

Assignment2Q2

July 3, 2018

```
In [1]: %run main.py --env-id Pendulum-v0 --nb-epochs 50
```

Logging to C:\Users\LINDAI~1\AppData\Local\Temp\openai-2018-07-02-23-26-41-735687

WARN: gym.spaces.Box autodetected dtype as <class 'numpy.float32'>. Please provide explicit dtype

WARN: gym.spaces.Box autodetected dtype as <class 'numpy.float32'>. Please provide explicit dtype

rank 0: seed=0, logdir=C:\Users\LINDAI~1\AppData\Local\Temp\openai-2018-07-02-23-26-41-735687

scaling actions by [2.] before executing in env

setting up param noise

```
param_noise_actor/dense/kernel:0 <- share/dense/kernel:0 + noise
param_noise_actor/dense/bias:0 <- share/dense/bias:0 + noise
param_noise_actor/LayerNorm/beta:0 <- share/LayerNorm/beta:0
param_noise_actor/LayerNorm/gamma:0 <- share/LayerNorm/gamma:0
param_noise_actor/dense_1/kernel:0 <- actor/dense/kernel:0 + noise
param_noise_actor/dense_1/bias:0 <- actor/dense/bias:0 + noise
param_noise_actor/LayerNorm_1/beta:0 <- actor/LayerNorm/beta:0
param_noise_actor/LayerNorm_1/gamma:0 <- actor/LayerNorm/gamma:0
param_noise_actor/dense_2/kernel:0 <- actor/dense_1/kernel:0 + noise
param_noise_actor/dense_2/bias:0 <- actor/dense_1/bias:0 + noise
adaptive_param_noise_actor/dense/kernel:0 <- share/dense/kernel:0 + noise
adaptive_param_noise_actor/dense/bias:0 <- share/dense/bias:0 + noise
adaptive_param_noise_actor/LayerNorm/beta:0 <- share/LayerNorm/beta:0
adaptive_param_noise_actor/LayerNorm/gamma:0 <- share/LayerNorm/gamma:0
adaptive_param_noise_actor/dense_1/kernel:0 <- actor/dense/kernel:0 + noise
adaptive_param_noise_actor/dense_1/bias:0 <- actor/dense/bias:0 + noise
adaptive_param_noise_actor/LayerNorm_1/beta:0 <- actor/LayerNorm/beta:0
adaptive_param_noise_actor/LayerNorm_1/gamma:0 <- actor/LayerNorm/gamma:0
adaptive_param_noise_actor/dense_2/kernel:0 <- actor/dense_1/kernel:0 + noise
adaptive_param_noise_actor/dense_2/bias:0 <- actor/dense_1/bias:0 + noise
```

setting up actor optimizer

```
actor shapes: [[3, 64], [64], [64], [64], [64, 64], [64], [64], [64], [64, 1], [1]]
actor params: 4737
```

setting up critic optimizer

```
regularizing: share/dense/kernel:0
regularizing: critic/dense/kernel:0
regularizing: critic/dense_1/kernel:0
applying l2 regularization with 0.01
critic shapes: [[3, 64], [64], [64], [64], [65, 64], [64], [64], [64], [64, 1], [1]]
```

```

critic params: 4801
WARNING:tensorflow:From c:\users\lin daiwei\baselines\baselines\ddpg\ddpg.py:29: calling reduce_
Instructions for updating:
keep_dims is deprecated, use keepdims instead
setting up target updates ...
target_actor/dense/kernel:0 <- share/dense/kernel:0
target_actor/dense/bias:0 <- share/dense/bias:0
target_actor/LayerNorm/beta:0 <- share/LayerNorm/beta:0
target_actor/LayerNorm/gamma:0 <- share/LayerNorm/gamma:0
target_actor/dense_1/kernel:0 <- actor/dense/kernel:0
target_actor/dense_1/bias:0 <- actor/dense/bias:0
target_actor/LayerNorm_1/beta:0 <- actor/LayerNorm/beta:0
target_actor/LayerNorm_1/gamma:0 <- actor/LayerNorm/gamma:0
target_actor/dense_2/kernel:0 <- actor/dense_1/kernel:0
target_actor/dense_2/bias:0 <- actor/dense_1/bias:0
setting up target updates ...
target_critic/dense/kernel:0 <- share/dense/kernel:0
target_critic/dense/bias:0 <- share/dense/bias:0
target_critic/LayerNorm/beta:0 <- share/LayerNorm/beta:0
target_critic/LayerNorm/gamma:0 <- share/LayerNorm/gamma:0
target_critic/dense_1/kernel:0 <- critic/dense/kernel:0
target_critic/dense_1/bias:0 <- critic/dense/bias:0
target_critic/LayerNorm_1/beta:0 <- critic/LayerNorm/beta:0
target_critic/LayerNorm_1/gamma:0 <- critic/LayerNorm/gamma:0
target_critic/dense_2/kernel:0 <- critic/dense_1/kernel:0
target_critic/dense_2/bias:0 <- critic/dense_1/bias:0
Using agent with the following configuration:
dict_items([('obs0', <tf.Tensor 'obs0:0' shape=(?, 3) dtype=float32>), ('obs1', <tf.Tensor 'obs1
-----
| obs_rms_mean          | -0.583   |
| obs_rms_std           | 1.48     |
| param_noise_stddev    | 0.167    |
| reference_Q_mean      | -58.2    |
| reference_Q_std       | 12.3     |
| reference_action_mean | 0.0971   |
| reference_action_std  | 0.937    |
| reference_actor_Q_mean | -57.6    |
| reference_actor_Q_std | 13       |
| reference_perturbed_... | 0.535    |
| rollout/Q_mean        | -25.1    |
| rollout/actions_mean  | -0.183   |
| rollout/actions_std   | 0.708    |
| rollout/episode_steps | 200      |
| rollout/episodes      | 10       |
| rollout/return        | -1.49e+03 |
| rollout/return_history | -1.49e+03 |
| total/duration        | 13.5     |
| total/episodes        | 10       |

