

```
Last login: Fri Apr 10 13:05:51 on ttys000
Run-Mac:~ mac$ cd ~/.ssh
Run-Mac:~.ssh mac$ ssh -i "Runzhe.pem" ubuntu@ec2-18-208-204-26.compute-1.amazonaws.com
The authenticity of host 'ec2-18-208-204-26.compute-1.amazonaws.com (18.208.204.26)' can't be established.
ECDSA key fingerprint is SHA256:1YgeH485B8lAc24NSJ5PTwVcA1G5W3X0xqyxh7guhck.
Are you sure you want to continue connecting (yes/no)? yes
Warning: Permanently added 'ec2-18-208-204-26.compute-1.amazonaws.com,18.208.204.26' (ECDSA) to the list of known hosts.
Welcome to Ubuntu 18.04.3 LTS (GNU/Linux 4.15.0-1060-aws x86_64)
```

```
* Documentation:  https://help.ubuntu.com
* Management:    https://landscape.canonical.com
* Support:       https://ubuntu.com/advantage
```

System information as of Fri Apr 10 17:11:46 UTC 2020

```
System load:  1.21           Processes:      227
Usage of /:   28.2% of 30.96GB Users logged in: 0
Memory usage: 1%            IP address for ens5: 172.31.2.198
Swap usage:   0%
```

```
* Kubernetes 1.18 GA is now available! See https://microk8s.io for docs or
  install it with:
```

```
sudo snap install microk8s --channel=1.18 --classic
```

```
* Multipass 1.1 adds proxy support for developers behind enterprise
  firewalls. Rapid prototyping for cloud operations just got easier.
```

```
https://multipass.run/
```

```
* Canonical Livepatch is available for installation.
  - Reduce system reboots and improve kernel security. Activate at:
    https://ubuntu.com/livepatch
```

```
89 packages can be updated.
38 updates are security updates.
```

```
Last login: Fri Apr  3 19:45:17 2020 from 107.13.161.147
ubuntu@ip-172-31-2-198:~$ export openblas_num_threads=1; export OMP_NUM_THREADS=1; python EC2.py
13:13, 04/10; num of cores:16sd_u_0_10
```

```
Basic setting:[rep_times, sd_0, sd_D, sd_u_0, w_0, w_A, u_0_u_D, sd_R_range, t_func] = [16, None, None, 10, 0.5, 1, 0, [0, 10, 20], None
]
```

```
-----
[pattern_seed, day, sd_R] = [2, 7, 0]
```

```
max(u_0) = 122.9
0_threshold = 90
number of reward locations: 20
0_threshold = 100
number of reward locations: 9
0_threshold = 110
number of reward locations: 3
target 1 in 3 DONE!
target 2 in 3 DONE!
target 3 in 3 DONE!
```

```
-----
Value of Behaviour policy:67.182
0_threshold = 90
MC for this TARGET:[76.906, 0.089]
[DR/QV/IS]; [DR_NO_MARL, DR_NO_MF, V_behav]
bias:[[1.42, 1.23, 0.66]][[2.42, -76.91, -9.72]]
std:[[0.69, 0.69, 0.34]][[0.32, 0.0, 0.24]]
MSE:[[1.58, 1.41, 0.74]][[2.44, 76.91, 9.72]]
MSE(-DR):[[0.0, -0.17, -0.84]][[0.86, 75.33, 8.14]]
***
```

```
=====
0_threshold = 100
MC for this TARGET:[70.073, 0.074]
[DR/QV/IS]; [DR_NO_MARL, DR_NO_MF, V_behav]
bias:[[-0.37, -0.41, -0.98]][[-0.96, -70.07, -2.89]]
std:[[0.49, 0.48, 0.37]][[0.3, 0.0, 0.24]]
MSE:[[0.61, 0.63, 1.05]][[1.01, 70.07, 2.9]]
MSE(-DR):[[0.0, 0.02, 0.44]][[0.4, 69.46, 2.29]]
***
```

```
=====
0_threshold = 110
MC for this TARGET:[72.699, 0.076]
[DR/QV/IS]; [DR_NO_MARL, DR_NO_MF, V_behav]
bias:[[-5.65, -5.64, -5.77]][[-9.93, -72.7, -5.52]]
std:[[0.75, 0.79, 0.32]][[0.27, 0.0, 0.24]]
MSE:[[5.7, 5.7, 5.78]][[9.93, 72.7, 5.53]]
MSE(-DR):[[0.0, 0.0, 0.08]][[4.23, 67.0, -0.17]]
***
=====
```

```
[[ 1.58  1.41  0.74  2.44 76.91  9.72]
 [ 0.61  0.63  1.05  1.01 70.07  2.9 ]
 [ 5.7   5.7   5.78  9.93 72.7   5.53]]
```

