Majorization and Extreme Points: Economic Applications

Andreas Kleiner, Benny Moldovanu, and Philipp Strack

June 2020

Recent Project on Majorization and its Applications to Economics

- “*Auctions with Endogenous Valuations*”, joint with Alex Gershkov, Philipp Strack and Mengxi Zhang (2019).
- “*Revenue Maximization in Auctions with Dual Risk Averse Bidders: Myerson Meets Yaari*”, joint with Alex Gershkov, Philipp Strack and Mengxi Zhang (2020)
- “*Majorization and Extreme Points: Economic Applications*”, joint with Andreas Kleiner and Philipp Strack (2020)

Main Results of the Present Paper

1. Extreme-points characterization for sets of non-decreasing functions that are either majorized by - or majorize a given non-decreasing function.
2. Applications:
   a. Feasibility and optimality for multi-unit auction mechanisms.
   b. BIC-DIC equivalence.
   c. Welfare/revenue comparisons for matching schemes in contests.
   d. Equivalence between optimal delegation and Bayesian persuasion + new insights into their solutions.
   e. Rank-dependent utility, risk aversion and portfolio choice.

Majorization Preliminaries

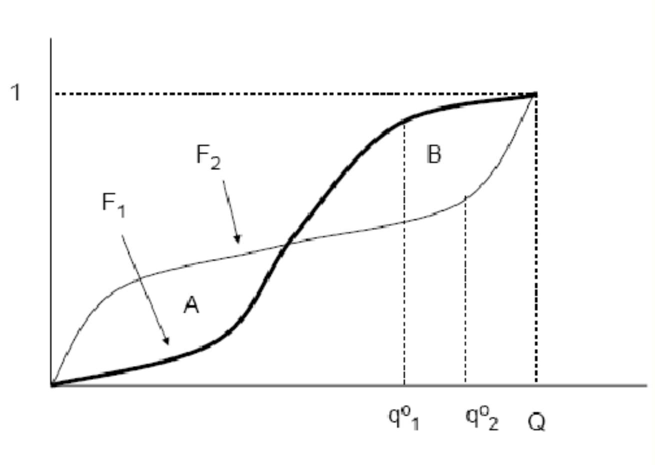
- We consider **only** non-decreasing functions $f, g : [0,1] \to \mathbb{R}$ such that $f, g \in L^1(0,1)$.
- We say that f *majorizes* g, denoted by $g \prec f$ if :

$$1. \int_x^1 g(s)ds \leq \int_x^1 f(s)ds \text{ for all } x \in [0,1]$$

$$2. \int_0^1 g(s)ds = \int_0^1 f(s)ds.$$

- We say that f *weakly majorizes* g, denoted by $g \prec_w f$ if 1 holds, but 2 not necessarily.

Majorization Preliminaries

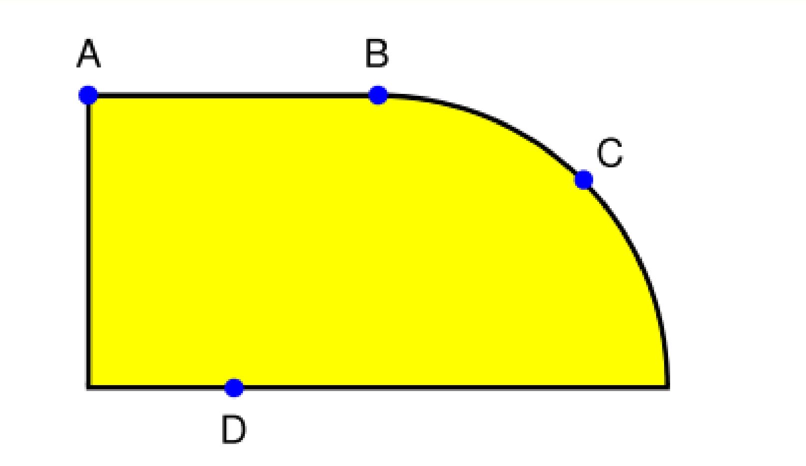


- $F_2 \prec F_1 \Leftrightarrow X_{F_2} \leq_{ssd} X_{F_1}$ and $\mathbb{E}[X_{F_1}] = \mathbb{E}[X_{F_2}] \Leftrightarrow F_1^{-1} \prec F_2^{-1}$

