1 Introduction

Let $X(t)$ be a discrete-time homogeneous Markov decision process; let μ represent its unique stationary measure. Consider the following optimization problem over a parameter domain D:

$$\begin{aligned}
&\underset{(u,b)\in D}{\arg\min}\, C(u,b) \text{ subject to } G(u,b) \le \alpha \in [0,1] \\
&\text{Where } G(u,b) \overset{\text{def}}{=} \int_\Omega G(x)\ d\mu(x,u,b) \\
&\text{and } C(u,b) \overset{\text{def}}{=} \int_\Omega C(x)\ d\mu(x,u,b)
\end{aligned}$$

The contribution of this paper is to deal with this specific case of the general problem:

1. D is a bounded domain inside $\mathbb{R}^+ \times \mathbb{Z}^+$.
2. $G(\cdot, b)$ is unimodal and continuously differentiable for any fixed b.
3. $G(u, \cdot)$ is monotonic decreasing for any fixed u.
4. $C(u,b) \approx b$.

First, the optimization procedure for the discrete case is described. Next, the problem will be expanded to include a stochastic noise for the function $G(u,b)$ and its derivative, and an algorithm is determined for the stochastic case. The convergence of this algorithm will be established by examining the ODE limit of the algorithm. The efficiency and efficacy of the stochastic optimization will be determined by use of confidence intervals. Finally, a worked case study to illustrate the application of these procedures is provided.

2 Deterministic Case

In the deterministic case, we assume that all of the function evaluations are completely accessible and have no noise. In other words, we assume both $G(u,b)$ and its partial derivative $G_u(u,b)$ can be evaluated at any $(u,b) \in D$. Since $G(u,b)$ is monotonic decreasing for a fixed u, select $b_{\min}$ be a value of b sufficiently small and $b_{\max}$ be sufficiently large where $G(u, b_{\min})$ does not satisfy the constraint for all u and $G(u_{\max}, b_{\max})$ does satisfy the constraint for some value $u_{\max}$. In this case, a numerical algorithm can be employed which utilizes binary search on the discrete domain and gradient descent on the continuous domain.

For a fixed value of b, for the right choice of stepsize sequence $\{\epsilon_{n,b}\}$, the algorithm converges to either a point (u,b) satisfying the constraint of the problem or the global minimum (over u) of the function $G(\cdot, b)$ (which may or may not satisfy the constraint).

Proof. It is already established that for the gradient descent part of the algorithm, the proper choice of stepsize sequence will lead to the algorithm converging to a local minimum of $G(u,b)$ for a fixed value of b. Since $G(\cdot, b)$ is unimodal for any fixed b, then this local

Algorithm 1 Deterministic case algorithm (bisection and gradient descent)

Input: $u = u_{\max} \geq 0, b_{\min} := 1, b_{\max} > 1$

```
while b_min ≠ b_max − 1 do                                      ▷ Bisection method on b.
    b = ⌊(b_max + b_min)/2⌋
    n ← 0
    while G_u(u, b) ≠ 0 and G(u, b) > α and u > 0 do            ▷ Gradient descent on u.
        u ← u − ϵ_{n,b} G_u(u, b)
        n ← n + 1
    end while
    if G_u(u, b) ≤ α then                                       ▷ (u, b) is a solution.
        b_max ← b
    else                                                        ▷ No solutions for this value of b.
        b_min ← b
    end if
end while
```

Output: $b_{\max}$

minimum is the global minimum (over u); it will either be at the boundary (i.e. $u = 0$) or it will be somewhere with $G_u(u, b) = 0$. □

We also have that for any fixed value of b, $\min_u G(u, b) \geq \min_u G(u, b+1)$:

Proof. Let u^* be the value u such that $G(u^*, b) = \min_u G(u, b)$. Since $G(u, \cdot)$ is monotonic decreasing, it must be that $\min_u G(u, b) = G(u^*, b) \geq G(u^*, b+1)$. However, by definition it must be that $G(u^*, b+1) \geq \min_u G(u, b+1)$; thus $\min_u G(u, b) \geq \min_u G(u, b+1)$. □

Now it remains to show that the algorithm will always terminate at the solution for any input:

