

Find Closest Values Along Grids

Fan Wang

2021-01-10

Contents

1	Find Closest Neighbor on Grid	1
1.1	Closest Neighbor on Grid	1

1 Find Closest Neighbor on Grid

Go to the [RMD](#), [R](#), [PDF](#), or [HTML](#) version of this file. Go back to [fan's REconTools Package](#), [R Code Examples Repository \(bookdown site\)](#), or [Intro Stats with R Repository \(bookdown site\)](#).

Using the [pivot_wider](#) function in [tidyr](#) to reshape panel or other data structures

1.1 Closest Neighbor on Grid

There is a dataframe that provides $V(coh, a, cev)$ levels. There is another dataframe with $\hat{V}(coh, a)$, for each coh, a , find the cev that such that the difference between $\hat{V}(coh, a)$ and $V(coh, a, cev)$ is minimized.

V and \hat{V} information are stored in a dataframe in the [csv folder](#) in the current directory. In fact, we have one V surface, but multiple \hat{V} files, so we want to do the find closest neighbor exercise for each one of the several \hat{V} files.

The structure is as follows: (1) Load in the V file, where coh, a, cev are all variable attributes. (2) Merge with one \hat{V} file. (3) Take the difference between the V and \hat{V} columns, and take the absolute value of the difference. (4) Group by coh, a , and sort to get the smallest absolute difference among the cev possibilities, and slice out the row for the smallest. (5) Now We have $V(coh, a, cev^*(coh, a))$. (6) Do this for each of the several \hat{V} files. (7) Stack the results from 1 through 6 together, generate a column that identifies which simulation/exercise/counterfactual each of the \hat{V} file comes from. (8) Visualize by plotting as subplot different a, coh is x-axis, different \hat{V} outcome are different lines, and $cev^*(coh, a, \hat{V})$ is the y-axis outcome.

First, load the CEV file.

```
# folder
spt_root <- c('C:/Users/fan/R4Econ/panel/join/_file/csv')
# cev surface file, the V file
snm_cev_surface <- 'e_19E1NEp99r99_ITG_PE_cev_subsettest.csv'
mt_cev_surface <- read.csv(file = file.path(spt_root, snm_cev_surface))
tb_cev_surface <- as_tibble(mt_cev_surface) %>%
  rename(EjVcev = EjV)
```

Second, loop over the V hat files, join V with V hat:

```
ls_tb_cev_surfhat = vector(mode = "list", length = 4)
for (it_simu_counter in c(1,2,3,4)) {
```

```

# conditionally change file names
if (it_simu_counter == 1) {
  st_counter <- '19E1NEp99r99'
} else if (it_simu_counter == 2) {
  st_counter <- '19E1NEp02r99'
} else if (it_simu_counter == 3) {
  st_counter <- '19E1NEp02per02ger99'
} else if (it_simu_counter == 4) {
  st_counter <- '19E1NEp02r02'
}
snm_v_hat <- paste0('e_', st_counter, '_ITG_PE_subsettest.csv')

# Overall path to files
mt_v_hat <- read.csv(file = file.path(spt_root, snm_v_hat))
tb_v_hat <- as_tibble(mt_v_hat) %>%
  select(prod_type_lvl, statesid, EjV)

# Merge file using key
tb_cev_surfhathat <- tb_cev_surface %>%
  left_join(tb_v_hat, by=(c('prod_type_lvl'='prod_type_lvl',
                           'statesid'='statesid')) %>%
    arrange(statesid, prod_type_lvl, cev_lvl) %>%
    mutate(counter_policy = st_counter)

# Store to list
ls_tb_cev_surfhathat[[it_simu_counter]] <- tb_cev_surfhathat
}

# Display
kable(ls_tb_cev_surfhathat[[1]][seq(1, 40, 5),]) %>% kable_styling_fc_wide()

