A Computational Method for Multidimensional Continuous-choice Dynamic Problems

Xiaolu Zhou[*]

School of Economics

Xiamen University

January 31, 2018

Abstract

This paper proposes a new approximation method for solving the continuous choice dynamic problems. This method reduces the computational burden caused by the "curse of dimensionality" for multidimensional action space. I show that the distribution of optimal actions can be analytically calculated on quasi-Monte Carlo points in the action space via the first order condition and a change of variables. Therefore, the value function iteration is able to be updated by the weighted average of action-specific value functions. I apply this method to a finite-horizon dynamic pricing problem of inventories with stochastic demand, and use it to illustrate the computational advantage of my method.

Keywords: Dynamic Programming, Continuous Choice, Curse of Dimensionality

JEL Codes:

[*] `xzz109@psu.edu`

1 Introduction

Over the past decades, there has been considerable activities in studying stochastic continuous-choice dynamic problems in economics, such as, dynamic searching and learning-by-doing in macroeconomics (Adda and Cooper (2003)), dynamic labor supply model in labor economics (Blundell and Macurdy (1999)), R&D investment, and airlines' or retailers' dynamic pricing in industrial organization and revenue management literatures (Gallego and van Ryzin (1994)).

However, such dynamic problems are hard to be solved and estimated in practice. Because the continuous-choice problems are analytically intractable, the solution requires numerically computing the policy function within the estimation loop.[1] If the action enters the problem in an nonlinear way, policy functions have to be numerically solved from nonlinear problems. This optimization step is costly for value function iteration substantially, and also slows down the estimation substantially (Judd (1999) Section 12.3). Even if methods that accelerate the optimization step have been proposed (see Judd (1999) Section 12.4), the solution of such problems still faces challenges about "curse of dimensionality" for multidimensional action spaces.

The existing method of solving optimal actions nonlinearly is to discretize the action space. Although this method avoids the calculation of the value function outside the grid points, and thus reduces the computational time, but at the same time, it requires a small step for the action space grid points. The discretization method is still computationally expensive for problems with a large action space. Also, this method becomes intractable for multi-dimensional actions, as the number of the grid points grows exponentially with the action space dimension, which is also known as the "curse of dimensionality".

Thus motivated, I propose a new computational method to solve the continuous-choice dynamic programming. My method not only avoids numerically solving policy functions from nonlinear problems, but also ameliorates the computational burden caused by the "curse of dimensionality". Thus, it is powerful for multidimensional actions. I believe this method extends the applicability of the continuous-choice dynamic programming in empirical works.

[1] In discrete choice dynamic problem as Rust (1987), if the value function is separable with the independent shock, and the shock follows an extreme value distribution, the choice probability and value function both have functional forms.

In stochastic continuous-choice dynamic problems, an individual know the states and see an *i.i.d* shock that is only realized at the beginning of each period. Based on the states and shock, she want to choose an action to maximize the long-run utilities. Her action affects current period's utility and the expected value afterwards through the state transition. Such continuous-choice dynamic models are usually solved from the Bellman equation, a mapping from the next period's value function to current period's value function, through value function iteration for infinite-horizon problems, or backward induction for finite-horizon problems.

Same with many other methods, I still work on discrete or discretized state space. However, instead of discretizing the action space with a small step, I generate quasi-Monte Carlo points from the action space. I then calculate the probability that each action point is optimal conditional on states and next period's value functions, i.e. the distribution of optimal actions conditional on states. The conditional distribution of optimal action can be computed analytically if we exert some functional form assumptions on the *i.i.d* shocks. Because, with such assumptions, the shock corresponding to each action point conditional on states can be calculated with an analytical form, the distribution of optimal action is thus known from the distribution of shocks with a change of variable. With the possibilities that each action point is optimal, the value function is then updated by the weighted average of the action-specific value functions using these possibilities as weights. The dynamic programming is thus solved when the value function iteration stops at a small movement.

The first contribution of this paper is to propose a new approximation way for the policy function without numerically solving nonlinear problems. My method is based on solving the distribution of optimal actions on some points in the action space. Specifically, given discrete or discretized states and a guess for the continuous value, I calculate the "optimal" shock corresponding to the point in the action space from the first order condition. This mapping is easy to be solved analytically if we allow shocks entering in the problem in a certain way, for example, in the the first order polynomial. Then, by having a functional form distribution on shocks, the distribution of optimal actions is therefore able to be computed via a change of variables, and the value function is updated by the weighted average of the action-specific value functions.

My method is built on solving the distribution of actions on a set of points in the action space. Thus, it is important to choose the set of action points with a majority coverage of the optimal actions space. I explain how to use economic intuitions and data to predict the optimal action space. With the knowledge, we can generate points points from the space. The points can be evenly distributed grid points or quasi-Monte Carlo points.

The second contribution is that my method reduces the computational burden for problems with multidimensional actions. The existing method to accelerate the solution include policy function iteration, policy and value function approximation (see Judd (1999) Section 12.4). However, all of above methods suffer from the "curse of dimensionality". For example, in the discretization method, the number of grid points increases exponentially with the dimension of the action space. Think about a firm with k products, who want to set optimal prices p for each product, we need n^k grid points for the k-dimensional action space with n points for each dimension. Similarly, the approximation for policy function is not accuracy if the action is high dimension. However, my method can work on quasi-Monte Carlo points that do not grow exponentially with the dimension. Actually, the computationally burden increases in a square rate in my method, because it needs a k-by-k Jacobian matrix to calculate the probability of optimal actions via a change variables for each action point.

Lastly, I apply my method on a finite-horizon dynamic pricing problem with stochastic demand. This example help to compare the computational performance of my method with the discretization method. Results show that my method uses much less points in the action space, but has little differences with discretizing the action space in approximating the value function. This difference increases with the dimension of the action space. My method saves 48% and 71% computational time relative to the discretization method for one-dimensional and three-dimensional actions respectively.

