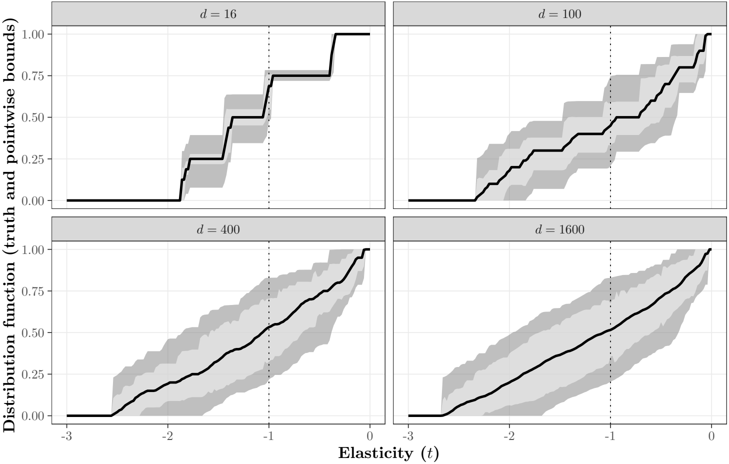


Figure: Dark: W with 4 support points, Lighter: W with 16 support points

Simulation Design

The General Problem

$$\beta(P) = Ax \text{ for some } x \geq 0$$

In this Design

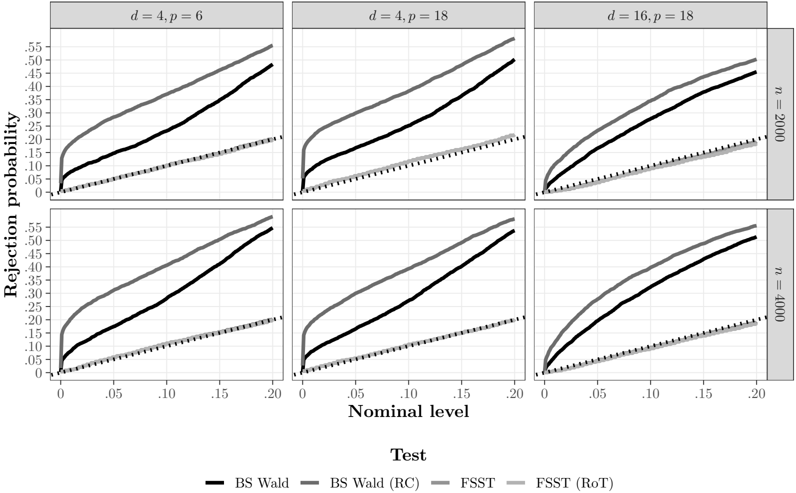
- $x \in \mathbf{R}^d$ is the unknown probabilities $\{\pi(v) : v \in \mathcal{V}\}$.
- $\beta(P) \in \mathbf{R}^p$, first $p-2$ coordinates correspond to $P(Y=1|W=w)$.
- The $p-1$ coordinate of $\beta(P)$ equals 1 ($\sum_{v \in \mathcal{V}} \pi(v) = 1$).
- The p coordinate of $\beta(P)$ equals hypothesized value for $F(-1|1)$.

Bandwidth Selection

- Law of iterated logarithm: $\lambda_n^{\mathrm{r}} = (\log(e \vee p)\log(e \vee \log(e \vee n)))^{-1/2}$.
- Bootstrap: Set $1/\lambda_n^{\mathrm{b}}$ to be $1 - (\log(e \vee \log(e \vee n)))^{-1/2}$ quantile of

$$\sup_{s \in \hat{\mathcal{V}}_n^{\mathrm{i}}} \langle A^\dagger s, A^\dagger \hat{\mathbb{G}}_n^{\mathrm{i}} \rangle$$

(Almost) Identified Case



Null Rejection: Bootstrap Bandwidth

Table: Null Hypothesis that $F_\epsilon(-1|1)$ equals lower bound of identified set.

		d						
n	p	100	400	1600	4900	100^2	225^2	317^2
1000	6	.036	.034	.034	.037	.038	.036	.036
	18	.040	.035	.036	.041	.039	.038	.036
2000	6	.042	.042	.049	.046	.047	.052	.061
	18	.031	.028	.032	.032	.030	.030	.028
	38	.053	.046	.051	.052	.052	.067	.053
4000	6	.045	.048	.049	.054	.058	.051	.065
	18	.028	.031	.029	.028	.030	.038	.035
	38	.031	.034	.039	.036	.040	.035	.037
	51	.042	.051	.051	.040	.047	.047	.030
8000	6	.049	.055	.056	.048	.054	.055	.073
	18	.034	.035	.036	.030	.032	.040	.041
	38	.033	.035	.035	.037	.037	.025	.047
	51	.034	.043	.035	.040	.037	.035	.038
	83	.043	.042	.050	.048	.042	.054	.046

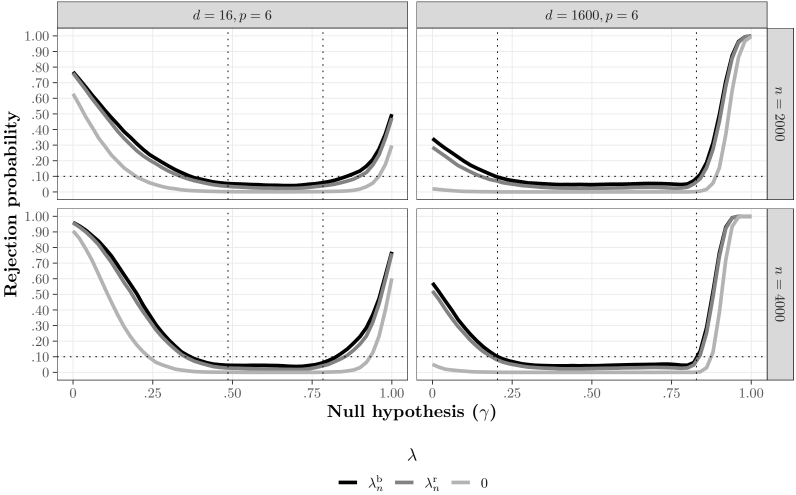
Null Rejection: RoT Bandwidth

Table: Null Hypothesis that $F_\epsilon(-1|1)$ equals lower bound of identified set.

		d						
n	p	100	400	1600	4900	100^2	225^2	317^2
1000	6	.020	.019	.021	.021	.022	.019	.021
	18	.037	.029	.029	.033	.033	.031	.030
2000	6	.030	.025	.033	.032	.033	.027	.039
	18	.023	.021	.028	.027	.025	.027	.020
	38	.048	.039	.043	.045	.047	.062	.046
4000	6	.034	.034	.038	.042	.046	.035	.058
	18	.023	.026	.024	.022	.025	.032	.028
	38	.026	.029	.033	.032	.035	.032	.033
	51	.038	.044	.045	.034	.042	.041	.027
8000	6	.040	.046	.048	.040	.046	.050	.061
	18	.028	.028	.032	.025	.027	.032	.034
	38	.027	.029	.030	.032	.032	.021	.043
	51	.029	.036	.028	.034	.033	.030	.031
	83	.038	.035	.046	.041	.034	.048	.042

Power Curves

Figure: Power for 10% nominal level test



Conclusion

Summary

- Mapped problems of interest into tests of $\beta(P) = Ax$ for some $x \geq 0$.
- Obtained new geometric characteristic of the null hypothesis.
- Derived test that can be evaluated by solving linear programs.
- Alternative tests also follow from geometric characterization.
- Immediate extension to (some) alternative sampling frameworks.