Optimal Search from Multiple Distributions with Infinite Horizon

Jean-Michel Benkert, Igor Letina and Georg Nöldeke

September 2017
Optimal Search from Multiple Distributions with Infinite Horizon

Jean-Michel Benkert\textsuperscript{a}, Igor Letina\textsuperscript{b,∗}, Georg Nöldeke\textsuperscript{c}

\textsuperscript{a}Department of Economics and UBS International Center of Economics in Society, University of Zurich, Switzerland
\textsuperscript{b}Department of Economics, University of Bern, Switzerland
\textsuperscript{c}Faculty of Business and Economics, University of Basel, Switzerland

Abstract

With infinite horizon, optimal rules for sequential search from a known distribution feature a constant reservation value that is independent of whether recall of past options is possible. We extend this result to the case when there are multiple distributions to choose from: it is optimal to sample from the same distribution in every period and to continue searching until a constant reservation value is reached.

\textit{JEL classification:} D83

\textit{Keywords:} Optimal Search, Search Intensity, Infinite Horizon, Recall

\textsuperscript{∗}Corresponding author. Igor Letina, University of Bern, Schanzenbeckstrasse 1, 3001 Bern, Switzerland. Tel.: +41 31 631 80 76. Email: igor.letina@vwi.unibe.ch

September 20, 2017
1. Introduction

In the standard optimal search problem (e.g. Lippman and McCall, 1976), the choice is to either stop searching and consume the best option available or to continue the costly search. Continuation yields a draw of a single observation from some known distribution. With infinite horizon, the optimal search rule is simple and independent of the possibility of recall: continue searching until a constant reservation value, uniquely determined by the distribution, has been reached.

A natural extension of the standard problem is to allow the searcher to also choose the search intensity. With infinite horizon, Morgan (1983, Proposition 1) shows that with no-recall it is optimal to search with constant intensity until stopping. With full-recall, Morgan (1983, Proposition 5) only establishes that the optimal intensity is weakly decreasing, leaving open the possibility that the searcher might reduce intensity after a favorable draw. In an interesting application to delegated R&D, Poblete and Spulber (2017, Lemma 3) show optimality of constant intensity with full recall while assuming existence, uniqueness and reservation value strategies. We strengthen these results and show that the optimal intensity is constant — regardless of the possibility of recall and without restricting the class of admissible search rules.

The choice of search intensity is a special case of a more general problem in which the searcher chooses a distribution from which to draw the observation. For example, the searcher could be choosing the riskiness of search. Similarly, no-recall and full-recall are special cases of a more general problem in which recall may be stochastic. We incorporate both of these generalizations in an otherwise standard search model with infinite horizon and show that it is
optimal to sample from the same distribution in every period and to continue searching until the constant reservation value associated with this distribution has been achieved.

2. Model

While search is ongoing, a searcher decides in each period \( t = 0, 1 \ldots \) whether to sample from one of \( n \) available distributions (take the action \( a_t \in A = \{1, \ldots, n\} \)) or to stop \((a_t = 0)\). Stopping yields a payoff \( x_t \in X \) in the current period, where \( x_t \) is the option the searcher has in hand at the beginning of period \( t \) and \( X \subset \mathbb{R}_+ \) is a finite set. Once search has been stopped, no further decisions can be taken and no further payoffs accrue. If the searcher samples from distribution \( a_t \), the state variable \( x_t \) transitions to \( x_{t+1} \in X \) with probability \( p(x_{t+1}|x_t, a_t) \), where \( x_{t+1} \) is the option the searcher has in hand at the beginning of period \( t + 1 \). The cost of sampling from distribution \( a \) is \( c(a) > 0 \). The per-period-discount factor is \( \delta \in (0, 1] \). The searcher’s problem is to maximize her expected discounted payoff (for every initial condition \( x_0 \in X \)) by choosing a search rule \( \mu : X \rightarrow A \cup \{0\} \) that specifies for each state whether search should be continued by sampling from distribution \( a \) (\( \mu(x) = a \)) or be stopped (\( \mu(x) = 0 \)).\(^1\)

The above model can be embedded into a Markov decision process framework (Bertsekas, 1995) by (i) appending a terminal state \( x = 0 \) that is reached with probability one whenever the stopping action \( a = 0 \) is taken and (ii) supposing that the only available action (at zero cost) in the terminal state

\(^1\)The restriction to such stationary search rules is without loss of generality; see the sources cited in the next paragraph.
is $a = 0$. With discounting ($\delta < 1$) standard results (e.g. Bertsekas, 1995, Chapter 1.2) ensure that the Bellman equations

$$v^*(x) = \max \left\{ x, \max_{a \in A} \left\{ \delta \sum_{y \in X} v^*(y)p(y|x,a) - c(a) \right\} \right\}$$

