A Lagrangian Approach to Optimal Lotteries in Non-Convex Economies

Chengfeng Shen[1] Felix Kubler[2] Yucheng Yang[2] Zhennan Zhou[3]

BSE Summer Forum

June 2025

[1]Peking University
[2]University of Zurich
[3]Westlake University

Introduction

- Many economic models involve non-convex optimization problems
  1. **1.** Examples: models with private information $\Rightarrow$ nonconvex incentive constraints.
  2. **2.** Mathematically challenging to characterize, numerically difficult to solve.
- Lottery/Randomization/Mixed Strategy solutions to non-convex economies (Myerson, 1982; Prescott & Townsend, 1984; Arnott & Stiglitz, 1988):
  1. **1.** Linear (convex) problem in the probability space.
  2. **2.** May increase the value of the objective function.
- Main difficulty for lottery problems: linear programming in high dimensional space.
- This paper: a new Lagrangian iteration algorithm to solve for optimal lotteries as weighted average of deterministic solutions.

This paper

- A new Lagrangian iteration algorithm to efficiently solve for optimal lotteries.
  1. Bridge pure strategy and lottery systems through Lagrangian iteration.
  2. Lottery solution is weighted average of deterministic solutions along the iterations.
- Theoretical guarantee: correctness and convergence (sub-gradient descent).
- Complexity estimate: orders of magnitude better than the linear programming approach.
- Applications: (1) Moral hazard problem, (2) Optimal income tax with hidden types.
- From applications: (1) much faster and memory-saving than conventional methods; (2) new insights when the randomized tax scheme is welfare improving.

Literature

Lagrangian Iteration Method

Illustrative Example: Moral Hazard Problem

A continuum of representative agents take unobserved action $a \in A$, which affects output $q \in Q$ via $p(q|a)$. A and Q are finite sets. consumption $c \in C$ (compact).

Deterministic solution: the planner choose allocation $c(q)$ and recommend the agent to choose a, to solve

$$\max_{a,c(q)} \sum_q p(q|a)u(c(q), a), \tag{1}$$

subject to resource constraint & incentive compatibility constraint:

$$\begin{aligned} &\sum_q p(q|a)(c(q) - q) \leq 0; \\ &\sum_q p(q|a)u(c(q), a) \geq \sum_q p(q|\hat{a})u(c(q), \hat{a}), \ \forall \hat{a} \in A \end{aligned} \tag{2}$$

Problem can be highly non-convex: challenging to solve.

Lottery Solution to Moral Hazard Problem

In the lottery problem, the planner chooses $x \in \mathcal{P}(A \times C^{|Q|})$, $x = x(a, c(q_1), \cdots, c(q_{|Q|}))$ to:

$$\begin{aligned}
&\max \sum_{a \in A} \int_{c \in C^{|Q|}} x(a, dc) \sum_{q \in Q} p(q|a) u(c(q), a), \\
\text{s.t.} &\sum_{a \in A} \int_{c \in C^{|Q|}} x(a, dc) \sum_{q \in Q} p(q|a)(c(q) - q) \leq 0 \\
&\int_{c \in C^{|Q|}} x(a, dc) \sum_{q} p(q|a) u(c(q), a) \geq \int_{c \in C^{|Q|}} x(a, dc) \sum_{q} p(q|\hat{a}) u(c(q), \hat{a}), \ \forall (a, \hat{a}) \in A \times A.
\end{aligned} \tag{3}$$

This problem: linear in the probability space $\mathcal{P}(A \times C^{|Q|}) \Rightarrow$ directly solve with large-scale linear programming tools.

Challenge: dimension of $\mathcal{P}(A \times C^{|Q|})$ is very high even for simple problem!