```

X	cev_st	cev_lvl	prod_type_st	prod_type_lvl	statesid	cash_tt	EjVcev	EjV	counter_policy
1	cev-2000	-0.2000	A0	0	526	32.84747	-1.0479929	-0.7957419	19E1NEp99r99
1501	cev-947	-0.0947	A0	0	526	32.84747	-0.9079859	-0.7957419	19E1NEp99r99
3001	cev105	0.0105	A0	0	526	32.84747	-0.7880156	-0.7957419	19E1NEp99r99
4501	cev1157	0.1157	A0	0	526	32.84747	-0.6803586	-0.7957419	19E1NEp99r99
51	cev-2000	-0.2000	A2504	2504	526	32.90371	-1.0002921	-0.7504785	19E1NEp99r99
1551	cev-947	-0.0947	A2504	2504	526	32.90371	-0.8613743	-0.7504785	19E1NEp99r99
3051	cev105	0.0105	A2504	2504	526	32.90371	-0.7423281	-0.7504785	19E1NEp99r99
4551	cev1157	0.1157	A2504	2504	526	32.90371	-0.6354620	-0.7504785	19E1NEp99r99

Third, sort each file, and keep only the best match rows that minimize the absolute distance between *EjV* and *EjVcev*.

```

ls_tb_cev_matched = vector(mode = "list", length = 4)
for (it_simu_counter in c(1,2,3,4)) {

  # Load merged file
  tb_cev_surfhathat <- ls_tb_cev_surfhathat[[it_simu_counter]]

  # Difference Column
  tb_cev_surfhathat <- tb_cev_surfhathat %>%
    mutate(EjVcev_gap = abs(EjVcev - EjV))

  # Group by, Arrange and Slice, get lowest gap

```

```

tb_cev_matched <- tb_cev_surfhat %>%
  arrange(statesid, prod_type_lvl, EjVcev_gap) %>%
  group_by(statesid, prod_type_lvl) %>%
  slice_head(n=1)

# Store to list
ls_tb_cev_matched[[it_simu_counter]] <- tb_cev_matched
}

```

```

# Display
kable(ls_tb_cev_matched[[2]][seq(1, 30, 1),]) %>% kable_styling_fc_wide()