The rest of this paper is organized as follows. Section 2 describes a framework for continuous-choice dynamic problems. Section 3 explains the computational method, and section 4 gives two examples. Section 5 illustrates the estimation procedure. Section 6 demonstrates the performance of my method with an numerical example.

2 A Framework for Continuous-choice Dynamic Problems

Consider a dynamic programming problem in which a firm makes a series of J-dimensional continuous choices $a_t = (a_{1t}, \ldots, a_{Jt})$ over his lifetime $t \in \{1, \ldots, T\}$. A firm can be described by a vector of characteristics (z_t, ϵ_t), where $\epsilon_t = (\epsilon_{1t}, \ldots, \epsilon_{Jt})$ is a J-dimensional shock that is only realized at the beginning of period t. Suppose that the shock ϵ_t is independent and identically distributed over time with continuous support and distribution function $G_\epsilon(\epsilon_t)$, and the vector z_t evolves as a Markov process, depending stochastically on the choices of the firm. The probability of z_{t+1} conditional on being in z_t and making choice a_t at time t is given by $f(z_{t+1}|z_t, a_t)$. At the beginning of each period t, the firm observes $(z_t, \epsilon_t(a_t))$. The firm makes a continuous choice a_t to sequentially maximize the expected discounted sum of profits

$$\max_{a_t} \sum_{t=1}^{T} \beta^{t-1} \pi(z_t, a_t, \epsilon_t) \tag{1}$$

where $\pi(z_t, a_t, \epsilon_t)$ denotes the current profit of a firm with characteristics z_t from choosing a_t. The discount factor is denoted by $\beta \in (0, 1)$, and the state z_t is updated at the end of each period.

Let $V(z_t)$ be the expected value of lifetime profit at date t as a function of the current state z_t:

$$V(z_t) = E_{\substack{z_{t+1},\ldots,z_T|z_t \\ \epsilon_t,\ldots,\epsilon_T}} \left\{ \max_{a_t} \sum_{\tau=t}^{T} \beta^{\tau-t} \pi(z_\tau, a_\tau, \epsilon_\tau)|z_t \} \right\} \tag{2}$$

where the expectation is taken on all of future states $(z_{t+1}, \ldots, z_T)$ conditional on current state z_t, and the shock from the current period until the final period $(\epsilon_t, \ldots, \epsilon_T)$. The final period T goes to infinity for infinite-horizon problems, and T is a finite number for finite-horizon problems. Write above problem with the recursive form:

$$V(z_t) = \int_{\epsilon_t} \left\{ \max_{a_t} \pi(z_t, a_t, \epsilon_t) + \beta \sum_{z_{t+1}} V(z_{t+1}) f(z_{t+1}|z_t, a_t) \right\} dG_\epsilon(\epsilon_t) \tag{3}$$

The value functions are given by current period profit plus the expected value of future utility. Let $\widetilde{a}^* = \widetilde{a}(z_t, \epsilon_t, V(z_{t+1}))$ be the optimal continuous choice the firm would like to choose conditional on state z_t after observing ϵ_t given the next period's value function $V(z_{t+1})$.

$$\widetilde{a}^* = \arg\max_{a \in A} \left\{ \pi(z_t, a, \epsilon_t) + \beta \sum_{z_{t+1}} V(z_{t+1}) f(z_{t+1}|z_t, a) \right\} \tag{4}$$

where A is the action space. Write the Bellman equation as a contraction mapping from next period's value function to current period's value function conditional on the state z_t, and the optimal action $\widetilde{a}^* = \widetilde{a}(z_t, \epsilon_t, V(z_{t+1}))$:

$$V_t(z_t) = \int_{\epsilon} \left\{ \pi(z_t, \widetilde{a}^*, \epsilon_t) + \beta \sum_{z_{t+1}} V(z_{t+1}) f(z_{t+1}|z_t, \widetilde{a}^*) \right\} dG_{\epsilon}(\epsilon_t). \tag{5}$$

This defines a mapping from the last period's value function to the first period's value function

$$V_t = T(V_{t+1}; \widetilde{a}^*) \quad \text{for } t = T-1, \ldots, 1 \tag{6}$$

where $V_{t+1} = \{V(z_{t+1})\}$ is a vector of the value function at different states. For infinite-horizon problems, equation 7 becomes a contraction mapping for the stationary value functions

$$V = T(V; \widetilde{a}^*) \tag{7}$$

where $V = \{V(z)\}$ is also a vector of the value function at different states.

The solution of the continuous-choice dynamic problem $V(z_t)$ can be computed by iterating over 6 until the time period goes to the initial period in finite-horizon problems, or over 7 until the value function converges in infinite-horizon problems. In the dynamic discrete choice model, the Bellman equations in 5 have functional forms if the action-specific shock linearly enter the value function and follows the generalized extreme value distributions. However, in continuous-choice problems, each iteration over 5 has to be computed numerically. The process is to simulate shocks ϵ_t from its distribution $G_\epsilon(\epsilon_t)$, calculate the optimal action $\widetilde{a}(z_t, \epsilon_t, V(z_{t+1}))$ at each state z_t and each shock ϵ_t given next period's value function, and integrate current value function over the simulated shocks. This means that we have to solve the policy function for each value function iteration and each parameter. However, if the action nonlinearly enter the current period profit function or the transition probabilities, numerically solving the optimal action for each iteration and each paraemter is intractable.

One approximation method for the policy function is to discretize the action space, calculate the action-specific value function at each grid point, and choose the point generating the maximum value as the optimal action. This method avoids solving optimal actions from maximization problems or nonlinear first order conditions. However, it is still computational expensive for a large action space, because the grid points' step size has to be small to precisely approximate the continuous action. Also, it is even intractable for multidimensional actions, as the number of grid points grow exponentially with the dimension of the action space. For example, a firm with 5 products want to set a 5-dimensional action to maximize the long-run profits. If there are 10 number of grid points for each dimension, the total number of points is 10^5.

