1 DYNAMIC DISCRETE CHOICE MODELS IN THEIR ORIGINAL FORMULATION

In this section, we formulate dynamic discrete choice models (DDMs) using the original formulation \cite{Rust1987}, and discuss its connection with the IRL formulation in Section 2.1. Note that the setup in this section is an alternative to the IRL formulation which our main results are based on and just is provided for completeness and comparison. SamQ does not require the assumptions listed in this section.

1.1 MODEL

Agents choose actions according to a Markov decision process described by the tuple \( \{S, E, A, r, \gamma, P\} \), where

- \( \{S, E\} \) denotes the space of state variables;
- \( A \) represents a set of \( n_a \) actions;
- \( r \) represents an agent utility function;
- \( \gamma \in [0, 1) \) is a discount factor;
- \( P \) represents the transition distribution.

At time \( t \), agents observe state \( S_t \) taking values in \( S \), and \( \epsilon_t \) taking values in \( E \) to make decisions. While \( S_t \) is observable to researchers, \( \epsilon_t \) is observable to agents but not to researchers. The action is defined as a \( n_a \times 1 \) indicator vector, \( A_t \), satisfying

- \( \sum_{j=1}^{n_a} A_{tj} = 1 \),
- \( A_{tj} \) takes value in \( \{0, 1\} \).

In other words, at each time point, agents make a distinct choice over \( n_a \) possible actions. Meanwhile, \( \epsilon_t \) is also a \( n_a \times 1 \) representing the potential shock of taking a choice.

The agent’s control problem has the following value function:

\[
V(s, \epsilon) = \max_{\{a_t\}_{t=0}^\infty} E \left[ \sum_{t=0}^{\infty} \gamma^t r(S_t, \epsilon_t, A_t) \mid s, \epsilon \right],
\]

where the expectation is taken over realizations of \( \epsilon_t \), as well as transitions of \( S_t \) and \( \epsilon_t \) as dictated by \( P \). The utility function \( r(s_t, \epsilon_t, a_t) \) can be further decomposed into

\[
r(s_t, \epsilon_t, a_t) = u(s_t, a_t) + a_t^T \epsilon_t,
\]

where \( u \) represents the deterministic part of the utility function. Agents, but not researchers, observe \( \epsilon_t \) before making a choice in each time period.
1.2 ASSUMPTIONS AND DEFINITIONS

We study DDMs under the following common assumptions.

**Assumption 1.** The transition from \( S_t \) to \( S_{t+1} \) is independent of \( \epsilon_t \)

\[
P(S_{t+1} \mid S_t, \epsilon_t, A_t) = P(S_{t+1} \mid S_t, A_t).
\]

**Assumption 2.** The random shocks \( \epsilon_t \) at each time point are independent and identically distributed (IID) according to a type-I extreme value distribution.

Assumption 1 ensures that unobservable state variables do not influence state transitions. This assumption is common, since it drastically simplifies the task of identifying the impact of changes in observable versus unobservable state variables. In our setting, Assumption 2 is convenient but not necessary, and \( \epsilon_t \) could follow other parametric distributions. As pointed out by Arcidiacono and Ellickson [2011], Assumptions 1 and 2 are nearly standard for applications of dynamic discrete choice models. Such a formulation is proved to be equivalent to the IRL formulation in Section 2.1 by Geng et al. [2020], Fu et al. [2018], Ermon et al. [2015].

2 PROOF OF THEOREM 1

**Proof.** By definition of \( L \) and \( \tilde{L} \), we can derive

\[
L(D; \theta^*) = \sum_{(s,a) \in D} \left[ Q^\theta^*(s,a) - \tilde{Q}^\theta^*(\Pi(s),a) \right]
+ \log \left( \sum_{a' \in A} \exp(\tilde{Q}^\theta^*(\Pi(s),a')) \right) - \log \left( \sum_{a' \in A} \exp(Q^\theta^*(s,a')) \right)
\]

\[
\leq \sum_{(s,a) \in D} \left[ Q^\theta^*(s,a) - \tilde{Q}^\theta^*(\Pi(s),a) \right] + \max_{a' \in A} \left| Q^\theta^*(s,a') - \tilde{Q}^\theta^*(\Pi(s),a') \right|
\]

where the first inequality is due to the fact that the log sum exp function is Lipschitz continuous with constant 1. Then, we take \( f \) in Lemma 2 as \( Q^\theta^*(s,a) \), and derive

\[
\max_{(s,a) \in S \times A} \left| Q^\theta^*(s,a) - \tilde{Q}^\theta^*(\Pi(s),a) \right| \leq \frac{2}{1 - \gamma} \max_{(s,a) \in S \times A} \left| Q^\theta^*(s,a) - Q^\theta^*(\Pi(s),a) \right|. \quad (3)
\]

