OPTIMAL ONLINE SELECTION OF AN ALTERNATING SUBSEQUENCE: A CENTRAL LIMIT THEOREM

ALESSANDRO ARLOTTO,* *Duke University*

J. MICHAEL STEELE,** *University of Pennsylvania*

Abstract

We analyze the optimal policy for the sequential selection of an alternating subsequence from a sequence of n independent observations from a continuous distribution F, and we prove a central limit theorem for the number of selections made by that policy. The proof exploits the backward recursion of dynamic programming and assembles a detailed understanding of the associated value functions and selection rules.

Keywords: Bellman equation; online selection; Markov decision problem; dynamic programming; alternating subsequence; central limit theorem; nonhomogeneous Markov chain

2010 Mathematics Subject Classification: Primary 60C05; 60G40; 90C40
Secondary 60F05; 90C27; 90C39

1. Introduction

In the problem of online selection of an alternating subsequence, a decision maker observes a sequence of independent random variables $\{X_1, X_2, \ldots, X_n\}$ with common continuous distribution F, and the task is to select a subsequence such that

$$X_{\theta_1} < X_{\theta_2} > X_{\theta_3} < \cdots \gtrless X_{\theta_k},$$

where the indices $1 \leq \theta_1 < \theta_2 < \theta_3 < \cdots < \theta_k \leq n$ are stopping times with respect to the σ-fields $\mathcal{F}_i = \sigma\{X_1, X_2, \ldots, X_i\}$, $1 \leq i \leq n$. In other words, at time i when the random variable X_i is first observed, the decision maker has to choose to accept X_i as a member of the alternating sequence that is under construction, or choose to reject X_i from any further consideration.

We call such a sequence of stopping times a *feasible policy*, and we denote the set of all such policies by Π. For any $\pi \in \Pi$, we then let $A_n^o(\pi)$ denote the number of selections made by π for the realization $\{X_1, X_2, \ldots, X_n\}$, i.e.

$$A_n^o(\pi) = \max\{k \colon X_{\theta_1} < X_{\theta_2} > \cdots \gtrless X_{\theta_k} \text{ and } 1 \leq \theta_1 < \theta_2 < \cdots < \theta_k \leq n\}.$$

It was found in Arlotto *et al.* (2011) that for each n there is a unique policy $\pi_n^* \in \Pi$ such that

$$\mathbb{E}[A_n^o(\pi_n^*)] = \sup_{\pi \in \Pi} \mathbb{E}[A_n^o(\pi)],$$

and it was proved that the optimal mean $\mathbb{E}[A_n^o(\pi_n^*)]$ can be tightly estimated. Specifically, we have

$$\mathbb{E}[A_n^o(\pi_n^*)] = (2 - \sqrt{2})n + O(1). \tag{1}$$

Here, our main goal is to show that $A_n^o(\pi_n^*)$ satisfies a central limit theorem.

Theorem 1. (Central limit theorem for optimal online alternating selection.) *There is a constant* $0 < \sigma^2 < \infty$ *for which we have the convergence in distribution such that*

$$\frac{A_n^o(\pi_n^*) - (2 - \sqrt{2})n}{\sqrt{n}} \xrightarrow{\text{D}} N(0, \sigma^2) \quad \text{as } n \to \infty.$$

The exact value of σ^2 is not known, but σ^2 has a representation as an infinite series and Monte Carlo calculations suggest that $\sigma^2 \sim 0.3096$. (Numerical estimates are obtained discretizing the state space with a grid size of 10^{-4} and performing 5×10^5 repetitions. The standard error for the estimate of σ^2 is 6.19×10^{-4}.) The determination of a closed-form expression for σ^2 remains an open problem. It may even be a tractable problem, though it is unlikely to be easy.

1.1. Motivation: history and connections

The theory of alternating sequences has ancient roots. It began with the investigations of Euler on alternating permutations, and, through a long evolution, it has become an important part of combinatorial theory (cf. Stanley (2010)). The probability theory of alternating sequences is much more recent, and its main problems fit into two basic categories: problems of global selection and problems of sequential selection.

In a problem of global selection (or an *offline* problem), we see the whole sequence $\{X_1, X_2, \ldots, X_n\}$, and the typical challenge is to understand the distribution of length of the longest alternating subsequence under various probability models. For example, when $\{X_1, X_2, \ldots, X_n\}$ is a random permutation of the integers $[1\colon n]$, explicit bivariate generating functions were used by Widom (2006), Pemantle (cf. Stanley (2007, p. 568)), and Stanley (2008) to obtain central limit theorems. Simpler probabilistic derivations of these results were then developed by Houdré and Restrepo (2010) and Romik (2011). These authors exploited the close connection between the length of the longest alternating subsequence and the number of local extrema of a sequence, a link that is also relevant to local minima problems studied in computer science (see e.g. Bannister and Eppstein (2012)) and to similar structures in the theory of turning point tests (see e.g. Brockwell and Davis (2006, p. 313) or Hua (2010, Section 1.2)).

The theory of online alternating subsequences is of more recent origin, but it is closely tied to some classic themes of applied probability. In the typical online decision problem, a decision maker considers n random values in sequential order and must decide whether to accept or reject each presented value at the time of its first presentation. In the most famous problem of this type, the decision maker gets to make only a single choice, and his goal is to maximize the probability that the selected value is the best out of all n values. Cayley (1875) considered a problem of this kind, but the modern development of the theory began in earnest with notable studies by Lindley (1961) and Dynkin (1963). Samuels (1991) gave a thoughtful survey of the work on related problems through the 1980s, and connections to more recent work were given by Krieger and Samuel-Cahn (2009), Buchbinder *et al.* (2010), and Bateni *et al.* (2010).

In more complex problems, the decision maker typically makes multiple sequential selections from the sequence of presented values, and the objective is to maximize the expected number of selected elements, subject to a combinatorial constraint. For example, we can consider the optimal sequential selection of a *monotone* subsequence. This online selection problem was

introduced by Samuels and Steele (1981), and was analyzed more recently by Gnedin (1999), (2000a), (2000b), Baryshnikov and Gnedin (2000), Bruss and Delbaen (2001), and Arlotto and Steele (2011).

The present investigation is particularly motivated by Bruss and Delbaen (2004), where a central limit theorem was proved for the sequential selection of a monotone subsequence when the number, N, of values offered to the decision maker is a Poisson random variable that is independent of the sequence of the offered values. The methodology of Bruss and Delbaen (2004) is tightly bound with the theory of Markov processes and Dynkin's formula, while the present method leans heavily on the Bellman equation and explicit estimates of the decision functions.

1.2. Organization of the analysis

The proof of Theorem 1 rests on a sustained investigation of the value functions that are determined by the Bellman equation for the alternating selection problem. The optimal policy π_n^* is determined in turn by the time-dependent threshold functions $\{g_n, g_{n-1}, \ldots, g_1\}$ that tell us when to accept or reject a newly presented value. Inferences from the Bellman equation almost inevitably require inductive arguments, and the numerical calculations summarized in Figure 1, below, are a great help in framing appropriate induction hypotheses.

In Section 2, we frame the selection problem as a dynamic program, and we summarize a few results from earlier work. The main observation is that, by symmetry, we can transform the natural Bellman equation into an equivalent recursion that is much simpler. We also note that the value functions determined by the reduced recursion have a useful technical feature, which we call the property of *diminishing returns*.

Sections 3–6 develop the geometry of the value and threshold functions. Even though the alternating subsequence problem is rather special, there are generic elements to its analysis, and our intention is to make these elements as visible as possible. Roughly speaking, we frame concrete hypotheses based on the suggestions of Figure 1 (or its analog), and we prove these hypotheses by inductions that are driven by the Bellman equation. While the specific inferences are unique to the problem of alternating selections, there is still some robustness to the pattern of the proof.

Sections 7 and 8 exploit the geometrical characterization of the threshold functions to obtain information about the distribution of $A_n^o(\pi_n^*)$, the number of selections made by the optimal policy for the problem with time horizon n. The main step here is the introduction of a horizon-independent policy π_∞ that is determined by the limit of the threshold functions that define π_n^*. It is relatively easy to check that the number of selections $A_n^o(\pi_\infty)$ made by this policy is a Markov additive functional of a stationary, uniformly ergodic, Markov chain. Given this observation, we can use off-the-shelf results to confirm that the central limit theorem holds for $A_n^o(\pi_\infty)$, provided that we show that the variance of $A_n^o(\pi_\infty)$ is not $o(n)$. We then complete the proof of Theorem 1 by showing that there is a coupling under which $A_n^o(\pi_n^*)$ and $A_n^o(\pi_\infty)$ are close in L^2; specifically, we show that we have $\|A_n^o(\pi_n^*) - A_n^o(\pi_\infty) - \mathbb{E}[A_n^o(\pi_n^*) - A_n^o(\pi_\infty)]\|_2 = o(\sqrt{n})$.

2. Dynamic programming formulation

We first note that, since the distribution F is continuous and since the problem is unchanged if we replace X_i by $U_i = F^{-1}(X_i)$, we can assume without loss of generality that the X_is are uniformly distributed on $[0, 1]$. The main task now is to exploit the symmetries of the problem to obtain a tractable version of the Bellman equation.

We proceed recursively and, for $1 \leq i \leq n$, we let S_i denote the value of the last member of the subsequence selected up to and including time i. We also set $R_i = 0$ if S_i is a local minimum of $\{S_0, S_1, \ldots, S_i\}$, and we set $R_i = 1$ if S_i is a local maximum. Finally, to initialize our process, we set $S_0 = 1$ and $R_0 = 1$, and we note that the process $\{(S_i, R_i) \colon 0 \leq i \leq n\}$ is Markov.

At time i, the decision to accept or reject the new observation X_i depends only on the following two quantities:

1. the state of the selection process before the presentation of X_i; this is represented by the pair (S_{i-1}, R_{i-1}),

2. the number of observations k that were yet to be seen before the presentation of X_i, i.e. $k = n - i + 1$.

We can now characterize the optimal policy π_n^* through these state variables and an associated dynamic programming equation (or Bellman equation) for the value function. We let $v_k(s, r)$ denote the expected number of optimal alternating selections when the number of observations yet to be seen is k, and the state of the selection process is given by the pair (s, r). If $k = 0$ then we set $v_0(s, r) \equiv 0$ for all $(s, r) \in [0, 1] \times \{0, 1\}$. Otherwise, for $1 \leq i \leq n$, states $S_{i-1} = s$ and $R_{i-1} = r$ and residual time $k = n - i + 1$, we have the Bellman equation

$$
v_k(s, r) = \begin{cases} s v_{k-1}(s, 0) + \displaystyle\int_s^1 \max\{v_{k-1}(s, 0), 1 + v_{k-1}(x, 1)\} \, \mathrm{d}x & \text{if } r = 0, \\ (1 - s) v_{k-1}(s, 1) + \displaystyle\int_0^s \max\{v_{k-1}(s, 1), 1 + v_{k-1}(x, 0)\} \, \mathrm{d}x & \text{if } r = 1. \end{cases} \tag{2}
$$

To see why this equation holds, first consider the case when $r = 0$ (so the next selection needs to be a local maximum). With probability s, we are presented with a value, X_i, that is less than the previously selected value. In this case, we do not have the opportunity to make a selection, we reduce the number of observations yet to be seen to $k - 1$, and this contributes the term $s v_{k-1}(s, 0)$ to our equation.