3 The Solution Method

Under some conditions, the continuous-choice dynamic programming can be solved without directly solving optimal actions from nonlinear problems. This approach comes from the assumptions:

Assumption 1 *The action space is bounded.*

Assumption 2 *Current period's profit function $\pi(z_t, a, \epsilon_t)$ and transition probabilities $f(z_{t+1}|z_t, a)$ are both differentiable with respect to actions a, and the maximizers a^* of the dynamic optimization problem are interior solutions.*

Assumption 3 *The derivative of current profit function with respect to actions is a simple function of the action specific shock. Let $\pi_a(a_t, z_t, \epsilon_t)$ be the derivative of current period's profit function with respect to the action a_t. For any number of x that*

$$\pi_a(a_t, z_t, \epsilon_t) = x, \tag{8}$$

there is an analytical form of its inverse function:

$$\epsilon_t = \pi_a^{-1}(a_t, z_t, x) \tag{9}$$

The third assumption can be satisfied if, for example, the derivative of current period's profit function is a linear function of shocks. Such functional form assumption is applicable to a variety of problems as the examples shown in section 4.

With the above assumption, we can solve an analytical form of the inverse policy function from actions to independent shocks conditional on states and next period's value functions from first order conditions. The first order condition at optimal action a_t^* and state (z_t, ϵ_t) given next period's value function $V(z_{t+1})$ is

$$\pi_a(a_t^*, z_t, \epsilon_t) + \beta \sum_{z_{t+1}} V(z_{t+1}) f_a(z_{t+1}|z_t, a_t^*) = 0 \tag{10}$$

where $f_a(a_{t+1}|a_t, z_t)$ is the derivative of transition probabilities with respect to the action a_t. The inverse policy function from actions to shocks conditional on states can be written as

$$\epsilon(a_t^*, z_t, V_{t+1}) = \pi_a^{-1}\left(a_t^*, z_t, -\beta \sum_{z_{t+1}} V_{t+1} f_a(z_{t+1}|z_t, a_t^*)\right) \tag{11}$$

where $V_{t+1} = V(z_{t+1})$ is a vector of value functions at next period's states.

Remember that shocks distributes with the function of $G_\epsilon(\epsilon_t)$. The distribution of optimal actions can be derived through a change of variables:

$$g_a^*(a_t^*; z_t, V_{t+1}) = g_\epsilon\left(\epsilon(a_t^*, z_t, V_{t+1})\right) \left| det \left\{ \frac{\partial \epsilon}{\partial a}(a_t^*, z_t, V_{t+1}) \right\} \right|. \tag{12}$$

With the distribution of optimal actions, the ex-ante value function can be integrated over the mass of optimal actions rather than shocks. Through a change of variable, the ex-ante value function in equation 5 can be written as

$$V(z_t) = \int_{a^*} \left\{ \pi(z_t, \epsilon(a^*, z_t, V_{t+1}), a^*) + \beta \sum_{z_{t+1}} V(z_{t+1}) f(z_{t+1}|z_t, a_t^*) \right\} g_a^*(a_t^*; z_t, V_{t+1}) \; da_t^*, \qquad (13)$$

where $g_a^*(a_t^*; z_t, V_{t+1})$ is derived in 12.

With some points $\{a^i\}_{i=1}^n$ from the action space, current period's value function given next period's ones can be calculated by taking the weighted average of values at these points. The weight is the probability that each point a^i is the optimal action satisfying the first order condition. I will illustrate the choice of the action points at the end of this section. With the action points $\{a^i\}_{i=1}^n$, the mapping from next period's value function to current period's value function in equation 5 can be written as

$$V(z_t) = \sum_{i=1}^{n} \left\{ \pi(a^i, z_t, \epsilon(a^i, z_t, V_{t+1})) + \beta \sum_{z_{t+1}} V(z_{t+1}) f(z_{t+1}|z_t, a^i) \right\} w(a^i; z_t, V_{t+1}), \qquad (14)$$

where

$$w(a^i; z_t, V_{t+1}) = \frac{g_a^* \left(\epsilon(a^i, z_t, V_{t+1}) \right)}{\sum_{r=1}^{n} g_a^* \left(\epsilon(a^r, z_t, V_{t+1}) \right)}, \qquad (15)$$

and $g_a^*(a_t^* = a^i; z_t, V_{t+1})$ is in 12.[2]

3.1 The Procedure

The above method is applicable to both finite-horizon and infinite-horizon problems with little differences.

Finite Horizon

With a finite horizon problem, the ex-ante value functions are solved backwards. At the final period T, the maximization problem becomes static. It is easy to solve the inverse policy function $\epsilon(a_T, z_T)$ exempt from the next period's value function. Then put the inverse policy function into

[2]The probabilities of optimal actions contains the Jacobian matrix. This can be approximated by the one-side difference. The procedure is to shift the action a with a small number $a + h$, calculate the action-specific shock at these two points, $\epsilon(a^i + h, z)$ and $\epsilon(a^i, z)$, and approximate the derivative with the change rate, $\partial\epsilon/\partial a = (\epsilon(a^i + h, z) - \epsilon(a^i, z))h$. If the action space is J-dimensional, the derivative is approximated by $\partial\epsilon/\partial a_j = (\epsilon((a_1^i, \ldots, a_j^i + h, \ldots, a_J^i; z) - \epsilon((a_1^i, \ldots, a_j^i, \ldots, a_J^i; z))/h$ for $\forall j = 1, \ldots, J$.

the distribution of shocks, and calculate the final period's value function through

$$V(z_T) \quad = \quad \sum_{i=1}^{n} \left\{ \pi(a^i, z_T, \epsilon(a^i, z_T)) \right\} w_i(a^i; z_T). \tag{16}$$

Then goes to one period ahead. With the final period's value function, we can solve the the inverse policy function $\epsilon(a_t, z_t, V_T)$. Similarly, derive the distribution of optimal actions from the inverse policy function, and then calculate the ex-ante value function and policy function as equation 14 described. The process goes through

1 Generate n points $\{a^i\}_{i=1}^n$ from the bounded action space.

2 Start from the final period $t = T$, where $V_{t+1} = V(z_{T+1})$ has terminal values.

   2.1 calculate $\epsilon(a_t^* = a^i, z_t, V_{t+1})$ in 11 for all $\{a^i\}_{i=1}^n$ and z_t;

   2.2 calculate $g_a^*(a^i, z_t, V_{t+1})$ in 12 for all $\{a^i\}_{i=1}^n$ and z_t;

   2.3 calculate $V(z_t)$ in 14 for all z_t.