time spent until now: 31.1 mins

13:44, 04/10

```
-----
[pattern_seed, day, sd_R] = [2, 7, 10]
```

```
max(u_0) = 122.9
0_threshold = 90
number of reward locations: 20
0_threshold = 100
number of reward locations: 9
0_threshold = 110
number of reward locations: 3
target 1 in 3 DONE!
target 2 in 3 DONE!
target 3 in 3 DONE!
```

```
-----
Value of Behaviour policy:67.167
0_threshold = 90
MC for this TARGET:[76.922, 0.147]
[DR/QV/IS]; [DR_NO_MARL, DR_NO_MF, V_behav]
bias:[[1.31, 1.13, 0.65]][[2.45, -76.92, -9.75]]
std:[[0.83, 0.85, 0.52]][[0.37, 0.0, 0.19]]
MSE:[[1.55, 1.41, 0.83]][[2.48, 76.92, 9.75]]
MSE(-DR):[[0.0, -0.14, -0.72]][[0.93, 75.37, 8.2]]
***
=====
0_threshold = 100
MC for this TARGET:[70.089, 0.139]
[DR/QV/IS]; [DR_NO_MARL, DR_NO_MF, V_behav]
bias:[[-0.26, -0.31, -0.97]][[-0.93, -70.09, -2.92]]
std:[[0.64, 0.63, 0.35]][[0.28, 0.0, 0.19]]
MSE:[[0.69, 0.7, 1.03]][[0.97, 70.09, 2.93]]
MSE(-DR):[[0.0, 0.01, 0.34]][[0.28, 69.4, 2.24]]
***
=====
0_threshold = 110
MC for this TARGET:[72.715, 0.131]
[DR/QV/IS]; [DR_NO_MARL, DR_NO_MF, V_behav]
bias:[[-5.36, -5.37, -5.79]][[-9.93, -72.72, -5.55]]
std:[[0.93, 1.0, 0.38]][[0.28, 0.0, 0.19]]
MSE:[[5.44, 5.46, 5.8]][[9.93, 72.72, 5.55]]
MSE(-DR):[[0.0, 0.02, 0.36]][[4.49, 67.28, 0.11]]
***
=====
[[ 1.58  1.41  0.74  2.44 76.91  9.72]
 [ 0.61  0.63  1.05  1.01 70.07  2.9 ]
 [ 5.7   5.7   5.78  9.93 72.7   5.53]]
```

```
[[ 1.55  1.41  0.83  2.48 76.92  9.75]
 [ 0.69  0.7   1.03  0.97 70.09  2.93]
 [ 5.44  5.46  5.8   9.93 72.72  5.55]]
```

time spent until now: 62.1 mins

14:15, 04/10

```
-----
[pattern_seed, day, sd_R] = [2, 7, 20]
```

```
max(u_0) = 122.9
0_threshold = 90
number of reward locations: 20
0_threshold = 100
number of reward locations: 9
0_threshold = 110
number of reward locations: 3
target 1 in 3 DONE!
target 2 in 3 DONE!
^CProcess Process-44:
Traceback (most recent call last):
  File "EC2.py", line 81, in <module>
Process Process-35:
Process Process-33:
Process Process-34:
  with_MF = with_MF, with_NO_MARL = with_NO_MARL, with_IS = with_IS,
  File "/home/ubuntu/simu_funs.py", line 63, in simu
    value_reps = parmap(once, range(OPE_rep_times), n_cores)
  File "/home/ubuntu/_uti_basic.py", line 80, in parmap
    [q_in.put((None, None)) for _ in range(nprocs)]
```

```

File "/home/ubuntu/_uti_basic.py", line 80, in <listcomp>
    [q_in.put((None, None)) for _ in range(nprocs)]
File "/home/ubuntu/anaconda3/lib/python3.7/multiprocessing/queues.py", line 82, in put
Process Process-43:
Process Process-45:
Process Process-38:
Process Process-40:
Process Process-47:
Process Process-42:
Process Process-37:
Process Process-39:
Traceback (most recent call last):
  File "/home/ubuntu/anaconda3/lib/python3.7/multiprocessing/process.py", line 297, in _bootstrap
    self.run()
  File "/home/ubuntu/anaconda3/lib/python3.7/multiprocessing/process.py", line 99, in run
    self._target(*self._args, **self._kwargs)
  File "/home/ubuntu/_uti_basic.py", line 67, in fun
    q_out.put((i, f(x)))
  File "/home/ubuntu/simu_funs.py", line 61, in once
    inner_parallel = inner_parallel)
  File "/home/ubuntu/simu_funs.py", line 213, in simu_once
    inner_parallel = inner_parallel)
  File "/home/ubuntu/main.py", line 157, in V_DR
    r = arr([getOneRegionValue(i) for i in range(N)])
  File "/home/ubuntu/main.py", line 157, in <listcomp>
    r = arr([getOneRegionValue(i) for i in range(N)])
  File "/home/ubuntu/main.py", line 79, in getOneRegionValue
    CV_QV = CV_QV, penalty_range = penalty, spatial = True)
  File "/home/ubuntu/main.py", line 302, in computeQV
    validation_set = valid_tuples)
  File "/home/ubuntu/main.py", line 429, in computeQV_basic
    left = (ECKQ1.T.dot(ECKQ1) + np.vstack((np.hstack((T * lam * KQ, zeros((2 * T, 1)))), zeros((1, 2 * T + 1)))))) # Left part of (\hat{\alpha}, \hat{\eta})
KeyboardInterrupt
Traceback (most recent call last):
  File "/home/ubuntu/anaconda3/lib/python3.7/multiprocessing/process.py", line 297, in _bootstrap
    self.run()
  File "/home/ubuntu/anaconda3/lib/python3.7/multiprocessing/process.py", line 99, in run
    self._target(*self._args, **self._kwargs)
  File "/home/ubuntu/_uti_basic.py", line 67, in fun
    q_out.put((i, f(x)))
  File "/home/ubuntu/simu_funs.py", line 61, in once
    inner_parallel = inner_parallel)
  File "/home/ubuntu/simu_funs.py", line 213, in simu_once
    inner_parallel = inner_parallel)
Traceback (most recent call last):
Traceback (most recent call last):
Traceback (most recent call last):
Traceback (most recent call last):
  File "/home/ubuntu/main.py", line 157, in V_DR
    r = arr([getOneRegionValue(i) for i in range(N)])
Traceback (most recent call last):
Traceback (most recent call last):
  File "/home/ubuntu/main.py", line 157, in <listcomp>
    r = arr([getOneRegionValue(i) for i in range(N)])
Traceback (most recent call last):
  File "/home/ubuntu/main.py", line 113, in getOneRegionValue
    spatial = False)
  File "/home/ubuntu/main.py", line 261, in getWeight
    epsilon = epsilon, spatial = spatial, mean_field = mean_field)
  File "/home/ubuntu/weight.py", line 301, in train
    self.policy_ratio2: policy_ratio2
  File "/home/ubuntu/anaconda3/lib/python3.7/multiprocessing/process.py", line 297, in _bootstrap
    self.run()
  File "/home/ubuntu/anaconda3/lib/python3.7/multiprocessing/process.py", line 297, in _bootstrap
    self.run()
  File "/home/ubuntu/anaconda3/lib/python3.7/multiprocessing/process.py", line 297, in _bootstrap
    self.run()
  File "/home/ubuntu/.local/lib/python3.7/site-packages/tensorflow/python/client/session.py", line 950, in run
    run_metadata_ptr)
  File "/home/ubuntu/anaconda3/lib/python3.7/multiprocessing/process.py", line 297, in _bootstrap
    self.run()
  File "/home/ubuntu/anaconda3/lib/python3.7/multiprocessing/process.py", line 99, in run
    self._target(*self._args, **self._kwargs)
  File "/home/ubuntu/anaconda3/lib/python3.7/multiprocessing/process.py", line 99, in run
    self._target(*self._args, **self._kwargs)
  File "/home/ubuntu/anaconda3/lib/python3.7/multiprocessing/process.py", line 297, in _bootstrap
    self.run()
  File "/home/ubuntu/anaconda3/lib/python3.7/multiprocessing/process.py", line 99, in run
    self._target(*self._args, **self._kwargs)
  File "/home/ubuntu/.local/lib/python3.7/site-packages/tensorflow/python/client/session.py", line 1173, in _run
    feed_dict_tensor, options, run_metadata)
  File "/home/ubuntu/anaconda3/lib/python3.7/multiprocessing/process.py", line 297, in _bootstrap
    self.run()
  File "/home/ubuntu/anaconda3/lib/python3.7/multiprocessing/process.py", line 297, in _bootstrap
    self.run()
  File "/home/ubuntu/anaconda3/lib/python3.7/multiprocessing/process.py", line 99, in run
    self._target(*self._args, **self._kwargs)
  File "/home/ubuntu/_uti_basic.py", line 67, in fun