Next, consider the case when $r = 0$ but $s < X_i \leq 1$. In this case, we must decide to select $X_i = x$, or to reject it. If we do not select the value $X_i = x$, then the expected number of subsequent selections equals $v_{k-1}(s, 0)$. If we do select $X_i = x$, then we account for the selection of x plus the expected number of optimal subsequent selections, which together equal $1 + v_{k-1}(x, 1)$. Since X_i is uniformly distributed in $[s, 1]$, the expected optimal contribution is given by the second term of our Bellman equation (2) (the first line). The proof of the second line of our Bellman equation (2) is completely analogous.

One benefit of indexing the value functions $v_k(\cdot, \cdot)$ by the 'time-to-go' parameter k is that, by the optimality principle of dynamic programming, the selection problem for a sequence of size n embeds automatically into the selection problem for a sequence of size $n + 1$. As a consequence, we can consider the infinite sequence of value functions $\{v_k(\cdot, \cdot), \, 1 \leq k < \infty\}$. It is also useful to observe that these value functions satisfy an intuitive symmetry property:

$$
v_k(s, 0) = v_k(1 - s, 1) \quad \text{for all } 1 \leq k < \infty \text{ and all } s \in [0, 1], \tag{3}
$$

so we can define the *single-variable value function* $v_k(y)$, $1 \leq k < \infty$, by setting

$$
v_k(y) \equiv v_k(y, 0) = v_k(1 - y, 1) \quad \text{for all } 1 \leq k < \infty \text{ and all } y \in [0, 1].
$$

(A formal proof of (3) is given in Arlotto *et al.* (2011, Lemma 3).) Now, when we replace the bivariate value function $v_k(\cdot, \cdot)$ in the earlier Bellman equation with the corresponding value of the univariate value function $v_k(\cdot)$, we obtain a much nicer recursion:

$$v_k(y) = yv_{k-1}(y) + \int_y^1 \max\{v_{k-1}(y), 1 + v_{k-1}(1 - x)\}\, dx. \tag{4}$$

Here, we have $v_0(y) \equiv 0$ for all $y \in [0, 1]$, and we note that the map $y \mapsto v_k(y)$ is continuous and differentiable on $[0, 1]$, and it satisfies the boundary condition $v_k(1) = 0$ for all $1 \le k < \infty$. In this reduced setting, the state of the selection process is simply given by the value y, rather than the pair (s, r).

The key benefit of the reduced Bellman equation (4) is that it leads to a simple rule for the optimal acceptance or rejection of a newly presented observation. Specifically, if we set

$$g_k(y) = \inf\{x \in [y, 1] : v_{k-1}(y) \le 1 + v_{k-1}(1 - x)\}, \tag{5}$$

then a value x is an optimal selection if and only if $g_k(y) \le x$. For $1 \le k < \infty$, we then call the function $g_k \colon [0, 1] \to [0, 1]$ the *optimal threshold function* if it satisfies the variational characterization given by (5). We will see shortly that the value function $y \mapsto v_k(y)$ is strictly decreasing on $[0, 1]$, a fact that will imply that $g_k(y)$ is uniquely determined for each $y \in [0, 1]$.

The optimal threshold functions $\{g_n, g_{n-1}, \ldots, g_1\}$ give us a useful representation for the number $A_n^o(\pi_n^*)$ of selections made by the optimal policy π_n^*. Specifically, if we set $Y_0 \equiv 0$ and define the sequence $Y_1, Y_2, \ldots,$ by the recursion

$$Y_i = \begin{cases} Y_{i-1} & \text{if } X_i < g_{n-i+1}(Y_{i-1}), \\ 1 - X_i & \text{if } X_i \ge g_{n-i+1}(Y_{i-1}), \end{cases}$$

then we have

$$A_n^o(\pi_n^*) = \sum_{i=1}^n \mathbf{1}(X_i \ge g_{n-i+1}(Y_{i-1})) \tag{6}$$

and, moreover, by the principle of optimality of dynamic programming, we have

$$\mathbb{E}[A_n^o(\pi_n^*)] = v_n(0) \quad \text{for each } n \ge 1.$$

The representation (6) also tells us that $A_n^o(\pi_n^*)$ is a sum of functions of a time inhomogeneous Markov chain. The analysis of this inhomogeneous additive functional calls for a reasonably detailed understanding of both the threshold functions $\{g_k(\cdot) \colon 1 \le k < \infty\}$ and the value functions $\{v_k(\cdot) \colon 1 \le k < \infty\}$.

A technical fact that will be needed shortly is that, for each $1 \le k < \infty$, the value function $v_k(\cdot)$ satisfies the bound

$$v_{k-1}(u) - v_{k-1}(1 - y) \le v_k(u) - v_k(1 - y) \quad \text{for all } y \in \left[0, \tfrac{1}{2}\right] \text{ and } u \in [y, 1 - y]. \tag{7}$$

This bound reflects a restricted *principle of diminishing returns*; a proof of (7) is given by Arlotto *et al.* (2011, Lemma 4).

Given the dynamic programming formulation provided here, the results in this paper can be read independently of Arlotto *et al.* (2011). Still, for the purpose of comparison, we should note that the notation used here simplifies our earlier notation in some significant ways. For example, we now take k to be the number of observations *yet to be seen*, and this gives us the pleasing formulation (4) of the Bellman equation. We also write $g_k(y)$ for the optimal threshold function when there are k observations yet to be seen, and this replaces the earlier, more cumbersome, notation $f_{n-k+1,n}^*(y)$.

3. Geometry of the value and threshold functions

Figure 1 gives a highly suggestive picture of the individual threshold functions $g_k(\cdot)$, and it foretells much of the story about how they behave as $k \to \infty$. Analytical confirmation of these suggestions is the central challenge. The path to understanding the threshold functions goes through the value functions, and we begin by proving the very plausible fact that the value functions are strictly decreasing.

Lemma 1. (Strict monotonicity of the value functions.) *For each $1 \leq k < \infty$, the value function $y \mapsto v_k(y)$ defined by the Bellman recursion (4) is strictly decreasing on $[0, 1]$.*

This assertion is certainly intuitive and we may not feel any need for a proof. Nevertheless, there is something to be gained from a formal proof; specifically, we see in a simple context how the Bellman equation can be used to propagate a sequence of induction hypotheses.

Proof of Lemma 1. We consider the sequence of hypotheses:

$$\mathbb{H}_k : v_k(y + \varepsilon) < v_k(y) \quad \text{for all } y \in [0, 1) \text{ and all } \varepsilon > 0 \text{ such that } y + \varepsilon \leq 1.$$

Since $v_1(y) = 1 - y$, $\mathbb{H}_1$ is true. For $k \geq 2$, we note that, by the Bellman recursion (4), we have

$$
\begin{aligned}
v_k(y + \varepsilon) - v_k(y) &= (y + \varepsilon)v_{k-1}(y + \varepsilon) + \int_{y+\varepsilon}^{1} \max\{v_{k-1}(y + \varepsilon), 1 + v_{k-1}(1 - x)\} \, dx \\
&\quad - yv_{k-1}(y) - \int_{y}^{1} \max\{v_{k-1}(y), 1 + v_{k-1}(1 - x)\} \, dx \\
&\leq (y + \varepsilon)v_{k-1}(y + \varepsilon) + \int_{y+\varepsilon}^{1} \max\{v_{k-1}(y), 1 + v_{k-1}(1 - x)\} \, dx \\
&\quad - (y + \varepsilon)v_{k-1}(y) - \int_{y+\varepsilon}^{1} \max\{v_{k-1}(y), 1 + v_{k-1}(1 - x)\} \, dx \\
&= (y + \varepsilon)\{v_{k-1}(y + \varepsilon) - v_{k-1}(y)\} \\
&< 0,
\end{aligned}
$$

where the first inequality of the chain follows from

$$\varepsilon \, v_{k-1}(y) \leq \int_{y}^{y+\varepsilon} \max\{v_{k-1}(y), 1 + v_{k-1}(1 - x)\} \, dx$$

and the second inequality follows from $\mathbb{H}_{k-1}$. This completes the proof of $\mathbb{H}_k$ and of the lemma.

Figure 1 further suggests that the threshold functions have a long interval of fixed points; the following lemma partially confirms this.

Lemma 2. (Range of fixed points.) *For all $k \geq 1$ and $y \in [0, 1]$, we have*

$$v_k(y) - v_k\left(\tfrac{2}{3}\right) \leq v_k(0) - v_k\left(\tfrac{2}{3}\right) \leq 1. \tag{8}$$

In particular, for all $k \geq 1$, we have

$$g_k(y) = y \quad \text{for all } y \in \left[\tfrac{1}{3}, 1\right], \tag{9}$$

and

$$g_k(y) \leq \tfrac{1}{3} \quad \text{for all } y \in \left[0, \tfrac{1}{3}\right]. \tag{10}$$
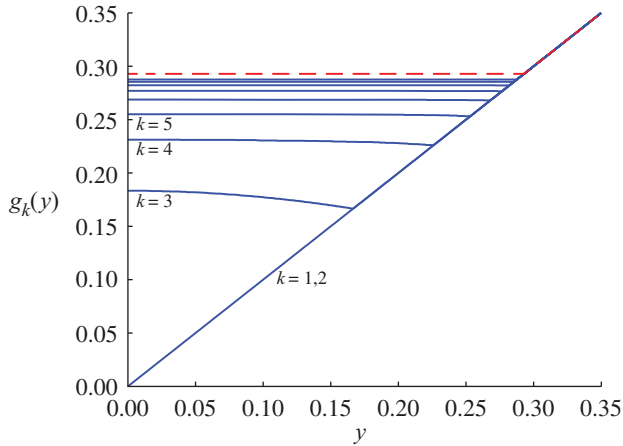
FIGURE 1: Threshold functions $g_k(y)$, $1 \le k \le 10$, and their limit as $k \to \infty$ for $y \in [0, 0.35]$. Solid lines plot the threshold functions g_k, $1 \le k \le 10$, for values of y in the range $[0, 0.35]$. We have $g_1(y) = g_2(y) = y$ for all $y \in [0, 1]$, and the piecewise smooth graphs of g_3, g_4, and g_5 are explicitly labeled with their index placed just below the curve. For $k = 6, 7, \ldots, 10$, the g_k are indicated without labels. Each g_k meets the diagonal line at some point, and we have $g_k(y) = y$ on the rest of the interval $[0, 1]$. The plot suggests most of the analytical properties of the sequence $\{g_k : 1 \le k < \infty\}$ that are needed for the proof of the central limit theorem. In particular, for each fixed $y \in [0, 1]$ the sequence $k \mapsto g_k(y)$ is monotone nondecreasing. The dashed line represents the limit of $g_k(y)$ as $k \to \infty$; this limit is piecewise linear.