3 Stop if $t = 1$. Otherwise, go to $t' = t-1$, and update $V_{t'+1} = V(z_t)$, and continue step 2.1-2.3.

Infinite Horizon

For an infinite-horizon problem, the ex-ante value function can be solved iteratively from the Bellman equation. The procedure goes through

1 Generate n points $\{a^i\}_{i=1}^n$ from the bounded action space, and guess an initial value function $V^0(z)$.

2 Given $\{a^i\}_{i=1}^n$ and $V^r(z) = V^0(z)$,

   2.1 calculate $\epsilon(a^i; z, V^r(z))$ in 11 for all a^i and state z;

   2.2 calculate $g_a^*(a_t^* = a^i, z, V^r(z))$ in 12 for all a^i and state z;

   2.3 solve $V^{r+1}(z)$ in 14 for all z.

3 Check convergence: $||V^{r+1}(z) - V^r(z)|| < \tau$, where τ is a small number. If not, update $V^r(z) = V^{r+1}(z)$ and continue step 2.1-2.3.

3.2 The Choice of Action Space Points

The advantage of my method is to avoid using the grid points with a small step. The action points $\{a^i\}_{i=1}^n$ can be quasi random Monte-Carlo points from the action data. Thus, for multidimensional actions, the number of action space points will not grow exponentially as the discretization method. My method actually borrows the idea from the quasi Monte Carlo integration. Thus, this requires a bounded action space. The boundedness can be satisfied in various cases. Given model parameters, the boundary of the action space is usually known directly from the model. For example, in dynamic pricing, optimal prices needs to be higher than zero and lower For example, in a dynamic investment problem, data contains investment measures over time. Then, the quasi Monte Carlo action points can be chosen from a larger box than the range of investment measures in the data.

4 Examples

4.1 Example 1: Dynamic Pricing with Limited Inventory

A retailer has z_t inventory for a product at period t. She want to set prices to maximize the expected discounted profits over a finite or infinite horizon. This problem is prevalent in airline, theater, hotels and retailing industries. For example, airline have limited number of seats and want to sell out its tickets before the departure date. A theatre has limited number of seats and want to sell out the tickets before the concert. Retailers want to sell out the limited inventory of seasonal products before the next season. Also, hotels have limited number of rooms, and want to set prices contingent on the number of rooms left.

The demand for the product at the price p_t is $Q(p_t)$, where demand decreases with prices $\partial Q(p_t)/\partial p_t < 0$ for $\forall z_t$. Suppose the retailer's current profit has a constant marginal cost, and the marginal cost function contains an action-specific shock.

$$\pi(z_t, p_t, \epsilon_t) = p_t \min\{Q(p_t), z_t\} - c(\epsilon_t) \min\{Q(p_t), z_t\} \tag{17}$$

where $c(\epsilon_t)$ is the marginal cost at the action-specific shock ϵ_t. Based on the distribution function of the action-specific shock $G_\epsilon(\epsilon)$, the distribution of marginal cost is known. Let $G_c(c)$ be the cumulative distribution, and $g_c(c)$ is the probability of the marginal cost is c.

The cumulative probability that the next period's inventory is less than $z \geq 0$ conditional on current period's inventory z_t and price p_t is

$$\begin{aligned} F(z_{t+1} = z|z_t, p_t) &= Pr(z_{t+1} = \max\{z_t - Q(p_t), 0\} \leq z|z_t, p_t) \quad (18) \\ &= 1 - \exp(-\lambda(z_t - Q(p_t) - z)) \quad (19) \end{aligned}$$

This functional form of the state transition means that $\max\{z_t - Q(p_t), 0\} - z_{t+1}$ is a random variable and is distributed as an exponential function with the parameter of λ. This assumption tells that the retailer's inventory depreciate over time.

After observing the inventory z_t and shocks ϵ_t, the retailer want to sequentially choose prices p_t to maximize the expected discounted profit over an infinite time horizon. Omit the time notation t, the retailer's problem becomes

$$V(z) = \int_c \max_{p \in \{\tilde{p}|Q(\tilde{p}) \leq z\}} (p-c)Q(p) + \beta \sum_{z'} V(z') f(z'|z,p) \;\; dG_c(c). \quad (20)$$

The model maximize the expected discounted profits under the constraint $Q(p) \leq z$. To solve the optimal prices, let us consider the case with no constraint first. For the price $p^0 = p^0(z, c, V')$ satisfying the first order condition without the constraint, the corresponding marginal cost is

$$c(p^0, z, V') = \left(\frac{\partial Q(p^0)}{\partial p}\right)^{-1} \left(Q(p^0) + \frac{\partial Q(p^0)}{\partial p} p^0 + \beta \sum_{z'} V(z') \frac{\partial f(z'|z,p^0)}{\partial p}\right) \quad (21)$$

The distribution of the optimal price without the condition $Q(p) \leq z$ can be derived from the distribution of costs as equation 12 describes. Let $f_p^0(p; z, V')$ be the distribution of optimal prices without the constraint $Q(p) \leq z$. From $c(p^0, z, V')$ in equation 21, $f_p^0(p, z, V')$ can be computed through a change of variables as 12 shows.

