Solving MDPs using the MDP package in R (exercises)

Lars Relund Nielsen*

Department of Genetics and Biotechnology, University of Aarhus, P.O. Box 50, DK-8830 Tjele, Denmark, lars@relund.dk.

Anders Ringgaard Kristensen

Department of Large Animal Sciences, University of Copenhagen, Groennegaardsvej 2, DK-1870 Frederiksberg C, Denmark.

July 26, 2009

Exercise 1

The goal of this exercise is get familiar with R and the MDP package. You may use a simple text editor (e.g. notepad) to write your R code. Afterwards you can simply copy the code and paste it into R. It is assumed that you have a basic knowledge to R. If not use some time on an introduction, e.g. the one supplied with the data files (see link below).

1. Create a folder for your exercises on your computer.
2. Download the data files needed for the exercises at http://www.research.relund.dk/?p=150 and extract the files in the folder you just have created.
3. Open R GUI and set the working directory to the folder you just created (File->Change dir...).
4. Load the MDP package.
5. Check that you have version 1.0 of the MDP package installed (?MDP - the version number appear in the bottom). If not install the package using `MDP_1.0.zip` just downloaded. Close R, open it again, set the working directory and run:

```
> install.packages("MDP_1.0.zip")
> library(MDP)
```

6. Consider the sow replacement problem in [2]. Generate the MDP in R. Remember that you can copy text from a pdf file if you read it using Acrobat reader, so you

*Corresponding author

do not have to type in the code fragments yourself.

7. Have a look at the data frames with information about the states and actions. What is the transition probability for a low litter size given a big litter size in the current stage and action `keep`?
8. Find the optimal policy under the expected discounted reward criterion using an interest rate of 5%. What is the interpretation of the weights?
9. Find the optimal policy under the expected discounted reward criterion using an interest rate of 2.5%. Compare the optimal values with the values under interest rate 10% and 5%. What is the trend in the numbers and why?

Exercise 2

We consider a sheep replacement problem. The model is formulated as an MDP with 2 levels. At the founder level we only have a single state representing the current sheep. The stage length is the lifespan of the sheep. Since we have an infinite time-horizon at the founder level we model the current sheep and all its successors.

At level one each stage represent the reproductive cycle of the sheep (parity = 1 year). States are a combination of the current litter size and last parity litter size of the sheep. Two actions are possible `keep` and `replace`. Net rewards are generated from selling the lamb and slaughtering the sheep minus feeding costs.

1. The model is supplied in the file `sheep.hmp`. Convert the MDP to binary format (hint see `?convertHMP2Binary`) and load the MDP into memory.
2. Get an overview over the model (the number of stages, states, actions etc.).
3. Calculate an optimal replacement policy under the following three criteria of optimality: (a) Maximum present value (discounting) using a discount rate of 10%. (b) Maximum average net returns per sheep per year. (c) Maximum average net returns per lamb produced.
4. Compare the three optimal policies and explain the differences (if any).
5. For the optimal policy maximizing average net returns per sheep per year, calculate the following technical and economical key figures. (a) Average lifetime of an sheep. (b) Average number of lambs per sheep per year. (c) Average number of lambs produced over the lifetime of an sheep.

Exercise 3

We consider a model of mate desertion in Cooper's Hawks [1]. An finite-horizon MDP was developed for a single nesting season to study the females' brood-rearing strategies. In this model a strategy consisted of combinations of staying at the nest, hunting, and deserting, i.e. we consider the actions `stay`, `hunt` and `desert`. The nestling season is divided into 5 stages:

(0) early nestling period (weeks 1-2)

(1) late nestling period (weeks 3-5)

(2) early fledgling-dependency period (weeks 6-7)

(3) late fledgling-dependency period (weeks 8-10)

(4) end of the observation period (week 10)

The states in the model are defined as the cross product of 2 state variables: the physical condition of the female and nestlings. Both are a numbers between 1 and 7 where 1 corresponds to the lowest body condition. If the physical condition of the female is 1 then the hawk must desert the nest and if equals 1 for the brood we assume that the nestlings have died.

