The Ranch

Amirreza Ghanizadeh Zarandi
Alireza Ghazanfarianpour

August 2023

Introduction

In this work we will analyze the population control problem in a ranch. the problem goes as this: lets say that at the beginning of the tth year we possess a population $\mathbf{X}_t$ of some livestock, which we have kept for a year, and each of them costs us a constant amount of money $\mathbf{c}$.
now we can sell a number $\mathbf{\Delta}_t$ of them for a constant price $\mathbf{p}$ (with $\mathbf{p} \geq \mathbf{c}$) so we will earn an amount $\mathbf{R}_{t+1} = \mathbf{p}\mathbf{\Delta}_t - \mathbf{c}\mathbf{X}_t$ of money and we will be left with $\mathbf{X}_t - \mathbf{\Delta}_t$ heads of livestock.
although we can not know exactly how many offsprings each head of the remaining population will have until the next year, we assume that this number would be random variable from a known distribution $\mathbf{Rep}$, therefore the population of the ranch through the years would best be represented with a sequence $\{\mathbf{X}_t\}_t$ of random variables such that the distribution of $\mathbf{X}_{t+1}$ would depend on $\mathbf{X}_t$ and $\mathbf{\Delta}_t$ only. moreover in order to simplify the analysis we assume that each head of the remaining population shall perish after it has reproduced, therefore at the next year the population would consist only of the offsprings of the remaining population.
To get this exactly, let's say that we enumerate the heads of livestock remaining after the sell as $\alpha = 1, ..., \mathbf{X}_t - \mathbf{\Delta}_t$ and let $\mathbf{Rep}_{t,\alpha}$ denote the reproduction of the αth head during the proceeding year and that $\{\mathbf{Rep}_{t,\alpha}\}_{t,\alpha}$ be an $i.i.d$ sequence of random variables from the distribution $\mathbf{Rep}$, so we have:

$$\mathbf{X}_{t+1} = \sum_{\alpha=1}^{\mathbf{X}_t - \mathbf{\Delta}_t} \mathbf{Rep}_{t,\alpha} \tag{1}$$

Now our question is to find the best choice for $\mathbf{\Delta}_t$ that maximizes our profit from this business.
To be precise, we define:

$$\mathbf{G}_t = \sum_{k=0}^{\infty} \gamma^k \mathbf{R}_{t+1+k} \tag{2}$$

where $\gamma \in [0, 1]$ is an arbitrary parameter that determines the relative desirability of a reward in the next year to the same amount in the present year.

So our goal is to find the proper sequence of sells $\{\mathbf{\Delta}_t\}_t$ that maximizes G_0.
As it appears, condition (1) implies that the formalism of *Markov-decision-processes* suit our needs.
To put the problem in this form, we take the population at time t, $\mathbf{X}_t$ as our state, which is in a properly defined state space $\mathcal{S}$.
We take $\mathbf{\Delta}_t$ as our action. At each state $x \in \mathcal{S}$ we assume that $\delta \in \mathcal{D}_x$ where $\mathcal{D}_x$ is a properly defined action space(the obvious constraints are $\mathcal{S} \subseteq \mathbb{N}_0$ and $\mathcal{D}_x \subseteq \{0, ..., x\}$).
And finally, our reward at time t would be determined deterministically as a function of the state we are in and the action we take, i.e., $\mathbf{R}_{t+1} = \mathbf{R}(\mathbf{X}_t, \mathbf{\Delta}_t)$ where $\mathbf{R}(x, \delta) = \mathbf{p}\delta - \mathbf{c}x$.
So as the theory of *MDPs* suggests, to find the best action, we need to solve the *Bellman optimality equation* for the unknown function $\mathbf{V}_*$ over the state space $\mathcal{S}$, which is in the form:

$$\mathbf{V}_*(x) = \max_{\delta \in \mathcal{D}_x} \left\{ \mathbf{R}(x, \delta) + \gamma \sum_{x'} \mathbb{P}(x'|x, \delta)\mathbf{V}_*(x') \right\} \quad ; \forall x \in \mathcal{S} \tag{3}$$

where by $\mathbb{P}(x'|x, \delta)$ we mean $\mathbb{P}(\mathbf{X}_{t+1} = x'|\mathbf{X}_t = x, \mathbf{\Delta}_t = \delta)$.
$\mathbf{V}_*$ denotes the value function under the optimal policy and the optimal action at each state x would be the $\delta \in \mathcal{D}_x$ that maximizes the right hand side of equation (3). Moreover, to simplify further references, we take the right hand side of (3) as the *Bellman operator* $\mathbf{B} : \mathbf{R}^{\mathcal{S}} \to \mathbf{R}^{\mathcal{S}}$ defined as:

$$\mathbf{B}(\mathbf{V})(x) = \max_{\delta \in \mathcal{D}_x} \left\{ \mathbf{R}(x, \delta) + \gamma \sum_{x'} \mathbb{P}(x'|x, \delta)\mathbf{V}(x') \right\} \quad ; \forall x \in \mathcal{S} \tag{4}$$

Therefore equation (3) would become:

$$\mathbf{V}_* = \mathbf{B}(\mathbf{V}_*) \tag{5}$$

i.e., $\mathbf{V}_*$ is the fixed point of the *Bellman operator* $\mathbf{B}$.
Our general approach for solving (5) is to use the following iterative method:

> first, we choose an arbitrary state function $v_0 \in \mathbf{R}^{\mathcal{S}}$; then we define the sequence $\{v_n\}_{n=0}^{\infty}$ recursively as $v_{n+1} = \mathbf{B}(v_n)$ for $n \in \mathbb{N}_0$.
> Then it can be shown that under the condition $\gamma < 1$, the sequence would approach $\mathbf{V}_*$.

Now we shall go through the details of our code and present our results for several cases that we have considered.