```

total/epochs	1	
total/steps	2000	
total/steps_per_second	148	
train/loss_actor	52.1	
train/loss_critic	15.4	
train/param_noise_di...	0.842	

obs_rms_mean	-0.0671	
obs_rms_std	1.97	
param_noise_stddev	0.137	
reference_Q_mean	-115	
reference_Q_std	21.5	
reference_action_mean	-0.229	
reference_action_std	0.928	
reference_actor_Q_mean	-114	
reference_actor_Q_std	22.2	
reference_perturbed_...	0.493	
rollout/Q_mean	-45.1	
rollout/actions_mean	-0.0265	
rollout/actions_std	0.729	
rollout/episode_steps	200	
rollout/episodes	20	
rollout/return	-1.41e+03	
rollout/return_history	-1.41e+03	
total/duration	26.5	
total/episodes	20	
total/epochs	2	
total/steps	4000	
total/steps_per_second	151	
train/loss_actor	98.1	
train/loss_critic	60.3	
train/param_noise_di...	0.514	

obs_rms_mean	-0.165	
obs_rms_std	1.95	
param_noise_stddev	0.112	
reference_Q_mean	-152	
reference_Q_std	28.9	
reference_action_mean	0.17	
reference_action_std	0.935	
reference_actor_Q_mean	-149	
reference_actor_Q_std	28.9	
reference_perturbed_...	0.885	
rollout/Q_mean	-67	

rollout/actions_mean	-0.143	
rollout/actions_std	0.739	
rollout/episode_steps	200	
rollout/episodes	30	
rollout/return	-1.4e+03	
rollout/return_history	-1.4e+03	
total/duration	39.3	
total/episodes	30	
total/epochs	3	
total/steps	6000	
total/steps_per_second	153	
train/loss_actor	126	
train/loss_critic	104	
train/param_noise_di...	0.514	

obs_rms_mean	-0.265	
obs_rms_std	1.94	
param_noise_stddev	0.092	
reference_Q_mean	-173	
reference_Q_std	38	
reference_action_mean	0.244	
reference_action_std	0.938	
reference_actor_Q_mean	-171	
reference_actor_Q_std	39.5	
reference_perturbed_...	0.868	
rollout/Q_mean	-79.5	
rollout/actions_mean	-0.146	
rollout/actions_std	0.738	
rollout/episode_steps	200	
rollout/episodes	40	
rollout/return	-1.35e+03	
rollout/return_history	-1.35e+03	
total/duration	52.3	
total/episodes	40	
total/epochs	4	
total/steps	8000	
total/steps_per_second	153	
train/loss_actor	140	
train/loss_critic	66.8	
train/param_noise_di...	0.924	

obs_rms_mean	-0.374	
obs_rms_std	1.92	
param_noise_stddev	0.0754	

reference_Q_mean	-191	
reference_Q_std	51	
reference_action_mean	0.189	
reference_action_std	0.948	
reference_actor_Q_mean	-187	
reference_actor_Q_std	53.1	
reference_perturbed_...	0.812	
rollout/Q_mean	-89	
rollout/actions_mean	-0.151	
rollout/actions_std	0.745	
rollout/episode_steps	200	
rollout/episodes	50	
rollout/return	-1.31e+03	
rollout/return_history	-1.31e+03	
total/duration	65	
total/episodes	50	
total/epochs	5	
total/steps	10000	
total/steps_per_second	154	
train/loss_actor	142	
train/loss_critic	90.4	
train/param_noise_di...	0.626	

