{capn} and groundwater exercise

Eli Fenichel, Yale University 2019

The results from Fenichel et al. 2016 PNAS and the state-wide results from Addicott & Fenichel 2019 JEEM can be reproduced using capn. There is a slight difference between to the two results because of some additional data cleaning. We now use the data associated with Addicott & Fenichel. The necessary data file to download is groundwater_capn.R.

To download the file go to
https://github.com/efenichel/capn_stuff/blob/master/groundwater_capn.R
above the code there is a button that says "raw" , right click and then click "save link as…"
enter the path where you want to save the file.

Getting ready

1. Install R and RStudio, instructions are here https://www.rstudio.com/products/rstudio/download/, if you have not done so already.
2. On the left-hand side of RStudio go to File -> Open File and navigate to open groundwater_capn.R.
3. Verify all is installed correctly by running groundwater_capn.R. Press Ctrl+Shift+Enter run the whole script or step through using Ctrl+Enter.
4. You should see a price curve and intertemporal welfare curve in the lower right Plots window. Type psim(pC,21.5)$shadowp into the console window, this is the shadow price at 21.5 acre feet of water. Verify the shadow price is $17.45 or about $17.
5. Once you have verified these piece click the brooms in the Global Environment and Viewer tabs to clear all results.

Understanding the code

1. Highlight lines 8 through 20 and click "run" or Ctrl+Enter. This loads the necessary packages and provides some directory information for Eli Fenichel's Github page.
2. Now we will get the data. Highlight lines 34-37 and click "run" or Ctrl+Enter. In the Global Environment you will ksdata appear under data. Click on this data structure. You will see it is actually many data structures. One for each groundwater management district (GMDX), and Outgroup, and State. We will focus on State. If you click State you will see state is gmdnum 7. There are number of dataframes and variables associated with a state wide analysis. These include coefficients from a multinomial logit regression and the selection model regression model mentioned in the papers. Also the means for the data are there, as are cropprices, planting costs and other necessary data.
3. On line 39 to 53 the data structure is explained.
4. Next we need to organize the data into a usable form. We want to take the raw regression results and organize them to focus on the groundwater depth or things that will be functions of groundwater depth like planting and water withdrawal choices. Run line 57.
5. In the Global Environment you will now see a section for functions, with a function called "datasetup." Click "datasetup" to view the function. There are three cases, they all do basically the same thing but have to run separately because of some dimensionality issues (case 1-6 are subsets of the state of Kansas). Look through the notes in the first case.

6. Now run lines 63-69. This creates a new object called gw.data. This object holds all the data needed to run the groundwater systems model.
7. Run lines 76-85. These lines set user defined variables such as the discount rate and recharge rate that will be used. They also set some parameters for the approximation that capn will use.
8. Run line 92. This pulls in all the functions for the system model for Kansas groundwater. You will have to create these functions for your system, just like you will need to collect your own data.
9. Now we are truly ready to start using capn. Run lines 108 and 109 to set the approximation space and lay out a grid of approximation nodes.
10. Run line 112-121 to loop through and create the matrix simuData that has columns nodes, change in stock (sdot), change in sdot with respect to s, change in profit with respect to s, and profit.
11. Run line 125 to approximate the Chebyshev coefficients and save them as the object pC.
12. Run lines 127-133 to use the Chebyshev coefficients to compute a shadow price, a wealth = shadow price times quantity, the value function, stock, and profit for each node.
13. Run lines 136-148 to plot the shadow price curve and check the results.
14. Run lines 150-167 to plot the value function, using ggplot. If you have not used ggplot before take a minute to examine the syntax.

SOME EXERCISES

First, let's explore the data and model (suggested weighting when used as a problem set, 20% equally weighted). The simulation model is built on two statistical models: one for water withdrawal (a two-stage regression linear regression model that controls for selection bias) and one for crop mix (a multinomial logit model). The coefficients from these regressions are in the ksdata object, which as divided by region. We focus on "region 7," which is a state level aggregation. This is saved in the object region_data. Recharge is considered constant, so the physical model is just recharge minus water withdrawal. However, capn can certainly use more complicated biophysical models. This is the section you will have to provide for you own natural asset valuation.