```

X	cev_st	cev_lvl	prod_type_st	prod_type_lvl	statesid	cash_tt	EjVcev	EjV	counter_policy	EjVcev_gap
3001	cev105	0.0105	A0	0	526	32.847471	-0.7880156	-0.7928034	19E1NEp02r99	0.0047878
3051	cev105	0.0105	A2504	2504	526	32.903714	-0.7423281	-0.7480617	19E1NEp02r99	0.0057336
3101	cev105	0.0105	A4145	4145	526	32.948970	-0.7082006	-0.7145418	19E1NEp02r99	0.0063412
3151	cev105	0.0105	A5633	5633	526	32.996952	-0.6753576	-0.6818996	19E1NEp02r99	0.0065420
3201	cev105	0.0105	A7274	7274	526	33.058832	-0.6368297	-0.6431710	19E1NEp02r99	0.0063413
3251	cev105	0.0105	A9779	9779	526	33.175241	-0.5711706	-0.5774648	19E1NEp02r99	0.0062942
3002	cev105	0.0105	A0	0	555	53.346587	-0.2985944	-0.3041922	19E1NEp02r99	0.0055978
3052	cev105	0.0105	A2504	2504	555	53.815772	-0.2617572	-0.2680026	19E1NEp02r99	0.0062454
3102	cev105	0.0105	A4145	4145	555	54.193302	-0.2340822	-0.2406142	19E1NEp02r99	0.0065320
3152	cev105	0.0105	A5633	5633	555	54.593579	-0.2067964	-0.2134634	19E1NEp02r99	0.0066670
3202	cev105	0.0105	A7274	7274	555	55.109790	-0.1740126	-0.1806320	19E1NEp02r99	0.0066194
3252	cev105	0.0105	A9779	9779	555	56.080888	-0.1169470	-0.1236111	19E1NEp02r99	0.0066641
3603	cev526	0.0526	A0	0	905	1.533025	-5.2530406	-5.2486887	19E1NEp02r99	0.0043519
3353	cev315	0.0315	A2504	2504	905	1.714498	-4.5517474	-4.5408560	19E1NEp02r99	0.0108913
3403	cev315	0.0315	A4145	4145	905	1.860521	-4.1039608	-4.1072736	19E1NEp02r99	0.0033128
3453	cev315	0.0315	A5633	5633	905	2.015341	-3.7465733	-3.7611842	19E1NEp02r99	0.0146109
3503	cev315	0.0315	A7274	7274	905	2.215003	-3.4101025	-3.4235413	19E1NEp02r99	0.0134388
3553	cev315	0.0315	A9779	9779	905	2.590608	-2.9413469	-2.9535570	19E1NEp02r99	0.0122101
3004	cev105	0.0105	A0	0	953	20.125381	-1.3249909	-1.3290865	19E1NEp02r99	0.0040957
3054	cev105	0.0105	A2504	2504	953	20.306854	-1.2476021	-1.2531860	19E1NEp02r99	0.0055839
3104	cev105	0.0105	A4145	4145	953	20.452876	-1.1916003	-1.1975215	19E1NEp02r99	0.0059211
3154	cev105	0.0105	A5633	5633	953	20.607697	-1.1383665	-1.1444048	19E1NEp02r99	0.0060383
3204	cev105	0.0105	A7274	7274	953	20.807359	-1.0766095	-1.0823344	19E1NEp02r99	0.0057250
3254	cev105	0.0105	A9779	9779	953	21.182964	-0.9729832	-0.9781408	19E1NEp02r99	0.0051576
3005	cev105	0.0105	A0	0	1017	63.774766	-0.1284542	-0.1342653	19E1NEp02r99	0.0058110
3055	cev105	0.0105	A2504	2504	1017	64.298911	-0.0967695	-0.1031112	19E1NEp02r99	0.0063417
3105	cev105	0.0105	A4145	4145	1017	64.720664	-0.0728485	-0.0793940	19E1NEp02r99	0.0065454
3155	cev105	0.0105	A5633	5633	1017	65.167829	-0.0490898	-0.0557238	19E1NEp02r99	0.0066341
3205	cev105	0.0105	A7274	7274	1017	65.744507	-0.0203378	-0.0269149	19E1NEp02r99	0.0065772
3255	cev105	0.0105	A9779	9779	1017	66.829359	0.0299397	0.0233507	19E1NEp02r99	0.0065890

Fourth, row_bind results together.

```

# Single dataframe with all results
tb_cev_matched_all_counter <- do.call(bind_rows, ls_tb_cev_matched)
# check size
print(dim(tb_cev_matched_all_counter))

```

```
## [1] 1200 11
```

Fifth, visualize results

```

# select four from the productivity types
ar_prod_type_lvl_unique <- unique(tb_cev_matched_all_counter %>% pull(prod_type_lvl))
ar_prod_type_lvl_selected <- ar_prod_type_lvl_unique[round(seq(1, length(ar_prod_type_lvl_unique), length=4))]
# graph
lineplot <- tb_cev_matched_all_counter %>%
  filter(prod_type_lvl %in% ar_prod_type_lvl_selected) %>%
  group_by(prod_type_st, cash_tt) %>%
  ggplot(aes(x=cash_tt, y=cev_lvl,

```

```

    colour=counter_policy, linetype=counter_policy, shape=counter_policy)) +
  facet_wrap( ~ prod_type_st) +
  geom_line() +
  geom_point() +
  labs(title = 'Visualizing the positions of matched values',
       x = 'Resource Levels',
       y = 'CEV',
       caption = paste0('https://fanwangecon.github.io/',
                        'R4Econ/panel/join/htmlpdf/fr/fs_join_compare.html'))
print(lineplot)

```