(1)

have a unique solution $v^* : X \to \mathbb{R}$ and that a search rule $\mu^*$ is optimal iff it satisfies

$$\mu^*(x) = \begin{cases} 
0 & \Rightarrow v^*(x) = x \\
a' \in A & \Rightarrow v^*(x) = \delta \sum_{y \in X} v^*(y)p(y|x,a') - c(a').
\end{cases}$$

(2)

In particular, an optimal search rule exists. The same conclusions hold without discounting ($\delta = 1$) because the Markov decision process formulation of our search problem is a special case of the stochastic shortest path problem analysed in Bertsekas and Tsitsiklis (1991).\(^2\)

We impose additional structure on the conditional probability distributions $p(\cdot|x,a)$ that accommodates the familiar cases of search with no or full recall but also allows for more general specifications.

**Assumption 1.** For all $a \in A$ there exists a probability distribution $p^n(\cdot|a)$ on $X$ such that $p(y|x,a) = p^n(y|a)$ holds for all $y > x \in X$.

\(^2\)Assumption 1 in Bertsekas and Tsitsiklis (1991) is satisfied: First, there exists an absorbing, cost- (and benefit)-free state (the terminal state 0). Second, there exists a proper stationary policy (choose the terminal action in each state). Third, improper stationary rules (policies for which there is a strictly positive probability that the terminal state is never reached) result in infinite expected cost because $c(a) > 0$ for all $a \in A$. Assumption 2 in Bertsekas and Tsitsiklis (1991) holds because $A$ is finite. Our claim then follows from Proposition 2 in Bertsekas and Tsitsiklis (1991).
If the condition \( p(y|x,a) = p^n(y|a) \) holds for all \( x \) and \( y \) in \( X \) we have a search problem without recall: the only option available to the searcher in period \( t + 1 \) after choosing \( a \) in period \( t \) is the realized draw from the distribution \( p^n(\cdot|a) \). Assumption 1 goes beyond this no-recall case by allowing for perfect or stochastic recall: In the perfect-recall case, the option \( x_t \) remains available in period \( t + 1 \), which is modeled by setting \( p(y|x,a) = p^n(y|a) \) for \( y > x \), \( p(x|x,a) = \sum_{y \leq x} p^n(y|a) \), and \( p(y|x,a) = 0 \) for \( y < x \). Stochastic-recall may be modeled by taking \( p(\cdot|x,a) \) to be convex combinations of the probability distributions describing the no-recall and the full-recall case.³

3. Optimal Search Rules

Define for each action \( a \in A \) the reservation value \( s(a) \) as the unique solution (Step 1, proof of Proposition 1) to the equation

\[
\delta \sum_{y \in X} \max\{y, s(a)\} p^n(y|a) - s(a) = c(a).
\]  

(3)

When there is only distribution \( a \) to sample from, it is well-known (DeGroot, 1970; Lippman and McCall, 1976) that both in the no-recall and the full-recall case the optimal rule is to continue searching until the current option \( x_t \) exceeds the reservation value \( s(a) \). The same kind of reservation value rule is optimal under Assumption 1 when searching from multiple distributions. Moreover, the searcher optimally samples from the same distribution

³The weights in these convex combinations can be both state- and action-dependent without invalidating Assumption 1. For instance, recall could be possible when sampling from some distributions but not when sampling from others.
Proposition 1. Let Assumption 1 hold and let $a^* \in \arg\max_s s(a)$. The unique solution to the Bellman equations (1) is given by $v^*(x) = \max\{x, s(a^*)\}$ and the rule

$$
\mu^*(x) = \begin{cases} 
0 & \text{if } x > s(a^*) \\
 a^* & \text{if } x \leq s(a^*) 
\end{cases}
$$

(4)
is optimal.

The proof of Proposition 1, which does not build on the known results for the case of sampling from one distribution but obtains these as a special case, is in the appendix.

For the no-recall case, the intuition for Proposition 1 is straightforward: As the continuation value of search is state independent, the searcher will find it optimal to sample from the same distribution whenever it is optimal to continue searching. Thus, attention can be restricted to reservation value rules that always sample from the same distribution. Among such rules the one which induces the highest reservation value is the best. What is more surprising is that the same rule remains optimal whenever Assumption 1 holds. As the proof of Proposition 1 demonstrates, this follows because it can only be better to sample from a distribution different from $a^*$ when $x_t > s(a^*)$ holds and in that case it remains optimal to stop searching.

4. Conclusion

We characterize the optimal search rule when the searcher can choose from which distribution to sample an observation. This optimal search rule
is remarkably simple: choose the same distribution in each period until the
constant reservation value has been achieved.