1 Finite capacity

First, we consider the case that our ranch has a finite capacity $\mathcal{C}$, therefore $\mathcal{S} = \{0, 1, ..., \mathcal{C}\}$. but if we let $\mathcal{K}$ be the maximum number of offspring that a given livestock can have, then for states $x \in \mathcal{S}$ with $\mathcal{C} < \mathcal{K}x$, there exists selling amounts $\delta \in \{0, ..., x\}$ such that the next year population has a positive probability of getting larger than the capacity of the ranch. So we should think about how we can keep the population within the allowed range at all times. We have thought of two ideas:

1. when we have a population x, and we sell a number δ of them, we will be left with $x - \delta$ heads of livestock, and the next year's population cannot exceed $\mathcal{K}(x-\delta)$, so one way is to put a constraint on δ to be large enough to ensure $\mathcal{K}(x-\delta) \leq \mathcal{C}$, therefore we will have $\mathcal{D}_x = \{\max\{0, x - \lfloor \frac{\mathcal{C}}{\mathcal{K}} \rfloor\}, ..., x\}$.

2. another way to resolve this is to simply spare the overflow but keep in mind that we shall collect no profit on the amount spared. the result of this would be to modify (1) as follows:

$$\mathbf{X}_{t+1} = \min \left\{ \sum_{\alpha=1}^{\mathbf{X}_t - \Delta_t} \mathbf{Rep}_{t,\alpha}, \mathcal{C} \right\} \tag{1*}$$

   in this case we have $\mathcal{D}_x = \{0, ..., x\}$.

Moreover, we note that both (1) and (1*) imply that the probability function $\mathbb{P}(x'|x, \delta)$ can be written as $f(x'|x - \delta)$ since the distribution of the next generation population would depend on x, δ only through $x - \delta$.

1.1 With no overflow

Now we shall discuss how to implement the aforementioned iterative method, here we note that since the state space $\mathcal{S}$ is finite, then any state function $v \in \mathbb{R}^{\mathcal{C}}$ could be easily represented as a $\mathcal{C} + 1$-tuple of real numbers; on the other hand representing the *Bellman operator* $\mathbf{B}$, might seem to be challenging at the first glance due to the presence of the distribution function $f(x'|r) = \mathbb{P}(\mathbf{X}_{t+1} = x'|\mathbf{X}_t - \mathbf{\Delta}_t = r)$ since despite the relatively simple form of equations (1) and (1*), the function $f(x'|r)$ might not be so easily represented in a compact form. So our first approach to overcome this, was to use an estimator $\tilde{f}_I$ of the distribution function instead of the exact one.
To do this, let's assume that $\{\mathbf{Rep}_{\alpha,i}\}_{\alpha=1,i=1}^{r,I}$ is an *i.i.d* sample from the distribution of $\mathbf{Rep}$ for some sampling frequency I, then we can estimate $f(x'|r)$ as:

$$\tilde{f}_I(x'|r) = \frac{1}{I} \# \left\{ i | \sum_{\alpha=1}^{r} \mathbf{Rep}_{\alpha,i} = x' \right\} \tag{6}$$

and this would yield us an estimation $\tilde{\mathbf{B}}$ of $\mathbf{B}$. Moreover, we should note that in the hope of reducing the error caused by this estimation in our iterative method, at each iteration, we use a new sample to estimate $\mathbf{B}$.

The python implementation of this approach can be found in the file named The_Ranch_finite_capacity_with_no_overflow_estimated, and the results for $\mathcal{C} = 20, \mathbf{p} = 6, \mathbf{c} = 1$(here we have chosen a small number for $\mathcal{C}$ because this approach takes a long time to complete for large capacities; later on, we shall investigate larger capacities with a different but faster approach.) and the reproduction distribution as:

$$\mathbb{P}(\mathbf{Rep} = m) = \begin{cases} 0 & ; m = 0 \\ 0.3 & ; m = 1 \\ 0.7 & ; m = 2 \end{cases} \tag{7}$$

and for three different discounts, 0.65, 0.7 , and 0.75 are as depicted in Figure 1.

As we can see there, different discounts cause different behaviors in both the value function and the optimal sell. To see why this classification happens and what is the significance of these discount values, we shall turn to a rather theoretical analysis of equation (3) later on.

But before that, we note that although the value function for the first and the second plots of Figure 1 look the same, their optimal sell appears to be different, and more importantly, the result for this part varies between different runs of the same code, while this variation is not observed for discounts away from 0.7!

These observations suggest that perhaps near 0.7 lies some critical value for the discount, and while the estimation-based approach for solving (3) shows to be adequate for discounts away from 0.7, the discount values near 0.7 must be dealt with more care.

It would be good if we could somehow find the exact distribution function $f(x'|r)$. One idea might be to use a recursive approach. First note that:

$$f(x'|0) = \begin{cases} 1 & ; x' = 0 \\ 0 & ; x' \neq 0 \end{cases} \tag{8}$$

moreover we have:

$$\begin{aligned} f(x'|r+1) &= \mathbb{P}\left(\sum_{\alpha=1}^{r+1} \mathbf{Rep}_\alpha = x'\right) \\ &= \sum_{m \in supp(\mathbf{Rep}_{r+1})} \mathbb{P}\left(\sum_{\alpha=1}^{r+1} \mathbf{Rep}_\alpha = x', \mathbf{Rep}_{r+1} = m\right) \\ &= \sum_{m \in supp(\mathbf{Rep}_{r+1})} \mathbb{P}\left(\sum_{\alpha=1}^{r} \mathbf{Rep}_\alpha = x' - m\right) \mathbb{P}\left(\mathbf{Rep}_{r+1} = m\right) \\ &= \sum_{m \in supp(\mathbf{Rep})} f(x' - m|r) \mathbb{P}\left(\mathbf{Rep} = m\right) \end{aligned} \tag{9}$$