Let $p^* = p(z, c, V')$ be the optimal price at the state z and the marginal cost shock c. If $Q(p^0) \leq z$, the optimal price is to charge $p^* = p^0$. Otherwise, the optimal price is to charge the price at which the quantities sold equal to the inventory, $Q(p^*) = z$. Also, remember that $Q(p)$ decreases with p. The optimal price is thus $p^* = \max\{p^0, Q^{-1}(z)\}$. The distribution of optimal prices in this example is therefore

$$f_p^*(p, z, V') = \frac{f_p^0(p, z, V')}{1 - F_p^0(Q^{-1}(z), z, V')} \tag{22}$$

for $p \geq Q^{-1}(z)$, where $F_p^0(p, z, V')$ is the cumulative distribution of $f_p^0(p, z, V')$.

4.2 Example 2: Plant Production

This example is borrowed from Arcidiacono and Miller (2010). A plant has the condition of $z_t \in \{1, \ldots, Z\}$, where higher levels of z_t indicate that the plant is in worse condition. The plant has the current profit from state z_t as $\pi(z_t)$. At each period t, a plant manager chooses an input $c_t \in (0, \infty)$. The cost of choosing the factor c_t is a quadrature function in the logarithm of c_t:

$$C(c_t, z_t, \epsilon_t) = (\epsilon_t + \alpha_1 z_t)\log(c_t) + \alpha_2 z_t(\log(c_t))^2, \tag{23}$$

where ϵ_t is an independent and identically distributed action-specific shock, and (α_1, α_2) are model parameters. The density function and the cumulative distribution function of ϵ_t is $g_\epsilon(\epsilon_t)$ and $G_\epsilon(\epsilon_t)$ specifically. Also, increasing inputs c_t raises the probability that the machinery is in bad condition next period $t + 1$,

$$f(z_{t+1}|z_t, c_t) = \begin{cases} \gamma_0/(\gamma_0 + c_t^{\gamma_1}) & ,\text{if } \; z_{t+1} = z_t + 1 \\ 1 - \gamma_0/(\gamma_0 + c_t^{\gamma_1}) & ,\text{if } \; z_{t+1} = z_t \\ 0 & ,\text{otherwise} \end{cases} \tag{24}$$

where $\gamma_0, \gamma_1 > 0$ are parameters.

After observing the state z_t and an action-specific shock ϵ_t, the plant manager sequentially decides an input level to maximize the discounted expected profit.

$$V(z_t) = \int_\epsilon \max_{c_t} \pi(z_t) - C(c_t, z_t, \epsilon_t) + \beta \sum_{z_{t+1}} V(z_{t+1}) f(z_{t+1}|z_t, c_t) \;\; dG_\epsilon(\epsilon_t) \tag{25}$$

Let $c_t^* = c(z_t, \epsilon_t, V_{t+1})$ be the optimal input at state z_t and shock ϵ_t. From the first order condition, the corresponding shock is

$$\epsilon(c_t^*, z_t, V_{t+1}) = -(\alpha_1 z_t + 2\alpha_2 z_t \log(c_t^*)) + \beta c_t^* \{V(z_t) - V(z_t+1)\} \gamma_0\gamma_1 (c_t^*)^{\gamma_1 - 1}/(\gamma_0 + (c_t^*)^{\gamma_1})^2 \quad (26)$$

The distribution of optimal inputs is thus computed as equation 12.

5 Estimation

Model parameters can be estimated via simulated method of moments, where the moments are the probabilities of observing a action at a state in the data. This requires that the actions need to be observed in the data. However, the model can contain unobserved states, which will be explained at the end of this section.

Finite Horizon

Suppose the data contains the action and state observations $\{\{a_{it}, z_{it}\}_{i=1}^m\}_{t=1}^T$, where m is the total number of observations, T is the total number of periods. For any two numbers of a and z, the frequency of observing actions less than a at states less than z at period t is

$$m(a, z, t) = \frac{1}{m}\sum_{i=1}^{m} 1\{a_{it} \le a\} 1\{z_{it} \le z\} \quad (27)$$

Using the distribution of optimal actions described in section 3, the moment can be calculated from the model. The cumulative probability of states can be calculated by aggregating the probability of paths arriving to state $z_t = z$ at period t. The unconditional distribution of state transition can be calculated as

$$f(z_t = z | z_{t-1}; \theta) = \sum_{i=1}^{n} f(z_t = z | z_{t-1}, a_{t-1} = a^i; \theta) g_a^*(a_{t-1} = a^i; z_{t-1}, \theta) \quad (28)$$

where $\{a^i\}_{i=1}^n$ are points from the action space as section 3 described, $F(z_t | z_{t-1}, a_{t-1})$ is known by the model assumption, and $g_a^*(a_{t-1} = a^i; z_{t-1})$ is calculated in section 3. Suppose all states start from an initial state z_1, i.e. $f(z_1)$ is known by the model assumption. The probability of observing the state $z_t = z$ at period t is

$$f(z_t = z; \theta) = \sum_{z_1} \sum_{z_2} \dots \sum_{z_{t-1}} f(z_t = z | z_{t-1}; \theta) f(z_{t-1} | z_{t-2}; \theta) \dots f(z_2 | z_1; \theta) f(z_1). \tag{29}$$

With the unconditional distribution of states, the probability of observing actions less than a at period t is

$$\widetilde{m}(\theta; a, z, t) = \sum_{z_t} G_a^*(a_t^* = a; z_t, \theta) f(z_t; \theta) 1\{z_t \leq z\} \tag{30}$$

Calculate the above moments in equation 27 and 30 for a vector of numbers $\{a, z\}$ at each period t. Parameters are thus estimated through matching the model-generated moments with the moment in the data:

$$\min_{\theta} \left(\frac{1}{T} \sum_{t=1}^{T} \big(m(a, z, t) - \widetilde{m}(\theta; a, z, t) \big) \right)' W \left(\frac{1}{T} \sum_{t=1}^{T} \big(m(a, z, t) - \widetilde{m}(\theta; a, z, t) \big) \right) \tag{31}$$

where W is the weighting matrix.