Proof. The first inequality of (8) is trivial since the map $y \mapsto v_k(y)$ is strictly decreasing in y. Also, the identities (9) and (10) are immediate from the variational characterization (5) and the bound (8).

The real task is to prove the second inequality of (8). This time we use induction on the hypotheses given by

$$\mathbb{H}_k : v_k(0) - v_k\left(\tfrac{2}{3}\right) \le 1 \quad \text{for } 1 \le k < \infty.$$

As before, $v_1(y) = 1 - y$, so $\mathbb{H}_1$ is trivially true. Now, when we apply the Bellman recursion (4) with $y = 0$ and $y = \tfrac{2}{3}$, we get

$$v_k(0) - v_k\left(\tfrac{2}{3}\right) = \int_0^1 \max\{v_{k-1}(0), 1 + v_{k-1}(1 - u)\}\, du$$

$$- \tfrac{2}{3} v_{k-1}\left(\tfrac{2}{3}\right) - \int_{2/3}^1 \max\left\{v_{k-1}\left(\tfrac{2}{3}\right), 1 + v_{k-1}(1 - u)\right\} du,$$

from which a change of variable gives

$$v_k(0) - v_k\left(\tfrac{2}{3}\right) = \int_0^{1/3} I_1(u)\, du + \int_{1/3}^1 I_2(u)\, du, \tag{11}$$

where $I_1(u)$ and $I_2(u)$ are defined by

$$I_1(u) \equiv \max\{v_{k-1}(0), 1 + v_{k-1}(u)\} - \max\left\{v_{k-1}\left(\tfrac{2}{3}\right), 1 + v_{k-1}(u)\right\}$$

and

$$I_2(u) \equiv \max\left\{v_{k-1}(0) - v_{k-1}\left(\tfrac{2}{3}\right), 1 + v_{k-1}(u) - v_{k-1}\left(\tfrac{2}{3}\right)\right\}.$$

For the first integrand, $I_1(u)$, we note that

$$I_1(u) = \max\left\{v_{k-1}(0) - v_{k-1}\left(\tfrac{2}{3}\right), 1 + v_{k-1}(u) - v_{k-1}\left(\tfrac{2}{3}\right)\right\} \\ - \max\left\{0, 1 + v_{k-1}(u) - v_{k-1}\left(\tfrac{2}{3}\right)\right\}. \tag{12}$$

The induction assumption $\mathbb{H}_{k-1}$ then tells us that

$$v_{k-1}(0) - v_{k-1}\left(\tfrac{2}{3}\right) \le 1,$$

and the strict monotonicity of the value function $v_{k-1}(\cdot)$ on $[0, 1]$ yields

$$1 \le 1 + v_{k-1}(u) - v_{k-1}\left(\tfrac{2}{3}\right) \quad \text{for all } u \in \left[0, \tfrac{1}{3}\right].$$

Thus, both the first and the second addend in (12) equal the right maximand and

$$I_1(u) = 0 \quad \text{for all } u \in \left[0, \tfrac{1}{3}\right],$$

so the first integral in (11) vanishes.

To estimate $I_2(u)$, we note that $\mathbb{H}_{k-1}$ and the monotonicity of $y \mapsto v_{k-1}(y)$ tell us that

  (i) if $u \in [\tfrac{1}{3}, \tfrac{2}{3}]$ then

$$I_2(u) = 1 + v_{k-1}(u) - v_{k-1}\left(\tfrac{2}{3}\right) \le 1 + v_{k-1}(0) - v_{k-1}\left(\tfrac{2}{3}\right) \le 2,$$

  (ii) if $u \in [\tfrac{2}{3}, 1]$ then

$$I_2(u) = \max\left\{v_{k-1}(0) - v_{k-1}\left(\tfrac{2}{3}\right), 1 + v_{k-1}(u) - v_{k-1}\left(\tfrac{2}{3}\right)\right\} \le 1.$$

Now we just calculate

$$v_k(0) - v_k\left(\tfrac{2}{3}\right) = \int_{1/3}^{1} I_2(u)\,\mathrm{d}u \le \int_{1/3}^{2/3} 2\,\mathrm{d}u + \int_{2/3}^{1} 1\,\mathrm{d}u = 1,$$

and thus we complete the proof of (8).

From Lemma 2, we know that a threshold function g_k has many fixed points; in particular, $g_k(y) = y$ if $y \in [\tfrac{1}{3}, 1]$. Figure 1 further suggests that much of the geometry of g_k is governed by its *minimal* fixed point:

$$\xi_k \equiv \inf\{y \colon g_k(y) = y\}. \tag{13}$$

The value ξ_k also has a useful policy interpretation. If the value y of the last observation selected is bigger than ξ_k, then the decision maker follows a greedy policy; he accepts *any* feasible arriving observation. On the other hand, if $y < \xi_k$, the decision maker acts conservatively; his choices are governed by the value of the threshold $g_k(y)$. Finally, if $y = \xi_k$, the greedy policy and the optimal policy agree. This interpretation of ξ_k is formalized in the following lemma, where we also prove that the sequence $\{\xi_k \colon k = 1, 2, \dots\}$ is nondecreasing.

Lemma 3. (Characterization of the minimal fixed point.) *For $k \ge 3$, the minimal fixed point $\xi_k \equiv \inf\{y \colon g_k(y) = y\}$ is the unique solution to the equation*

$$v_{k-1}(y) - v_{k-1}(1 - y) = 1.$$

Moreover, the minimal fixed points form a nondecreasing sequence, so we have

$$\xi_k \le \xi_{k+1} \quad \text{for all } k \ge 1.$$

Proof. From the variational characterization of $g_k(\cdot)$, we have

$$g_k(y) = \inf\{x \in [y, 1]: v_{k-1}(y) \leq 1 + v_{k-1}(1 - x)\},$$

so if we set $\delta_k(y) \equiv v_{k-1}(y) - v_{k-1}(1 - y)$, then we have

$$g_k(y) = y \quad \text{if and only if} \quad \delta_k(y) \leq 1. \tag{14}$$

The Bellman equation (4) for $v_k(\cdot)$ and Lemma 1 tell us that the map $y \mapsto v_{k-1}(y)$ is continuous and strictly decreasing with $v_1(y) = 1 - y$ and $v_2(y) = \frac{3}{2}(1 - y^2)$. Then, the function δ_k is continuous and strictly decreasing, and for $k \geq 3$ we have $\delta_k(0) = v_{k-1}(0) \geq v_2(0) = \frac{3}{2} > 1$, and $\delta_k(1) = -v_{k-1}(0) < 0$, so, there is a unique value y^* such that

$$\delta_k(y^*) \equiv v_{k-1}(y^*) - v_{k-1}(1 - y^*) = 1.$$

Since the map $y \mapsto \delta_k(y)$ is strictly decreasing, we can also write y^* as

$$y^* = \inf\{y: v_{k-1}(y) - v_{k-1}(1 - y) \leq 1\} = \inf\{y: g_k(y) = y\} = \xi_k,$$

where the second equality follows from (14) and the third equality comes from the definition of ξ_k.

To prove the monotonicity property $\xi_k \leq \xi_{k+1}$ for all $k \geq 1$, we first note that, since $v_0(y) \equiv 0$ and $v_1(y) \equiv 1 - y$, $\xi_1 = \xi_2 = 0$. Also, by Lemma 2 we have for $k \geq 3$ that there is always a value $0 \leq y \leq \frac{1}{3}$ such that $g_k(y) = y$, so

$$\begin{aligned}
\xi_k &= \inf\{y \in [0, \tfrac{1}{3}]: g_k(y) = y\} \\
&= \inf\{y \in [0, \tfrac{1}{3}]: \delta_k(y) \equiv v_{k-1}(y) - v_{k-1}(1 - y) \leq 1\} \\
&\leq \inf\{y \in [0, \tfrac{1}{3}]: \delta_{k+1}(y) \equiv v_k(y) - v_k(1 - y) \leq 1\} \\
&= \inf\{y \in [0, \tfrac{1}{3}]: g_{k+1}(y) = y\} \\
&= \xi_{k+1},
\end{aligned} \tag{15}$$

where the inequality in (15) follows from the diminishing return property (7).

4. A second property of diminishing returns

The value functions have a second property of diminishing returns that provides some crucial help. Specifically, we need it to show that the threshold functions $g_k(\cdot)$ increase with $1 \leq k < \infty$. This monotonicity moves us a long way toward an exhaustive understanding of the asymptotic behavior of the threshold functions.

Proposition 1. (Second property of diminishing returns.) *For all $k \geq 3$, the value functions defined by the Bellman recursion (4) satisfy the bound*

$$v_{k-1}(y) - v_{k-1}(1 - x) \leq v_k(y) - v_k(1 - x) \quad \text{for all } y \leq \xi_k \text{ and } x \in [y, g_k(y)]. \tag{16}$$

Proof. We again use induction to exploit the Bellman equation, and this time the sequence of hypotheses is given by

$$\mathbb{H}_k: v_{k-1}(y) - v_{k-1}(1 - x) \leq v_k(y) - v_k(1 - x) \quad \text{for all } y \leq \xi_k \text{ and } x \in [y, g_k(y)].$$

We first prove $\mathbb{H}_3$, which we then use as the base case for our induction. We recall that $v_1(y) = 1 - y$ and, if we use the Bellman recursion (4), $v_2(y) = \frac{3}{2}(1 - y^2)$. In turn, this implies $g_3(y) = \max\{1 - (\frac{2}{3} + y^2)^{1/2}, y\}$ and $\xi_3 = \frac{1}{6}$. To calculate $v_3(y)$ we apply the Bellman recursion one more time, and we obtain a messier but still tractable formula:

$$v_3(y) = \begin{cases} \frac{3}{2}(1 - y^2) + 3^{-3/2}(2 + 3y^2)^{3/2} & \text{if } y \leq \frac{1}{6}, \\ \frac{1}{2}(1 - y)(4 + 5y + 2y^2) & \text{if } y \geq \frac{1}{6}. \end{cases}$$

Thus, for $y \leq \xi_3 = \frac{1}{6}$, we need to show

$$v_2(y) - v_2(1 - x) \leq v_3(y) - v_3(1 - x) \quad \text{for all } x \in [y, g_3(y)],$$

where $g_3(y) = 1 - (\frac{2}{3} + y^2)^{1/2}$. From our explicit formulae for $v_2(\cdot)$ and $v_3(\cdot)$, we have

$$v_3(1 - x) - v_2(1 - x) = \frac{5}{2}x - 3x^2 + x^3$$

and

$$v_3(y) - v_2(y) = 3^{-3/2}(2 + 3y^2)^{3/2} \geq \left(\frac{2}{3}\right)^{3/2} \approx 0.5443.$$

Calculus shows that $\frac{5}{2}x - 3x^2 + x^3$ increases on $0 \leq x \leq 1 - (\frac{2}{3})^{1/2}$ and attains an endpoint maximum of $\frac{1}{18}(9 - \sqrt{6}) \approx 0.3640$. Thus, we find

$$v_3(1 - x) - v_2(1 - x) \leq \frac{1}{18}(9 - \sqrt{6}) < \left(\frac{2}{3}\right)^{3/2} \leq v_3(y) - v_2(y)$$

for all $y \leq \frac{1}{6}$ and $y \leq x \leq 1 - (\frac{2}{3} + y^2)^{1/2}$, completing the proof of $\mathbb{H}_3$.

