Notes *Nonparametric Analysis of Random Utility Models: Computational Tools for Statistical Testing*

Bart Smeulders, Laurens Cherchye, Bram De Rock

Paper link ; Repo link

1 Overview

This paper develops computationally tractable methods for implementing nonparametric statistical tests of the **Random Utility Model (RUM)**. Building on McFadden and Richter (1991) and Kitamura and Stoye (2018), the authors propose new optimization formulations that make testing feasible in large-dimensional choice environments.

Key goals:

- Reformulate RUM consistency as a *cone membership problem* in linear algebraic form.
- Translate stochastic rationalizability into a set of linear inequalities.
- Develop efficient linear and quadratic programming algorithms.
- Apply these computational tools to empirical testing and welfare inference.

2 Random Utility Framework

2.1 Setup

Let X be a finite set of alternatives with $|X| = J$. An individual facing a menu $A \subseteq X$ chooses one element $x \in A$. Let $\mathcal{D}$ denote the collection of observed menus.

For each (A, x) define the population probability

$$p(x, A) = \Pr(x \text{ is chosen from } A), \qquad \text{with } \sum_{x \in A} p(x, A) = 1.$$

Stacking all these probabilities yields the vector

$$\pi = (p(x, A))_{x,A} \in \mathbb{R}^{J_\mathcal{D}}, \quad J_\mathcal{D} = \sum_{A \in \mathcal{D}} |A|.$$

2.2 Deterministic Choice Types and Matrix A

Every deterministic type $r = 1, \ldots, R$ corresponds to a strict preference ordering $\succ_r$ on X. Given A, this type chooses

$$c_r(A) = \arg\max_{x \in A} u_r(x),$$

the most preferred element under $\succ_r$. Each type defines a deterministic choice pattern over menus:

$$A_{(x,A),r} = \begin{cases} 1 & \text{if } c_r(A) = x, \\ 0 & \text{otherwise.} \end{cases}$$

Stacking these $A_{(x,A),r}$ gives the binary matrix

$$A \in \{0,1\}^{J_{\mathcal{D}} \times R},$$

whose columns represent deterministic types and rows correspond to (A, x) pairs.

The population is described by mixing weights $\nu = (\nu_1, \ldots, \nu_R)'$, $\nu_r \geq 0$, $\mathbf{1}'\nu = 1$, so that

$$\pi = A\nu. \tag{1}$$

2.3 Rationalizability

Equation (1) implies that π must lie in the convex hull of the deterministic choice types:

$$\mathcal{C} = \{A\nu : \nu \geq 0,\, \mathbf{1}'\nu = 1\}.$$

Thus, π is *stochastically rationalizable* $\iff \pi \in \mathcal{C}$.

Geometric interpretation. $\mathcal{C}$ is a closed convex cone in $\mathbb{R}^{J_{\mathcal{D}}}$. Each column of A is an extreme ray corresponding to a deterministic preference ordering.

3 Testing the RUM Hypothesis

Given empirical frequencies $\hat{\pi}$, we test

$$H_0 : \hat{\pi} \in \mathcal{C}.$$

Under H_0, there exists ν satisfying (1). Rejection implies observed choice frequencies cannot be generated by any mixture of utility-maximizing agents.

3.1 Distance to the Cone

Define the minimum weighted squared distance of $\hat{\pi}$ from $\mathcal{C}$:

$$J_n = n \min_{\nu \geq 0,\, \mathbf{1}'\nu = 1} (\hat{\pi} - A\nu)'\Omega(\hat{\pi} - A\nu), \tag{2}$$

where Ω is a positive definite weighting matrix (often diagonal with inverse variance weights). This is the test statistic of Kitamura and Stoye (2018).

If $\hat{\pi} \in \mathcal{C}$, $J_n = 0$; larger values indicate larger violations.