Proof. For all values of b, the inner loop will find the global minimum of $G(\cdot, b)$ or a point (u, b) such that $G(u, b) \leq \alpha$. Since $\min_u G(u, b) \geq \min_u G(u, b+1)$, it must be that there is one unique value b^* where $\min_u G(u, b) > \alpha$ for all $b < b^*$ and $\min_u G(u, b) \leq \alpha$ for all $b \geq b^*$. A solution to the problem must therefore exist on the level set $(\cdot, b^*)$. Since the algorithm starts with $b_{\min}$ where $G(u, b_{\min}) > \alpha$ for all u and $b_{\max}$ where $G(u, b_{\max}) \leq \alpha$ for some u then bisection will determine for which values of b the inner loop will run on. The bisection technique will achieve the correct solution due to there being only one value b^*. Since the inner loop will find the global minimum (or a point satisfying the constraint) for any starting value of u and any value of b, then the whole algorithm will work until it finds the unique value b^*; from there, gradient search will guarantee that a point will be found satisfying the constraint (since $\min_u G(u, b^*) \leq \alpha$). □

For reference, a generalized version of the algorithm that also takes advantage of the monotonic behavior of $G(u, \cdot)$ is to change b according to some step sequence $\{\xi_k\}$. In the case of bisection search, this step sequence is exactly $\lceil b_{\max} - b_{\min} \rceil 2^{-(k+1)}$. In this generalized discrete search, we may choose a different sequence of $\{\xi_k\}$; this will prove useful in the stochastic case, where $G(u, b)$ is only estimated and thus there is no clean split of b values between ones that may satisfy the constraint and ones that cannot. With this step sequence and the same assumptions from earlier, the generalized algorithm is as follows:

Algorithm 2 Deterministic case algorithm (discrete stepping and gradient descent)

```
Input: u_1 = u_max, u_2 = u_max ≥ 0, b_min := 1, b_max > 1
  k ← 0
  while b_min ≠ b_max − 1 do                          ▷ Discrete step on both ends of b.
    b_1 = b_min + ξ_k, b_2 = b_max − ξ_k
    n ← 0
    while G_u(u_1, b_1) ≠ 0 and G(u_1, b_1) > α and u_1 > 0 do   ▷ Determine if b_1 is in a solution by gradient descent.
      u_1 ← u_1 − ε_{n,b_1} G_u(u_1, b_1)
      n ← n + 1
    end while
    if G_u(u_1, b_1) ≤ α then                          ▷ (u_1, b_1) is a solution.
      b_max ← b_1
    else                                               ▷ No solutions for this value of b_1.
      b_min ← b_1
    end if
    n ← 0
    while G_u(u_2, b_2) ≠ 0 and G(u_2, b_2) > α and u_2 > 0 do   ▷ Determine if b_2 is in a solution by gradient descent.
      u_2 ← u_2 − ε_{n,b_2} G_u(u_2, b_2)
      n ← n + 1
    end while
    if G_u(u_2, b_2) ≤ α then                          ▷ (u_2, b_2) is a solution.
      b_max ← b_2
    else                                               ▷ No solutions for this value of b_2.
      b_min ← b_2
    end if
    k ← k + 1
  end while
```

Theorem 1. *Let $\mathcal{F}$ be the set of all unimodal functions on I; let x^* represent the extrema of f. Let S_N be a search strategy using N observations; let I_{S_N} be the interval containing x^* that the search strategy can guarantee. For all $\epsilon > 0$ and $N \geq 2$ the Fibonacci search strategy S_N^* satisfies*

$$\sup_{f \in \mathcal{F}} \left| I_{S_N^*} \right| \leq \inf_{\text{all } S_N} \sup_{f \in \mathcal{F}} \left| I_{S_N} \right|$$

Moreover, $\sup_{f \in \mathcal{F}} \left| I_{S_N^*} \right| = \dfrac{1}{F_{N+1}}$.

Proof. Without loss of generality, we can consider only the class of unimodal functions with a maximum on $[0, 1]$. The proof is by induction on N: for $N = 2$ the best any strategy can do over the space of all possible unimodal functions is to reduce the search space by half.