Infinite Horizon

Suppose the data contains the action and state observations $\{a_i, z_i\}_{i=1}^{m}$, where m is the total number of observations. The frequency of observing actions less than a at states less than z is

$$m(a, z) = \sum_{i=1}^{m} 1\{a_i \leq a\} 1\{z_i \leq z\} \tag{32}$$

Using the solution method described in section 3, the unconditional state transition can be calculated through aggregating over the distribution of optimal actions:

$$f(z'|z) = \sum_{i=1}^{n} f(z'|z, a = a^i) g_a^*(a = a^i; z) \tag{33}$$

where $\{a^i\}_{i=1}^{n}$ are points from the action space as section 3 described, $F(z'|z, a)$ is known from the model assumption, and $g_a^*(a = a^i; z)$ is calculated in section 3. Let $f^0(z)$ be the stationary distribution of states. From the transition probabilities, the stationary distribution of states can be calculated from

$$f^0(z) = f(z'|z)f^0(z) \tag{34}$$

In practice, the stationary distribution of states can be calculated by multiply the unconditional state transition with an initial guess for many times.[3] With the stationary distribution of states and the distribution of optimal actions conditional on states, the model-generated moment is

$$\widetilde{m}(\theta; a, z) = \sum_z G_a^*(a; z)f^0(z) \tag{35}$$

The model estimation is through simulated method of moments, where the objective function is similar to equation 31.

6 Application

The above computational method can approximate smoother value/policy function than the method of discretizing the action space. Also, the method proposed in this paper is able to solve a multi-dimensional action. To demonstrate the advantage, I consider an numerical example of a multi-product firm's dynamic pricing problem as example 1 in section 4.1.

Suppose the retailer has J number of products. Each product has z_j number of inventory. Suppose consumer i's utility of purchasing product $j \in \widetilde{J}$ is

$$u_{ij} = \delta_j - \alpha p_j + \epsilon_{ij} \tag{36}$$

where δ_j is the mean utility of product j, α is the price coefficient, and ϵ_{ij} is the idiosyncratic shock. Also, normalize consumer i's utility from choosing the outside option as zero: $u_{i0} = \epsilon_{i0}$. With the assumption that idiosyncratic shocks are independent and follow the type I extreme value

[3] For example, for an arbitrary guess of the distribution f^r, iterate the calculation $f^{r+1} = f(z'|z)f^r$ for $r = 1, \ldots, R$ times until the distribution converges $|f^{r+1} - f^r| < \tau$, where τ is a small number.

distribution, product j's market share as a function of price p and inventory z is

$$Q_j(p,z) = \begin{cases} \frac{N\exp(\delta_j - \alpha p_j)}{1+\sum_{k\in\tilde{J}}\exp(\delta_k-\alpha p_k)}, & \text{if } \; z_j > 0 \\ 0 & \text{otherwise} \end{cases} \tag{37}$$

where N is the total number of potential consumers.

Discretize the state space $z_j \in \{s_j^1, \ldots, s_j^k, \ldots, s_j^K\}$, where $s^k \leq s^{k+1}$ for $\forall k = 1, \ldots, K-1$. Let $f_{mk}(j)$ be the probability that the inventory of product j transit from $z_{jt} = s^m$, to the inventory z_{jt+1} where $s_j^k < z_{jt+1} \leq s_j^{k+1}$, is

$$f_{mk}(j) = \begin{cases} H(z_j - Q_j(p) - s_j^{k-1}) - H(z_j - Q_j(p) - s_j^k) & \text{, if } s^k > 0 \\ 1 - F_\epsilon(z_j - Q_j(p) - s_j^k) & \text{, otherwise} \end{cases}, \tag{38}$$

where $H()$ is the cumulative distribution function of exponential distribution. The probability that the vector of inventory $z_t = s^m = (s_1^m, \ldots, s_J^m)$ to the inventory $z_{t+1} = s^k = (s_1^k, \ldots, s_J^k)$ is

$$f_{mk} = \Pi_j f_{mk}(j) \tag{39}$$

With the above functional form assumptions, I can solve costs corresponding optimal prices conditional on states and next period's value function from the first order condition as equation 21. Since the retailer has multiple products, the derivative of demand with respect to prices is a J-by-J matrix, and the derivative of state transition with respect to prices is a J-by-1 vector. From the demand function in equation 37, the element at the jth row and lth column of the market share derivative $\nabla Q(s,p)$ is

$$(\nabla Q(s,p))_{jl} = \frac{\partial Q_l(p)}{\partial p_j} = \begin{cases} -\alpha sh_j(p)(1-sh_j(p))N & \text{if } j = l \\ \alpha sh_j(p)sh_l(p)N & \text{if } j \neq l \end{cases} \tag{40}$$

From the inventory transition in equation 39, the jth element of the state transition derivatives is

$$(\nabla f_{mk})_j = \frac{\partial f_{mk}}{\partial p_j} = \sum_l \frac{f_{mk}}{f_{mk}(l)} \frac{\partial f_{mk}(l)}{\partial p_j} \tag{41}$$

The numerical example used in this paper has parameters: $\alpha = -0.08, \beta = 0.99$. The marginal cost of product j has the function $c_j = \bar{c}_j + \epsilon_j$, where ϵ_j is independent and identically distributed as $N(0, 9)$. Thus, the distribution of marginal cost c is $G_c = N(\bar{c}, \Sigma)$, where $\Sigma_{ij} = 9I$, and I is an identity matrix. The action space is $[20, 40]^J$, where J is the total number of available products. In the method of discretizing the action space, I draw 5-25 number of quasi-newton points from the distribution of cost shocks $G_c(c)$ depending on the number of products.

Table 1 shows the performance for these two solution methods, where Panel A, B, C display the result for a retailer with one, two and three products respectively. In the case of single-product retailer, I use 1001 state points $z_t \in \{0, 1, \ldots, 1000\}$. To discretize the action space, I use 1000 grid points from the action space with the step of 0.05, and 5 quasi-newton points from the distribution function of marginal costs. To use the new method, I use 20 price points $\{a^i\}_{i=1}^{20}$, where the price points are evenly distributed along the action space $[20, 40]$. The result shows that my method saves 37 seconds (48%) of the computational time for one solution. The value/policy function computed with my method is within two digits, i.e. -4.7×10^{-3} and -1.5×10^{-3}, relative to the discretization method.