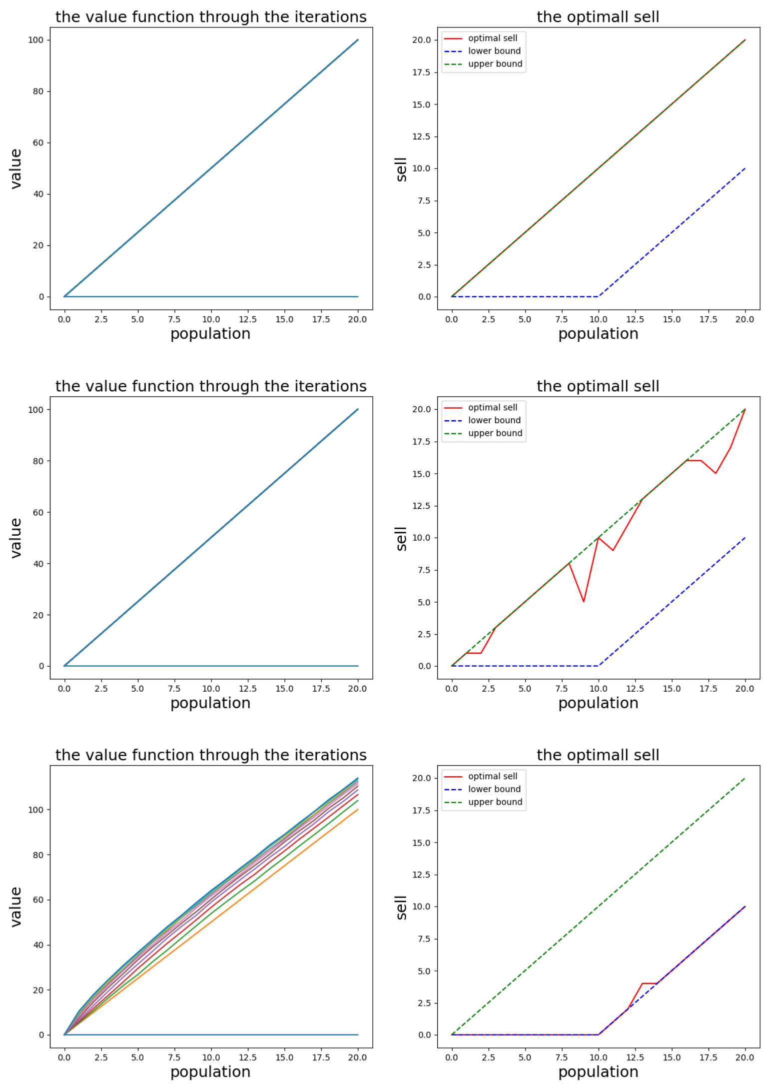

Figure 1: value function and optimal sell calculated using the estimation approach for three different discounts = 0.65, 0.70, 0.75 from above in the case of no overflow.

By this way we can calculate exactly all the probabilities we need.
The python implementation of this approach can be found in the file named The_Ranch_finite_capacity_with_no_overflow_exact and the results for the same values for parameters but with $\mathcal{C} = 200$ are depicted in Figure 2.
and as we can see, the variation that we observed before, when using the esti-

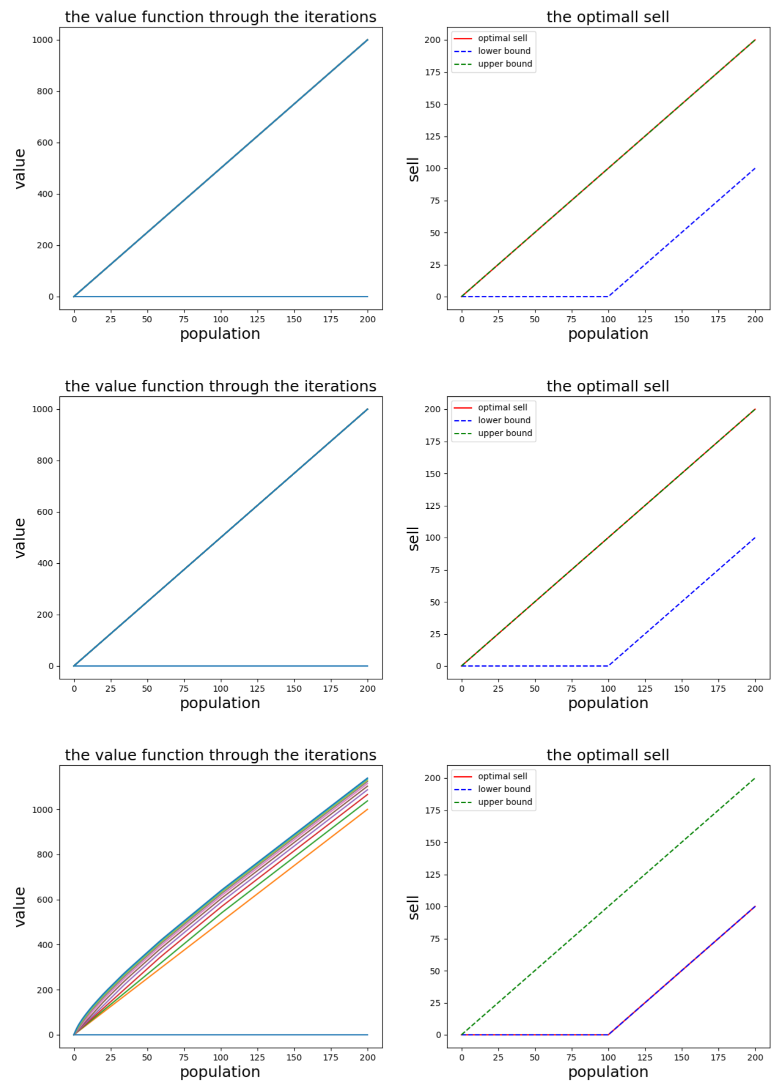

Figure 2: value function and optimal sell calculated using the exact approach for three different discounts = 0.65, 0.70, 0.75 from above in the case of no overflow.

mation approach for 0.7 discount in the optimal sell disappears.
Now it's time to see whether we can gain more insight into the dependency on the discount value.