Convex Sets and their Extreme Points

- An *extreme point* of a convex set A is an element $x \in A$ that **cannot** be represented as a convex combination of two other elements in A.
- The *Krein–Milman Theorem* (1940): any convex and compact set A in a locally convex space is the closed, convex hull of its extreme points. In particular, such a set has extreme points.
- *Bauer's Maximum Principle* (1958): a convex, upper-semicontinuous functional on a non-empty, compact and convex set A of a locally convex space attains its maximum at an extreme point of A.

Convex Sets and their Extreme Points



An element x of a convex set A is *exposed* if there exists a linear functional that attains its maximum on A uniquely at x. B is **not** exposed.

Orbits and Choquet's (1960) Integral Representation

- Let $\Omega_m(f)$ denote the (monotonic) *orbit of* f :

$$\Omega_m(f) = \{g \mid g \prec f\}$$

- Let $\Phi_m(f)$ to be the (monotonic) *anti-orbit* of f :

$$\Phi_m(f) = \{g \mid f(0_+) \leq g \leq f(1_-) \text{ and } g \succ f\}$$

Theorem

The sets $\Omega_m(f)$ and $\Phi_m(f)$ are convex and compact in the L^1-norm topology. For any $g \in \Omega_m(f)$ there exists a probability measure λ_g supported on the set of extreme points of $\Omega_m(f)$, $ext\Omega_m(f)$, such that

$$g = \int_{ext\Omega_m(f)} h \, d\lambda_g(h)$$

and analogously for $g \in \Phi_m(f)$.

Orbits and their Extreme Points

Theorem

A non-decreasing function g is an extreme point of $\Omega_m(f)$ if and only if there exists a countable collection of disjoint intervals $\{[\underline{x}_i, \overline{x}_i)\}_{i\in I}$ such that a.e.

$$g(x) = \begin{cases} f(x) & \text{if } x \notin \cup_{i\in I}[\underline{x}_i, \overline{x}_i) \\ \frac{\int_{\underline{x}_i}^{\overline{x}_i} f(s)ds}{\overline{x}_i - \underline{x}_i} & \text{if } x \in [\underline{x}_i, \overline{x}_i). \end{cases}$$

Corollary

Every extreme point is exposed.

Orbits and their Extreme Points: Illustration

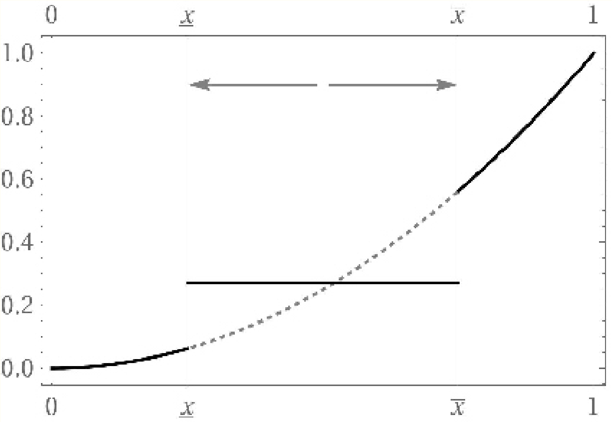


Figure: *1. Majorized Extreme Point*

Anti-Orbits and their Extreme Points

Theorem

A non-decreasing function g is an extreme point of $\Phi_m(f)$ if and only if there exists a collection of intervals $\{[\underline{x}_i, \overline{x}_i)\}_{i \in I}$ and (potentially empty) sub-intervals $[\underline{y}_i, \overline{y}_i) \subset [\underline{x}_i, \overline{x}_i)$ such that a.e

$$g(x) = \begin{cases} f(x) & \text{if } x \notin \bigcup_{i \in I} [\underline{x}_i, \overline{x}_i) \\ f(\underline{x}_i) & \text{if } x \in [\underline{x}_i, \underline{y}_i) \\ v_i & \text{if } x \in [\underline{y}_i, \overline{y}_i) \\ f(\overline{x}_i) & \text{if } x \in [\overline{y}_i, \overline{x}_i) \end{cases}$$

where v_i satisfies:

$$(\overline{y}_i - \underline{y}_i) v_i = \int_{\underline{x}_i}^{\overline{x}_i} f(s)\, ds - f(\underline{x}_i)(\underline{y}_i - \underline{x}_i) - f(\overline{x}_i)(\overline{x}_i - \overline{y}_i)$$