This result can be extended in several directions. First, under standard
regularity conditions (Bertsekas and Tsitsiklis, 1991; James and Collins,
2006) the set of distributions $A$ can be compact or countably infinite rather
than finite and the support of the probability distributions $p(\cdot|x,a)$ can be
compact as in Poblete and Spulber (2017). Second, as long as the obvious
counterpart to Assumption 1 is imposed, the transition probabilities can be
made dependent on a finite history of past actions and options rather than
just the pair $(a_t,x_t)$. For example, our result holds in a model where recall
is only possible as long as the searcher samples from the same distribution.
Third, as recall has no value in our setting, it is immediate that a costly
option to recall past values (Saito, 1998) will never be exercised.

Acknowledgments

We thank Julia Grünseis for suggestions.

Appendix: Proof of Proposition 1

**Step 1**: The function $g : A \times \mathbb{R} \to \mathbb{R}$ defined by

$$g(s,a) = \delta \sum_{y \in X} \max\{y, s\} p^n(y|a) - s$$

(5)

is (strictly) decreasing in $s$ for all $a \in A$ (if $s < \max\{X\}$). This follows from
rewriting

$$g(s,a) = \delta \sum_{y \in X} \max\{y - s, 0\} p^n(y|a) - (1 - \delta)s$$
and noting (DeGroot, 1970, Sec. 11.8) that $\sum_{y \in X} \max\{y - s, 0\} p^n(y|a)$ is (strictly) decreasing in $s$ (whenever it is strictly positive).

Observing that (3) is equivalent to

$$g(s(a), a) = c(a)$$

and that $c(a) > 0$ for all $a \in A$, it is immediate from the above monotonicity properties that (3) can have at most one solution $s(a)$. Existence of a solution follows upon observing that $g(s, a)$ is continuous, satisfies $g(\max\{X\}, a) = 0$, and $\lim_{s \to -\infty} g = \infty$.

**Step 2:** Let $v^*(x) = \max\{x, s(a^*)\}$.

Consider $x \leq s(a^*)$. We then have for all $a \in A$:

$$\delta \sum_{y \in X} v^*(y)p(y|x, a) - c(a) = \delta \sum_{y \in X} \max\{y, s(a^*)\} p(y|x, a) - c(a)$$

$$= \delta \sum_{y \in X} \max\{y, s(a^*)\} p^n(y|a) - c(a)$$

$$= g(s(a^*), a) - c(a) + s(a^*)$$

$$= g(s(a^*), a) - g(s(a), a) + s(a^*)$$

$$\leq s(a^*)$$

where the first equality is from the specification of $v^*$, the second uses Assumption 1 to infer that for $y \geq s^*(a)$ the condition $x \leq s^*(a)$ implies $p(y|x, a) = p^n(y|x)$, the third is from the definition of $g$ in (5), and the fourth is from (6). The inequality in the last line follows from Step 1 and $s(a^*) \geq s(a)$. It holds as an equality for $a = a^*$. Therefore

$$s(a^*) = \delta \sum_{y \in X} v^*(y)p(y|x, a^*) - c(a^*) = \max_{a \in A} \left\{ x, \max_{a \in A} \left\{ \delta \sum_{y \in X} v^*(y)p(y|x, a) - c(a) \right\} \right\}$$

(7)
holds for all \( x \leq s(a^*) \).

Consider \( x > s(a^*) \). We then have for all \( a \in A \):

\[
\delta \sum_{y \in X} v^*(y)p(y|x, a) - c(a) = \delta \sum_{y \in X} \max\{y, s(a^*)\}p(y|x, a) - c(a)
\]

\[
\leq \delta \sum_{y \in X} \max\{y, x\}p(y|x, a) - c(a)
\]

\[
= \delta \sum_{y \in X} \max\{y, x\}p^n(y|a) - c(a)
\]

\[
= g(x, a) - c(a) + x
\]

\[
= g(x, a) - g(s(a), a) + x
\]

\[
\leq x,
\]

where the equality in the first line again uses \( v^*(x) = \max\{x, s(a^*)\} \), the inequality in the second line uses \( x > s^*(a) \), the equality in the third line is from Assumption 1, and the remaining lines are obtained in the same way as in the case \( x \leq s(a^*) \), using that \( x > s(a^*) \) implies \( x > s(a) \) for all \( a \in A \) to obtain the final inequality. Therefore,

\[
x = \max \left\{ x, \max_{a \in A} \left\{ \delta \sum_{y \in X} v^*(y)p(y|x, a) - c(a) \right\} \right\}
\]

holds for all \( x > s(a^*) \).

Combining (7) and (8), it is immediate that \( v^*(x) = \max\{x, s(a^*)\} \) solves the Bellman equations (1) and that \( \mu^*(x) \) as defined in (4) satisfies the optimality conditions (2).

References