```

```

    q_out.put((i, f(x)))
File "/home/ubuntu/_uti_basic.py", line 67, in fun
    q_out.put((i, f(x)))
File "/home/ubuntu/anaconda3/lib/python3.7/multiprocessing/process.py", line 99, in run
    self._target(*self._args, **self._kwargs)
File "/home/ubuntu/_uti_basic.py", line 67, in fun
    q_out.put((i, f(x)))
File "/home/ubuntu/.local/lib/python3.7/site-packages/tensorflow/python/client/session.py", line 1350, in _do_run
    run_metadata)
Traceback (most recent call last):
File "/home/ubuntu/anaconda3/lib/python3.7/multiprocessing/process.py", line 99, in run
    self._target(*self._args, **self._kwargs)
File "/home/ubuntu/anaconda3/lib/python3.7/multiprocessing/process.py", line 99, in run
    self._target(*self._args, **self._kwargs)
File "/home/ubuntu/_uti_basic.py", line 67, in fun
    q_out.put((i, f(x)))
File "/home/ubuntu/simu_funs.py", line 61, in once
    inner_parallel = inner_parallel)
File "/home/ubuntu/simu_funs.py", line 61, in once
    inner_parallel = inner_parallel)
File "/home/ubuntu/_uti_basic.py", line 67, in fun
    q_out.put((i, f(x)))
File "/home/ubuntu/simu_funs.py", line 61, in once
    inner_parallel = inner_parallel)
File "/home/ubuntu/.local/lib/python3.7/site-packages/tensorflow/python/client/session.py", line 1356, in _do_call
    return fn(*args)
File "/home/ubuntu/_uti_basic.py", line 67, in fun
    q_out.put((i, f(x)))
File "/home/ubuntu/_uti_basic.py", line 67, in fun
    q_out.put((i, f(x)))
File "/home/ubuntu/simu_funs.py", line 61, in once
    inner_parallel = inner_parallel)
File "/home/ubuntu/simu_funs.py", line 213, in simu_once
    inner_parallel = inner_parallel)
File "/home/ubuntu/simu_funs.py", line 213, in simu_once
    inner_parallel = inner_parallel)
File "/home/ubuntu/simu_funs.py", line 213, in simu_once
    inner_parallel = inner_parallel)
File "/home/ubuntu/simu_funs.py", line 61, in once
    inner_parallel = inner_parallel)
File "/home/ubuntu/simu_funs.py", line 213, in simu_once
    inner_parallel = inner_parallel)
File "/home/ubuntu/.local/lib/python3.7/site-packages/tensorflow/python/client/session.py", line 1341, in _run_fn
    options, feed_dict, fetch_list, target_list, run_metadata)
File "/home/ubuntu/anaconda3/lib/python3.7/multiprocessing/process.py", line 297, in _bootstrap
    self.run()
File "/home/ubuntu/simu_funs.py", line 61, in once
    inner_parallel = inner_parallel)
File "/home/ubuntu/simu_funs.py", line 61, in once
    inner_parallel = inner_parallel)
File "/home/ubuntu/simu_funs.py", line 213, in simu_once
    inner_parallel = inner_parallel)
File "/home/ubuntu/main.py", line 157, in V_DR
    r = arr([getOneRegionValue(i) for i in range(N)])
File "/home/ubuntu/main.py", line 157, in V_DR
    r = arr([getOneRegionValue(i) for i in range(N)])
File "/home/ubuntu/simu_funs.py", line 213, in simu_once
    inner_parallel = inner_parallel)
File "/home/ubuntu/main.py", line 157, in V_DR
    r = arr([getOneRegionValue(i) for i in range(N)])
File "/home/ubuntu/.local/lib/python3.7/site-packages/tensorflow/python/client/session.py", line 1429, in _call_tf_sessionrun
    run_metadata)
File "/home/ubuntu/anaconda3/lib/python3.7/multiprocessing/process.py", line 99, in run
    self._target(*self._args, **self._kwargs)
File "/home/ubuntu/simu_funs.py", line 213, in simu_once
    inner_parallel = inner_parallel)
File "/home/ubuntu/simu_funs.py", line 213, in simu_once
    inner_parallel = inner_parallel)
File "/home/ubuntu/main.py", line 157, in V_DR
    r = arr([getOneRegionValue(i) for i in range(N)])
File "/home/ubuntu/main.py", line 157, in <listcomp>
    r = arr([getOneRegionValue(i) for i in range(N)])
File "/home/ubuntu/main.py", line 157, in <listcomp>
    r = arr([getOneRegionValue(i) for i in range(N)])
File "/home/ubuntu/main.py", line 157, in V_DR
    r = arr([getOneRegionValue(i) for i in range(N)])
File "/home/ubuntu/main.py", line 157, in <listcomp>
    r = arr([getOneRegionValue(i) for i in range(N)])
File "/home/ubuntu/main.py", line 157, in <listcomp>
    r = arr([getOneRegionValue(i) for i in range(N)])
File "/home/ubuntu/_uti_basic.py", line 67, in fun
    q_out.put((i, f(x)))
KeyboardInterrupt
File "/home/ubuntu/main.py", line 157, in V_DR
    r = arr([getOneRegionValue(i) for i in range(N)])
File "/home/ubuntu/main.py", line 157, in V_DR
    r = arr([getOneRegionValue(i) for i in range(N)])
File "/home/ubuntu/main.py", line 157, in <listcomp>
    r = arr([getOneRegionValue(i) for i in range(N)])
File "/home/ubuntu/main.py", line 113, in getOneRegionValue
    spatial = False)
File "/home/ubuntu/main.py", line 113, in getOneRegionValue
    spatial = False)