By taking (3) to (2),

\[
L(D; \theta^*) - \tilde{L}(D; \theta^*) \leq \frac{4}{1 - \gamma} \max_{(s,a) \in S \times A} \left| Q^\theta^*(s,a) - Q^\theta^*(\Pi(s),a) \right|.
\]

Finally, by Lemma 1

\[
\epsilon_{asy} \leq \frac{4}{c_H (1 - \gamma)} \max_{(s,a) \in S \times A} \left| Q^\theta^*(s,a) - Q^\theta^*(\Pi(s),a) \right| = \epsilon_Q,
\]

which finishes the proof.

**Lemma 1.** Under Assumption 1 and Assumption 2

\[
\left\| \tilde{\theta} - \theta^* \right\|^2 \leq \frac{E[L(D; \theta^*) - \tilde{L}(D; \theta^*)]}{c_H}.
\]

*Proof.** By the definition of \( \tilde{\theta} \),

\[
0 \leq E[\tilde{L}(D; \tilde{\theta}) - \tilde{L}(D; \theta^*)] \leq E[L(D; \theta^*) - \tilde{L}(D; \theta^*)]. \quad (4)
\]
Further, by Taylor expansion, we have
\[
\mathbb{E}[\hat{L}(\mathbb{D}; \hat{\theta}) - \hat{L}(\mathbb{D}; \theta^*)] = (\hat{\theta} - \theta^*)^T \mathbb{E} \left[ -\frac{\partial^2 \hat{L}(\mathbb{D}; \theta)}{\partial \theta^2} \right] (\hat{\theta} - \theta^*),
\]
where \( \hat{\theta} = k\theta^* + (1-k)\tilde{\theta} \) with some \( k \in [0, 1] \). Note that the first order term is zero, since \( \tilde{\theta} \) maximizes \( \mathbb{E}[\hat{L}(\mathbb{D}, \theta)] \). By Assumption \( \Pi \) we finish the proof.

\[
\mathbb{E}[\hat{L}(\mathbb{D}; \hat{\theta}) - \hat{L}(\mathbb{D}; \theta^*)] = (\hat{\theta} - \theta^*)^T \mathbb{E} \left[ -\frac{\partial^2 \hat{L}(\mathbb{D}; \tilde{\theta})}{\partial \theta^2} \right] (\hat{\theta} - \theta^*) \geq C_H \|\hat{\theta} - \theta^*\|^2.
\]

Lemma 2. For any projection function \( \Pi \) defined in Section 3.1 and its aggregated \( \hat{Q} \) function \( \hat{Q} \), the following inequality is true:
\[
\max_{(s,a) \in S \times A} |Q^\theta(s, a) - \hat{Q}^\theta(\Pi(s), a)| \leq \frac{2}{1 - \gamma} \min_f \max_{(s,a) \in S \times A} |Q^\theta(s, a) - f(\Pi(s), a)|,
\]
where \( f(s, a) : S \times A \rightarrow \mathbb{R} \) is any function.

Proof. The proof follows Theorem 3 of [Tsitsiklis and Van Roy 1996].

3 PROOF OF THEOREM

3.1 TECHNICAL LEMMAS FOR THEOREM

Lemma 3. Given \( \theta \in \Theta \), for any \( \delta \in (0, 1) \), we provide the following probabilistic bound for the estimated aggregated likelihood \( \hat{L} \)
\[
P \left( \left| \hat{L}(\mathbb{D}; \theta) - \mathbb{E}[\hat{L}(\mathbb{D}; \theta)] \right| \leq \frac{2(R_{\max} + 1)}{1 - \gamma} \sqrt{\frac{\log(\frac{4}{\delta})}{2N}} + \frac{R_{\max} + 1}{1 - \gamma} \sqrt{\frac{\log(S|S||A|)}{2N}} (C_{uni} - C_{uni}^2) \right) \geq 1 - \delta,
\]
where the expectation is over the sample \( \mathbb{D} \).