3.2 Equivalent Optimization Forms

The quadratic program in (2) can be equivalently written as:

$$\min_{\nu \geq 0} \frac{1}{2}\nu'(A'\Omega A)\nu - (A'\Omega\hat{\pi})'\nu \quad \text{s.t. } \mathbf{1}'\nu = 1.$$

KKT conditions yield

$$A'\Omega(A\nu - \hat{\pi}) + \lambda\mathbf{1} - \mu = 0, \tag{3}$$

$$\lambda(\mathbf{1}'\nu - 1) = 0, \tag{4}$$

$$\nu'\mu = 0, \quad \nu, \mu \geq 0. \tag{5}$$

4 Asymptotic Theory

Under standard regularity conditions,

$$\sqrt{n}(\hat{\pi} - \pi_0) \to_d Z \sim N(0, \Sigma),$$

with Σ the sampling covariance. The limit distribution of J_n is

$$J_n \to_d \min_{t \in T(\pi_0)} (Z - t)'\Omega(Z - t),$$

where $T(\pi_0)$ is the *tangent cone* of $\mathcal{C}$ at π_0. Because $T(\pi_0)$ depends on which inequalities in $\nu \geq 0$ bind, the limit distribution is nonstandard.

5 Modified Bootstrap

To obtain valid inference, Cherchye et al. adopt the regularized (modified) bootstrap of Kitamura–Stoye.

Algorithm.

(1) Generate a bootstrap sample $\hat{\pi}^*$ by resampling individuals.

(2) Compute the regularized projection

$$\hat{\eta}_{\tau_n} = \arg \min_{\nu \geq \mathbf{1}\tau_n/H} (\hat{\pi} - A\nu)'\Omega(\hat{\pi} - A\nu),$$

where $\tau_n = \sqrt{\log n/n}$ ensures numerical stability.

(3) Recenter the bootstrap draw:

$$\hat{\pi}^*_{\tau_n} = \hat{\pi}^* + \hat{\eta}_{\tau_n} - \hat{\pi}.$$

(4) Compute

$$J_n^* = n \min_{\nu \geq \mathbf{1}\tau_n/H} (\hat{\pi}^*_{\tau_n} - A\nu)'\Omega(\hat{\pi}^*_{\tau_n} - A\nu).$$

(5) The bootstrap p–value is the proportion of replications with $J_n^* \geq J_n$.

This approach achieves asymptotically correct size and good finite-sample power.

6 Computational Reformulations

The paper's main contribution is to show that the RUM testing problem can be efficiently computed through several equivalent formulations.

6.1 Cone Projection as a Quadratic Program

Projecting $\hat{\pi}$ onto $\mathcal{C}$ is equivalent to solving:

$$\min_{\nu \geq 0,\, \mathbf{1}'\nu = 1} \|L(\hat{\pi} - A\nu)\|_2^2,$$

where L is the Cholesky factor of Ω ($\Omega = L'L$). This projection defines the *closest stochastically rationalizable vector* $\hat{\eta} = A\hat{\nu}$.

6.2 Normal Equations and Dual Problem

The first-order condition gives:

$$A'\Omega A\nu = A'\Omega\hat{\pi} + \lambda\mathbf{1} - \mu,$$

with complementary slackness. The dual program maximizes the quadratic form in the residuals $s = \hat{\pi} - A\nu$:

$$\max_{s \in \mathcal{C}^*} -\tfrac{1}{2}s'\Omega^{-1}s + \hat{\pi}'s,$$

where $\mathcal{C}^* = \{A'\Omega s \geq 0\}$ is the dual cone.

6.3 Linear Programming Approximation

When $\Omega = I$, the distance to the cone simplifies to Euclidean projection:

$$J_n = n\|\hat{\pi} - P_{\mathcal{C}}(\hat{\pi})\|_2^2.$$

They show that for large-scale problems, a linear programming relaxation with slack variables e,

$$\min_{e,\nu \geq 0} \mathbf{1}'e \quad \text{s.t.} \quad -e \leq \hat{\pi} - A\nu \leq e,$$

approximates the quadratic distance with high accuracy but lower computational cost.