We now suppose that $\mathbb{H}_k$ holds, and we seek to show $\mathbb{H}_{k+1}$. First, from the variational characterization of $g_k(\cdot)$ and the definition of ξ_k, recall that

$$1 \leq v_{k-1}(y) - v_{k-1}(1 - x) \quad \text{for } y \leq \xi_k \text{ and } x \in [y, g_k(y)],$$

which, together with the induction assumption $\mathbb{H}_k$, implies

$$1 \leq v_{k-1}(y) - v_{k-1}(1 - x) \leq v_k(y) - v_k(1 - x) \quad \text{for } y \leq \xi_k \text{ and } x \in [y, g_k(y)]. \quad (17)$$

The second inequality in (17) and the variational characterization (5) give us

$$g_k(y) \leq g_{k+1}(y) \quad \text{for all } y \leq \xi_k.$$

Moreover, if $x \in [g_k(y), g_{k+1}(y)]$, the variational characterization of $g_{k+1}(\cdot)$ also gives

$$v_{k-1}(y) - v_{k-1}(1 - x) \leq 1 \leq v_k(y) - v_k(1 - x) \quad \text{for } y \leq \xi_k \text{ and } x \in [g_k(y), g_{k+1}(y)],$$

which combines with (17) to give the crucial inequality

$$v_{k-1}(y) - v_{k-1}(1 - x) \leq v_k(y) - v_k(1 - x) \quad \text{for } y \leq \xi_k \text{ and } x \in [y, g_{k+1}(y)]. \quad (18)$$

From an application of the Bellman recursion (4) for $y \leq \xi_k$ and $x \in [y, g_{k+1}(y)]$, we obtain

$$v_k(y) - v_k(1 - x)$$
$$= y(v_{k-1}(y) - v_{k-1}(1 - x))$$
$$+ \int_y^{1-x} \max\{v_{k-1}(y) - v_{k-1}(1 - x), 1 + v_{k-1}(1 - u) - v_{k-1}(1 - x)\} \, du. \quad (19)$$

If we now change variable in the last integral by replacing u with $1 - u$, then the range of integration changes to $[x, 1 - y]$ and we can rewrite (19) as

$$
v_k(y) - v_k(1 - x) = y(v_{k-1}(y) - v_{k-1}(1 - x))
$$
$$
+ \int_x^{1-x} \max\{v_{k-1}(y) - v_{k-1}(1 - x), 1 + v_{k-1}(u) - v_{k-1}(1 - x)\} \, du
$$
$$
+ \int_{1-x}^{1-y} \max\{v_{k-1}(y) - v_{k-1}(1 - x), 1 + v_{k-1}(u) - v_{k-1}(1 - x)\} \, du.
$$

In this last equation, we see that we can use our crucial inequality (18) to bound the first addend and the left maximand of the other two addends. Moreover, since $x \le g_{k+1}(y) \le \frac{1}{3}$, we can appeal to the diminishing return property (7) to bound the right maximand of the second addend. In doing so, we obtain

$$
v_k(y) - v_k(1 - x) \le y(v_k(y) - v_k(1 - x))
$$
$$
+ \int_x^{1-x} \max\{v_k(y) - v_k(1 - x), 1 + v_k(u) - v_k(1 - x)\} \, du
$$
$$
+ \int_{1-x}^{1-y} \max\{v_k(y) - v_k(1 - x), 1 + v_{k-1}(u) - v_{k-1}(1 - x)\} \, du.
$$
$$
\tag{20}
$$

We now observe that the monotonicity property of the map $u \mapsto v_{k-1}(u)$, for $u \in [1-x, 1-y]$, and the variational characterization of $g_{k+1}(\cdot)$ combine to give

$$
1 + v_{k-1}(u) - v_{k-1}(1 - x) \le 1 \le v_k(y) - v_k(1 - x)
$$

for all $y \le \xi_k$ and $x \in [y, g_{k+1}(y)]$. Hence, the second integrand in (20) satisfies the equality

$$
\max\{v_k(y) - v_k(1 - x), 1 + v_{k-1}(u) - v_{k-1}(1 - x)\} = v_k(y) - v_k(1 - x),
$$

and an analogous monotonicity argument for $u \in [1 - x, 1 - y]$ also yields

$$
\max\{v_k(y) - v_k(1 - x), 1 + v_k(u) - v_k(1 - x)\} = v_k(y) - v_k(1 - x).
$$

When we use the last two observations in (20) we obtain

$$
v_k(y) - v_k(1 - x) \le v_{k+1}(y) - v_{k+1}(1 - x) \quad \text{for all } y \le \xi_k \text{ and } x \in [y, g_{k+1}(y)].
$$

We now conclude our argument by considering values $y \in [\xi_k, \xi_{k+1}]$. From the variational characterization of $g_{k+1}(\cdot)$ and the definition of ξ_k, we obtain

$$
v_{k-1}(y) - v_{k-1}(1 - x) \le 1 \le v_k(y) - v_k(1 - x) \quad \text{for } y \in [\xi_k, \xi_{k+1}] \text{ and } x \in [y, g_{k+1}(y)],
$$

which can be used instead of (18) to construct an argument similar to the earlier one, so we conclude that

$$
v_k(y) - v_k(1 - x) \le v_{k+1}(y) - v_{k+1}(1 - x) \quad \text{for } y \in [\xi_k, \xi_{k+1}] \text{ and } x \in [y, g_{k+1}(y)],
$$

just as needed to complete the proof of (16).

The usefulness of the property of diminishing returns in Proposition 1 shows itself simply, but clearly, in the following corollary.

Corollary 1. (Monotonicity of optimal thresholds.) *For all $y \in [0, 1]$, the threshold functions satisfy*

$$g_k(y) \leq g_{k+1}(y) \quad \text{for all } k \geq 1, \tag{21}$$

$$\tfrac{1}{6} \leq g_k(y) \quad \text{for all } k \geq 3. \tag{22}$$

Proof. For $k = 1, 2$, we have $v_0(y) = 0$ and $v_1(y) = 1 - y$, so that

$$g_1(y) = g_2(y) = y.$$

For $k = 3$, we have already noticed in the course of proving Proposition 1 that we have $g_3(y) = \max\{1 - (\tfrac{2}{3} + y^2)^{1/2}, y\}$, so, in particular, $g_3(y) \geq \tfrac{1}{6}$ for $y \in [0, 1]$. Finally, for $k > 3$, the bound (16) and the variational characterization (5) of the threshold function give us (1), and this confirms the lower bound (22).

We now pursue two further suggestions from Figure 1. Specifically, we show that the limit function g_∞ has exactly the piecewise linear shape that the figure suggests, and we also show that the convergence to g_∞ is uniform. The proof of these facts requires some additional regularity properties that are discussed in Section 5.

5. Regularity of the value and threshold functions

The minimal fixed points give us a powerful guide to the geometry of the value function and its derivatives. The connection begins with the Bellman recursion (4) and the variational characterization (5) which together give the identity

$$v_k(y) = g_k(y)v_{k-1}(y) + \int_{g_k(y)}^{1} \{1 + v_{k-1}(1 - x)\} \, \mathrm{d}x.$$

If we now differentiate both sides with respect to y, we obtain the recursion for the first derivative, i.e.

$$v_k'(y) = g_k'(y)v_{k-1}(y) + g_k(y)v_{k-1}'(y) - g_k'(y)\{1 + v_{k-1}(1 - g_k(y))\}.$$

The definition of the minimal fixed point (13) and the variational characterization (5) then gives us

$$v_{k-1}(y) = 1 + v_{k-1}(1 - g_k(y)) \quad \text{if } y \leq \xi_k, \tag{23}$$

so our recursion for $v_k'(\cdot)$ can be written more informatively as

$$v_k'(y) = \begin{cases} g_k(y)v_{k-1}'(y) & \text{if } y \leq \xi_k, \\ v_{k-1}(y) - 1 - v_{k-1}(1 - y) + yv_{k-1}'(y) & \text{if } y \geq \xi_k. \end{cases} \tag{24}$$

These relations underscore the importance of the minimal fixed points to the geometry of the value function, and they also lead to useful regularity properties.

Lemma 4. (Monotonicity properties of the derivatives.) *For all $k \geq 1$, we have*

$$-1 \leq v_k'(y) \leq v_{k+1}'(y) \leq 0 \quad \text{for } y \in \left[0, \tfrac{1}{3}\right], \tag{25}$$

$$v_{k+1}'(y) \leq v_k'(y) \leq -1 \quad \text{for } y \in \left[\tfrac{1}{2}, 1\right]. \tag{26}$$

Proof. We already know from Lemma 1 that $y \mapsto v_k(y)$ is strictly decreasing, so $v_k'(y)$ is nonpositive on $[0, 1]$. Since $0 \le g_k(y) \le 1$, the top line of (24) tells us that

$$v_{k-1}'(y) \le g_k(y)v_{k-1}'(y) = v_k'(y) \quad \text{for } y \le \xi_k. \tag{27}$$

To cover the rest of the range in (25), we use induction on the sequence of hypotheses

$$\mathbb{H}_k: v_{k-1}'(y) \le v_k'(y) \quad \text{for all } y \in \left[\xi_k, \tfrac{1}{3}\right] \text{ and } 2 \le k < \infty.$$

For the base case $\mathbb{H}_2$, we have $\xi_2 = 0$, $v_1(y) = 1 - y$, and $v_2(y) = \tfrac{3}{2}(1 - y^2)$. So

$$v_1'(y) = -1 \le -3y = v_2'(y) \quad \text{if and only if} \quad y \le \tfrac{1}{3},$$

just as needed. Now taking $\mathbb{H}_k$ as our induction assumption we seek to prove $\mathbb{H}_{k+1}$.