Proof. By inserting \( \hat{L}(\mathbb{D}; \theta) \), we have
\[
\left| \hat{L}(\mathbb{D}; \theta) - \mathbb{E}[\hat{L}(\mathbb{D}; \theta)] \right| \leq \left| \hat{L}(\mathbb{D}; \theta) - \hat{L}(\mathbb{D}; \theta^*) \right| + \left| \hat{L}(\mathbb{D}; \theta^*) - \mathbb{E}[\hat{L}(\mathbb{D}; \theta)] \right|.
\]

First term on the RHS of (5) To start with, we consider \( \left| \hat{L}(\mathbb{D}; \theta) - \hat{L}(\mathbb{D}; \tilde{\theta}) \right| \). To this end, we aim to bound
\[
\max_{(s,a) \in S \times A} \left| \hat{Q}^\theta(s, a) - \hat{Q}^\theta(\tilde{\theta}, a) \right|.
\]
We insert \( \tilde{T}(\hat{Q}^\theta(s, a)) \):
\[
\hat{Q}^\theta(s, a) - \hat{Q}^\theta(\tilde{\theta}, a) = \tilde{T}(\hat{Q}^\theta(s, a)) - \tilde{T}(\hat{Q}^\theta(\tilde{\theta}, a)) + \tilde{T}(\hat{Q}^\theta(\tilde{\theta}, a)) - \tilde{T}(\hat{Q}^\theta(s, a)).
\]
Since \( \tilde{T} \) is a contraction with \( \gamma \), we further derive
\[
\left| \hat{Q}^\theta(s, a) - \hat{Q}^\theta(\tilde{\theta}, a) \right| \leq \frac{\tilde{T}(\hat{Q}^\theta(\tilde{\theta}, a)) - \tilde{T}(\hat{Q}^\theta(s, a))}{1 - \gamma}.
\]

(6)
By the definition of $\hat{T}$ and $\tilde{T}$, it can be seen that $\hat{T}(\tilde{Q}(\tilde{s}, \tilde{a}))$ is a sample average estimation to $\tilde{T}(\tilde{Q}(\tilde{s}, \tilde{a}))$. Therefore, we aim to bound the difference between the two by concentration inequalities. Specifically, by assumption 6 and Hoeffding’s inequality, we have

$$P\left(\sum_{i=1,2,\ldots,N} 1_{\{\Pi(s_i) = \tilde{s}, a_i = \tilde{a}\}} \geq NC_{uni} - \sqrt{-\frac{1}{2} N \log(\frac{\delta}{2})}\right) \geq 1 - \frac{\delta}{2}. \quad (7)$$

Further, conditional on the event $\left\{\sum_{i=1,2,\ldots,N} 1_{\{\Pi(s_i) = \tilde{s}, a_i = \tilde{a}\}} \geq NC_{uni} - \sqrt{-N \log(\frac{\delta}{2})}\right\}$, by Hoeffding’s inequality and Assumption 7, for any $(\tilde{s}, \tilde{a}) \in \tilde{S} \times \tilde{A}$

$$P\left(\left|\tilde{T}(\tilde{Q}(\tilde{s}, \tilde{a})) - \tilde{T}(\tilde{Q}(\tilde{s}, \tilde{a}))\right| \leq \frac{R_{max} + 1}{1 - \gamma} \sqrt{\frac{\log(\frac{4|\tilde{S}|\tilde{A}|}{\delta})}{2N}} \leq 1 - \frac{\delta}{C_{uni} - \sqrt{\frac{\log(\frac{2}{\delta})}{2N}}}\right) \geq 1 - \frac{\delta}{2}. \quad (8)$$

Combining (7) and (8), for a given $(\tilde{s}, \tilde{a}) \in \tilde{S} \times \tilde{A}$, for any $\delta \in (0, 1)$

$$P\left(\left|\tilde{T}(\tilde{Q}(\tilde{s}, \tilde{a})) - \tilde{T}(\tilde{Q}(\tilde{s}, \tilde{a}))\right| \leq \frac{R_{max} + 1}{1 - \gamma} \sqrt{\frac{\log(\frac{4|\tilde{S}|\tilde{A}|}{\delta})}{2N}} \leq 1 - \frac{\delta}{C_{uni} - \sqrt{\frac{\log(\frac{2}{\delta})}{2N}}}\right) \geq 1 - \frac{\delta}{2}. \quad (9)$$