It is assumed that the female hawks maximizes her reproductive fitness, defined as the weighted average of the expected probability of survival of her current offspring and her expected future reproduction.

The transition probabilities[1] and the fitness at the end of the nesting season are given i 2 csv files.

```
> states<-read.csv(file="hawk_states.csv", as.is=T)
> head(states)
```

```
  sIdx sF sB label endFitness
1    0  1  1   1,1     0.0000
2    1  2  1   2,1     0.1375
3    2  3  1   3,1     0.1475
4    3  4  1   4,1     0.1625
5    4  5  1   5,1     0.1800
6    5  6  1   6,1     0.1875
```

```
> transPr<-read.csv(file="hawk_transPr.csv", as.is=T)
> head(transPr)
```

```
  d a sNow sNext           pr
1 0 0    0     0 8.051497e-01
2 0 0    0     1 1.796532e-01
3 0 0    0     2 1.474683e-02
4 0 0    0     3 4.453157e-04
5 0 0    0     4 4.947011e-06
6 0 0    0     5 2.021730e-08
```

[1] We use simulated transition probabilities and not the ones from the paper which was not given explicitly.

The vector `states$endFitness` provide us with the terminal values used at the last stage. All other weights are set to zero. Columns `sF` and `sB` in states provide us with the physical condition of the femal and the brood, respectively.

A function is provided for easy reading of the transition probabilities in `transPr`:

```
> getTransPr<-function(d,s,a) {
+   pr<-cbind(1,transPr[transPr$d==d & transPr$sNow==s & transPr$a==a,c('sNext','pr')])
+   pr<-pr[pr[,3]>1e-5,]
+   pr[,3]<-pr[,3]/sum(pr[,3])
+   as.numeric(t(pr))
+ }
> getTransPr(d=0,s=3,a=2)     # trans pr (early nestling period, state with sIdx=3,
action desert)
```

```
 [1] 1.0000000000 0.0000000000 0.0001344445 1.0000000000 1.0000000000 0.0044521889 1.0000000000 2.0000000000
 [9] 0.0542387649 1.0000000000 3.0000000000 0.2430812800 1.0000000000 4.0000000000 0.4007732768 1.0000000000
[17] 5.0000000000 0.2430812800 1.0000000000 6.0000000000 0.0542387649
```

1. Generate the model. The following uncomplete code may be used:

```
> prefix="hawk_"
> w<-binaryMDPWriter(prefix)
> w$setWeights("Fitness")
> w$process()
>   for (d in 0:4) {
+     w$stage()
+       for (s in states$sId) {
+         sF<-states$sF[s+1]
+         sB<-states$sB[s+1]
+         w$state(label=states$label[s+1])
+           if (d<4) {
+             if (sF!=1 & sB!=1) {
+               # <Put action stay and hunt here>
+             }
+             # <Put action desert here>
+           }
+         w$endState()
+       }
+     w$endStage()
+   }
> w$endProcess()
> w$closeWriter()
```

2. Solve the MDP using value iteration. Hint: remember to set the terminal values of the MDP to `states$endFitness`.

3. A plot of the optimal policy can be created by:

```
> # assume that you have created a data frame of the policy
> policy$sF<-as.numeric(substr(policy$label,1,1))
> policy$sB<-as.numeric(substr(policy$label,3,3))
> library(lattice)
> policy<-subset(policy,n0<4)
> xyplot(sF~sB | paste("period",n0),
+   group=aLabel,
```

```
+    par.settings=list(superpose.symbol=list(pch=c(22,23,24),cex=c(1.7,1.6))),
+    auto.key = list(columns=3,points=T,lines=F),
+    as.table=T,
+    data=policy)
```

What is the optimal action in the early fledgling-dependency period when the the physical condition of the female and the brood is 2 and 4, respectively?

References

[1] E.J. Kelly and P.L. Kennedy. A dynamic state variable model of mate desertion in cooper's hawks. *Ecology*, 74(2):351–366, 1993.

[2] Lars Relund Nielsen. *Solving MDPs using the MDP package in R.* Department of Genetics and Biotechnology, University of Aarhus, P.O. Box 50, DK-8830 Tjele, Denmark, July 2009.