1. Create a new script.
2. Take a look at the regression results and means. In your script type
    a. region_data['watercoeff'], run it and describe what you see.
    b. region_data['mlogitcoeff'] , run it and describe what you see, why are there many columns?
3. Take a look at the summary statistic means.
    a. region_data['wwdmeans']
    b. region_data['mlogitmeans']
    c. Notice there is still some cleaning to do. Of course after this cleaning you could change wdep_ (which stands for water depth) to see how withdrawal or planting decisions change as a function of water depth. This is the economic program.
4. Take a look at datasetup. Look specifically at lines 169-177 (these are in the state wide case). Notice, some variables are being removed and online 174. Everything not removed is being evaluated at the mean. What is Alpha and which variables do you think are being held back? (hint: look at the output from region_data['wwdmeans'] or ksdata$State$wwdmeans).
5. Can you identify the lines doing something similar for crop choice model?

The organized data are in gw.data, which you created on line 68 of groundwater_capn.R.

6. Plot water withdrawals as a function of water depth on the representative acre. The water withdrawal function is Wwd(water, parameters).
   a. Click the Wwd or water withdrawal function to see what it does.
   b. Try Wwd function at 21.5 feet of water. The arguments are Wwd(21.5, gw.data).
   c. Simulate the model using a for loop.
      i. Make a matrix to hold the output (called pre-allocation), by typing `data.ww <- matrix(0, nrow = 100, ncol = 2)`. This says make a matrix of zeros that is 100 rows long and two columns wide.
      ii. Name these columns by adding a line `colnames(data.ww) <- c("water", "water.withdrawal")`
      iii. Now run the simulation with the following block of code:

```
for(j in 1:100){
        data.ww[j,1] <- j
        data.ww[j,2] <-Wwd(j,gw.data)
}
```

   Can you explain what this does?
      iv. Check your result by adding a line `View(data.ww)`
      v. Change your data.ww to a dataframe, by typing `data.ww <- as.data.frame(data.ww)`
   d. Let's plot the data using the ggplot2 package.
      i. The ggplot2 package should already be installed. Make sure the package is installed, try adding code that says `library(ggplot2)`
         1. If it runs, then great.
         2. If it does not run, add `install.packages("ggplot2")` and try again.
      ii. Copy the following

```
ggplot() +
 geom_line(data = data.ww, aes(x = water, y = water.withdrawal),
            color = 'blue') +
 labs(
  x= "Stored groundwater",
  y = "Water Withdrawal")  +
 theme(  #http://ggplot2.tidyverse.org/reference/theme.html
  axis.line = element_line(color = "black"),
  panel.background = element_rect(fill = "transparent",color = NA),
  plot.background = element_rect(fill = "transparent",color = NA)
 )
```

   The first line sets up a place for the plot. The second line draws the plot. The third line changes the line color to blue. The four – sixth lines label the axes. The last four lines control the color of the axis lines and the background of the plot.

7. View the cropFwater(water, parameters) function. This function determines how much land is planted with the representative well as a function of water depth. It returns a horizontal vector of acres planted [alfalfa, corn, sorghum, soy, wheat, unirrigated or fallow]. Continue working in your script.

a. Use a similar approach as before to simulate acres planted and their shares. You might want to create a new variable, e.g., data.cs. Also, your storage matrix will need to be 7 columns wide, one for water level, then one for each of the six crops (including fallow).
b. You can sum across the crops columns by using the command `rowSums(data.cs[,2:7])`. This will give you the total area planted for irrigated crops. Use this result to plot the irrigated acres as a function of water. The command cbind(x,y) puts columns x and y together horizontally.
c. Create some pie charts at various water depths, e.g., 2, 5, 15, 20, 50, 90 to see you the share of crops changes. Use the following code

```
my.pie<- data.frame(
  crops = c("alfalfa", "corn", "sorghum", "soy", "wheat","fallow"),
  fracs = cropFwater(20 ,gw.data)
)

ggplot(my.pie, aes(x = factor(1), y=fracs, fill=factor(crops)) )+
  geom_bar(width = 1,stat="identity")+coord_polar(theta = "y") +
  labs(
    x= "",
    y = "") +
  theme(axis.ticks = element_blank(),
        panel.grid  = element_blank(),
        axis.text = element_blank(),
        panel.background = element_rect(fill = "transparent",color = NA),
        plot.background = element_rect(fill = "transparent",color = NA)
        )+
  guides(fill=guide_legend(title="crop fractions"))
```

d. What needs to be changed to try different levels?