```

```

File "/home/ubuntu/main.py", line 157, in <listcomp>
    r = arr([getOneRegionValue(i) for i in range(N)])
File "/home/ubuntu/main.py", line 86, in getOneRegionValue
    epsilon = epsilon)
File "/home/ubuntu/simu_funs.py", line 61, in once
    inner_parallel = inner_parallel)
File "/home/ubuntu/main.py", line 157, in <listcomp>
    r = arr([getOneRegionValue(i) for i in range(N)])
File "/home/ubuntu/main.py", line 157, in <listcomp>
    r = arr([getOneRegionValue(i) for i in range(N)])
File "/home/ubuntu/main.py", line 113, in getOneRegionValue
    spatial = False)
File "/home/ubuntu/main.py", line 261, in getWeight
    epsilon = epsilon, spatial = spatial, mean_field = mean_field)
File "/home/ubuntu/main.py", line 261, in getWeight
    epsilon = epsilon, spatial = spatial, mean_field = mean_field)
File "/home/ubuntu/main.py", line 113, in getOneRegionValue
    spatial = False)
File "/home/ubuntu/main.py", line 261, in getWeight
    epsilon = epsilon, spatial = spatial, mean_field = mean_field)
File "/home/ubuntu/simu_funs.py", line 213, in simu_once
    inner_parallel = inner_parallel)
File "/home/ubuntu/main.py", line 86, in getOneRegionValue
    epsilon = epsilon)
File "/home/ubuntu/main.py", line 106, in getOneRegionValue
    CV_QV = CV_QV, penalty_range = penalty, spatial = False)
File "/home/ubuntu/main.py", line 261, in getWeight
    epsilon = epsilon, spatial = spatial, mean_field = mean_field)
File "/home/ubuntu/weight.py", line 301, in train
    self.policy_ratio2: policy_ratio2
File "/home/ubuntu/weight.py", line 301, in train
    self.policy_ratio2: policy_ratio2
File "/home/ubuntu/main.py", line 261, in getWeight
    epsilon = epsilon, spatial = spatial, mean_field = mean_field)
File "/home/ubuntu/weight.py", line 301, in train
    self.policy_ratio2: policy_ratio2
File "/home/ubuntu/main.py", line 157, in V_DR
    r = arr([getOneRegionValue(i) for i in range(N)])
File "/home/ubuntu/main.py", line 261, in getWeight
    epsilon = epsilon, spatial = spatial, mean_field = mean_field)
File "/home/ubuntu/main.py", line 302, in computeQV
    validation_set = valid_tuples)
File "/home/ubuntu/weight.py", line 301, in train
    self.policy_ratio2: policy_ratio2
File "/home/ubuntu/.local/lib/python3.7/site-packages/tensorflow/python/client/session.py", line 950, in run
    run_metadata_ptr)
File "/home/ubuntu/.local/lib/python3.7/site-packages/tensorflow/python/client/session.py", line 950, in run
    run_metadata_ptr)
File "/home/ubuntu/weight.py", line 301, in train
    self.policy_ratio2: policy_ratio2
File "/home/ubuntu/.local/lib/python3.7/site-packages/tensorflow/python/client/session.py", line 950, in run
    run_metadata_ptr)
File "/home/ubuntu/main.py", line 157, in <listcomp>
    r = arr([getOneRegionValue(i) for i in range(N)])
File "/home/ubuntu/weight.py", line 301, in train
    self.policy_ratio2: policy_ratio2
File "/home/ubuntu/main.py", line 432, in computeQV_basic
    alpha_eta = np.linalg.lstsq(left, np.expand_dims(right,1))[0]
File "/home/ubuntu/.local/lib/python3.7/site-packages/tensorflow/python/client/session.py", line 950, in run
    run_metadata_ptr)
File "/home/ubuntu/.local/lib/python3.7/site-packages/tensorflow/python/client/session.py", line 1173, in _run
    feed_dict_tensor, options, run_metadata)
File "/home/ubuntu/.local/lib/python3.7/site-packages/tensorflow/python/client/session.py", line 1173, in _run
    feed_dict_tensor, options, run_metadata)
File "/home/ubuntu/.local/lib/python3.7/site-packages/tensorflow/python/client/session.py", line 950, in run
    run_metadata_ptr)
File "/home/ubuntu/.local/lib/python3.7/site-packages/tensorflow/python/client/session.py", line 1173, in _run
    feed_dict_tensor, options, run_metadata)
File "/home/ubuntu/main.py", line 79, in getOneRegionValue
    CV_QV = CV_QV, penalty_range = penalty, spatial = True)
File "/home/ubuntu/.local/lib/python3.7/site-packages/tensorflow/python/client/session.py", line 950, in run
    run_metadata_ptr)
File "/home/ubuntu/anaconda3/lib/python3.7/site-packages/numpy/linalg/linalg.py", line 2236, in lstsq
    x, resids, rank, s = gufunc(a, b, rcond, signature=signature, extobj=extobj)
File "/home/ubuntu/.local/lib/python3.7/site-packages/tensorflow/python/client/session.py", line 1173, in _run
    feed_dict_tensor, options, run_metadata)
File "/home/ubuntu/.local/lib/python3.7/site-packages/tensorflow/python/client/session.py", line 1350, in _do_run
    run_metadata)
File "/home/ubuntu/.local/lib/python3.7/site-packages/tensorflow/python/client/session.py", line 1350, in _do_run
    run_metadata)
File "/home/ubuntu/.local/lib/python3.7/site-packages/tensorflow/python/client/session.py", line 1173, in _run
    feed_dict_tensor, options, run_metadata)
File "/home/ubuntu/.local/lib/python3.7/site-packages/tensorflow/python/client/session.py", line 1350, in _do_run
    run_metadata)
File "/home/ubuntu/main.py", line 302, in computeQV
    validation_set = valid_tuples)
File "/home/ubuntu/.local/lib/python3.7/site-packages/tensorflow/python/client/session.py", line 1173, in _run
    feed_dict_tensor, options, run_metadata)
File "/home/ubuntu/.local/lib/python3.7/site-packages/tensorflow/python/client/session.py", line 1350, in _do_run