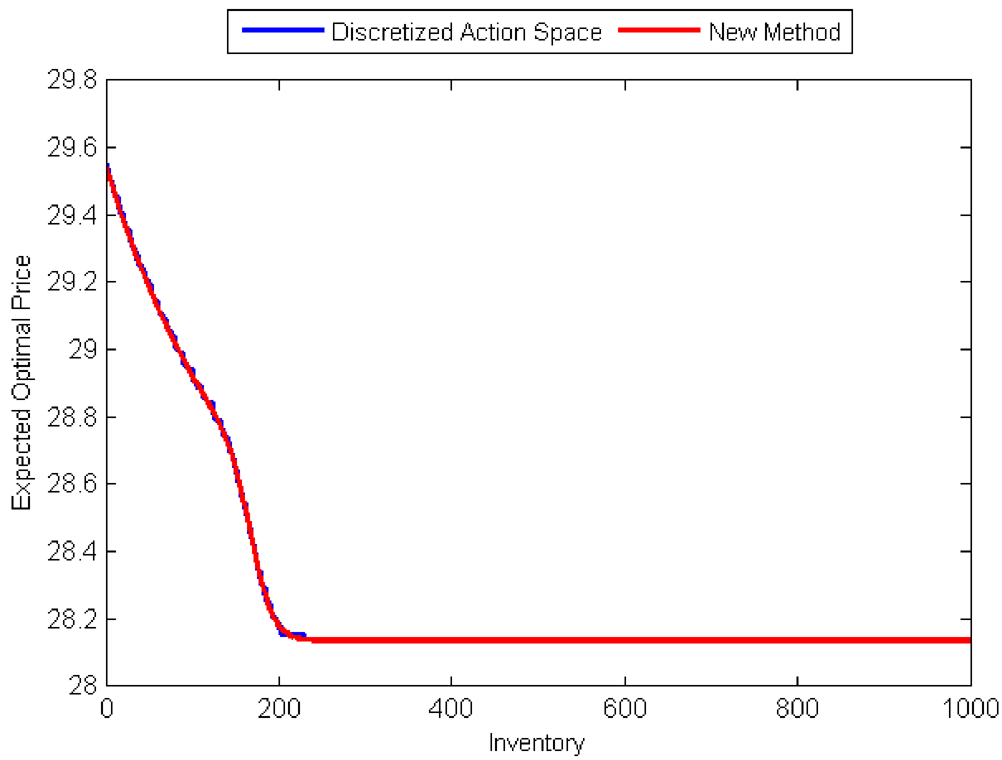
Figure 1 displays the difference of policy functions under these two solution methods at the parameters $\lambda = 0.5$ and $\lambda = 0.1$ respectively. With the parameter of $\lambda = 0.5$, these two methods compute similar values at different inventory levels. However, when the parameter is $\lambda = 0.1$, my method computes a smooth policy function, but the discretization method computes a step function. This is because, using discrete action space, the optimal action may allocate within the intervals between grid points, but the computed optimal action is chosen from the grid points. In the two-products case, my method still saves computational time to solve the model, and the ex-ante policy function shown in figure 2 is smoother in my method relative to the other method.

Table 1: Performance of two solution methods

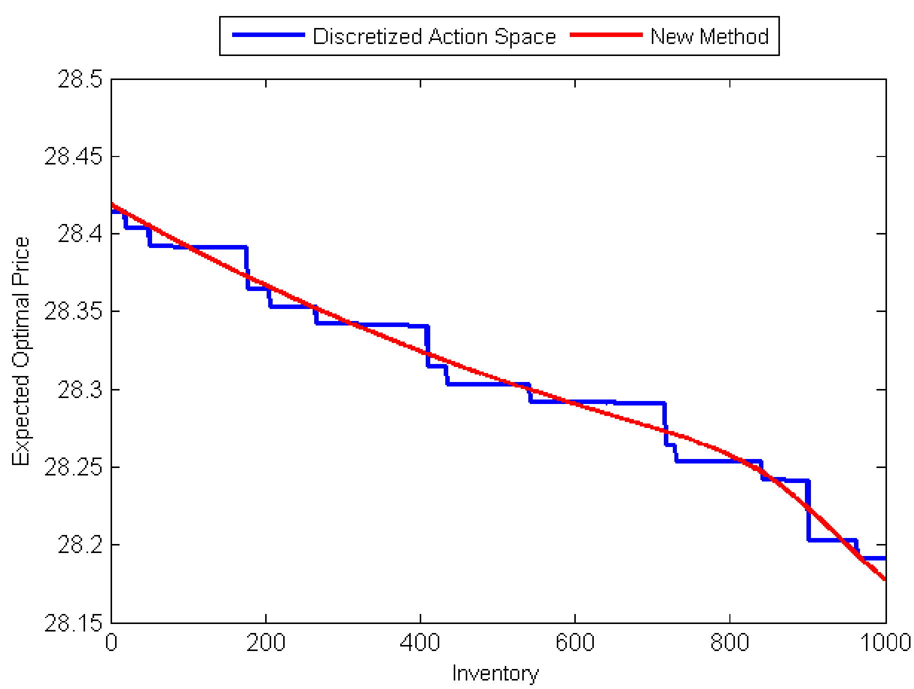
	Discretization	New Method	Differences
Panel A: $J = 1$ ($\lambda = 0.5$)			
$\#s$	1001	1001	-
$\#a$ (step size)	401 (0.05)	20	-
$\#\epsilon$	5	-	-
Computational Time	76s	39s	-37s (-48%)
Accuracy			
$\overline{V(s)}$	214.1455	214.1407	-4.7$\times 10^{-3}$
$\overline{p(s)}$	28.2649	28.2634	-1.5$\times 10^{-3}$
Panel B: $J = 2$ ($\lambda = 3$)			
$\#s$	961	961	-
$\#a$ (step size)	441 (1)	75	-
$\#\epsilon$	25	-	-
Computational Time	164s	82s	-81s (-49%)
Accuracy			
$\overline{V(s)}$	336.8751	336.5058	-0.3693
$\overline{p_j(s)}$	30.3181	30.3338	15.7$\times 10^{-3}$
Panel C: $J = 3$ ($\lambda = 0.5$)			
$\#s$	961	961	-
$\#a$ (step size)	1331 (2)	256	-
$\#\epsilon$	9	-	-
Computational Time	609s	174s	-435s (-71%)
Accuracy			
$\overline{V(s)}$	30.9359	30.5946	-0.3413
$\overline{p_j(s)}$	29.6069	27.9979	-1.6091