This is because the first observation cannot provide any information about the minimum of a function by itself; thus the second observation can only provide in worst-case scenario a binary determination.
The inductive hypothesis is that for $N \leq n$ the theorem holds. Put explicitly, the inductive hypothesis is:

$$\forall N \leq n \colon \frac{1}{F_{N+1}} = \sup_{f \in \mathcal{F}} \left| I_{S_N^*} \right| \leq \inf_{\text{all } S_N} \sup_{f \in \mathcal{F}} \left| I_{S_N} \right| \tag{1}$$

Now we show that it is true for $N = n+1$ by contradiction: suppose there exists some strategy $\overline{S}$ such that after $n+1$ observations

$$\sup_{f \in \mathcal{F}} \left| I_{\overline{S}_{n+1}} \right| < \sup_{f \in \mathcal{F}} \left| I_{S_{n+1}^*} \right| \leq \frac{1}{F_{n+2}} |I| \tag{2}$$

The strategy $\overline{S}_{n+1}$ will end in an interval of length a, $[b, b+a]$, with $1-(b+a) = c$; $a > 0$. Similarly the Fibonacci search S_{n+1}^* ends in an interval of length e, with $[d, d+e]$ with $1-(d+e) = 1 - F_{n+1}/F_{n+2} = F_n/F_{n+2} = d$. The following lemma must hold between the intervals: the interval from $\overline{S}_{n+1}$ must be contained within the interval from S_{n+1}^*:

Lemma 1. *$b \geq d$, $b + a \leq d + e$*

Proof. This fact will be shown by contradiction as well. The proof for both end points will be roughly similar: if $\overline{S}_{n+1}$ is not contained inside the interval from S_{n+1}^*, then there is an *even better* procedure S' that will outperform S^* using only n observations. The following will be for the right endpoint; the proof for the left endpoint will be similar.

Suppose not:

$$a + b > d + e \tag{3}$$

Thus a search procedure S' can be constructed with the interval of uncertainty $I_{S_n'}$ for any function $f \in \mathcal{F}$ from S' after n observations satisfying the property:

$$\sup_{f \in \mathcal{F}} \left| I_{S_n'} \right| < \sup_{f \in \mathcal{F}} \left| I_{S_n^*} \right|$$

contradicting the inductive hypothesis.
The construction is as follows: for any $f \in \mathcal{F}$, define g as follows:

$$g(x) = \begin{cases} \exp\left(f\left(\dfrac{x}{a+b}\right)\right) & \text{if } 0 \leq x < a+b, \\ -x & \text{if } a+b \leq x \leq 1 \end{cases}$$

First, it must be shown $g \in \mathcal{F}$. g is well-defined over $[0,1]$, and the derivative of g can be directly computed:

$$g'(x) = \begin{cases} \exp\left(f\left(\dfrac{x}{a+b}\right)\right) \cdot f'\left(\dfrac{x}{a+b}\right) \cdot \dfrac{1}{a+b} & \text{if } 0 \leq x < a+b, \\ -1 & \text{if } a+b \leq x \leq 1 \end{cases}$$

Thus $g'(x) = 0$ only when $f'\left(\frac{x}{a+b}\right) = 0$; since we have that f only has one extrema, x^f, it must be the case that g only has one extrema, x^g with $x^f = \frac{x^g}{a+b}$. The behavior of the sign of g' can be determined in two parts:

- $0 \leq x < x^g$: From the definition of g', it is seen that $g'(x) > 0$ only when $f'\left(\frac{x}{a+b}\right) > 0$. Since f has the maximum at x^f it must be that $f'\left(\frac{x}{a+b}\right) > 0$ for all $0 \leq \frac{x}{a+b} < x^f = \frac{x^g}{a+b}$; in other words, $g'(x) > 0$ for $0 \leq x < x^g$.

- $x^g < x \leq 1$: Similarly, $g'(x) < 0$ only when $f'\left(\frac{x}{a+b}\right) < 0$. This means $\frac{x^g}{a+b} = x^f < \frac{x}{a+b} \leq 1$. In other words, $g'(x) < 0$ for $x^g < x \leq 1$.

So $g \in \mathcal{F}$; it bears repeating that $x^g = (a+b)x^f$.
Now, let x_k denote the evaluation points for f under the strategy S'; let y_k denote the evaluation points for g under the strategy $\overline{S}$. From the contradiction of the inductive hypothesis, $\overline{S}$ already has the observations $y_0, y_1, ..., y_{n+1}$, with interval of uncertainty $[b, b+a]$ for the function g. The strategy S' on function f is as follows:

1. The first evaluation point x_0 of f under S' is at $\frac{b}{a+b}$. Use this information to define the first observation y_0 of g under $\overline{S}$ at b. For $\overline{S}$, let the second observation y_1 be at $b+a$; by definition of g, this is just the value $-a-b$. This does not require any observation of f. Since $x^f \in [0,1]$ it must be that $x^g = (a+b)x^f \in [0, a+b]$.