Lagrangian Approach: Lagrangians of the Lottery vs Deterministic Problems

$$
\begin{aligned}
L(x;\lambda,\gamma) := & \sum_{a\in A}\int_{c\in C^{|Q|}} x(a,dc)\sum_{q\in Q} p(q|a)u(c(q),a) - \lambda\left(\sum_{a\in A}\int_{c\in C^{|Q|}} x(a,dc)\sum_{q\in Q} p(q|a)(c(q)-q)\right) \\
& - \sum_{\hat{a}\in A}\sum_{a\in A}\gamma_{\hat{a},a}\left(\int_{c\in C^{|Q|}} x(a,dc)\sum_{q} p(q|\hat{a})u(c(q),\hat{a}) - \int_{c\in C^{|Q|}} x(a,dc)\sum_{q} p(q|a)u(c(q),a)\right) \quad , \\
= & \sum_{a\in A}\int_{c\in C^{|Q|}} x(a,dc)\mathcal{L}(a,c;\lambda,\mu)
\end{aligned}
\tag{4}
$$

which is an “average” of the Lagrangian of the deterministic problem:

$$
\begin{aligned}
\mathcal{L}(a,c;\lambda,\gamma) = & \sum_{q} p(q|a)u(c(q),a) - \lambda\sum_{q} p(q|a)(c(q)-q) \\
& - \sum_{\hat{a}}\gamma_{\hat{a},a}\left[\sum_{q}\big(p(q|\hat{a})u(c(q),\hat{a}) - p(q|a^k)u(c(q),a)\big)\right].
\end{aligned}
$$

Idea: construct lottery solution as weighted average of deterministic solutions.

Lagrangian Iteration for the Lottery Problem

Initial guess of: λ^0, γ^0. In the k-th iteration,

- Update optimal allocation $(a^k, c^k) = \arg\max_{a,c} \mathcal{L}(a, c; \lambda^k, \gamma^k)$.
- Update Lagrangian multipliers λ^k and γ^k, with learning rate μ_k, e.g. $\mu_k = 1/k$:

$$\lambda^{k+1} = \max\left\{\lambda^k + \mu^k \sum_q p(q|a^k)(c^k(q) - q), 0\right\}.$$

$$\gamma^{k+1}_{\hat{a},a^k} = \max\left\{\gamma^k_{\hat{a},a^k} + \mu^k \left[\sum_q p(q|\hat{a})u(c^k(q),\hat{a}) - p(q|a^k)u(c^k(q),a^k)\right], 0\right\} \quad \forall \hat{a}$$

Intuition: Update Lagrangian multipliers according to how “close” the current allocation satisfies the inequality constraints.

Lagrangian Iteration for the Lottery Problem

Initial guess of: λ^0, γ^0. In the k-th iteration,

- Update optimal allocation $(a^k, c^k) = \arg\max_{a,c} \mathcal{L}(a, c; \lambda^k, \gamma^k)$.
- Update Lagrangian multipliers λ^k and γ^k, with learning rate μ_k, e.g. $\mu_k = 1/k$:

$$\lambda^{k+1} = \max\left\{\lambda^k + \mu^k \sum_q p(q|a^k)(c^k(q) - q), 0\right\}.$$

$$\gamma^{k+1}_{\hat{a},a^k} = \max\left\{\gamma^k_{\hat{a},a^k} + \mu^k \left[\sum_q p(q|\hat{a})u(c^k(q), \hat{a}) - p(q|a^k)u(c^k(q), a^k)\right], 0\right\} \quad \forall \hat{a}$$

Intuition: Update Lagrangian multipliers according to how "close" the current allocation satisfies the inequality constraints.

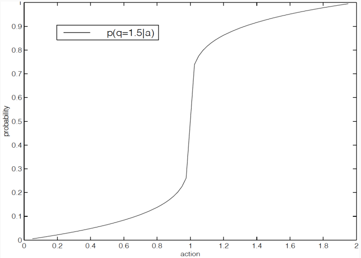
Final lottery solution:

$$x^N = \frac{1}{\sum_{k=1}^N \mu^k} \sum_{k=1}^N \mu^k \delta_{(a^k, c^k)}.$$

Calibration (Prescott, 1998)

- $u(a, c) = \sqrt{c} + 0.8\sqrt{2-a}$.
- Consumption set: $C = [0, 2]$, output set $Q = \{0.5, 1.5\}$, action set: $A = 0.05 : \Delta a : 1.95$, with $\Delta a = 0.025$. Output distribution as function of action:

$$p(q = 1.5 \mid a) = \begin{cases} \frac{1-(1-a)^{0.2}}{2}, & \text{if } a < 1, \\ \frac{1+(a-1)^{0.2}}{2}, & \text{if } a \geq 1. \end{cases}$$