The action space is $[20, 40]^J$. The starting point of value function iterations is $(0, \ldots, 0)$. The convergence criteria for the ex-ante value function is $\sum_s (V^r(s) - V^{r-1}(s))^2 / \sum_s (V^{r-1}(s))^2 < 10^{-6}$.

Figure 1: Policy Function for $J = 1$

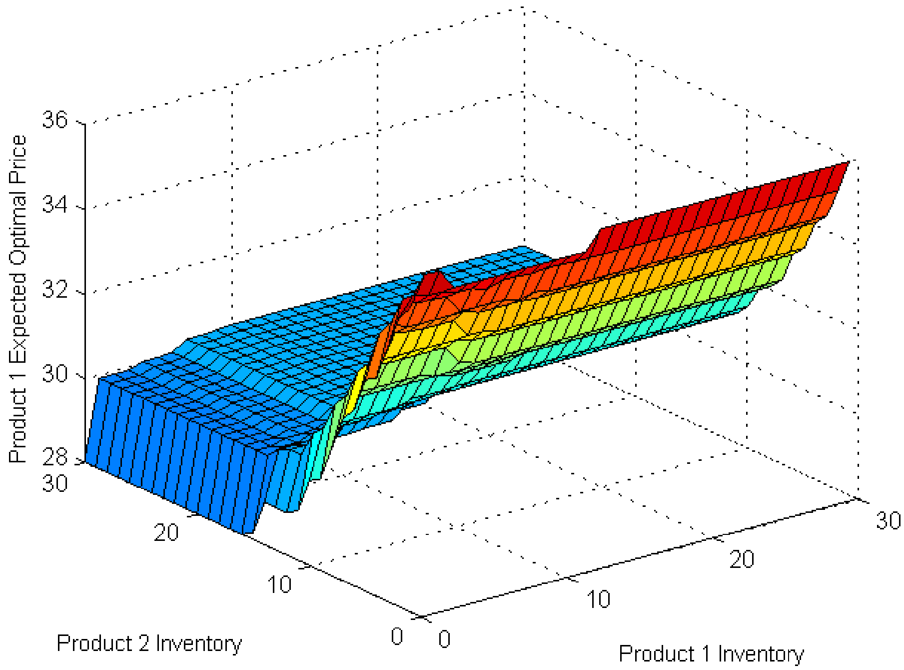


(a) $\lambda = 0.5$

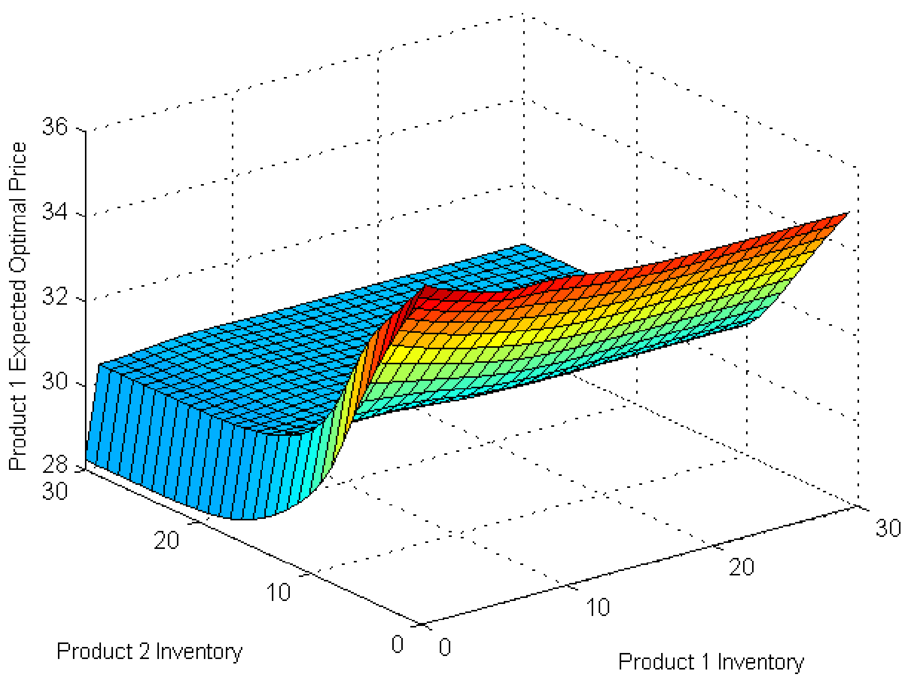


(b) $\lambda = 0.1$

Figure 2: Policy Function for $J = 2$



(a) Discretization



(b) New Method

References

Adda, J. and R. W. Cooper (January 2003). *Dynamic Economics: Quantitative Methods and Applications* (1st Edition ed.). MIT Press.

Blundell, R. and T. Macurdy (1999). *Handbook of Labor Economics*, Volume 3.

Gallego, G. and G. van Ryzin (1994). Optimal dynamic pricing of inventories with stochastic demand over finite horizons. *Management Science 40*(8), 999–1020.

Judd, K. (1999). *Numerical methods in economics* (1st Edition ed.). MIT Press.

Rust, J. (September 1987). Optimal replacement of gmc bus engines: An empirical model of harold zurcher. *Econometrica 55*(5), 999–1033.