From the first two plots in Figure 1 or 2, we see that at least in some cases, the solution to (3) has a linear form, let's say $\mathbf{V}_*(x) = \theta x$ for some constant θ. then the right hand side of (3) becomes:

$$\max_{\delta \in \mathcal{D}_x} \left\{ \mathbf{p}\delta - \mathbf{c}x + \gamma\theta \sum_{x'} \mathbb{P}(x'|x,\delta)x' \right\} \tag{10}$$

although we do not have $\mathbb{P}(x'|x,\delta)$ in a closed form, using (1), we can calculate the sum in (10) as:

$$\begin{aligned} \sum_{x'} \mathbb{P}(x'|x,\delta)x' &= \mathbb{E}[\mathbf{X}_{t+1}|\mathbf{X}_t = x, \mathbf{\Delta}_t = \delta] \\ &= \sum_{\alpha=1}^{x-\delta} \mathbb{E}[\mathbf{Rep}_{t,\alpha}] = (x-\delta)\mathbb{E}[\mathbf{Rep}] \end{aligned} \tag{11}$$

and so if we denote $\mathbb{E}[\mathbf{Rep}]$ by μ, then (10) becomes:

$$\begin{aligned} &\max_{\delta \in \mathcal{D}_x} \{(\mathbf{p} - \gamma\mu\theta)\delta + (\gamma\mu\theta - \mathbf{c})x\} \\ &= \begin{cases} (\mathbf{p}-\mathbf{c})x & ; \mathbf{p} - \gamma\mu\theta > 0 \\ (\gamma\mu\theta - \mathbf{c})x & ; \mathbf{p} - \gamma\mu\theta = 0 \\ (\mathbf{p} - \gamma\mu\theta)\max\{0, x - \lfloor \frac{\mathcal{C}}{\mathcal{K}} \rfloor\} + (\gamma\mu\theta - \mathbf{c})x & ; \mathbf{p} - \gamma\mu\theta < 0 \end{cases} \end{aligned} \tag{12}$$

therefore when $\mathbf{p} - \gamma\mu\theta \geq 0$, for $\theta = (\mathbf{p} - \mathbf{c})$ the linear form can solve (3). this condition for γ takes the form:

$$\gamma \leq \frac{1}{\mu} \frac{1}{1 - \mathbf{c}/\mathbf{p}} \tag{13}$$

So when the above inequality is strict, the optimal policy is to sell all the population and end the business; this means that we do not care about the future profit enough to stay in the business with these prices and costs.
But when (13) becomes equality, the argument to be maximized in (10) becomes independent of δ, implying that it does not matter what we do and how much we sell; the expectation of the goal stays the same. For the values we chose for price, cost, and the reproduction distribution, the threshold for discount(right-hand side of (13)) would be approximately equal to 0.705; therefore, the seemingly non-trivial results on the second graph of Figure 1 is merely the effect of using the estimated *Bellman operator* instead of the exact one. Now that we have found the exact critical value for the discount, we can see how our exact approach works at this critical state.

The results depicted in Figure 3 show some deviation from the 0.7 case in the optimal sell. However as we have shown, this has no significance and is merely caused by computational errors since, precisely at the critical discount value, there is no difference between different selling amounts.
But the fact that we can get a singular optimal selling value for each population from our code, tempts us to doubt our results for other cases, again our results there might have been only a result of numerical errors rather than actual

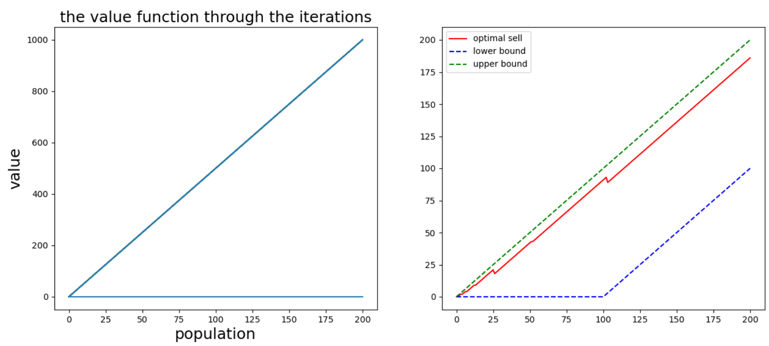

Figure 3: the value function and corresponding optimal sell at the critical discount value in the case of no overflow.

meaningful optimization.
To overrule this possibility we firs take a look at some results regarding the behavior of the expression to be maximized in (12), which is actually the *state-action value function*, $\mathbf{Q}_*(x, \delta)$. The results for two values for population are depicted in Figure 4.

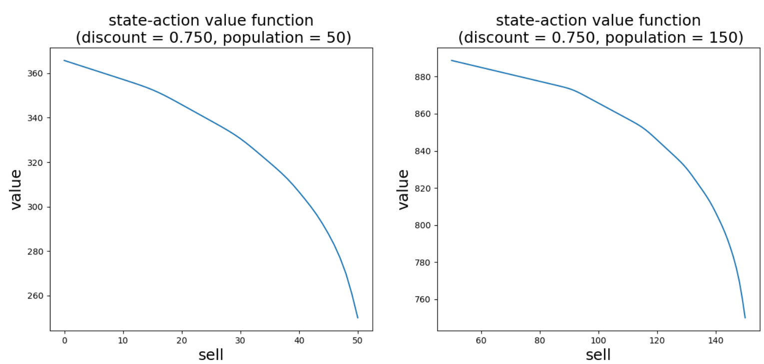

Figure 4: The *state-action value function* for 0.75 as the discount in the case of no overflow.

as we can see the *state-action value function* changes rather continuously and if we choose a reasonably low threshold for finding its maximum value(i.e., all the selling amounts with the state-action value function deviating from the maximum value by an amount less than this threshold), we can check significance of our results for this case. The results for this are depicted in Figure 5.

Confirming our previous claim that in the critical case it does not matter how many of the population we sell.