```

```

run_metadata)
File "/home/ubuntu/.local/lib/python3.7/site-packages/tensorflow/python/client/session.py", line 1356, in _do_call
return fn(*args)
KeyboardInterrupt
File "/home/ubuntu/.local/lib/python3.7/site-packages/tensorflow/python/client/session.py", line 1356, in _do_call
return fn(*args)
File "/home/ubuntu/.local/lib/python3.7/site-packages/tensorflow/python/client/session.py", line 1350, in _do_run
run_metadata)
File "/home/ubuntu/.local/lib/python3.7/site-packages/tensorflow/python/client/session.py", line 1356, in _do_call
return fn(*args)
File "/home/ubuntu/main.py", line 401, in computeQV_basic
gamma_q = 1 / (2 * np.median(pdist(Z_tilde[:,:(Z_tilde.shape[1]-2)]))**2)
File "/home/ubuntu/.local/lib/python3.7/site-packages/tensorflow/python/client/session.py", line 1350, in _do_run
run_metadata)
File "/home/ubuntu/.local/lib/python3.7/site-packages/tensorflow/python/client/session.py", line 1356, in _do_call
return fn(*args)
File "/home/ubuntu/.local/lib/python3.7/site-packages/tensorflow/python/client/session.py", line 1341, in _run_fn
options, feed_dict, fetch_list, target_list, run_metadata)
File "/home/ubuntu/.local/lib/python3.7/site-packages/tensorflow/python/client/session.py", line 1341, in _run_fn
options, feed_dict, fetch_list, target_list, run_metadata)
File "/home/ubuntu/.local/lib/python3.7/site-packages/tensorflow/python/client/session.py", line 1356, in _do_call
return fn(*args)
File "/home/ubuntu/.local/lib/python3.7/site-packages/tensorflow/python/client/session.py", line 1341, in _run_fn
options, feed_dict, fetch_list, target_list, run_metadata)
File "/home/ubuntu/anaconda3/lib/python3.7/site-packages/scipy/spatial/distance.py", line 2066, in pdist
pdist_fn(X, dm, **kwargs)
File "/home/ubuntu/.local/lib/python3.7/site-packages/tensorflow/python/client/session.py", line 1356, in _do_call
return fn(*args)
File "/home/ubuntu/.local/lib/python3.7/site-packages/tensorflow/python/client/session.py", line 1341, in _run_fn
options, feed_dict, fetch_list, target_list, run_metadata)
File "/home/ubuntu/.local/lib/python3.7/site-packages/tensorflow/python/client/session.py", line 1429, in _call_tf_sessionrun
run_metadata)
File "/home/ubuntu/.local/lib/python3.7/site-packages/tensorflow/python/client/session.py", line 1429, in _call_tf_sessionrun
run_metadata)
File "/home/ubuntu/.local/lib/python3.7/site-packages/tensorflow/python/client/session.py", line 1341, in _run_fn
options, feed_dict, fetch_list, target_list, run_metadata)
File "/home/ubuntu/.local/lib/python3.7/site-packages/tensorflow/python/client/session.py", line 1429, in _call_tf_sessionrun
run_metadata)
File "/home/ubuntu/.local/lib/python3.7/site-packages/tensorflow/python/client/session.py", line 1341, in _run_fn
options, feed_dict, fetch_list, target_list, run_metadata)
File "/home/ubuntu/.local/lib/python3.7/site-packages/tensorflow/python/client/session.py", line 1429, in _call_tf_sessionrun
run_metadata)
KeyboardInterrupt
File "/home/ubuntu/.local/lib/python3.7/site-packages/tensorflow/python/client/session.py", line 1429, in _call_tf_sessionrun
run_metadata)
KeyboardInterrupt
KeyboardInterrupt
KeyboardInterrupt
KeyboardInterrupt
KeyboardInterrupt
if not self._sem.acquire(block, timeout):
KeyboardInterrupt
ubuntu@ip-172-31-2-198:~$ export openblas_num_threads=1; export OMP_NUM_THREADS=1; python EC2.py
14:42, 04/10; num of cores:16
vary_T