Anti-Orbits and their Extreme Points: Illustration

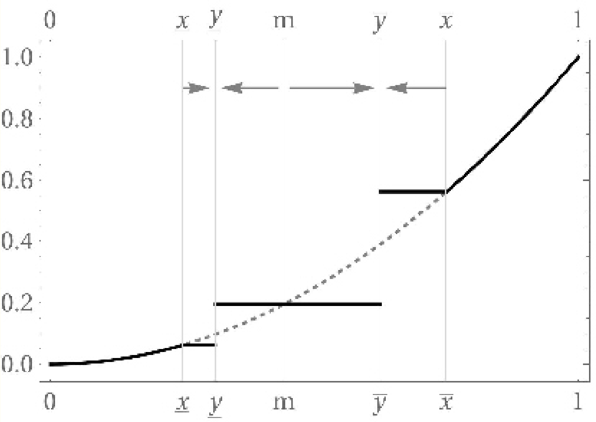


Figure: 2. *Majorizing Extreme Point*

Application: The SIPV Ranked-Item Allocation Model

- *SIPV* model with N agents. Types distributed on $[0, 1]$ according to F, with bounded density $f > 0$.
- W.l.o.g. N objects with qualities $0 \leq q_1 \leq \ldots \leq q_N = 1$. Each agent wants at most one object.
- If agent i with type θ_i receives object with quality q_m and pays t for it, then his utility is given by $\theta_i q_m - t$.
- Let Π be the set of *doubly sub-stochastic* $N \times N$-matrices.
- An *allocation* rule is given by $\boldsymbol{\alpha} : [0, 1]^N \to \Pi$, where $\alpha_{ij}(\theta_i, \theta_{-i})$ is the probability with which agent i obtains the object with quality j.

The SIPV Ranked-Item Allocation Model II

- Let $\boldsymbol{\alpha}^* : [0,1]^N \to \Pi$ denote the *assortative allocation* of objects to agents (highest type gets highest quality, etc.) with ties broken by fair randomization.
- Let
$$\varphi_i(\theta_i) = \int_{[0,1]^{N-1}} [\boldsymbol{\alpha}_i(\theta_i, \theta_{-i}) \cdot \mathbf{q}] \, f_{-i}(\theta_{-i}) \, d\theta_{-i}.$$
denote agent i's *interim allocation* (conditional on type) and let
$$\psi_i(s_i) = \varphi_i(F^{-1}(s_i))$$
be the *interim quantile allocation.*

Feasibility and BIC-DIC Equivalence

Theorem

1. *A symmetric and monotonic interim allocation rule φ is feasible if and only if its associated quantile interim allocation $\psi(s) = \varphi(F^{-1}(s))$ satisfies*

$$\psi \prec_w \psi^*$$

   where ψ^* is the quantile interim allocation generated by the assortative allocation α^*.

2. *For any symmetric, BIC mechanism there exists an equivalent, symmetric DIC mechanism that yields all agents the same interim utility, and that creates the same social surplus.*

The Fan-Lorentz (1954) Integral Inequality

- A functional $V : L^1(0,1) \to \mathbb{R}$ that is monotonic with respect to the majorization order is called *Schur-concave*.

Theorem

Let $K : [0,1] \times [0,1] \to \mathbb{R}$ *. Then*

$$V(f) = \int_0^1 K(f(t), t)\, dt$$

is Schur-concave if and only if $K(u,t)$ *is convex in* u *and super-modular in* (u,t)*.*

- Under twice-differentiability, the FL conditions become:

$$\frac{\partial^2 K}{\partial u^2} \geq 0 \;\; ; \;\; \frac{\partial^2 K}{\partial u \partial t} \geq 0$$

Application: Rank-Dependent Utility and Risk Aversion

- Utility with *rank-dependent* assessments of probabilities:

$$U(F) = \int_0^1 v(s)d(g \circ F)(s)$$

  where F is a distribution on $[0,1]$, $v : [0,1] \to R$ is continuous, strictly increasing and bounded, and $g : [0,1] \to [0,1]$ is strictly increasing, continuous and onto.