Finally, we find it noteworthy that it is only for high enough discount val-

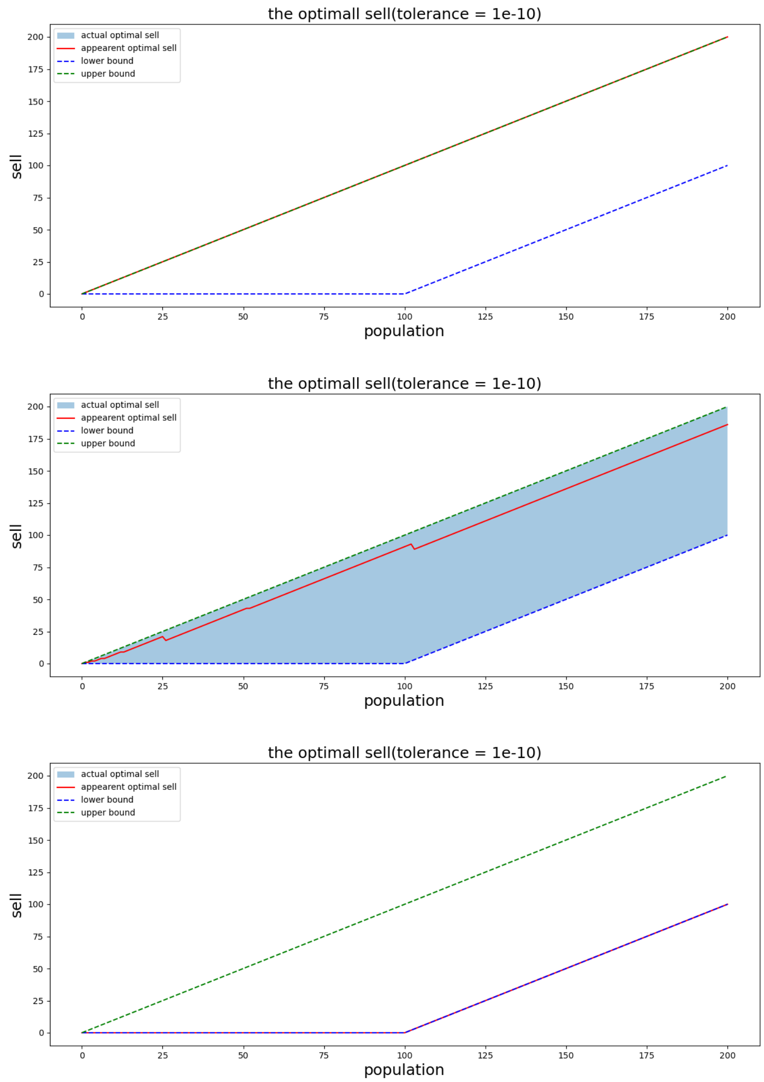

Figure 5: The significant optimal sells calculated using the exact approach for three different discounts = 0.65, 0.70, 0.75 from above in the case of no overflow.

ues(strictly above the critical value) that it is profitable to stay in business and in that case, the optimal policy would be:

> to not sell at all until the population reaches the desired value $\lfloor \frac{\mathcal{C}}{\mathcal{K}} \rfloor$, and after that, we should only sell the additional population to this value.

The significance of this desired population is that it is the largest population

that does not necessitate any selling to ensure that the next generation can be contained in the ranch.
Moreover, since discounts are always taken to be less than or equal to 1, (13) states that there are ratios of price to the cost that no matter how high a discount we choose, staying in the business is not profitable!
To be more precise, we should have:

$$\frac{\mathbf{p}}{\mathbf{c}} > \frac{1}{1 - 1/\mu} \tag{14}$$

to stand a chance of having a profitable business.

1.2 With sparing the overflow

In this section, we shall study the second approach to confining the population, which is to spare the overflow when it happens. here we use the same recursive method as in (9) to find $f(x'|r)$ but keeping in mind that it comes from (1*), so the inductive step should be as:

$$\begin{aligned} f(x'|r+1) &= \mathbb{P}\left(\min\left\{\sum_{\alpha=1}^{r+1}\mathbf{Rep}_\alpha, \mathcal{C}\right\} = x'\right) \\ &= \sum_{m \in supp(\mathbf{Rep}_{r+1})} \mathbb{P}\left(\min\left\{\sum_{\alpha=1}^{r+1}\mathbf{Rep}_\alpha, \mathcal{C}\right\} = x', \mathbf{Rep}_{r+1} = m\right) \end{aligned} \tag{9*}$$

so for $x' < \mathcal{C}$, it becomes:

$$\begin{aligned} &\sum_{m \in supp(\mathbf{Rep}_{r+1})} \mathbb{P}\left(\sum_{\alpha=1}^{r+1}\mathbf{Rep}_\alpha = x', \mathbf{Rep}_{r+1} = m\right) \\ &= \sum_{m \in supp(\mathbf{Rep}_{r+1})} \mathbb{P}\left(\sum_{\alpha=1}^{r}\mathbf{Rep}_\alpha = x' - m\right)\mathbb{P}\left(\mathbf{Rep}_{r+1} = m\right) \\ &= \sum_{m \in supp(\mathbf{Rep})} f(x' - m|r)\mathbb{P}(\mathbf{Rep} = m) \end{aligned} \tag{9.1*}$$

and when $x' = \mathcal{C}$, we have:

$$
\begin{aligned}
&\sum_{m \in supp(\mathbf{Rep}_{r+1})} \mathbb{P}\left(\sum_{\alpha=1}^{r+1} \mathbf{Rep}_\alpha \geq \mathcal{C}, \mathbf{Rep}_{r+1} = m\right) \\
&= \sum_{m \in supp(\mathbf{Rep}_{r+1})} \mathbb{P}\left(\sum_{\alpha=1}^{r} \mathbf{Rep}_\alpha \geq \mathcal{C} - m, \mathbf{Rep}_{r+1} = m\right) \\
&= \sum_{m \in supp(\mathbf{Rep}_{r+1})} \mathbb{P}\left(\sum_{\alpha=1}^{r} \mathbf{Rep}_\alpha \geq \mathcal{C} - m\right) \mathbb{P}(\mathbf{Rep}_{r+1} = m) \\
&= \sum_{m \in supp(\mathbf{Rep}_{r+1})} \sum_{x''=\mathcal{C}-m}^{\mathcal{C}} \mathbb{P}\left(\sum_{\alpha=1}^{r} \mathbf{Rep}_\alpha = x''\right) \mathbb{P}(\mathbf{Rep}_{r+1} = m) \\
&= \sum_{m \in supp(\mathbf{Rep}_{r+1})} \sum_{x''=\mathcal{C}-m}^{\mathcal{C}} \mathbb{P}\left(\min\left\{\sum_{\alpha=1}^{r} \mathbf{Rep}_\alpha, \mathcal{C}\right\} = x''\right) \mathbb{P}(\mathbf{Rep}_{r+1} = m) \\
&= \sum_{m \in supp(\mathbf{Rep}_{r+1})} \sum_{x''=\mathcal{C}-m}^{\mathcal{C}} f(x''|r) \mathbb{P}\left(\mathbf{Rep}_{r+1} = m\right) \\
&= \sum_{m \in supp(\mathbf{Rep})} \sum_{x''=\mathcal{C}-m}^{\mathcal{C}} f(x''|r) \mathbb{P}(\mathbf{Rep} = m)
\end{aligned}
\tag{9.2*}
$$

The python implementation of this approach can be found in the file named The_Ranch_finite_capacity_with_sparing_the_overflow_exact and the results for the same values of cost, selling price, capacity, and reproduction as in the case of Figure 2 are depicted in Figure 6.