```

Basic setting:[rep_times, sd_0, sd_D, sd_u_0, w_0, w_A, u_0_u_D, sd_R_range, t_func] = [16, None, None, 20, 0.5, 1, 0, [40], None]

[pattern_seed, day, sd_R] = [2, 3, 40]

```

max(u_0) = 145.8
0_threshold = 100
number of reward locations: 9
0_threshold = 105
number of reward locations: 7
0_threshold = 110
number of reward locations: 6
0_threshold = 115
number of reward locations: 3
target 1 in 4 DONE!
target 2 in 4 DONE!
target 3 in 4 DONE!
target 4 in 4 DONE!

```

Value of Behaviour policy:64.586
0_threshold = 100
MC for this TARGET:[70.833, 0.737]
[DR/QV/IS]; [DR_NO_MARL, DR_NO_MF, V_behav]
bias:[[-1.05, -1.23, -1.33]][[-1.92, -70.83, -6.25]]
std:[[2.91, 2.85, 2.01]][[1.82, 0.0, 0.8]]
MSE:[3.09, 3.1, 2.41][[2.65, 70.83, 6.3]]
MSE(-DR):[[0.0, 0.01, -0.68]][[-0.44, 67.74, 3.21]]
=====

```

0_threshold = 105
MC for this TARGET:[71.839, 0.734]
  [DR/QV/IS]; [DR_NO_MARL, DR_NO_MF, V_behav]
bias:[[-3.8, -3.89, -3.65]][[-4.58, -71.84, -7.25]]
std:[[3.22, 3.19, 2.11]][[1.54, 0.0, 0.8]]
MSE:[4.98, 5.03, 4.22]][[4.83, 71.84, 7.29]]
MSE(-DR):[0.0, 0.05, -0.76]][[-0.15, 66.86, 2.31]]
=====
0_threshold = 110
MC for this TARGET:[70.941, 0.734]
  [DR/QV/IS]; [DR_NO_MARL, DR_NO_MF, V_behav]
bias:[[-3.81, -3.85, -3.69]][[-5.36, -70.94, -6.36]]
std:[[3.97, 3.89, 2.19]][[1.63, 0.0, 0.8]]
MSE:[5.5, 5.47, 4.29]][[5.6, 70.94, 6.41]]
MSE(-DR):[0.0, -0.03, -1.21]][[0.1, 65.44, 0.91]]
**
=====
0_threshold = 115
MC for this TARGET:[71.856, 0.716]
  [DR/QV/IS]; [DR_NO_MARL, DR_NO_MF, V_behav]
bias:[[-6.06, -6.02, -5.9]][[-10.23, -71.86, -7.27]]
std:[[5.29, 5.18, 3.28]][[1.63, 0.0, 0.8]]
MSE:[8.04, 7.94, 6.75]][[10.36, 71.86, 7.31]]
MSE(-DR):[0.0, -0.1, -1.29]][[2.32, 63.82, -0.73]]
**
=====
[[ 3.09  3.1   2.41  2.65 70.83  6.3 ]
 [ 4.98  5.03  4.22  4.83 71.84  7.29]
 [ 5.5   5.47  4.29  5.6  70.94  6.41]
 [ 8.04  7.94  6.75 10.36 71.86  7.31]]

```

time spent until now: 34.6 mins

15:17, 04/10

[*pattern_seed*, *day*, *sd_R*] = [2, 5, 40]

```

max(u_0) = 145.8
0_threshold = 100
number of reward locations: 9
0_threshold = 105
number of reward locations: 7
0_threshold = 110
number of reward locations: 6
0_threshold = 115
number of reward locations: 3
target 1 in 4 DONE!
target 2 in 4 DONE!
target 3 in 4 DONE!
target 4 in 4 DONE!