First, for $y \in [\xi_k, \tfrac{1}{3}]$, the second line of (24) gives us $v_k'(\cdot)$. By the diminishing return property (7), the monotonicity $\xi_k \le \xi_{k+1}$, and the induction assumption $\mathbb{H}_k$, we see for $y \in [\xi_{k+1}, \tfrac{1}{3}]$ that

$$\begin{aligned}
v_k'(y) &= v_{k-1}(y) - 1 - v_{k-1}(1 - y) + yv_{k-1}'(y) \\
&\le v_k(y) - 1 - v_k(1 - y) + yv_k'(y) \\
&= v_{k+1}'(y),
\end{aligned}$$

completing the proof of $\mathbb{H}_{k+1}$. To complete the proof of (25), we just need to note that the lower bound $-1 \le v_k'(y)$ now follows from $v_1'(y) = -1$ together with (27) and $\mathbb{H}_k$.

To prove (26), we again use induction, but this time the sequence of hypotheses is given by

$$\mathbb{H}_k: v_k'(y) \le v_{k-1}'(y) \quad \text{for } y \in \left[\tfrac{1}{2}, 1\right] \text{ and } 2 \le k < \infty.$$

As before, $v_1(y) = 1 - y$ and $v_2(y) = \tfrac{3}{2}(1 - y^2)$, so $v_1'(y) = -1$ and $v_2'(y) = -3y$. For $y \ge \tfrac{1}{2}$, we then have

$$v_2'(y) \le -\tfrac{3}{2} \le -1 = v_1'(y),$$

proving $\mathbb{H}_2$. As tradition demands, we again take $\mathbb{H}_k$ as our induction assumption, and we seek to prove $\mathbb{H}_{k+1}$.

Since $y \in [\tfrac{1}{2}, 1]$, we have $1 - y \le \tfrac{1}{2} \le y$, so the diminishing return property (7) gives us

$$v_{k-1}(1 - y) - v_{k-1}(y) \le v_k(1 - y) - v_k(y). \tag{28}$$

Next, recall the identity of the bottom line of (24), but, as you do so, replace k by $k + 1$. We can then directly apply (28) and $\mathbb{H}_k$ to get

$$\begin{aligned}
v_{k+1}'(y) &= v_k(y) - 1 - v_k(1 - y) + yv_k'(y) \\
&\le v_{k-1}(y) - 1 - v_{k-1}(1 - y) + yv_{k-1}'(y) \\
&= v_k'(y).
\end{aligned}$$

This inequality completes the proof of $\mathbb{H}_{k+1}$ and confirms the lower bound of (26). For the upper bound of (26), $v_k'(y) \le -1$ on $[\tfrac{1}{2}, 1]$, we just need to note that it follows from the fact $v_1'(y) = -1$ and the validity of $\mathbb{H}_k$ for all $k \ge 1$.

The smoothness of the value functions converts easily into a very useful Lipschitz equicontinuity property of the threshold functions.

Lemma 5. (Lipschitz equicontinuity of threshold functions.) *For all $k \geq 1$, we have*

$$|g_k(y) - g_k(z)| \leq |y - z| \quad \text{for all } y, z \in [0, 1]. \tag{29}$$

Proof. We first consider $y \in [0, \xi_k]$. In this case, (23) holds, so, by its differentiation, we obtain

$$g_k'(y) = -\frac{|v_{k-1}'(y)|}{|v_{k-1}'(1 - g_k(y))|} \leq 0 \quad \text{for all } y \in [0, \xi_k]. \tag{30}$$

Moreover, since $y \in [0, \xi_k]$, we know that $y \leq \frac{1}{3}$, so by (10) we have $g_k(y) \leq \frac{1}{3}$; hence, by (26) we obtain $1 \leq |v_{k-1}'(1 - g_k(y))|$. Consequently, (30) gives us

$$|g_k'(y)| \leq |v_{k-1}'(y)| \quad \text{for all } y \in [0, \xi_k], \tag{31}$$

and (25) implies $|v_k'(y)| \leq 1$. Thus, at last, we have the uniform bound

$$|g_k'(y)| \leq 1 \quad \text{for all } y \in [0, \xi_k],$$

which confirms inequality (29) for $y, z \in [0, \xi_k]$. Also, for $y, z \in [\xi_k, 1]$, (29) trivially holds, so if we choose $y < \xi_k < z$, the triangle inequality gives us

$$|g_k(y) - g_k(z)| \leq |g_k(y) - g_k(\xi_k)| + |g_k(\xi_k) - g_k(z)| \leq |y - z|,$$

confirming that (29) holds in general.

6. The optimal policy at ∞

The minimal fixed points ξ_k, $1 \leq k < \infty$, are nondecreasing and bounded by $\frac{1}{3}$, so they have a limit

$$\lim_{k \to \infty} \xi_k =: \xi \leq \frac{1}{3}. \tag{32}$$

The threshold values $g_k(y)$, $1 \leq k < \infty$, are also nondecreasing and bounded, so they have a pointwise limit $g_\infty(y)$. The following proposition characterizes g_∞ and gives a crucial bound on the uniform rate of convergence to g_∞.

Proposition 2. (Characterization of limiting threshold.) *For the limit threshold g_∞, we have*

$$g_\infty(y) = \max\{\xi, y\} \quad \text{for all } y \in [0, 1].$$

Moreover, we have an exact measure of the uniform rate of convergence

$$\max_{0 \leq y \leq 1} |g_k(y) - g_\infty(y)| = \xi - \xi_k \quad \text{for all } k \geq 1. \tag{33}$$

Proof. We first fix m and $y \in [0, \xi_m]$. We then recall that $y \leq \xi_m \leq \frac{1}{3}$ implies that $g_j(y) \leq \frac{1}{3}$ for all $j \geq 1$. Now, given $k \geq m$, we can repeatedly apply the top line of (24) to obtain

$$|v_k'(y)| = |v_{m-1}'(y)| \left(\prod_{j=m}^{k} g_j(y) \right) \leq 3^{m-k} |v_{m-1}'(y)| \quad \text{for } y \in [0, \xi_m],$$

and by (25) we have $|v_{m-1}'(y)| \leq 1$ for all $y \in [0, \frac{1}{3}]$, so (31) gives us more simply

$$\max_{0 \leq y \leq \xi_m} |g_k'(y)| \leq 3^{m-k} \quad \text{for all } k \geq m.$$

Now, for any y, z in $[0, \xi_m]$ we have $|g_k(y) - g_k(z)| \le 3^{m-k}|y - z|$, so, letting $k \to \infty$, g_∞ is constant on $[0, \xi_m]$ for each $m \ge 1$. Since $\xi_m \uparrow \xi$, there is a constant c such that $g_\infty(y) = c$ for all $y \in [0, \xi)$.

As Figure 1 suggests $c = \xi$, and this is easy to confirm. Again we fix m, take $k \ge m$, and note that by the triangle inequality and the Lipschitz bound (29) on g_k we have

$$|g_\infty(\xi_m) - \xi_k| \le |g_\infty(\xi_m) - g_k(\xi_m)| + |g_k(\xi_m) - g_k(\xi_k)|$$
$$\le |g_\infty(\xi_m) - g_k(\xi_m)| + |\xi_m - \xi_k|.$$

When $k \to \infty$, $g_k(\xi_m)$ converges to $g_\infty(\xi_m)$ and ξ_k to ξ, so we have

$$|g_\infty(\xi_m) - \xi| \le |\xi_m - \xi|.$$

Since $g_\infty(\xi_m) = c$ does not depend on m and since $|\xi_m - \xi| \to 0$ as $m \to \infty$, we see that $g_\infty(\xi_m) = \xi$ for all $m \ge 1$; consequently, $g_\infty(y) = \xi$ for all $y \in [0, \xi]$. Finally, for all $m \ge 1$, we also have $g_m(y) = y$ for each $y \in [\xi, 1]$, so the proof of the formula for g_∞ is complete.

To prove (33), we first note that

$$g_\infty(y) - g_k(y) = \begin{cases} \xi - g_k(y), & y \in [0, \xi_k], \\ \xi - y, & y \in [\xi_k, \xi], \\ 0, & y \in [\xi, 1]. \end{cases}$$

By (30), $g_k(y)$ is strictly decreasing on $[0, \xi_k]$, so the gap $g_\infty(y) - g_k(y)$ is maximized when $y = \xi_k$. This gap decreases linearly over the interval $[\xi_k, \xi]$ and equals 0 at ξ; consequently, the maximal gap is exactly equal to $\xi - \xi_k$.

7. The central limit theorem for $A_n^o(\pi_\infty)$ is easy

We now recall that ξ denotes the limit (32) of the minimal fixed points, and we define a selection policy π_∞ for all $X_1, X_2, \ldots$ by taking the (time independent) threshold function to be

$$g_\infty(y) = \max\{\xi, y\} \equiv \xi \vee y.$$

If $A_n^o(\pi_\infty)$ counts the number of selections made by policy π_∞ up to and including time n, then we have the explicit formula

$$A_n^o(\pi_\infty) = \sum_{i=1}^n \mathbf{1}(X_i \ge \xi \vee Y'_{i-1}), \tag{34}$$

where we set $Y'_0 = 0$ and we define Y'_i for $i \ge 1$ recursively by

$$Y'_i = \begin{cases} Y'_{i-1} & \text{if } X_i < \xi \vee Y'_{i-1}, \\ 1 - X_i & \text{if } X_i \ge \xi \vee Y'_{i-1}. \end{cases} \tag{35}$$

Given the facts that have been accumulated, it turns out to be a reasonably easy task to prove a central limit theorem for $A_n^o(\pi_\infty)$. We just need to make the right connection to the known central limit theorems for Markov additive processes.

To make this connection explicit, we first recall that, at any given time $1 \le i \le n$, the decision maker knows the state of the selection process Y'_{i-1} prior to time i, and the decision

maker also knows the value X_i of the observation currently under consideration for selection. The bivariate random sequence

$$\{Z_i = (X_i, Y'_{i-1}): i = 1, 2, 3, \dots\}$$

then represents the state of knowledge immediately prior to the decision to accept or to reject X_i, and this sequence may be viewed as a Markov chain on the two-dimensional state space $\mathscr{S} \equiv [0, 1] \times [0, 1 - \xi]$. The Markov chain $\{Z_i : i = 1, 2, 3, \dots\}$ evolves over time according to a point-to-set transition kernel that specifies the probability of moving from an arbitrary state $(x, y) \in \mathscr{S}$ into a Borel set $C \subseteq \mathscr{S}$ in one unit of time. If we denote the transition kernel by $k((x, y), C)$, then we have the explicit formula

$$K((x, y), C) = \mathbb{P}((X_{i+1}, Y'_i) \in C \mid X_i = x, Y'_{i-1} = y)$$
$$= \int_0^1 [\mathbf{1}\{(u, 1 - x) \in C\}\mathbf{1}(x \geq \xi \vee y) + \mathbf{1}\{(u, y) \in C\}\mathbf{1}(x < \xi \vee y)]\,\mathrm{d}u,$$

where the first summand of the integrand governs the transition when X_i is chosen and the second summand governs the transition when X_i is rejected. Given this explicit formula, it is straightforward (but admittedly a little tedious) to check that a stationary probability measure for the kernel K is given by the uniform distribution γ on $\mathscr{S} = [0, 1] \times [0, 1 - \xi]$. We will confirm shortly that γ is also the unique stationary distribution.