Now we are ready to explore capN (80% equally weighted). Go back to the groundwater_capn.R script. You may want to review the capn documentation (https://cran.r-project.org/web/packages/capn/capn.pdf).

In this exercise we will focus on the “p” family of functions. Read through the code and see if you can understand what is happening. The two functions to pay close attention to are the paprox and psim functions.

Lines 73-85 let you control additional simulation parameters (discount rate, dr, and recharge rate per year, recharge). To change the recharge rate, put a # at the start of line 79 and remove the # at the start of line 80.

8. Explore the effects of the discount rate by changing it to 5% and then to 7%. How does the shadow price at 21.5 feet change relative to the default 3%? How about at 18 feet? Hint: Rerun lines 108- 125, then use

```
output<-psim(pcoeff = pC,
        stock = 21.5)
```

   to simulate a specific water depth. Then `output$shadowp` will give you the shadow price.
9. Now set the discount rate back to 3% and try changing the recharge rate. What if the recharge rate were half the assumed rate, how about double, or five times? Does it matter if you make comparisons at 21.5, 18, or 30 feet?
10. capn also reports inclusive wealth, using output above, the IW is `output$iw`. Show that this is the same as quantity of water times shadow price.

11. What really matters is change in IW.
    a. Draw figure by hand that helps you figure out the change in IW. Remember it is meant as a linear approximation of the change in area under the price curve (really in a multi-dimensional space, but we are only working with one stock).
    b. Is it possible to subtract two different capn IW values to get the change, why or why not?
    c. How could you more accurately compute the change in inclusive wealth? Demonstrate this approach to measuring the change in inclusive wealth going from 21.5 to 20.5 feet and then from 21.5 to 18.5 feet for the representative acre.
    d. capn also approximates the intertemporal welfare, but this is only meaningful for changes. This is called output$vfun. However, if you try this you will get an error. Re-run with some extra information

```
output<-psim(pcoeff = pC,
        stock = 21.5,
        wval = profit(21.5, gw.data),
        sdot = sdot(21.5, recharge, gw.data)
        )
```

    e. Now revisit part c and compare difference between the vfun value and the iw values and compare. It maybe helpful to consider the difference between vfun at 18.5 and 21.5 feet.
    f. Draw a picture (not using R) that explains why the difference between vfun and iw are different.

Let's consider a policy shift. What if crop subsidizes were used for water conservation. Corn is the most commonly grown crop in the system and is relatively water intensive. Wheat is also grown, and it is the least water intensive crop. Let's first consider a water motivated tax on corn and then a water motivated subsidy on wheat.

12. First, copy the data. This will let you go back to the original data set.
    a. `ksdata.alt <- ksdata`
13. Next, identify the corn price to do this type
    a. `ksdata.alt$State$mlogitmeans[14,2] <- 0.9*ksdata.alt$State$mlogitmeans[14,2]`
       state is here because we are focusing on the state subset of the data. We are changing the price in the crop choice model (the multinomial logit model) but not the profit model, because we assume the revenue is fully recycled to offset income effects.
14. The wheat price is the 17th price in the list, extract that to see 3.38.

Let's first consider a 10% tax on corn revenue.

15. We will do this by modifying ksdata with the following code
    a. `ksdata.alt$State$mlogitmeans[14,2] <- 0.9*ksdata.alt$State$mlogitmeans[14,2]`
16. Now create gw.data.alt following the way we did it earlier, with the datasetup function but add the option `dataset = ksdata.alt`.
17. How does the tax change affect the amount of water withdrawn if there is 21.5 acre feet of water?
18. How does the tax change to the total acreage planted for the representative well at 21.5 acre feet?
19. How does the tax change the share of crops planted and irrigated at 21.5 acre feet?
20. How does the tax change farm profit for the representative acre [use the function profit(water, parameters)]. Don't worry about adding back in the tax, focus on the private farm profit.

21. How does the tax affect the shadow price of water?
22. Does the tax increase the wealth stored in water? What about sustainability?
23. Choose a level to subsidize wheat, and repeat the exercise.