```

```

-----
Value of Behaviour policy:64.712
0_threshold = 100
MC for this TARGET:[70.784, 0.525]
  [DR/QV/IS]; [DR_NO_MARL, DR_NO_MF, V_behav]
bias:[[-1.01, -1.08, -1.73]][[-1.71, -70.78, -6.07]]
std:[[1.73, 1.74, 1.05]][[0.95, 0.0, 0.52]]
MSE:[2.0, 2.05, 2.02]][[1.96, 70.78, 6.09]]
MSE(-DR):[0.0, 0.05, 0.02]][[-0.04, 68.78, 4.09]]
=====
0_threshold = 105
MC for this TARGET:[71.788, 0.52]
  [DR/QV/IS]; [DR_NO_MARL, DR_NO_MF, V_behav]
bias:[[-3.28, -3.39, -3.85]][[-4.51, -71.79, -7.08]]
std:[[2.71, 2.59, 1.38]][[1.05, 0.0, 0.52]]
MSE:[4.25, 4.27, 4.09]][[4.63, 71.79, 7.1]]
MSE(-DR):[0.0, 0.02, -0.16]][[0.38, 67.54, 2.85]]
**
=====
0_threshold = 110
MC for this TARGET:[70.886, 0.522]
  [DR/QV/IS]; [DR_NO_MARL, DR_NO_MF, V_behav]
bias:[[-3.21, -3.26, -3.54]][[-5.22, -70.89, -6.17]]
std:[[2.28, 2.22, 1.38]][[1.07, 0.0, 0.52]]
MSE:[3.94, 3.94, 3.8]][[5.33, 70.89, 6.19]]
MSE(-DR):[0.0, 0.0, -0.14]][[1.39, 66.95, 2.25]]
**
=====
0_threshold = 115
MC for this TARGET:[71.8, 0.52]
  [DR/QV/IS]; [DR_NO_MARL, DR_NO_MF, V_behav]
bias:[[-6.42, -6.48, -5.98]][[-9.85, -71.8, -7.09]]
std:[[3.34, 3.3, 1.65]][[1.02, 0.0, 0.52]]
MSE:[7.24, 7.27, 6.2]][[9.9, 71.8, 7.11]]
MSE(-DR):[0.0, 0.03, -1.04]][[2.66, 64.56, -0.13]]
**

```

```
=====
[[ 3.09  3.1   2.41  2.65 70.83  6.3 ]
 [ 4.98  5.03  4.22  4.83 71.84  7.29]
 [ 5.5   5.47  4.29  5.6  70.94  6.41]
 [ 8.04  7.94  6.75 10.36 71.86  7.31]]
```

```
[[ 2.    2.05  2.02  1.96 70.78  6.09]
 [ 4.25  4.27  4.09  4.63 71.79  7.1 ]
 [ 3.94  3.94  3.8   5.33 70.89  6.19]
 [ 7.24  7.27  6.2   9.9  71.8   7.11]]
```

time spent until now: 69.7 mins

15:52, 04/10

```
-----
[pattern_seed, day, sd_R] = [2, 7, 40]
```

```
max(u_0) = 145.8
0_threshold = 100
number of reward locations: 9
0_threshold = 105
number of reward locations: 7
0_threshold = 110
number of reward locations: 6
0_threshold = 115
number of reward locations: 3
target 1 in 4 DONE!
target 2 in 4 DONE!
target 3 in 4 DONE!
target 4 in 4 DONE!
```

```
-----
Value of Behaviour policy:64.788
0_threshold = 100
MC for this TARGET:[70.839, 0.474]
[DR/QV/IS]; [DR_NO_MARL, DR_NO_MF, V_behav]
bias:[[-0.3, -0.38, -1.22]][[-1.44, -70.84, -6.05]]
std:[1.68, 1.74, 1.18]][[0.78, 0.0, 0.35]]
MSE:[1.71, 1.78, 1.7]][[1.64, 70.84, 6.06]]
MSE(-DR):[[0.0, 0.07, -0.01]][[-0.07, 69.13, 4.35]]
```

```
=====
0_threshold = 105
MC for this TARGET:[71.843, 0.469]
[DR/QV/IS]; [DR_NO_MARL, DR_NO_MF, V_behav]
bias:[[-1.97, -2.06, -3.63]][[-4.27, -71.84, -7.05]]
std:[2.41, 2.51, 1.0]][[0.77, 0.0, 0.35]]
MSE:[3.11, 3.25, 3.77]][[4.34, 71.84, 7.06]]
MSE(-DR):[[0.0, 0.14, 0.66]][[1.23, 68.73, 3.95]]
```

```
***
=====
0_threshold = 110
MC for this TARGET:[70.937, 0.469]
[DR/QV/IS]; [DR_NO_MARL, DR_NO_MF, V_behav]
bias:[[-2.28, -2.33, -3.31]][[-4.88, -70.94, -6.15]]
std:[1.98, 2.02, 1.14]][[0.82, 0.0, 0.35]]
MSE:[3.02, 3.08, 3.5]][[4.95, 70.94, 6.16]]
MSE(-DR):[[0.0, 0.06, 0.48]][[1.93, 67.92, 3.14]]
```

```
***
=====
0_threshold = 115
MC for this TARGET:[71.861, 0.467]
[DR/QV/IS]; [DR_NO_MARL, DR_NO_MF, V_behav]
bias:[[-5.66, -5.65, -5.75]][[-9.74, -71.86, -7.07]]
std:[2.18, 2.07, 1.75]][[0.86, 0.0, 0.35]]
MSE:[6.07, 6.02, 6.01]][[9.78, 71.86, 7.08]]
MSE(-DR):[[0.0, -0.05, -0.06]][[3.71, 65.79, 1.01]]
```

```
=====
[[ 3.09  3.1   2.41  2.65 70.83  6.3 ]
 [ 4.98  5.03  4.22  4.83 71.84  7.29]
 [ 5.5   5.47  4.29  5.6  70.94  6.41]
 [ 8.04  7.94  6.75 10.36 71.86  7.31]]
```

```
[[ 2.    2.05  2.02  1.96 70.78  6.09]
 [ 4.25  4.27  4.09  4.63 71.79  7.1 ]
 [ 3.94  3.94  3.8   5.33 70.89  6.19]
 [ 7.24  7.27  6.2   9.9  71.8   7.11]]
```

```
[[ 1.71  1.78  1.7   1.64 70.84  6.06]
 [ 3.11  3.25  3.77  4.34 71.84  7.06]
 [ 3.02  3.08  3.5   4.95 70.94  6.16]
 [ 6.07  6.02  6.01  9.78 71.86  7.08]]
```