- v transforms monetary payoffs; g transforms probabilities.

- $g(x) = x$ yields von Neumann-Morgenstern expected utility, while $v(x) = x$ yields Yaari's (1987) dual utility.

Theorem

(Machina, 1982, Hong, Karni, Safra, 1987) The agent with preferences represented by U is risk averse if and only if both v and g are concave.

Linear Objectives and Schur-Concavity

Theorem

(Riesz, 1907) For every continuous, linear functional V on $L^1(0,1)$, there exists a unique, essentially bounded function $c \in L^\infty(0,1)$ such that for every $f \in L^1(0,1)$

$$V(f) = \int_0^1 c(x) f(x)\, dx$$

Corollary

By the Fan-Lorentz Theorem, the kernel

$$K(f,x) = c(x) f(x)$$

yields a Schur-concave (convex) functional $V \Leftrightarrow K$ is super-modular (sub-modular) in $(f,x) \Leftrightarrow c$ is non-decreasing (non-increasing).

Maximizing a Linear Functional on Orbits

- Consider the problem

$$\max_{h \in \Omega_m(f)} \int c(x) h(x)\, dx.$$

1. If c is non-decreasing, then f itself is the solution for the optimization problem.
2. If c is non-increasing, then the solution for the optimization problem is the overall constant function $g = \int_0^1 f(x)\, dx$. This follows since $g \in \Omega_m(f)$ and $h \succ g$ for any $h \in \Omega_m(f)$.
3. If c is not monotonic, other extreme points of $\Omega_m(f)$ may be optimal. They are obtained by an *ironing* procedure.

Application: Revenue Maximization

- The revenue maximization problem becomes

$$\max_{\psi \in \Omega_{m,w}(\psi^*)} N \int_0^1 \left[F^{-1}(s_1) - \frac{1 - s_1}{f(F^{-1}(s_1))} \right] \psi(s_1)\, ds_1$$

  where ψ^* is the interim quantile function induced by assortative matching.
- **Result:** the optimal solution is an extreme point of $\Omega_m(\psi^* \cdot \mathbf{1}_{[\widehat{s}_1,1]})$ for some $\widehat{s}_1 \in [0,1]$.
- Assuming that the *virtual value* is non-decreasing, we obtain by the FL Theorem that the optimal allocation $\widehat{\psi}$ satisfies:

$$\widehat{\psi}(s_1) = \begin{cases} \psi^*(s_1) & \text{for } s_1 \geq \widehat{s}_1 \\ 0 & \text{otherwise} \end{cases}.$$

Application: Matching Contests

- Let F be the distribution of types, and G be the distribution of prizes, both defined on $[0, 1]$.
- If type θ obtains prize y and pays t, her utility is given by $\theta y - t$.
- The *assortative* matching $\psi(\theta) = G^{-1}(F(\theta))$ is implemented by:

$$t(\theta) = \theta\psi(\theta) - \int_0^{\theta} \psi(t)dt$$

- High match value and high waste.
- Damiano and Li (2007), Hoppe et al. (2009, 2012),Olszewski and Siegel (2018) among others: What about other schemes (*random, coarse*)?

Matching Contests II

- Individual Utility and Welfare:

$$U(\theta) = \int_0^\theta G_{ic}^{-1}(F(t))dt \; ; \; W = \int_0^1 G_{ic}^{-1}(t)(1-t)dF^{-1}(t)$$

Theorem

1. *An allocation is feasible and implementable if and only if the induced distribution of prizes G_{ic} satisfies $G_{ic}^{-1} \prec G^{-1}$.*
2. *Assume that the distribution of types F is convex. Then each type of the agent prefers random matching to any other scheme.*
3. *Random matching (assortative matching) maximizes the agents' welfare if F has an Increasing (Decreasing) Failure Rate.*
4. *If F has an Increasing Failure Rate, the designer's revenue is maximized by assortative matching.*