Next, by union bound again, we can extend the results to any $(\tilde{s}, \tilde{a}) \in \tilde{S} \times \tilde{A}$

$$P\left(\max_{\tilde{s} \in \tilde{S}, \tilde{a} \in \tilde{A}} \left|\tilde{T}(\tilde{Q}(\tilde{s}, \tilde{a})) - \tilde{T}(\tilde{Q}(\tilde{s}, \tilde{a}))\right| \leq \frac{R_{max} + 1}{(1 - \gamma)^2} \sqrt{\frac{\log(\frac{4|\tilde{S}|\tilde{A}|}{\delta})}{2N}} \leq 1 - \frac{\delta}{C_{uni} - \sqrt{\frac{\log(\frac{2}{\delta})}{2N}}}\right) \geq 1 - \delta. \quad (9)$$

Combined with (6), we derive:

$$P\left(\max_{(\tilde{s}, \tilde{a}) \in \tilde{S} \times \tilde{A}} \left|\tilde{Q}(\tilde{s}, \tilde{a}) - \tilde{Q}(\tilde{s}, \tilde{a})\right| \leq \frac{R_{max} + 1}{(1 - \gamma)^2} \sqrt{\frac{\log(\frac{4|\tilde{S}|\tilde{A}|}{\delta})}{2N}} \leq 1 - \frac{\delta}{C_{uni} - \sqrt{\frac{\log(\frac{2}{\delta})}{2N}}}\right) \geq 1 - \delta. \quad (9)$$

By the definition of $\tilde{L}$ in (6) and (2), we have

$$P\left(\left|\tilde{L}(\tilde{D}; \theta) - \tilde{L}(\tilde{D}; \theta)\right| \leq \frac{R_{max} + 1}{(1 - \gamma)^2} \sqrt{\frac{\log(\frac{4|\tilde{S}|\tilde{A}|}{\delta})}{2N}} \leq 1 - \frac{\delta}{C_{uni} - \sqrt{\frac{\log(\frac{2}{\delta})}{2N}}}\right) \geq 1 - \delta. \quad (9)$$

Second term on the RHS of (5) Now, we consider $\left|\tilde{L}(\tilde{D}; \theta) - E[\tilde{L}(\tilde{D}; \theta)]\right|$. By (3) and Assumption 1, $\tilde{L}(\tilde{D}; \theta)$ is bounded by $\frac{2(R_{max} + 1)}{1 - \gamma}$. Thus, by Hoeffding’s inequality, for any $\delta \in (0, 1)$

$$P\left(\left|\tilde{E}[\tilde{L}(\tilde{D}; \theta)] - \tilde{L}(\tilde{D}; \theta)\right| \leq \frac{2(R_{max} + 1)}{1 - \gamma} \sqrt{\frac{\log(\frac{2}{\delta})}{2N}} \geq 1 - \delta. \quad (9)$$

Therefore, by union bound, (5) can be bounded by

$$P\left(\left|\tilde{L}(\tilde{D}; \theta) - E[\tilde{L}(\tilde{D}; \theta)]\right| \leq \frac{2(R_{max} + 1)}{1 - \gamma} \sqrt{\frac{\log(\frac{2}{\delta})}{2N}} + \frac{R_{max} + 1}{(1 - \gamma)^2} \sqrt{\frac{\log(\frac{4|\tilde{S}|\tilde{A}|}{\delta})}{2N}} \geq 1 - \delta. \quad (9)$$
Lemma 4. Let \( \hat{\Theta} := \arg\max_{\theta \in \Theta} E[\hat{L}(\Theta; \hat{\Theta})] \). Then,

\[
\|\theta^* - \hat{\Theta}\| \leq \frac{4}{C_H(1-\gamma)} \left( \frac{R_{\max} + 1}{1-\gamma} \frac{4}{n_s^{\frac{1}{\gamma}} - 1} + 2\epsilon_Q + \epsilon_c \right),
\]

Proof. A Euclidean ball of radius \( R \) in \( \mathbb{R}^{n_s} \) can be covered by \( \left( \frac{4R + \delta}{\delta} \right)^{n_s} \) balls of radius \( \delta \) (see Lemma 2.5 of Van de Geer and van de Geer [2000]). Therefore, with \( n_s \) states after aggregation, by Assumption 4

\[
\epsilon(\Pi^*) \leq R_{\max} + 1 \frac{4}{1-\gamma} \frac{1}{n_s^{\frac{1}{\gamma}} - 1}.
\]

Further by Assumption 4 and Assumption 5

\[
\epsilon(\Pi) \leq \epsilon(\Pi^*) + 2\epsilon_Q + \epsilon_c \leq \frac{R_{\max} + 1}{1-\gamma} \frac{4}{n_s^{\frac{1}{\gamma}} - 1} + 2\epsilon_Q + \epsilon_c.
\]