time spent until now: 107.2 mins

16:30, 04/10

[pattern_seed, day, sd_R] = [2, 9, 40]

max(u_0) = 145.8
0_threshold = 100
number of reward locations: 9
0_threshold = 105
number of reward locations: 7
0_threshold = 110
number of reward locations: 6
0_threshold = 115
number of reward locations: 3
target 1 in 4 DONE!
target 2 in 4 DONE!
target 3 in 4 DONE!
target 4 in 4 DONE!

Value of Behaviour policy:64.743

0_threshold = 100
MC for this TARGET:[70.807, 0.394]
[DR/QV/IS]; [DR_NO_MARL, DR_NO_MF, V_behav]
bias:[[-0.95, -1.07, -1.24]][[-1.55, -70.81, -6.06]]
std:[[1.97, 1.89, 1.3]][[0.97, 0.0, 0.38]]
MSE:[2.19, 2.17, 1.8]][[1.83, 70.81, 6.07]]
MSE(-DR):[[0.0, -0.02, -0.39]][[-0.36, 68.62, 3.88]]

=====

0_threshold = 105
MC for this TARGET:[71.806, 0.393]
[DR/QV/IS]; [DR_NO_MARL, DR_NO_MF, V_behav]
bias:[[-3.19, -3.31, -3.53]][[-4.22, -71.81, -7.06]]
std:[2.83, 2.77, 1.49]][[0.88, 0.0, 0.38]]
MSE:[4.26, 4.32, 3.83]][[4.31, 71.81, 7.07]]
MSE(-DR):[[0.0, 0.06, -0.43]][[0.05, 67.55, 2.81]]

=====

0_threshold = 110
MC for this TARGET:[70.901, 0.391]
[DR/QV/IS]; [DR_NO_MARL, DR_NO_MF, V_behav]
bias:[[-3.55, -3.64, -3.41]][[-4.99, -70.9, -6.16]]
std:[2.72, 2.62, 1.29]][[0.91, 0.0, 0.38]]
MSE:[4.47, 4.48, 3.65]][[5.07, 70.9, 6.17]]
MSE(-DR):[[0.0, 0.01, -0.82]][[0.6, 66.43, 1.7]]

=====

0_threshold = 115
MC for this TARGET:[71.829, 0.397]
[DR/QV/IS]; [DR_NO_MARL, DR_NO_MF, V_behav]
bias:[[-7.34, -7.48, -6.15]][[-10.0, -71.83, -7.09]]
std:[1.96, 1.98, 1.16]][[0.93, 0.0, 0.38]]
MSE:[7.6, 7.74, 6.26]][[10.04, 71.83, 7.1]]
MSE(-DR):[[0.0, 0.14, -1.34]][[2.44, 64.23, -0.5]]

=====

| | | | | | | | |
|---|------|------|------|-------|-------|------|---|
| [| 3.09 | 3.1 | 2.41 | 2.65 | 70.83 | 6.3 |] |
| [| 4.98 | 5.03 | 4.22 | 4.83 | 71.84 | 7.29 |] |
| [| 5.5 | 5.47 | 4.29 | 5.6 | 70.94 | 6.41 |] |
| [| 8.04 | 7.94 | 6.75 | 10.36 | 71.86 | 7.31 |] |

| | | | | | | | |
|---|------|------|------|------|-------|------|---|
| [| 2. | 2.05 | 2.02 | 1.96 | 70.78 | 6.09 |] |
| [| 4.25 | 4.27 | 4.09 | 4.63 | 71.79 | 7.1 |] |
| [| 3.94 | 3.94 | 3.8 | 5.33 | 70.89 | 6.19 |] |
| [| 7.24 | 7.27 | 6.2 | 9.9 | 71.8 | 7.11 |] |

| | | | | | | | |
|---|------|------|------|------|-------|------|---|
| [| 1.71 | 1.78 | 1.7 | 1.64 | 70.84 | 6.06 |] |
| [| 3.11 | 3.25 | 3.77 | 4.34 | 71.84 | 7.06 |] |
| [| 3.02 | 3.08 | 3.5 | 4.95 | 70.94 | 6.16 |] |
| [| 6.07 | 6.02 | 6.01 | 9.78 | 71.86 | 7.08 |] |

| | | | | | | | |
|---|------|------|------|-------|-------|------|---|
| [| 2.19 | 2.17 | 1.8 | 1.83 | 70.81 | 6.07 |] |
| [| 4.26 | 4.32 | 3.83 | 4.31 | 71.81 | 7.07 |] |
| [| 4.47 | 4.48 | 3.65 | 5.07 | 70.9 | 6.17 |] |
| [| 7.6 | 7.74 | 6.26 | 10.04 | 71.83 | 7.1 |] |

time spent until now: 150.3 mins

17:13, 04/10

ubuntu@ip-172-31-2-198:~\$