6.4 Active-Set Algorithms

The authors design a customized active-set algorithm that:

- Iteratively identifies binding inequality constraints in $\nu \geq 0$,
- Solves reduced least-squares problems on active sets,
- Updates the active set until convergence.

This dramatically speeds up projection computations.

7 Extensions

7.1 Partial Observation and Matrix Compression

When not all menus are observed, many deterministic types yield identical columns in A on $\mathcal{D}$. Let $\tilde{A}$ be the matrix after merging identical columns. Then $\text{cone}(\tilde{A}) = \text{cone}(A)$ on $\mathcal{D}$, preserving all theoretical results but reducing dimension.

7.2 Regularization and Numerical Stability

The lower bound $\nu_r \geq \tau_n/H$ avoids degenerate faces of $\mathcal{C}$ in the bootstrap world and guarantees unique projections.

7.3 Welfare Bounds

Given utilities $u(x)$, expected welfare is $W = \pi' u = \nu'(A'u)$. To obtain welfare bounds consistent with RUM,

$$\underline{W} = \min_{\nu \geq 0,\, \mathbf{1}'\nu = 1} u' A \nu,$$
$$\overline{W} = \max_{\nu \geq 0,\, \mathbf{1}'\nu = 1} u' A \nu.$$

These are linear programs directly implementable with the same computational tools.

8 Empirical and Numerical Performance

The authors perform simulation exercises confirming:

- The new formulations replicate Kitamura–Stoye's results exactly.
- Runtime reductions of up to two orders of magnitude for realistic datasets.
- Excellent stability of the active-set solver and bootstrap implementation.

They also illustrate applications to consumer choice and risk preference data.

9 Mathematical Summary

$$
\begin{aligned}
\textbf{Feasible cone:}\ & \mathcal{C} = \{A\nu : \nu \geq 0, \mathbf{1}'\nu = 1\}. \\
\textbf{Test statistic:}\ & J_n = n\,(\hat{\pi} - A\hat{\nu})'\Omega(\hat{\pi} - A\hat{\nu}). \\
\textbf{Projection:}\ & \hat{\nu} = \arg\min_{\nu \geq 0, \mathbf{1}'\nu = 1} (\hat{\pi} - A\nu)'\Omega(\hat{\pi} - A\nu). \\
\textbf{KKT system:}\ & \begin{cases} A'\Omega(A\nu - \hat{\pi}) + \lambda\mathbf{1} - \mu = 0, \\ \nu, \mu \geq 0,\ \mathbf{1}'\nu = 1, \\ \nu'\mu = 0. \end{cases} \\
\textbf{Bootstrap:}\ & \begin{cases} \hat{\eta}_{\tau_n} = \arg\min_{\nu \geq \mathbf{1}\tau_n/H} (\hat{\pi} - A\nu)'\Omega(\hat{\pi} - A\nu), \\ \hat{\pi}^*_{\tau_n} = \hat{\pi}^* + \hat{\eta}_{\tau_n} - \hat{\pi}, \\ J^*_n = n \min_{\nu \geq \mathbf{1}\tau_n/H} (\hat{\pi}^*_{\tau_n} - A\nu)'\Omega(\hat{\pi}^*_{\tau_n} - A\nu). \end{cases}
\end{aligned}
$$

10 Key Takeaways

1. The RUM imposes linear restrictions: observed probabilities must lie in a convex cone generated by deterministic types.
2. Testing RUM reduces to computing the distance from $\hat{\pi}$ to this cone.
3. Cherchye–De Rock–Smeulders derive efficient optimization algorithms—active-set, LP relaxations, and matrix compression—that make this computation tractable.
4. Their framework provides a practical bridge between economic theory and statistical testing of revealed-preference models.