   For the remaining $n-1$ observations on S', employ $\overline{S}$ on g to determine S' as follows: at step k, use the observations of g on y_{k-1} and y_{k-2} in $\overline{S}$ to define y_k. From y_k, the entries of the sequence x_k can be constructed. Two possibilities exist for y_k:

2. $y_k \geq a+b$, then S' does nothing. Observe g at y_k for the purpose of using $\overline{S}$.

3. $y_k < a+b$, then S' performs an observation of f at $\frac{y_k}{a+b}$. Note that, by construction, this is only an exponentiation step away from an observation of g at y_k.

Once $\overline{S}$ has been used $n+1$ times, it will produce an interval of uncertainty $I_{\overline{S}_{n+1}} = [s,t]$. The earlier identity $x^g = (a+b)x^f$ means that $0 \leq s \leq (a+b)x^f \leq t \leq 1$; so x^f must satisfy $\frac{s}{a+b} \leq x^f \leq \min\left(1, \frac{t}{a+b}\right)$. Thus define the interval of uncertainty for S' as $I_{S'_n} = \left[\frac{s}{a+b}, \min\left(1, \frac{t}{a+b}\right)\right]$. S' will have at most n observations required for this interval. The length of this interval is $\left|I_{S'_n}\right| = \min\left(\frac{t}{a+b} - \frac{s}{a+b}, \frac{1}{a+b} - \frac{s}{a+b}\right)$. This quantity is bounded above by $\dfrac{t-s}{a+b} = \dfrac{\left|I_{\overline{S}_{n+1}}\right|}{a+b}$; thus:

$$(a+b)\left|I_{S'_n}\right| \leq \left|I_{\overline{S}_{n+1}}\right| \tag{4}$$

Finally, by assuming $b+a > d+e = \dfrac{F_{n+1}}{F_{n+2}}$, a contradiction is achieved:

$$\begin{aligned}
\left|I_{\overline{S}_{n+1}}\right| &\leq \frac{1}{F_{n+2}} \text{ from (2).} \\
(a+b)\left|I_{S'_n}\right| &\leq \left|I_{\overline{S}_{n+1}}\right| \text{ from (4).} \\
\frac{F_{n+1}}{F_{n+2}} &= d+e < a+b \text{ from (3).} \\
(d+e)\left|I_{S'_n}\right| < (a+b)\left|I_{S'_n}\right| &\leq \left|I_{\overline{S}_{n+1}}\right| \leq \frac{1}{F_{n+2}} = \frac{d+e}{F_{n+1}} \\
\left|I_{S'_n}\right| &< \frac{1}{F_{n+1}}
\end{aligned}$$

This is a contradiction with the inductive hypothesis, reprinted below:

$$\forall N \leq n\colon \frac{1}{F_{N+1}} = \sup_{f \in \mathcal{F}} \left|I_{S^*_N}\right| \leq \inf_{\text{all } S_N} \sup_{f \in \mathcal{F}} \left|I_{S_N}\right| \text{ from (1).}$$

$\square$

Thus, it must be the case that $b+a \geq d+e$. As mentioned earlier, a similar proof can be constructed for the $b \geq d$ case, using instead for g the following function:

$$g(x) = \begin{cases} x & \text{if } 0 \leq x \leq b, \\ \exp\left(f\left(\frac{x}{b}\right)\right) & \text{if } b < x \leq 1 \end{cases}$$

Now, returning to $\overline{S}_{n+1}$: with $d \leq b$ and $a+b \leq d+e$, a similar procedure for proving the lemma will allow the construction of a strategy S' that outperforms S^* using $n-1$ observations for both of them.

Again, for any function $f \in \mathcal{F}$, define g as follows:

$$g(x) = \begin{cases} \exp\left(f\left(\frac{x}{b}\right)\right) & \text{if } 0 \leq x < b, \\ -x & \text{if } b \leq x \leq 1 \end{cases}$$

As in the previous constructions, $g \in \mathcal{F}$ and $x^g = bx^f$. Similarly to before, the strategy S' for any function f will be constructed by way of $\overline{S}$ on g; let x_k denote the evaluation points for f under S' and let y_k denote the $n+1$ evaluation points for g under $\overline{S}$.