To more deeply understand the chain Z_i, $i = 1, 2, \dots$, we now consider the double chain $(Z_i, \bar{Z}_i)$, $i = 1, 2, \dots$, where $Z_1 = (x, y)$ is an arbitrary point of $\mathscr{S}$ and $\bar{Z}_1$ has the uniform distribution on $\mathscr{S}$. For $i = 1, 2, \dots$, the chains $\{Z_i = (X_i, Y'_{i-1})\}$ and $\{\bar{Z}_i = (X_i, \bar{Y}'_{i-1})\}$ share the same independent uniform sequence X_i, $i = 1, 2, \dots$, as their first coordinate, while their second coordinates Y'_{i-1} and $\bar{Y}'_{i-1}$ are both determined by the recursion (35). Typically these coordinates differ because of their differing initial values, but we will check that they do not differ for long.

To make this precise, we set $\nu = \min\{i \geq 1: X_i \geq 1 - \xi\}$, and we show that ν is a *coupling time* for $(Z_i, \bar{Z}_i)$ in the sense that

$$Z_i = \bar{Z}_i \quad \text{for all } i > \nu.$$

Since Y'_i and $\bar{Y}'_i$ both satisfy the recursion (35), we have

$$Y'_i \leq 1 - \xi \quad \text{and} \quad \bar{Y}'_i \leq 1 - \xi \qquad \text{for all } i = 1, 2, \dots,$$

so by the definition of ν, we must have

$$\max\{\xi \vee Y'_{\nu-1}, \xi \vee \bar{Y}'_{\nu-1}\} \leq X_\nu.$$

Recursion (35) then gives us

$$Y'_\nu = \bar{Y}'_\nu = 1 - X_\nu \quad \text{and} \quad Z_\nu = \bar{Z}_\nu.$$

By the construction of the double process $(Z_i, \bar{Z}_i)$, if we have $Z_i(\omega) = \bar{Z}_i(\omega)$ for some $i = i(\omega)$, then $Z_j(\omega) = \bar{Z}_j(\omega)$ for all $j \geq i(\omega)$, so ν is indeed a coupling time for $(Z_i, \bar{Z}_i)$.

The coupling inequality (see e.g. Lindvall (2002, p. 12)) then tells us that, for all Borel sets $C \subseteq \mathscr{S}$, we have the total variation bound

$$\|K^\ell((x, y), C) - \gamma(C)\|_{\mathrm{TV}} \leq \mathbb{P}(\nu > \ell) = (1 - \xi)^\ell, \tag{36}$$

where γ is the uniform stationary distribution on $\mathcal{S}$. The bound (36) has several useful implications. First, it implies that γ is the *unique* stationary distribution for the chain with kernel K. It also implies (see e.g. Meyn and Tweedie (2009, Theorem 16.0.1)) that the chain $\{Z_i : i = 1, 2, \dots\}$ is uniformly ergodic; more specifically, it is a ϕ-mixing chain with

$$\phi(\ell) \leq 2\rho^\ell \quad \text{and} \quad \rho = 1 - \xi.$$

If we set $z = (x, y)$ and $f(z) = \mathbf{1}(x \geq y \vee \xi)$, then representation (34) can also be written in terms of the chain $\{Z_i : i = 1, 2, \dots\}$ as

$$A_n^o(\pi_\infty) = \sum_{i=1}^{n} f(Z_i),$$

and this makes it explicit that $A_n^o(\pi_\infty)$ is a Markov additive process. Our coupling and the uniform ergodicity of $\{Z_i : i = 1, 2, \dots\}$ imply (see e.g. Meyn and Tweedie (2009, Theorem 17.5.3 and Lemma 17.5.1)) that there is a constant $\sigma^2 \geq 0$ such that

$$\lim_{n \to \infty} n^{-1} \operatorname{var}(A_n^o(\pi_\infty)) = \lim_{n \to \infty} n^{-1} \operatorname{var}_\gamma(A_n^o(\pi_\infty)) = \sigma^2, \tag{37}$$

where the first variance refers to the chain started at $Z_1 = (X_1, 0)$ and the second variance refers to the chain started at Z_1 with the stationary distribution γ (i.e. the uniform distribution on $\mathcal{S}$). The general theory also provides the series representation for the limit (37), i.e.

$$\sigma^2 = \operatorname{var}_\gamma[\mathbf{1}(X_1 \geq \{\xi \vee Y_0'\})]$$
$$+ 2 \sum_{i=2}^{\infty} \operatorname{cov}_\gamma[\mathbf{1}(X_1 \geq \{\xi \vee Y_0'\}), \mathbf{1}(X_i \geq \{\xi \vee Y_{i-1}'\})], \tag{38}$$

where the subscript γ again refers to the situation in which the chain starts with Z_1 having the stationary distribution.

The general representations (37) and (38) give us the existence of σ^2 but they do not automatically entail $\sigma^2 > 0$, so to prove a central limit theorem for $A_n^o(\pi_\infty)$ with the classical normalization, we must independently establish that $\sigma^2 > 0$. To show this, we first need an elementary lemma that provides a variance analog to the information processing inequality for entropy.

Lemma 6. (Information processing lemma.) *If a random variable X has values in $\{1, 2, \dots\}$ and $\mathbb{P}(X = 1) = p$, then $p(1 - p) \leq \operatorname{var}(X)$.*

Proof. Define a function f on the natural numbers $\mathbb{N}$ by setting $f(1) = 0$ and $f(k) = 1$ for $k > 1$. We then have $|f(x) - f(y)| \leq |x - y|$ for all $x, y \in \mathbb{N}$. If we take Y to be an independent copy of X, then we have

$$2p(1 - p) = \mathbb{E}[(f(X) - f(Y))^2] \leq \mathbb{E}[(X - Y)^2] = 2 \operatorname{var}(X).$$

Now we can state the main lemma of this section.

Lemma 7. *There are constants $\alpha > 0$ and $N_* < \infty$ such that*

$$\alpha n \leq \operatorname{var}(A_n^o(\pi_\infty)) \quad \text{for all } n \geq N_*.$$

Proof. We first set $v_0 \equiv 0$ and then define the stopping times

$$v_t = \inf\{i > v_{t-1} : X_i \geq 1 - \xi\}, \qquad t = 1, 2, \dots .$$

We also set $T(n) = \inf\{t : v_t \geq n\}$, and note that $T(n)$ is a stopping time with respect to the increasing sequence of σ-fields

$$\mathcal{G}_t = \sigma\{v_1, v_2, \dots, v_t\} \quad \text{for all } t \geq 1.$$

Next, we set

$$U_t = \sum_{i=v_{t-1}+1}^{v_t} \mathbf{1}(X_i \geq \xi \vee Y'_{i-1}) \quad \text{for } 1 \leq t \leq T(n), \tag{39}$$

and

$$V = \sum_{i=n+1}^{v_{T(n)}} \mathbf{1}(X_i \geq \xi \vee Y'_{i-1}),$$

so we have the representation

$$A_n^o(\pi_\infty) = A_{v_{T(n)}}^o(\pi_\infty) - V = \sum_{t=1}^{T(n)} U_t - V.$$

Here, the random variables U_t, $t = 1, 2, \dots$, are independent and identically distributed. We also have $V \leq v_{T(n)} - n$ and $v_{T(n)} = \inf\{i \geq n : X_i \geq 1 - \xi\}$, so the variance of V is bounded by a constant that depends only on ξ. The existence of the limit (37) and the Cauchy–Schwarz inequality then give us

$$\text{var}(A_n^o(\pi_\infty)) = \text{var}(A_{v_{T(n)}}^o(\pi_\infty)) + O(\sqrt{n}) \quad \text{as } n \to \infty, \tag{40}$$

so to prove the lemma it suffices to obtain a linear lower bound for $\text{var}(A_{v_{T(n)}}^o(\pi_\infty))$.

By the definition of $v_{T(n)}$ and U_t, $t = 1, 2, \dots$, we have

$$A_{v_{T(n)}}^o(\pi_\infty) = \sum_{t=1}^{T(n)} U_t,$$

so, by the conditional variance formula, the independence of the U_t, and the fact that $T(n)$ is $\mathcal{G}_{T(n)}$-measurable, we have the bound

$$\text{var}\left(\sum_{t=1}^{T(n)} U_t\right) \geq \mathbb{E}\left[\text{var}\left(\sum_{t=1}^{T(n)} U_t \,\Big|\, \mathcal{G}_{T(n)}\right)\right] = \mathbb{E}\left[\sum_{t=1}^{T(n)} \text{var}(U_t \mid \mathcal{G}_{T(n)})\right]. \tag{41}$$

We now note from (39) that U_t takes values in $\{1, 2, \dots, v_t - v_{t-1}\}$. Thus, if p is the probability that no X_i is selected for $i \in \{v_{t-1} + 1, \dots, v_t - 1\}$, then, setting $a = (1 - \xi)^{-1}\xi$, we have

$$p = \mathbb{P}(U_t = 1 \mid \mathcal{G}_{T(n)}) = \mathbb{P}(X_i < \xi \text{ for all } v_{t-1} + 1 \leq i \leq v_t - 1 \mid \mathcal{G}_{T(n)}) = a^{v_t - v_{t-1} - 1}.$$

Now, by applying Lemma 6 to the conditional expectation, we have

$$\text{var}(U_t \mid \mathcal{G}_{T(n)}) \geq a^{v_t - v_{t-1} - 1}(1 - a^{v_t - v_{t-1} - 1}),$$

so from (41) we have

$$\mathrm{var}\left(\sum_{t=1}^{T(n)} U_t\right) \geq \mathbb{E}\left[\sum_{t=1}^{T(n)} a^{\nu_t - \nu_{t-1} - 1}(1 - a^{\nu_t - \nu_{t-1} - 1})\right].$$

The summands are independent and identically distributed and $T(n)$ is a stopping time with respect to the increasing sequence of σ-fields $\mathcal{G}_t = \sigma\{\nu_1, \nu_2, \ldots, \nu_t\}$, $t \geq 1$, so by Wald's identity we have

$$\mathrm{var}\left(\sum_{t=1}^{T(n)} U_t\right) \geq \mathbb{E}[T(n)]\mathbb{E}[a^{\nu_1 - 1}(1 - a^{\nu_1 - 1})]. \tag{42}$$

For the stopping time $T(n)$, we have the alternative representation

$$T(n) = 1 + \sum_{i=1}^{n-1} \mathbf{1}(X_i \geq 1 - \xi),$$

so we have $\mathbb{E}[T(n)] = \xi n + O(1)$. Since ν_1 has the geometric distribution with success probability ξ, we also have $\mathbb{E}[a^{\nu_1 - 1}(1 - a^{\nu_1 - 1})] > 0$, so by (40) and (42) the proof of the lemma is complete.