As we can see there, the results for discounts less than the critical value are the same as before when we ensured prevention from overflow.
The results for the critical case seems to be the same as the previous case.
We find it curious that the optimal sell for discounts above the critical value shows the same general behavior, namely:

> to not sell until the population reaches some desired value, and after that, we should only sell the additional population to this value.

But with the difference that here *the selling point*, i.e., the value above which we should trim our population, has slightly increased from the previous value of $\lfloor \frac{\mathcal{C}}{\mathcal{K}} \rfloor$.
Figure 7 and 8 show the selling point versus respectively capacity and discount value.

But again as in the previous case, we should analyze the significance of our results. Figure 9 shows the behavior of the *state-action value function* for two populations and three discount values. As we can see there, in the critical

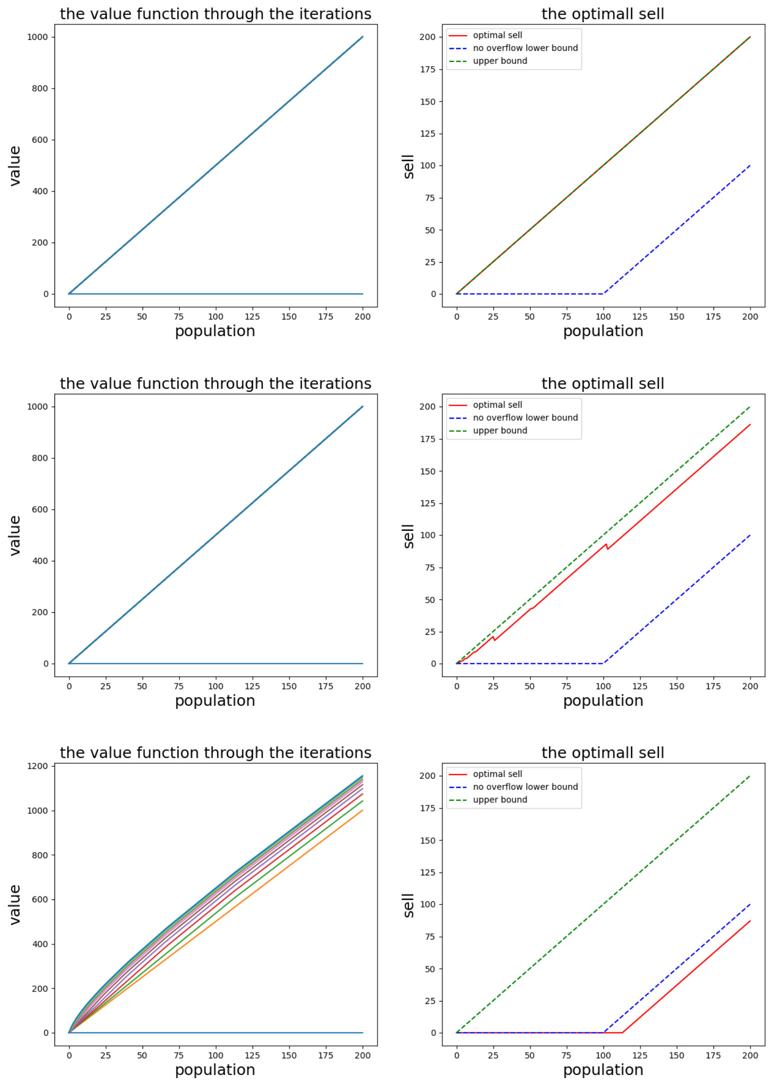

Figure 6: value function and optimal sell calculated using the exact approach for three different discounts 0.65, the critical value and 0.75 from above in the case of sparing the overflow.

case, there are a lot of selling values that have no significant difference in their *state-action value function* from its maximum value, so again by introducing a tolerance value we can find the actual optimal selling values. The results for this are depicted in Figure 10. As we can see there, the optimal sell depicted in Figure 6, is not a significant one, and in fact this case behaves exactly the same as

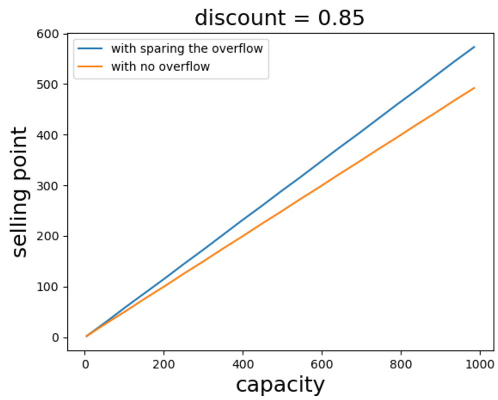

Figure 7: The selling versus capacity in the case of sparing the overflow.

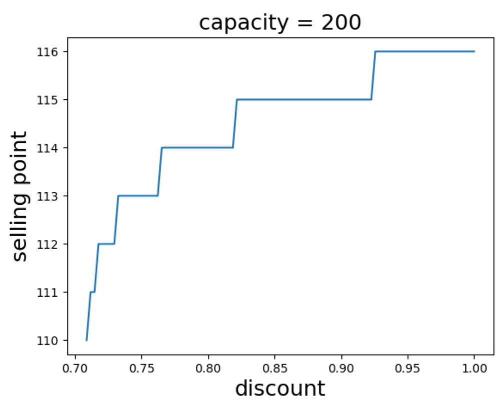

Figure 8: The selling point versus discount in the case of sparing the overflow.

the critical case of the previous scenario, when we prevented any overflow.But its important to note that in that case, the optimal sell included all of the action space, while here, there are selling values in our action space that are not optimal.