- Solution (Prescott, 1998):

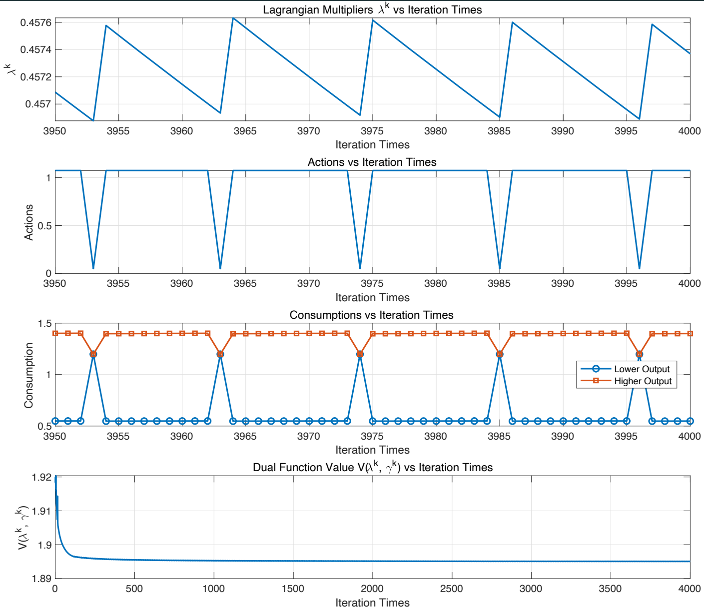
$$\pi(a = 0.050) = 0.0924,\ \pi(a = 1.075) = 0.9076$$

$$\pi(c = 1.20 \mid q = 0.5, a = 0.05) = 1,\ \pi(c = 1.19 \mid q = 1.5, a = 0.05) = 1.$$

$$\pi(c = 0.54 \mid q = 0.5, a = 1.075) = 0.5311,\ \pi(c = 0.55 \mid q = 0.5, a = 1.075) = 0.4689.$$

$$\pi(c = 1.40 \mid q = 1.5, a = 1.075) = 1.$$

Lagrangian multiplier λ^k, allocation a^k, c^k, dual value $V(\lambda^k, \gamma^k)$ along iteration



Theoretical Framework

General Framework

Deterministic problem with action $a \in A$ (finite), consumption $c \in C$ (compact), payoff function f:

$$\begin{aligned} \max_{a \in A,\, c \in C} \; & f(a,c), \\ \text{s.t. } \; & g_i(a,c) \le 0 \quad i \in \{1, \dots, m\}, \\ & h_j(a,c) \le 0 \quad j \in \{1, \dots, \ell\}, \end{aligned} \tag{5}$$

Lottery problem with probability $x(a, dc) \in \mathcal{P}(A \times C)$:

$$\begin{aligned} \max_{x \in \mathcal{P}(A \times C)} \; & \sum_{a \in A} \int_{c \in C} f(a,c) x(a,dc), \\ \text{s.t. } \; & \sum_{a \in A} \int_{c \in C} g_i(a,c) x(a,dc) \le 0 \quad \forall i \in \{1, \dots, m\}, \\ & \int_{c \in C} h_j(a,c) x(a,dc) \le 0 \quad \forall a \in A,\, j \in \{1, \dots, \ell\}. \end{aligned} \tag{6}$$

Lagrangian Iteration: General Setup

Given $\lambda_i^1 (i \in \{1, \cdots, m\})$, $\gamma_{j,a}^1 (a \in A,\, j \in \{1, \cdots, \ell\})$, $\mu^k \in \mathbb{R}_+$, $N \in \mathbb{N}_+$. For $k = 1 : N$,

Step 1. Solve the Lagrangian problem.

$$(a^k, c^k) \in \arg \max_{a \in A, c \in C} \mathcal{L}(a, c; \lambda^k, \gamma^k).$$

Step 2. Update the Lagrangian multipliers.

$$\lambda_i^{k+1} = \max\{\lambda_i^k + \mu^k g_i(a^k, c^k), 0\},\ \forall i \in \{1, \cdots, n\}.$$

$$\gamma_{j,a^k}^{k+1} = \max\{\gamma_{j,a^k}^k + \mu^k h_j(a^k, c^k), 0\},\ \forall j \in \{1, \cdots, \ell\}.$$

$$\gamma_{j,a}^{k+1} = \gamma_{j,a}^k,\ \forall j \in \{1, \cdots, \ell\}, a \neq a^k.$$