All of the pieces are now in place. By the central limit theorem for functions of uniformly ergodic Markov chains (see Meyn and Tweedie (2009, Theorem 17.5.3) or Jones (2004, Corollary 5)), we get our central limit theorem for $A_n^o(\pi_\infty)$.

Proposition 3. (Central limit theorem for $A_n^o(\pi_\infty)$.) As $n \to \infty$, we have the limit

$$\frac{A_n^o(\pi_\infty) - \mu n}{\sqrt{n}} \xrightarrow{\mathrm{D}} N(0, \sigma^2),$$

where $\mu = \mathbb{E}_\gamma[\mathbf{1}(X_1 \geq \{\xi \vee Y_0'\})]$, γ is the stationary distribution for the Markov chain $\{Z_i : i = 1, 2, \ldots\}$, and σ^2 is the constant defined by either the limits (37) or the sum (38).

By appealing to the known relation (1) that $\mathbb{E}[A_n^o(\pi_n^*)] = (2 - \sqrt{2})n + O(1)$, we can show with a bit of calculation that here we have $\mu = 2 - \sqrt{2}$. Since this identification is implicit in the calculations of the next section, there is no reason to belabor it here.

8. $A_n^o(\pi_n^*)$ and $A_n^o(\pi_\infty)$ are close in L^2

Proposition 3 tells us that the easy sum $A_n^o(\pi_\infty)$ obeys a central limit theorem, and now the task is to show that the harder sum $A_n^o(\pi_n^*)$ follows the same law. The essence is to show that, after centering, the random variables $A_n^o(\pi_n^*)$ and $A_n^o(\pi_\infty)$ are close in L^2 in the sense that $\|A_n^o(\pi_n^*) - A_n^o(\pi_\infty) - \mathbb{E}[A_n^o(\pi_n^*) - A_n^o(\pi_\infty)]\|_2 = o(\sqrt{n})$ as $n \to \infty$. For technical convenience, we work with the random variable

$$\Delta_n := A_{n-2}^o(\pi_n^*) - \mathbb{E}[A_{n-2}^o(\pi_n^*)] - A_{n-2}^o(\pi_\infty) + \mathbb{E}[A_{n-2}^o(\pi_\infty)].$$

The essential estimate of our development is given by the following lemma. In one way or another, the proof of the lemma calls on all of the machinery that has been developed.

Lemma 8. (L^2-estimate.) *There is a constant C such that, for all $n \geq 3$, we have*

$$\|\Delta_n\|_2^2 \leq C \sum_{k=3}^{n} (\xi - \xi_k);$$

so, in particular, we have the asymptotic estimate

$$\|\Delta_n\|_2 = o(\sqrt{n}) \quad \text{as } n \to \infty.$$

Proof. We first note that the threshold function lower bound (22) implies that $Y_i \leq \frac{5}{6}$ for all $1 \leq i \leq n-2$. Consequently, if $X_i \geq \frac{5}{6}$ then X_i is selected by both of the policies π_n^* and π_∞. At such a time i, we have a kind of 'renewal event', though we still have to be attentive to the nonhomogeneity of the selection process driven by π_n^*.

To formalize this notion, we set $\tau_0 = 0$ and, for $m \geq 1$, we define stopping times

$$\tau_m = \inf\{i > \tau_{m-1} : X_i \geq \tfrac{5}{6}\} \quad \text{and} \quad \tau_m' = \min\{\tau_m, n-2\};$$

so τ_m is the time at which the mth 'renewal' is observed. For each $1 \leq j \leq n-2$, we then set

$$N(j) = \sum_{i=1}^{j} \mathbf{1}\big(X_i \geq \tfrac{5}{6}\big),$$

so the time $\tau_{N(j)}$ is the time of the last renewal up to or equal to j, the time $\tau_{N(j)+1}$ is the time of the first renewal strictly after j, and we have the inclusion

$$\tau_{N(j)} \leq j < \tau_{N(j)+1}.$$

For $1 \leq j \leq n-2$, we then consider the martingale differences defined by

$$d_j = \mathbb{E}[A_{n-2}^o(\pi_n^*) - A_{n-2}^o(\pi_\infty) \mid \mathcal{F}_j] - \mathbb{E}[A_{n-2}^o(\pi_n^*) - A_{n-2}^o(\pi_\infty) \mid \mathcal{F}_{j-1}],$$

where $\mathcal{F}_0$ is the trivial σ-field and $\mathcal{F}_j = \sigma\{X_1, X_2, \ldots, X_j\}$ for $1 \leq j \leq n$. Using the counting variables

$$\eta_i \equiv \mathbf{1}(X_i \geq g_{n-i+1}(Y_{i-1})) \quad \text{and} \quad \eta_i' \equiv \mathbf{1}(X_i \geq \xi \vee Y_{i-1}'),$$

we have the tautology

$$d_j = \mathbb{E}\left[\sum_{i=j}^{\tau_{N(j)+1}'} (\eta_i - \eta_i') \,\bigg|\, \mathcal{F}_j \right] - \mathbb{E}\left[\sum_{i=j}^{\tau_{N(j)+1}'} (\eta_i - \eta_i') \,\bigg|\, \mathcal{F}_{j-1} \right] \tag{43}$$

$$+ \mathbb{E}\left[\sum_{i=\tau_{N(j)+1}'+1}^{n-2} (\eta_i - \eta_i') \,\bigg|\, \mathcal{F}_j \right] - \mathbb{E}\left[\sum_{i=\tau_{N(j)+1}'+1}^{n-2} (\eta_i - \eta_i') \,\bigg|\, \mathcal{F}_{j-1} \right],$$

and this becomes more interesting after we check that the last two terms cancel.

To confirm the cancelation, we first recall that, for $\tau_{N(j)+1} < n-2$, the value $X_{\tau_{N(j)+1}} \geq \frac{5}{6}$ is selected as a member of the alternating subsequence under both policies π_n^* and π_∞, so we also have

$$Y_{\tau_{N(j)+1}} = Y_{\tau_{N(j)+1}}' = 1 - X_{\tau_{N(j)+1}}.$$

Any difference in the selections that are made by the policies π_n^* and π_∞ after time $\tau_{N(j)+1}$ is measurable with respect to the σ-field

$$\mathcal{T}_j = \sigma\{X_{\tau_{N(j)+1}}, X_{\tau_{N(j)+1}+1}, \ldots, X_{n-2}\}.$$

Trivially, we have $j < \tau_{N(j)+1}$, so $\mathcal{F}_j$ is independent of $\mathcal{T}_j$, and the last two addends in (43) do cancel as claimed.

We can therefore write

$$d_j = \mathbb{E}\left[\sum_{i=j}^{\tau'_{N(j)+1}} (\eta_i - \eta'_i) \,\Big|\, \mathcal{F}_j \right] - \mathbb{E}\left[\sum_{i=j}^{\tau'_{N(j)+1}} (\eta_i - \eta'_i) \,\Big|\, \mathcal{F}_{j-1} \right] = W_j - I_{j-1}(W_j),$$

where W_j denotes the first summand and I_{j-1} is the projection onto $L^2(\mathcal{F}_{j-1})$. Denoting the identity by I, $I - I_{j-1}$ is an L^2 contraction, so

$$\mathbb{E}[d_j^2] \leq \mathbb{E}[W_j^2] = \mathbb{E}\left[\left(\sum_{i=j}^{\tau'_{N(j)+1}} (\eta_i - \eta'_i) \right)^2 \right], \tag{44}$$

and the remaining task is to estimate the last right-hand side.

For $1 \leq j \leq n-2$, we let $L(j)$ denote time from j since the last renewal preceding j; in other words, $L(j)$ is the *age* at time j. Analogously, we let $M(j)$ denote the time from j until the time of the next renewal or until time $n-2$; so $M(j)$ is the *residual life* at time j with truncation at time $n-2$. We then have

$$L(j) = j - \tau_{N(j)} \quad \text{and} \quad M(j) = \tau'_{N(j)+1} - j.$$

Our interarrival times are geometric, so $L(j)$ and $M(j)$ are independent, and for $p = \frac{1}{6}$ we have

$$\mathbb{P}(L(j) = \ell) = \begin{cases} p(1-p)^\ell & \text{if } 0 \leq \ell < j, \\ (1-p)^j & \text{if } \ell = j, \end{cases}$$

and

$$\mathbb{P}(M(j) = m) = \begin{cases} p(1-p)^{m-1} & \text{if } 1 \leq m < n-2-j, \\ (1-p)^{n-3-j} & \text{if } m = n-2-j. \end{cases}$$

We now introduce the *disagreement* set

$$D_j[\ell, m] = \{\omega : \text{there exists } i \in \{j-\ell+1, \ldots, j, \ldots, j+m\} : X_i(\omega) \in [\xi_{n-i+1}, \xi]\};$$

this is precisely the set of ω for which, if $Y_{j-\ell} = Y'_{j-\ell}$, then the policies π_∞ and π_n^* differ in at least one selection during the time interval $\{j-\ell+1, \ldots, j+m\}$, while on the complementary set $D_j^c[\ell, m]$ the selections all agree. Thus, by the crudest possible bound, we have

$$\left| \sum_{i=j}^{\tau'_{N(j)+1}} (\eta_i - \eta'_i) \right| \leq (L(j) + M(j))\mathbf{1}(D_j[L(j), M(j)]),$$

and when we square both sides and rearrange, we obtain

$$\left(\sum_{i=j}^{\tau'_{N(j)+1}} (\eta_i - \eta_i')\right)^2 \le (L(j) + M(j))^2 \mathbf{1}(D_j[L(j), M(j)])$$

$$= \sum_{m=1}^{n-2-j} \sum_{\ell=0}^{j} (\ell + m)^2 \mathbf{1}(D_j[\ell, m])\mathbf{1}(L(j) = \ell)\mathbf{1}(M(j) = m). \quad (45)$$

For each $1 \le j \le n - 2$, we now set

$$R_j[\ell, m] = \big\{\omega\colon X_i(\omega) < \tfrac{5}{6} \text{ for all } i \in \{j - \ell + 1, \ldots, j + m\}\big\},$$

so $R_j[\ell, m]$ is the event that no renewal takes place in $[j - \ell + 1, j]$ or in $[j + 1, j + m]$. By the definition of $L(j)$ and $M(j)$, we then have

$$\mathbf{1}(L(j) = \ell) = \mathbf{1}(R_j[\ell, 0])\mathbf{1}\big(X_{j-\ell} \ge \tfrac{5}{6} \text{ or } \ell = j\big) \quad \text{for } 0 \le \ell \le j,$$

and

$$\mathbf{1}(M(j) = m) \le \mathbf{1}(R_j[0, m - 1]) \quad \text{for } 1 \le m \le n - 2 - j.$$