Finally as we can see in Figure 10, the results depicted in Figure 6 for the sub-critical and the super-critical cases where already significant.

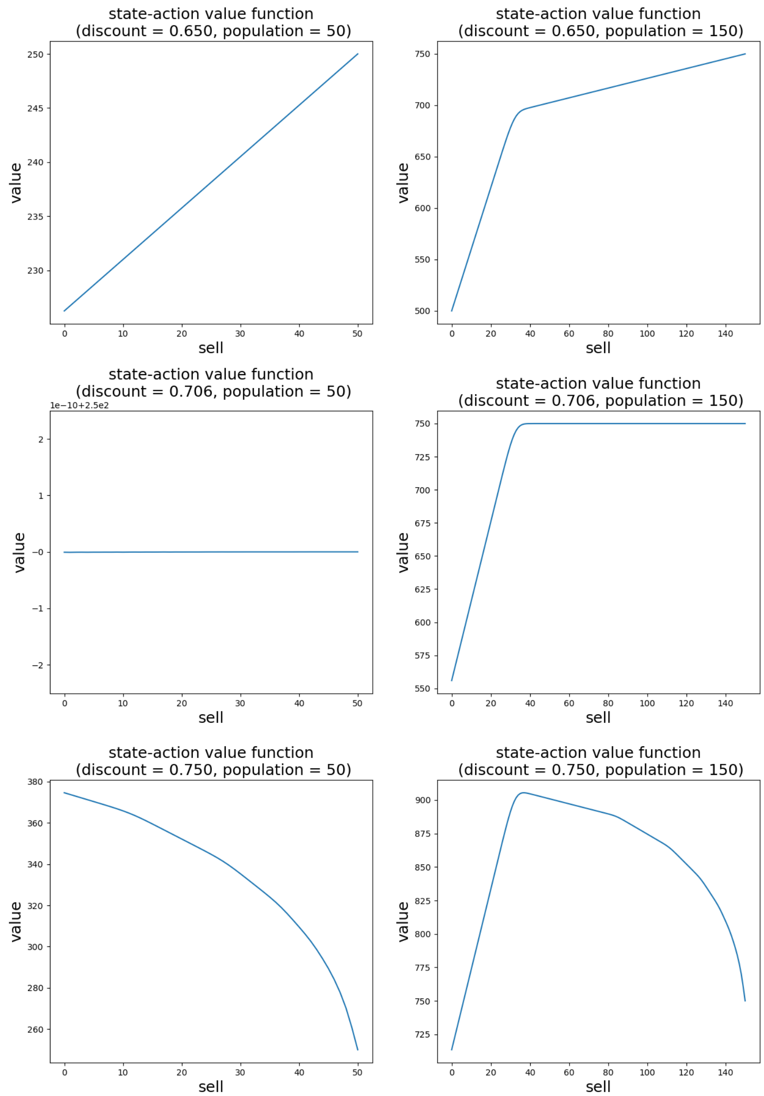

Figure 9: The *state-action value function* for three discount values, sub-critical, critical, and super-critical respectively from above. The plots on the left are typical for populations below the selling point and those on the right are typical of the populations above the selling point

2 Infinite capacity

Now we will investigate the population control problem when there is no natural bound on how large the population can grow. The *Bellman equation* would be the same as (3) with the difference that here $\mathcal{S} = \mathbb{N}_0$. Our approach, as before,

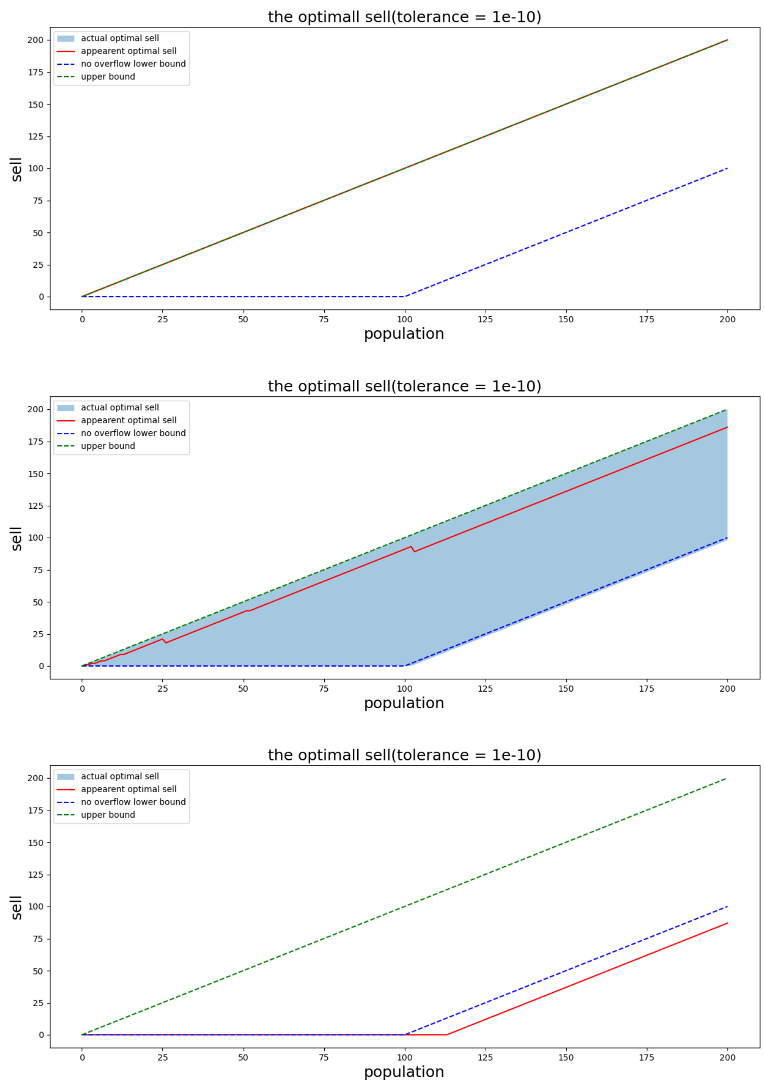

Figure 10: The actual optimal selling amounts for three discount values, sub-critical(0.65), critical, and super-critical(0.75), respectively from above.