Step 3. Construct the lottery solution with $\delta_{(a^k,c^k)}$ as $\delta-$measure at the point (a^k, c^k).

$$x^N := \frac{1}{\sum_{k=1}^N \mu^k} \sum_{k=1}^N \mu^k \delta_{(a^k,c^k)},$$

Main Theorem

Theorem
Suppose the sequence of learning rates $(\mu^k)_{k=1}^{\infty}$ *satisfies*

$$\sum_{k=1}^{\infty} \mu^k = \infty \text{ and } \sum_{k=1}^{\infty} (\mu^k)^2 < \infty.$$

Let x^* *be the solution to the lottery problem, and suppose the corresponding Lagrangian multipliers to* x^* *exist. Then for any* $\epsilon > 0$*, there exists* $N \in \mathbb{N}_+$*, such that when* $n > N$*,* x^n *obtained from the Algorithm is an* $\epsilon-$*optimal solution to the problem.*

Note: $\epsilon-$optimal solution is the solution that maximizes the same objective function subject to ϵ relaxation of the constraints such as $\sum_{a \in A} \int_{c \in C} g_i(a, c) x(a, dc) \leq \epsilon$.

Why Lagrangian Iteration? Some Theoretical Insights

1. Lottery solution only assigns positive probability to (unconstrained) deterministic allocations that maximize the Lagrangian
2. Value of the saddle point of the Lagrangian of the lottery problem $=$ value of the dual of the deterministic problem
   $\Rightarrow$ iteratively solve the dual of the deterministic problem to solve the lottery problem.
3. Frequency of different optimal points of the deterministic system along the iterations approximates the probability distribution of the optimal lottery solution

Above are the theoretical motivations to the Lagrangian iteration method, and we'll prove that the algorithm indeed converges to the lottery solution.

Thm 1: given λ, γ, deterministic/lottery Lagrangians have same optimal value

Lagrangian in the pure strategy space $A \times C$:

$$\mathcal{L}(a, c; \lambda, \gamma) := f(a, c) - \sum_{i=1}^{n} \lambda_i g_i(a, c) - \sum_{j=1}^{\ell} \gamma_{j,a} h_j(a, c). \tag{7}$$

Lagrangian in the probability space $\mathcal{P}(A \times C)$:

$L(x; \lambda, \gamma) :=$

$$\sum_{a \in A} \int_{c \in C} f(a, c) x(a, dc) - \sum_{i=1}^{n} \lambda_i \sum_{a \in A} \int_{c \in C} g_i(a, c) x(a, dc) - \sum_{j=1}^{\ell} \sum_{a \in A} \gamma_{j,a} \int_{c \in C} h_j(a, c) x(a, dc), \tag{8}$$

Given Lagrangian multipliers λ, γ, we prove:

$$\max_{x \in \mathcal{P}(A \times C)} L(x; \lambda, \gamma) = \max_{a \in A, c \in C} \mathcal{L}(a, c; \lambda, \gamma). \tag{9}$$

Thm 2: given λ, γ, optimal lottery only contains optimal deterministic solutions

Furthermore, if we define $Z = \arg\max_{a \in A, c \in C} \mathcal{L}(a, c; \lambda, \gamma)$, then

$$x^* \in \arg \max_{x \in \mathcal{P}(A \times C)} L(x; \lambda, \gamma)$$

if and only if

$$(a, c) \in Z \text{ a.s. with respective to the probability measure } x^*. \tag{10}$$

Thm 3: equivalence of lottery solution to the dual of deterministic problem

Theorem

Suppose x^* is the solution to the lottery problem, and the Lagrangian multipliers corresponding to x^* exist, denoted as $(\lambda^*, \gamma^*) \in \mathbb{R}^m \times \mathbb{R}^{\ell|A|}$. Then (λ^*, γ^*) is (part of) the solution to the dual problem of the deterministic problem, i.e.