obs_rms_mean	-0.419	
obs_rms_std	1.9	
param_noise_stddev	0.0618	
reference_Q_mean	-197	
reference_Q_std	61.1	
reference_action_mean	0.172	
reference_action_std	0.944	
reference_actor_Q_mean	-193	
reference_actor_Q_std	62.3	
reference_perturbed_...	0.843	
rollout/Q_mean	-90.9	
rollout/actions_mean	-0.165	
rollout/actions_std	0.748	
rollout/episode_steps	200	
rollout/episodes	60	
rollout/return	-1.24e+03	
rollout/return_history	-1.24e+03	
total/duration	77.9	
total/episodes	60	
total/epochs	6	
total/steps	12000	
total/steps_per_second	154	
train/loss_actor	137	

train/loss_critic	158	
train/param_noise_di...	0.682	

obs_rms_mean	-0.337	
obs_rms_std	1.88	
param_noise_stddev	0.0507	
reference_Q_mean	-202	
reference_Q_std	69.8	
reference_action_mean	0.309	
reference_action_std	0.878	
reference_actor_Q_mean	-198	
reference_actor_Q_std	71.3	
reference_perturbed_...	0.904	
rollout/Q_mean	-90.8	
rollout/actions_mean	-0.149	
rollout/actions_std	0.746	
rollout/episode_steps	200	
rollout/episodes	70	
rollout/return	-1.17e+03	
rollout/return_history	-1.17e+03	
total/duration	90.8	
total/episodes	70	
total/epochs	7	
total/steps	14000	
total/steps_per_second	154	
train/loss_actor	126	
train/loss_critic	122	
train/param_noise_di...	0.539	

obs_rms_mean	-0.235	
obs_rms_std	1.89	
param_noise_stddev	0.0415	
reference_Q_mean	-205	
reference_Q_std	77.6	
reference_action_mean	0.399	
reference_action_std	0.878	
reference_actor_Q_mean	-200	
reference_actor_Q_std	78.1	
reference_perturbed_...	0.869	
rollout/Q_mean	-89	
rollout/actions_mean	-0.119	
rollout/actions_std	0.753	
rollout/episode_steps	200	
rollout/episodes	80	

rollout/return	-1.11e+03	
rollout/return_history	-1.11e+03	
total/duration	104	
total/episodes	80	
total/epochs	8	
total/steps	16000	
total/steps_per_second	154	
train/loss_actor	110	
train/loss_critic	103	
train/param_noise_di...	0.629	

obs_rms_mean	-0.179	
obs_rms_std	1.86	
param_noise_stddev	0.034	
reference_Q_mean	-190	
reference_Q_std	77.2	
reference_action_mean	0.37	
reference_action_std	0.886	
reference_actor_Q_mean	-184	
reference_actor_Q_std	76.3	
reference_perturbed_...	0.902	
rollout/Q_mean	-85.4	
rollout/actions_mean	-0.0935	
rollout/actions_std	0.752	
rollout/episode_steps	200	
rollout/episodes	90	
rollout/return	-1.06e+03	
rollout/return_history	-1.06e+03	
total/duration	117	
total/episodes	90	
total/epochs	9	
total/steps	18000	
total/steps_per_second	154	
train/loss_actor	94.5	
train/loss_critic	85.8	
train/param_noise_di...	0.491	

obs_rms_mean	-0.0865	
obs_rms_std	1.86	
param_noise_stddev	0.0279	
reference_Q_mean	-183	
reference_Q_std	79.1	
reference_action_mean	0.352	
reference_action_std	0.897	

reference_actor_Q_mean	-177	
reference_actor_Q_std	79.5	
reference_perturbed_...	0.789	
rollout/Q_mean	-79.3	
rollout/actions_mean	-0.0574	
rollout/actions_std	0.75	
rollout/episode_steps	200	
rollout/episodes	100	
rollout/return	-1.02e+03	
rollout/return_history	-1.02e+03	
total/duration	129	
total/episodes	100	
total/epochs	10	
total/steps	20000	
total/steps_per_second	155	
train/loss_actor	71.6	
train/loss_critic	110	
train/param_noise_di...	0.437	

obs_rms_mean	-0.0212	
obs_rms_std	1.84	
param_noise_stddev	0.0229	
reference_Q_mean	-172	
reference_Q_std	81.4	
reference_action_mean	0.373	
reference_action_std	0.888	
reference_actor_Q_mean	-165	
reference_actor_Q_std	81.4	
reference_perturbed_...	0.884	
rollout/Q_mean	-71.2	
rollout/actions_mean	-0.0571	
rollout/actions_std	0.741	
rollout/episode_steps	200	
rollout/episodes	110	
rollout/return	-967	
rollout/return_history	-915	
total/duration	142	
total/episodes	110	
total/epochs	11	
total/steps	22000	
total/steps_per_second	155	
train/loss_actor	47.8	
train/loss_critic	79	
train/param_noise_di...	0.285	