is to use the iterative method of successive operations of $\mathbf{B}$ on a given initial state function v_0, which we choose to be the constant zero function. So for the first few iterations, we get:

$$v_1(x) = \mathbf{B}(v_0)(x) = \max_{\delta \in \mathcal{D}_x}\{\mathbf{p}\delta - \mathbf{c}x\} = (\mathbf{p} - \mathbf{c})x \tag{15}$$

$$
\begin{aligned}
v_2(x) = \mathbf{B}(v_1)(x) &= \max_{\delta \in \mathcal{D}_x} \{\mathbf{p}\delta - \mathbf{c}x \ + \gamma\mu(\mathbf{p}-\mathbf{c})(x-\delta)\} \\
&= \max_{\delta \in \mathcal{D}_x} \{\{\mathbf{p} - \gamma\mu(\mathbf{p}-\mathbf{c})\}\delta + \{\gamma\mu(\mathbf{p}-\mathbf{c}) - \mathbf{c}\}x\} \\
&= \begin{cases} (\mathbf{p}-\mathbf{c})x & ; \mathbf{p} - \gamma\mu(\mathbf{p}-\mathbf{c}) \geq 0 \\ \{\gamma\mu(\mathbf{p}-\mathbf{c}) - \mathbf{c}\}x & ; \mathbf{p} - \gamma\mu(\mathbf{p}-\mathbf{c}) < 0 \end{cases}
\end{aligned} \tag{16}
$$

Therefore in the former case, i.e., when $\gamma \leq \frac{1}{\mu}\frac{1}{1-\mathbf{c}/\mathbf{p}}$, we shall have $v_\infty(x) = (\mathbf{p}-\mathbf{c})x$ as in the case of finite capacity, which means that its not profitable to stay in business. But in the later one, we have:

$$
\begin{aligned}
v_3(x) = \mathbf{B}(v)(x) &= \max_{\delta \in \mathcal{D}_x} \{\{\mathbf{p} - \gamma\mu(\mathbf{p}-\mathbf{c})\}\delta + \gamma\mu\{\gamma\mu(\mathbf{p}-\mathbf{c}) - \mathbf{c}\}(x-\delta)\} \\
&= \max_{\delta \in \mathcal{D}_x} \{\{\mathbf{p} - \gamma\mu\{\gamma\mu(\mathbf{p}-\mathbf{c}) - \mathbf{c}\}\}\delta + \{\gamma\mu\{\gamma\mu(\mathbf{p}-\mathbf{c}) - \mathbf{c}\} - \mathbf{c}\}x\} \\
&= \{\gamma\mu\{\gamma\mu(\mathbf{p}-\mathbf{c}) - \mathbf{c}\} - \mathbf{c}\}x
\end{aligned} \tag{17}
$$

Therefore in this case, v_n can be written as $v_n(x) = k_n x$, where:

$$
\begin{aligned}
k_1 &= \mathbf{p} - \mathbf{c} \\
k_{n+1} &= \gamma\mu k_n - \mathbf{c} \quad ; n \geq 1
\end{aligned} \tag{18}
$$

From this we conclude that for $n \geq 1$, since $1 \leq \frac{1}{1-\mathbf{c}/\mathbf{p}} < \gamma\mu$ we have:

$$
\begin{aligned}
k_n &= (\gamma\mu)^{n-1}(\mathbf{p}-\mathbf{c}) - \frac{1-(\gamma\mu)^{n-1}}{1-\gamma\mu}\mathbf{c} \\
&= \frac{1}{\gamma\mu - 1}\left\{\mathbf{c} + \left\{\gamma\mu - \frac{1}{1-\mathbf{c}/\mathbf{p}}\right\}(\mathbf{p}-\mathbf{c})(\gamma\mu)^{n-1}\right\}
\end{aligned} \tag{19}
$$

So we have $k_\infty = \infty$, and hence $v_\infty(x) = \infty$, for $x \neq 0$ and $v_\infty(0) = 0$, thus our iterative algorithm diverges. We believe that, this happens as a result of the fact that in this case the *Bellman equation* does not have any finite solution. To see this, recall that for a given policy π we have:

$$
\mathbf{V}_\pi(x) = \mathbb{E}[\mathbf{G}_t | \mathbf{X}_t = x] = \sum_{k=0}^{\infty} \gamma^k \mathbb{E}[\mathbf{R}_{t+1+k} | \mathbf{X}_t = x] \tag{20}
$$

and despite the fact that at each t, $\mathbf{R}_t$ is bounded (by $(\mathbf{p}-\mathbf{c})\mathbf{X}_t$ from above and $-\mathbf{c}\mathbf{X}_t$ from below),they are not equibounded(since the state space is not bounded). So there is no guarantee for the convergence of the value function for every policy.

Therefore, we claim that in this case, which is ironically the only case where there might be a chance of getting a non-trivial solution, the problem is unsolvable.

Summary

At the end we can say that our solution for this problem depends on the discount value we choose:

- when discount is less than the critical value $\frac{1}{\mu}\frac{1}{1-\mathbf{c}/\mathbf{p}}$: in this case it is not profitable to stay in the business.
- when discount is equal to the critical value: it does not matter how many of the population we sell as long as we stay in the action space of the case with no overflow.
- when discount is greater than the critical value and:
  1. the ranch has a finite capacity: in this case there exists a selling point, below which we shouldn't sell at all, but above this value we should sell the additional population to this selling point.
     The value of this selling point may differ based on the approach we take to confine the population, when we prevent any overflow it's simply $\lfloor \frac{\mathcal{C}}{\mathcal{K}} \rfloor$. but when we spare the overflow, the selling point increase as depicted in Figure 5.
  2. the ranch has infinite capacity: in this case the problem is not solvable, i.e., there exists no optimal policy.