$$\max_{a \in A,\, c \in C} \mathcal{L}(a, c; \lambda^*, \gamma^*) = \inf_{(\lambda,\gamma) \in \mathbb{R}^m_+ \times \mathbb{R}^{\ell|A|}_+} \max_{a \in A,\, c \in C} \mathcal{L}(a, c; \lambda, \gamma), \tag{11}$$

Furthermore, the optimal objective value of the dual problem above is the same as the optimal objective value of the lottery problem:

$$\inf_{(\lambda,\gamma) \in \mathbb{R}^m_+ \times \mathbb{R}^{\ell|A|}_+} \max_{a \in A,\, c \in C} \mathcal{L}(a, c; \lambda, \gamma) = \sum_{a \in A} \int_{c \in C} f(a, c) x^*(a, dc). \tag{12}$$

Remark

- Above are three theoretical motivations (proved as theorems in the paper) for the Lagrangian iteration method.
- Now we'll prove that the algorithm indeed converges to the lottery solution.

Proof Step 1: given convergence, solution is correct (Step 1a)

Step 1a: given convergence, the solution satisfies the inequality constraints.

Want to show

$$\sum_{a\in A}\int_{c\in C} g_i(a,c)x^\epsilon(a,dc) \le \epsilon, \quad \text{for } i \in \{1,\cdots,m\}.$$

For $i \in \{1, ..., m\}$, we have

$$\sum_{a\in A}\int_{c\in C} g_i(a,c)x^n(a,dc) = \frac{1}{\sum_{k=1}^n \mu_k}\sum_{k=1}^n \mu_k g_i(a^k,c^k).$$

By the updating rule for λ_i, written

$$\lambda_i^{k+1} = \max\{\lambda_i^k + \mu^k g_i(a^k,c^k), 0\} \ge \lambda_i^k + \mu^k g_i(a^k,c^k), \quad k = 1,\cdots,n,$$

Proof Step 1: given convergence, solution is correct (Step 1a, Ct'd)

$$\sum_{k=1}^{n} \lambda_i^{k+1} \geq \sum_{k=1}^{n} \left[\lambda_i^k + \mu^k g_i(a^k, c^k) \right],$$

which can be simplified as

$$\lambda_i^{n+1} \geq \lambda_i^1 + \sum_{k=1}^{n} \mu_k g_i(a^k, c^k) = \lambda_i^1 + \left(\sum_{k=1}^{n} \mu^k \right) \sum_{a \in A} \int_{c \in C} g_i(a, c) x^n(a, dc).$$

Hence

$$\sum_{a \in A} \int_{c \in C} g_i(a, c) x^n(a, dc) \leq \frac{\lambda_i^{n+1} - \lambda_i^1}{\sum_{k=1}^{n} \mu^k}. \tag{13}$$

Proof Step 1: given convergence, solution is correct (Step 1b)

Step 1b: given convergence, the solution attains the optimum.

$$
\begin{aligned}
\frac{1}{\sum_{k=1}^{n}\mu^k}\sum_{k=1}^{n}\mu^k f(a^k,c^k) \geq \frac{1}{\sum_{k=1}^{n}\mu^k}\sum_{k=1}^{n}\mu^k \Bigg(& f(a,c) - \sum_{i=1}^{m}\lambda_i^k g_i(a,c) \\
& - \sum_{j=1}^{\ell}\sum_{\hat{a}\in A}\gamma_{j,\hat{a},a}^k h_j(a,c,\hat{a}) + \underbrace{\sum_{i=1}^{m}\lambda_i^* g_i(a^k,c^k) + \sum_{j=1}^{\ell}\sum_{\hat{a}\in A}\gamma_{j,\hat{a},a^k}^* h_j(a^k,c^k,\hat{a})}_{K_n^1:\text{approximately } \geq 0 \text{ according to Step 1a}} \\
& + \underbrace{\sum_{i=1}^{m}(\lambda_i^k-\lambda_i^*)g_i(a^k,c^k) + \sum_{j=1}^{\ell}\sum_{\hat{a}\in A}(\gamma_{j,\hat{a},a^k}^k-\gamma_{j,\hat{a},a^k}^*)h_j(a^k,c^k,\hat{a})}_{K_n^2:\text{approximately } =0 \text{ according to the convergence of } \lambda^k \text{ and } \gamma^k} \Bigg) \\
= f(a,c) - \frac{1}{\sum_{k=1}^{n}\mu^k}\sum_{k=1}^{n}\mu^k\sum_{i=1}^{m}\lambda_i^k g_i(a,c) & - \frac{1}{\sum_{k=1}^{n}\mu^k}\sum_{k=1}^{n}\mu^k\sum_{j=1}^{l}\sum_{\hat{a}\in A}\gamma_{j,\hat{a},a}^k h_j(a,c,\hat{a}) + K_n^1 + K_n^2
\end{aligned}
$$