Thus, if we define $\mathbf{1}(R_j[0, 0]) \equiv 1$, then we have the composite bound

$$\mathbf{1}(L(j) = \ell)\mathbf{1}(M(j) = m) \le \mathbf{1}(R_j[\ell, m - 1])\mathbf{1}\big(X_{j-\ell} \ge \tfrac{5}{6} \text{ or } \ell = j\big), \quad (46)$$

so by inserting (46) into (45) and recalling (44), we find

$$\mathbb{E}[d_j^2] \le \sum_{m=1}^{n-2-j} \sum_{\ell=0}^{j} (\ell + m)^2 \mathbb{E}\big[\mathbf{1}(D_j[\ell, m])\mathbf{1}(R_j[\ell, m - 1])\mathbf{1}\big(X_{j-\ell} \ge \tfrac{5}{6} \text{ or } \ell = j\big)\big]. \quad (47)$$

The expected value on the right-hand side of (47) accounts for the probability that policies π_n^* and π_∞ differ when one renewal has occurred at time $j - \ell$, and no renewal will occur until time $j + m$. For this to happen, we need at least one $i \in \{j - \ell + 1, \ldots, j + m\}$ such that $X_i \in [\xi_{n-i+1}, \xi]$. Since the X_is are uniformly distributed on $[0, 1]$, the probability that $X_i \in [\xi_{n-i+1}, \xi]$ equals $\xi - \xi_{n-i+1}$ and, by the monotonicity of the minimal fixed points in Lemma 3, we have the upper bound $\xi - \xi_{n-i+1} \le \xi - \xi_{n-(j+m)+1}$ for all $i \in \{j - \ell + 1, \ldots, j + m\}$. Then, we can estimate the right-hand side of (47) with Boole's inequality, and obtain a constant C such that

$$\mathbb{E}\big[\mathbf{1}(D_j[\ell, m])\mathbf{1}(R_j[\ell, m - 1])\mathbf{1}\big(X_{j-\ell} \ge \tfrac{5}{6} \text{ or } \ell = j\big)\big]$$
$$\le C(m - \ell)(\xi - \xi_{n-(j+m)+1})(1 - p)^{\ell+m-1}.$$

At this point, $C = \tfrac{6}{5}$ would suffice, but subsequently C denotes a Hardy-style constant that may change from line to line. If we use this last bound in (47), we obtain

$$\mathbb{E}[d_j^2] \le C \sum_{m=1}^{n-2-j} \sum_{\ell=0}^{j} (\ell + m)^3 (\xi - \xi_{n-(j+m)+1})(1 - p)^{\ell+m-1},$$

so, if we change variable by applying the transformation $r = j + m$, we have

$$\mathbb{E}[d_j^2] \le C \sum_{r=j+1}^{n-2} \sum_{\ell=0}^{j} (\ell + r - j)^3 (\xi - \xi_{n-r+1})(1 - p)^{\ell+r-j-1}.$$

If we now sum over $1 \leq j \leq n-2$, we obtain

$$\mathbb{E}[\Delta_n^2] = \sum_{j=1}^{n-2} \mathbb{E}[d_j^2] \leq C \sum_{j=1}^{n-2} \sum_{r=j+1}^{n-2} \sum_{\ell=0}^{j} (\ell + r - j)^3 (\xi - \xi_{n-r+1})(1 - p)^{\ell + r - j - 1},$$

so if we interchange the first with the second sum and rearrange, we have

$$\mathbb{E}[\Delta_n^2] \leq C \sum_{r=2}^{n-2} (\xi - \xi_{n-r+1}) \left\{ \sum_{j=1}^{r-1} \sum_{\ell=0}^{j} (\ell + r - j)^3 (1 - p)^{\ell + r - j - 1} \right\}.$$

At this point, it is elementary to check that for all r the last double sum is bounded by the constant $\sum_{u=1}^{\infty} u^4 (1 - p)^{u-1}$, and this completes the proof of our lemma.

9. Some perspective

We have pursued the proof of a specific central limit theorem, but some aspects of our analysis may have useful implications for a wider class of Markov decision problems (MDPs). For example, we took advantage here of the existence of a policy π_∞ that could be viewed heuristically as the 'optimal policy at infinity', and the temporal homogeneity of this policy then gave us access to the machinery of Markov additive processes. Many MDPs offer similar prospects.

To be sure, specialized efforts were needed to relate the finite horizon policy π_n^* to the limiting policy, but the pattern used here does offer some general guidance. In almost any MDP, the Bellman equation gives us good prospects for computing the value function, but to extract the full value of those functions we need to develop a deeper understanding of their geometry—and the geometry of the associated threshold functions. Here, the development of such an understanding would have been stymied without the guidance provided by Figure 1. If we view our analysis as a case study, then one message is that when facing a new MDP we would almost always be wise to begin with the best numerical work that the problem allows.

Finally, the Bellman equation grants a natural place for induction in the analysis of many MDPs, and here we have seen that such inductions can be greatly helped by various forms of diminishing returns. Without the special properties represented by (7) and (16) our inductions could not have moved forward. We anticipate that some aspect of this experience will be present in the analysis of many other MDPs.

Acknowledgements

References

ARLOTTO, A. AND STEELE, J. M. (2011). Optimal sequential selection of a unimodal subsequence of a random sequence. *Combin. Prob. Comput.* **20,** 799–814.

ARLOTTO, A., CHEN, R. W., SHEPP, L. A. AND STEELE, J. M. (2011). Online selection of alternating subsequences from a random sample. *J. Appl. Prob.* **48,** 1114–1132.

BANNISTER, M. J. AND EPPSTEIN, D. (2012). Randomized speedup of the Bellman–Ford algorithm. In *Proc. Meeting on Analytic Algorithmics and Combinatorics*. Society for Industrial and Applied Mathematics, Philadelphia, PA, pp. 41–47.

BARYSHNIKOV, Y. M. AND GNEDIN, A. V. (2000). Sequential selection of an increasing sequence from a multidimensional random sample. *Ann. Appl. Prob.* **10,** 258–267.

BATENI, M., HAJIAGHAYI, M. AND ZADIMOGHADDAM, M. (2010). Submodular secretary problem and extensions. In *Approximation, Randomization, and Combinatorial Optimization* (Lecture Notes Comput. Sci. **6302**), Springer, Berlin, pp. 39–52.

BROCKWELL, P. J. AND DAVIS, R. A. (2006). *Time Series: Theory and Methods*. Springer, New York.

BRUSS, F. T. AND DELBAEN, F. (2001). Optimal rules for the sequential selection of monotone subsequences of maximum expected length. *Stoch. Process. Appl.* **96,** 313–342.

BRUSS, F. T. AND DELBAEN, F. (2004). A central limit theorem for the optimal selection process for monotone subsequences of maximum expected length. *Stoch. Process. Appl.* **114,** 287–311.

BUCHBINDER, N., JAIN, K. AND SINGH, M. (2010). Secretary problems via linear programming. In *Integer Programming and Combinatorial Optimization* (Lecture Notes Comput. Sci. **6080**), Springer, Berlin, pp. 163–176.

CAYLEY, A. (1875). Mathematical questions and their solutions. *Educational Times* **22,** 18–19. (Available in *The Collected Mathematical Papers of Arthur Cayley* (1986), Vol. 10, Cambridge University Press, pp. 587–588.)

DYNKIN, E. B. (1963). Optimal choice of the stopping moment of a Markov process. *Dokl. Akad. Nauk SSSR* **150,** 238–240.

GNEDIN, A. V. (1999). Sequential selection of an increasing subsequence from a sample of random size. *J. Appl. Prob.* **36,** 1074–1085.

GNEDIN, A. V. (2000a). A note on sequential selection from permutations. *Combin. Prob. Comput.* **9,** 13–17.

GNEDIN, A. V. (2000b). Sequential selection of an increasing subsequence from a random sample with geometrically distributed sample-size. In *Game Theory, Optimal Stopping, Probability and Statistics* (IMS Lecture Notes Monogr. Ser. **35**), Institute of Mathematical Statistics, Beachwood, OH, pp. 101–109.

HOUDRÉ, C. AND RESTREPO, R. (2010). A probabilistic approach to the asymptotics of the length of the longest alternating subsequence. *Electron. J. Combin.* **17,** Research Paper 168, 19pp.

HUA, X. (2010). Testing regression models with residuals as data. Doctoral Thesis, Massachusetts Institute of Technology.

JONES, G. L. (2004). On the Markov chain central limit theorem. *Prob. Surv.* **1,** 299–320.

KRIEGER, A. M. AND SAMUEL-CAHN, E. (2009). The secretary problem of minimizing the expected rank: a simple suboptimal approach with generalizations. *Adv. Appl. Prob.* **41,** 1041–1058.

LINDLEY, D. V. (1961). Dynamic programming and decision theory. *Appl. Statist.* **10,** 39–51.

LINDVALL, T. (2002). *Lectures on the Coupling Method*. Dover, Mineola, NY.

MEYN, S. AND TWEEDIE, R. L. (2009). *Markov Chains and Stochastic Stability*, 2nd edn. Cambridge University Press.

ROMIK, D. (2011). Local extrema in random permutations and the structure of longest alternating subsequences. In *23rd International Conference on Formal Power Series and Algebraic Combinatorics (FPSAC 2011)*, Association of Discrete Mathematics and Theoretical Computer Science, Nancy, pp. 825–834.

SAMUELS, S. M. (1991). Secretary problems. In *Handbook of Sequential Analysis* (Statist. Textbooks Monogr. **118**), Dekker, New York, pp. 381–405.

SAMUELS, S. M. AND STEELE, J. M. (1981). Optimal sequential selection of a monotone sequence from a random sample. *Ann. Prob.* **9,** 937–947.

STANLEY, R. P. (2007). Increasing and decreasing subsequences and their variants. In *International Congress of Mathematicians*, Vol. I, European Mathematical Society, Zürich, pp. 545–579.

STANLEY, R. P. (2008). Longest alternating subsequences of permutations. *Michigan Math. J.* **57,** 675–687.

STANLEY, R. P. (2010). A survey of alternating permutations. In *Combinatorics and Graphs* (Contemp. Math. **531**), American Mathematical Society, Providence, RI, pp. 165–196.

WIDOM, H. (2006). On the limiting distribution for the length of the longest alternating sequence in a random permutation. *Electron. J. Combin.* **13,** Research Paper 25, 7pp.