1. The first two evaluation points y_0 and y_1 of $\overline{S}$ on g will be at b and $a+b$. The observation points x_k can be determined from y_k just as before:

2. $y_k \geq b$: S' does nothing. Observe g at y_k for using $\overline{S}$.

3. $y_k < b$: S' performs an observation of f at $\dfrac{y_k}{b}$. By construction, this will also provide an observation of g at y_k.

By this procedure S' requires at most $n-1$ evaluation points. Let $I_{\overline{S}_{n+1}} = [s,t]$; thus (in a process similar to before) $I_{S'_{n-1}} = \left[\dfrac{s}{b}, \min\left(1, \dfrac{t}{b}\right)\right]$. This quantity is bounded above by $\dfrac{t-s}{b} = \dfrac{\left|I_{\overline{S}_{n+1}}\right|}{b}$. Thus, the contradiction is achieved:

$$\begin{aligned}
b\left|I_{S'_{n-1}}\right| &\leq \left|I_{\overline{S}_{n+1}}\right| \\
\frac{F_n}{F_{n+2}} &= d \leq b \text{ from Lemma 1.} \\
\left|I_{\overline{S}_{n+1}}\right| &\leq \frac{1}{F_{n+2}} \text{ from (2).} \\
d\left|I_{S'_{n-1}}\right| \leq b\left|I_{S'_{n-1}}\right| &\leq \left|I_{\overline{S}_{n+1}}\right| < \frac{1}{F_{n+2}} \\
\left|I_{S'_{n-1}}\right| &< \frac{1}{d}\frac{1}{F_{n+2}} = \frac{F_{n+2}}{F_n}\frac{1}{F_{n+2}} = \frac{1}{F_n}
\end{aligned}$$

This contradicts the inductive hypothesis, reprinted below:

$$\forall N \leq n\colon \frac{1}{F_{N+1}} = \sup_{f\in\mathcal{F}} \left|I_{S^*_N}\right| \leq \inf_{\text{all } S_N} \sup_{f\in\mathcal{F}} \left|I_{S_N}\right| \text{ from (1).}$$

Thus it must be the case that for any fixed number of observations n the Fibonacci search strategy S^* produces the smallest interval of uncertainty. $\square$

3 Stochastic Case

We now relax the deterministic case and introduce a martingale noise term to observations of both $G(u,b)$ and $G_u(u,b)$. In this case, the only information at hand are the estimators $\widehat{G}(u,b)$ and $\widehat{G_u}(u,b)$.

Algorithm 3 Stochastic case algorithm (discrete stepping and gradient descent)

Input: $\sigma, \tau > 0, u_1 = u_{\max}, u_2 = u_{\max} \geq 0, b_{\min} := 1, b_{\max} > 1$

$k \leftarrow 0$
while $b_{\min} \neq b_{\max} - 1$ **do** $\triangleright$ Discrete step on both ends of b.
  $b_1 = b_{\min} + \xi_k,\ b_2 = b_{\max} - \xi_k$
  $n \leftarrow 0$
  while $|G_u(u_1, b_1)| > \tau$ and $|G(u_1, b_1) - \alpha| > \sigma$ and $u_1 > 0$ **do** $\triangleright$ Check b_1.
    $u_1 \leftarrow u_1 - \epsilon_{n,b_1} G_u(u_1, b_1)$
    $n \leftarrow n + 1$
  end while
  if $|G_u(u_1, b_1) - \alpha| \leq \sigma$ **then** $\triangleright$ (u_1, b_1) is a solution at this indifference level.
    $b_{\max} \leftarrow b_1$
  else $\triangleright$ No solutions at this indifference level for this value of b_1.
    $b_{\min} \leftarrow b_1$
  end if
  $n \leftarrow 0$
  while $|G_u(u_2, b_2)| > \tau$ and $|G(u_2, b_2) - \alpha| > \sigma$ and $u_2 > 0$ **do** $\triangleright$ Check b_2.
    $u_2 \leftarrow u_2 - \epsilon_{n,b_2} G_u(u_2, b_2)$
    $n \leftarrow n + 1$
  end while
  if $|G_u(u_2, b_2) - \alpha| \leq \sigma$ **then** $\triangleright$ (u_2, b_2) is a solution at this indifference level.
    $b_{\max} \leftarrow b_2$
  else $\triangleright$ No solutions at this indiffierence level for this value of b_2.
    $b_{\min} \leftarrow b_2$
  end if
  $k \leftarrow k + 1$
end while