Maximizing a Linear Functional on an Anti-Orbit

- Consider the problem

$$\max_{h \in \Phi_m(f)} \int c(x)h(x)\,\mathrm{d}x\,.$$

1. If c is non-increasing, then f solves this problem.
2. If c is non-decreasing, then the optimum is obtained at the step function g defined by

$$g(x) = \begin{cases} f(0_+) & \text{for } x < \bar{x} \\ f(1_-) & \text{for } x \geq \bar{x}, \end{cases}$$

   where $\bar{x}$ solves

$$\int_0^{\bar{x}} f(0_+)\,\mathrm{d}s + \int_{\bar{x}}^1 f(1_-)\,\mathrm{d}s = \int_0^1 f(s)\,\mathrm{d}s$$

   Indeed, it holds that $g \in \Phi_m(f)$ and $g \succ h$ for all $h \in \Phi_m(f)$.
3. If c is non-monotonic, other extreme points of $\Phi_m(f)$ may be optimal.

Application: Bayesian Persuasion

- The state of the world ω is distributed according to a prior F (common knowledge)
- Sender chooses a *signal* π : a signal realization space S and a family of distributions $\{\pi_\omega\}$ over S.
- Given π, a realization s induces a *posterior* F_s with mean x. Thus, a signal induces a distribution of posteriors, and hence a distribution of posterior means.

- The receiver first observes the choice of signal and the signal realization; then chooses an optimal action that depends on x, the expected value of the state .

- The sender's payoff v depends only on x (see *Dworczak and Martini* 2019, *Kolotilin* 2018).

Bayesian Persuasion II

- For any signal π, the prior F must be a *mean-preserving spread* of the generated distribution of posterior means G_π, i.e. $G_\pi \succ F$. Conversely, for any $G \succ F$ there exists a signal π such $G_\pi = G$.
- The sender's problem becomes:

$$\max_{G \in \Phi_m(F)} \int_0^1 v(x) dG(x)$$

Theorem

The optimal signal structure is a combination of three schemes:

1. *Reveal the state perfectly on an interval.*
2. *Pool all states in an interval so that only one signal realization is sent.*
3. *Send two different signal realizations on an interval.*

- Same result obtained independently by *Arieli et al* (2020)

Application: Optimal Delegation

- The state of the world θ is distributed according to F with support $[0, 1]$ and density f. Its realization is privately observed by an agent. The action space is the real line.
- The agent's and principal's Bernoulli utilities from a (deterministic) action a in state θ are given by

$$U_A(\theta, a) = -(\theta - a)^2, \; U_P(\theta, a) = -(\gamma(\theta) - a)^2$$

  where $\gamma : [0, 1] \to \mathbb{R}$ is bounded.

- A *direct mechanism* $M : [0, 1] \to \Delta(\mathbb{R})$ assigns to each agent's report a lottery over actions with finite variance.
- Denote by $\mu_M : [0, 1] \to \mathbb{R}$ the type-dependent mean action function and by $\sigma_M^2 : [0, 1] \to \mathbb{R}_+$ the type-dependent variance.

Optimal Delegation II

- Both the agent's and the principal's indirect utilities can be expressed as a function of μ_M and σ_M^2,

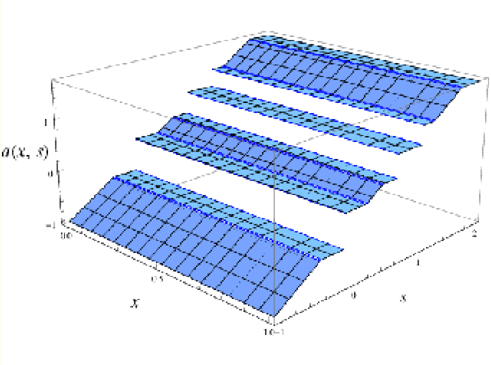
$$\begin{aligned} U_A(\theta) &= -(\theta - \mu_M(\theta))^2 - \sigma_M^2(\theta)\,, \\ U_P(\theta) &= -(\gamma(\theta) - \mu_M(\theta))^2 - \sigma_M^2(\theta), \end{aligned}$$

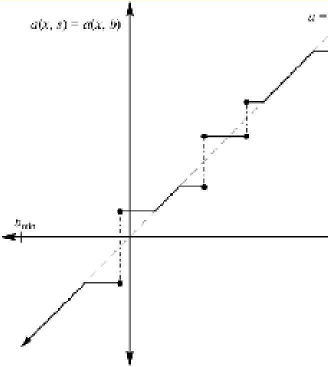
  and we write $M = (\mu_M, \sigma_M^2)$.