Proof Step 1: given convergence, solution is correct (Step 1b, Ct'd)

$$
\begin{aligned}
&\frac{1}{\sum_{k=1}^{n}\mu^k}\sum_{k=1}^{n}\mu^k f(a^k,c^k) \\
\geq &\sum_{a\in A}\int_{c\in C} f(a,c)x^*(a,dc) - \frac{1}{\sum_{k=1}^{n}\mu^k}\sum_{k=1}^{n}\mu^k\sum_{i=1}^{m}\lambda_i^k\sum_{a\in A}\int_{c\in C} g_i(a,c)x^*(a,dc) \\
&- \frac{1}{\sum_{k=1}^{n}\mu^k}\sum_{k=1}^{n}\mu^k\sum_{j=1}^{l}\sum_{\hat{a}\in A}\sum_{a\in A}\gamma_{j,\hat{a},a}^k\int_{c\in C} h_j(a,c,\hat{a})x^*(a,dc) - \epsilon \\
\geq &\sum_{a\in A}\int_{c\in C} f(a,c)x^*(a,dc) - \epsilon,
\end{aligned}
$$

Proof Step 2: our algorithm does converge

Proposition

Suppose the sequence $(\mu^k)_{k=1}^{\infty}$ *satisfies*

$$\sum_{k=1}^{\infty} \mu^k = \infty, \quad \textit{and} \quad \sum_{k=1}^{\infty} (\mu^k)^2 < \infty.$$

Let x^* *be the optimal solution to the lottery problem, and corresponding Lagrangian multipliers to* x^* *exist. Then* (λ^k, γ^k) *generated by our Algorithm converge to some* (λ^*, γ^*)*, where* (λ^*, γ^*) *is a minimizer of the dual problem*

$$\inf_{(\lambda,\gamma)\in\mathbb{R}_+^m\times\mathbb{R}_+^{\ell|A|}} V(\lambda, \gamma) := \max_{a\in A,\, c\in C} \mathcal{L}(a, c; \lambda, \gamma).$$

Proof follows the sub-gradient descent literature (Shor, 2012; Nedic and Bertsekas, 2001)

Wrap-up: Key Steps and Main Ideas of the Proof

1. Given the Lagrangian multipliers, the deterministic/lottery Lagrangians have the same optimal value.

2. The value of the Lagrangian of the lottery system must be the same as the value of the dual of the deterministic problem $\Rightarrow$ solve the dual of the deterministic problem to solve the lottery problem.

3. If the Lagrangian iteration method converges, the solution we construct must
   1. Satisfy all the constraints
   2. Achieve the optimal value

4. Following the sub-gradient descent literature, the algorithm we design must converge.

Computational Complexity Analysis

- Take $\mu^k \sim k^{\frac{1}{2}(1+\rho)}$ for $0 < \rho < 1$. Then for $\epsilon > 0$, the overall computational complexity for finding an ϵ-optimal lottery solution with Lagrangian iteration is
$$O\left(\left(\frac{\frac{M}{\rho}+\bar{\Lambda}}{\epsilon}\right)^{\frac{2}{1-\rho}} |A||\hat{C}|(m+\ell)^{1+\frac{1}{1-\rho}}\right)$$
where $\hat{C}$ is the discretized set of C.
- Comparing the complexity to that of the linear programming interior point method, we see in the case that $|\hat{C}| \sim |A| \sim \ell \gg m$,
$$|A||\hat{C}|(m+\ell)^{1+\frac{1}{1-\rho}} \sim |A|^{3+\frac{1}{1-\rho}} \ll |A|^7 \sim (|A||\hat{C}| + \ell|A| + m)^{3.5},$$
for ρ close to 0.
- Bigger computational advantage if optimal deterministic c can be solved by FOC.