Therefore, by Theorem 1

\[
\|\theta^* - \hat{\Theta}\| \leq \frac{4}{C_H(1-\gamma)} \left( \frac{R_{\max} + 1}{1-\gamma} \frac{4}{n_s^{\frac{1}{\gamma}} - 1} + 2\epsilon_Q + \epsilon_c \right).
\]

\[\square\]

3.2 PROOF

We first aim to bound \( E[\tilde{L}(\Theta; \hat{\Theta}) - \tilde{L}(\Theta; \hat{\Theta})] \), where the expectation is over \( \Theta \) only instead of \( \hat{\Theta} \). To this end, we insert \( L(\Theta; \hat{\Theta}) \) and \( \tilde{L}(\Theta; \hat{\Theta}) \):

\[
E[\tilde{L}(\Theta; \hat{\Theta}) - \tilde{L}(\Theta; \hat{\Theta})] \leq E[\tilde{L}(\Theta; \hat{\Theta}) - \tilde{L}(\Theta; \hat{\Theta})] + \tilde{L}(\Theta; \hat{\Theta}) - \hat{L}(\Theta; \hat{\Theta}) + \hat{L}(\Theta; \hat{\Theta}) - \tilde{L}(\Theta; \hat{\Theta}) - E[\tilde{L}(\Theta; \hat{\Theta})] - E[\tilde{L}(\Theta; \hat{\Theta})].
\]

By Lemma 3 and the union bound,

\[
P\left( \max_{\theta \in \Theta} |\tilde{L}(\Theta; \theta) - \tilde{L}(\Theta; \hat{\Theta})| \right) \leq \frac{2(R_{\max} + 1)}{1-\gamma} \sqrt{\frac{\log(\frac{4|\Theta|}{\alpha})}{2N}} + \frac{R_{\max} + 1}{(1-\gamma)^2} \sqrt{\frac{\log(\frac{8|\Theta||A|}{\delta})}{2N}} \left( C_{\text{uni}} - \sqrt{\frac{\log(\frac{4|\Theta||A|}{\delta})}{2N}} \right) \geq 1 - \delta.
\]

Therefore,

\[
P\left( E[\tilde{L}(\Theta; \hat{\Theta})] - \tilde{L}(\Theta; \hat{\Theta}) \right) \leq \frac{4R_{\max} + 1}{1-\gamma} \sqrt{\frac{\log(\frac{4|\Theta|}{\delta})}{2N}} + \frac{R_{\max} + 1}{(1-\gamma)^2} \sqrt{\frac{\log(\frac{8|\Theta||A|}{\delta})}{2N}} \left( C_{\text{uni}} - \sqrt{\frac{\log(\frac{4|\Theta||A|}{\delta})}{2N}} \right) \geq 1 - \delta.
\]

By Assumption 1 and a similar analysis as Lemma 1

\[
P\left( |\hat{\Theta} - \hat{\Theta}| \right) \leq \frac{4(R_{\max} + 1)}{(1-\gamma)C_H} \sqrt{\frac{\log(\frac{4|\Theta|}{\delta})}{2N}} + \frac{R_{\max} + 1}{(1-\gamma)^2C_H} \sqrt{\frac{\log(\frac{8|\Theta||A|}{\delta})}{2N}} \left( C_{\text{uni}} - \sqrt{\frac{\log(\frac{4|\Theta||A|}{\delta})}{2N}} \right) \geq 1 - \delta.
\]
Combined with Lemma 4,

$$P\left( \left| \hat{\theta} - \theta^* \right| \right) \leq \frac{4}{CH(1-\gamma)} \left( \frac{R_{\text{max}} + 1}{1-\gamma} \frac{4}{n_s^{\Theta}} - 1 \right) + 2\epsilon_Q + \epsilon_c + \frac{4(R_{\text{max}} + 1)}{(1-\gamma)CH} \sqrt{\frac{\log\left( \frac{4\left(\epsilon^2\right)}{\delta} \right)}{2N}}$$

$$+ \frac{R_{\text{max}} + 1}{(1-\gamma)^2C_H} \sqrt{\frac{\log\left( \frac{8n_s n_s^{\Theta}}{\delta} \right)}{2N}} \left( C_{uni} - \sqrt{\frac{4}{\log\left( \frac{8n_s n_s^{\Theta}}{\delta} \right)}} \right) \geq 1 - \delta.$$

References