- Let $\Lambda = \sup_{\theta \in [0,1]} |\theta - \gamma(\theta)|$ and define

$$[\underline{a}\ ,\ \overline{a}] = [-\sqrt{2Var(\gamma(\theta) + 2\Lambda^2)}\ ,\ 1 + \sqrt{2Var(\gamma(\theta) + 2\Lambda^2)}]$$

Optimal Delegation and Majorization





IC Delegation Mechanisms

- We call a mechanism *undominated* if there does not exist a mechanism where the set of actions is a singleton, that yields a higher utility for the principal.

Theorem

A (potentially randomized) undominated mechanism $M = (\mu_M, \sigma_M^2)$ is incentive compatible if and only if there exists an extension $(\mu_{\tilde{M}}, \sigma_{\tilde{M}}^2)$ of (μ_M, σ_M^2) to the interval $[\underline{a}, \overline{a}]$ such that:

1. $\mu_{\tilde{M}}(\underline{a}) = \underline{a},\ \mu_{\tilde{M}}(\overline{a}) = \overline{a}$, $\sigma_{\tilde{M}}^2(\underline{a}) = \sigma_{\tilde{M}}^2(\overline{a}) = 0$
2. $\mu_{\tilde{M}} \in \Phi_m(a^*)$ *where* $a^* : [\underline{a}, \overline{a}] \to [\underline{a}, \overline{a}]$ *is the Identity function*
3. $\sigma_{\tilde{M}}^2(\theta) = -(\mu_{\tilde{M}}(\theta) - \theta)^2 - 2\int_{\underline{a}}^{\theta} (\mu_{\tilde{M}}(s) - s)\, ds$ *for all* $\theta \in [\underline{a}, \overline{a}]$.

The Principal's Problem

Theorem

1. *The principal's expected utility in an undominated, IC mechanism* $M = (\mu_M, \sigma^2_M)$ *with appropriate extension* $(\mu_{\tilde{M}}, \sigma^2_{\tilde{M}})$ *is given by*

$$V_P(\mu_{\tilde{M}}) = 2\int_{\underline{a}}^{\overline{a}} J(\theta)\,\mu_{\tilde{M}}(\theta)\,d\theta + C\,,$$

$$\text{where } J(\theta) = \begin{cases} 1 & \text{for } \theta \in [\underline{a}, 0) \\ 1 - F(\theta) + (\gamma(\theta) - \theta) f(\theta) & \text{for } \theta \in [0, 1] \\ 0 & \text{for } \theta \in (1, \overline{a}] \end{cases}$$

2. *The principal's problem is thus given by*

$$\max_{\mu_{\tilde{M}} \in \Phi_m(a^*)} V_P(\mu_{\tilde{M}})$$

Equivalence between Persuasion and Delegation

- Both exercises can be reduced to a maximization of a linear functional over an anti-orbit Φ_m. Hence, the basic structure of their respective optimal mechanisms is identical.
- The equivalence is general: it extends to optimal signal structures for Bayesian persuasion that are **not** monotone partitional. Such structures correspond then to randomized optimal delegation mechanisms.
- This simple observation generalizes the insight obtained by *Kolotilin and Zapechelnyuk* (2019) who restricted attention to deterministic delegation mechanisms and to monotone partitional signals, respectively.

Conclusion

- Characterizations of the extreme points of the sets of all monotonic functions that are either majorized by- or themselves majorize a given function.
- Many well-known optimization exercises in Economics can be rephrased as maximizing a convex functional over such sets. Hence, a maximum must be attained at one of the extreme points.
- Together with the Choquet integral representation, the characterizations of extreme points directly imply many results, both novel and well-known.
- Open Question: analogous extreme point characterization for notions of *multivariate majorization* and applications to models where the state is naturally multi-dimensional.