Computational Performance for Moral Hazard Problem with Different Δa

Δa	Iterations	CPU time	LP CPU time	Size of LP		
				#variable	#EC	#IC
0.2	500	0.006	0.05	4020	21	91
0.1	1000	0.01	0.13	8040	41	381
0.05	2000	0.02	0.44	15678	79	1483
0.025	4000	0.05	1.33	30954	155	5853
0.0125	8000	0.16	6.84	61506	307	23257
0.00625	16000	0.98	-	122610	611	92721

Note: $\Delta a = 0.0125$, it takes LP method 6.84 s to solve (2-3 orders of magnitude slower). LP cannot handle $\Delta a < 0.0125$ on a laptop due to memory limit.

Application II: Optimal Income Tax

Optimal Income Tax with Hidden Types

Agent preference a la King, Plosser, and Rebelo (1988) with hidden types $\sigma_h, \kappa_h \in (0,1)$:

$$u_h(c,\ell) = (c^{1-\kappa_h}(1-\ell)^{\kappa_h})^{1-\sigma_h}.$$

A utilitarian social planner solves

$$\max_{x\in\mathcal{P}(C\times L)} \int_{(c,\ell)\in C\times L} \left(\sum_{h\in H} u_h(c_h,\ell_h)\right) x(dc,d\ell), \tag{14}$$

$$\text{s.t.} \int_{(c,\ell)\in C\times L} \left(u_h(c_h,\ell_h) - u_h(c_{h'},\ell_{h'})\right) x(dc,d\ell) \geq 0 \text{ for all } h,h' \in H \tag{15}$$

$$\int_{(c,\ell)\in C\times L} \left(\sum_h c_h - \sum_h \ell_h\right) x(dc,d\ell) \leq 0. \tag{16}$$

This is equivalent to optimal taxation with income tax schedule $T(\ell)$ and let each agent choose labor supply ℓ and consumption $c = \ell - T(\ell)$.

Optimal Income Tax with Hidden Types: Welfare Gain from Lotteries

Calibration: four households.

HH	(κ_h, σ_h)	**Lottery**	**Deterministic**	**Full-Info**	**Welfare Loss**	
					Lottery	Deterministic
1	(0.2, 0.1)	$\pi_1 = 0.76$: (2.69, 0.36) $\pi_2 = 0.23$: (0, 1)	(0.8, 0.8)	(1.94, 0.52)		
2	(0.2, 0.9)	(0.16, 0.96)	(0.8, 0.8)	(0.12, 0.97)	3.39%	27.99%
3	(0.8, 0.1)	$\pi_1 = 0.83$: (0.29, 0) $\pi_2 = 0.17$: (0, 1)	(0.2, 0.2)	(0.27, 0)		
4	(0.8, 0.9)	(0.04, 0.85)	(0.2, 0.2)	(0.03, 0.88)		

Welfare differences as Hicksian "compensating variation in resources" to achieve the same level of welfare as the full-information problem.

Lottery scheme generates large welfare gain when agents are very different in risk aversion.

Conclusion

- A new Lagrangian iteration algorithm to efficiently solve for optimal lotteries.
- Theoretical guarantee: correctness and convergence (sub-gradient descent).
- Complexity estimate: orders of magnitude better than the conventional approach.
- Applications: (1) Moral hazard problem, (2) Optimal income tax with hidden types.
- From applications: (1) much faster and memory-saving than conventional methods; (2) new insights when the randomized tax scheme is welfare improving.

Appendix

Related Literature

- Lottery solution to problems with non-convex constraints.
  1. Myerson (1982), Prescott and Townsend (1984a, 1984b), Arnott and Stiglitz (1988)
  2. Prescott (2004), Prescott and Townsend (2006), Doepke and Townsend (2006)
- Computational methods for moral hazard and optimal tax with hidden types:
  1. Su and Judd (2007), Armstrong et al. (2010), etc.
  2. Weiss (1976), Brito et al. (1995), Hellwig (2007), Gauthier and Laroque (2014), Judd et al (2017), among many others.
- Math literature on sub-gradient descent: Shor (2012); Nedic and Bertsekas (2001)

back