obs_rms_mean	0.0501	
obs_rms_std	1.83	
param_noise_stddev	0.0191	
reference_Q_mean	-156	
reference_Q_std	87.4	
reference_action_mean	0.378	
reference_action_std	0.884	
reference_actor_Q_mean	-149	
reference_actor_Q_std	86.9	
reference_perturbed_...	0.786	
rollout/Q_mean	-63.7	
rollout/actions_mean	-0.0403	
rollout/actions_std	0.736	
rollout/episode_steps	200	
rollout/episodes	120	
rollout/return	-937	
rollout/return_history	-842	
total/duration	155	
total/episodes	120	
total/epochs	12	
total/steps	24000	
total/steps_per_second	155	
train/loss_actor	24.3	
train/loss_critic	117	
train/param_noise_di...	0.445	

obs_rms_mean	0.152	
obs_rms_std	1.86	
param_noise_stddev	0.0163	
reference_Q_mean	-148	
reference_Q_std	93.9	
reference_action_mean	0.396	
reference_action_std	0.869	
reference_actor_Q_mean	-142	
reference_actor_Q_std	95.3	
reference_perturbed_...	0.93	
rollout/Q_mean	-58.2	
rollout/actions_mean	-0.00247	
rollout/actions_std	0.741	
rollout/episode_steps	200	
rollout/episodes	130	
rollout/return	-944	
rollout/return_history	-806	
total/duration	168	
total/episodes	130	

total/epochs	13	
total/steps	26000	
total/steps_per_second	155	
train/loss_actor	5.6	
train/loss_critic	62.7	
train/param_noise_di...	0.268	

obs_rms_mean	0.204	
obs_rms_std	1.85	
param_noise_stddev	0.0142	
reference_Q_mean	-139	
reference_Q_std	95.7	
reference_action_mean	0.414	
reference_action_std	0.869	
reference_actor_Q_mean	-132	
reference_actor_Q_std	96.2	
reference_perturbed_...	0.879	
rollout/Q_mean	-49.9	
rollout/actions_mean	0.00247	
rollout/actions_std	0.738	
rollout/episode_steps	200	
rollout/episodes	140	
rollout/return	-922	
rollout/return_history	-751	
total/duration	181	
total/episodes	140	
total/epochs	14	
total/steps	28000	
total/steps_per_second	155	
train/loss_actor	-19.1	
train/loss_critic	134	
train/param_noise_di...	0.205	

obs_rms_mean	0.23	
obs_rms_std	1.82	
param_noise_stddev	0.0121	
reference_Q_mean	-126	
reference_Q_std	105	
reference_action_mean	0.71	
reference_action_std	0.608	
reference_actor_Q_mean	-120	
reference_actor_Q_std	106	
reference_perturbed_...	0.803	
rollout/Q_mean	-39.3	

rollout/actions_mean	0.00628	
rollout/actions_std	0.727	
rollout/episode_steps	200	
rollout/episodes	150	
rollout/return	-886	
rollout/return_history	-673	
total/duration	194	
total/episodes	150	
total/epochs	15	
total/steps	30000	
total/steps_per_second	155	
train/loss_actor	-44.6	
train/loss_critic	94.1	
train/param_noise_di...	0.16	

obs_rms_mean	0.266	
obs_rms_std	1.81	
param_noise_stddev	0.0112	
reference_Q_mean	-118	
reference_Q_std	106	
reference_action_mean	0.679	
reference_action_std	0.639	
reference_actor_Q_mean	-113	
reference_actor_Q_std	107	
reference_perturbed_...	0.653	
rollout/Q_mean	-30.6	
rollout/actions_mean	0.00624	
rollout/actions_std	0.725	
rollout/episode_steps	200	
rollout/episodes	160	
rollout/return	-866	
rollout/return_history	-640	
total/duration	207	
total/episodes	160	
total/epochs	16	
total/steps	32000	
total/steps_per_second	155	
train/loss_actor	-64.8	
train/loss_critic	107	
train/param_noise_di...	0.253	

obs_rms_mean	0.275	
obs_rms_std	1.78	
param_noise_stddev	0.00991	

reference_Q_mean	-114	
reference_Q_std	112	
reference_action_mean	0.647	
reference_action_std	0.644	
reference_actor_Q_mean	-108	
reference_actor_Q_std	112	
reference_perturbed_...	0.811	
rollout/Q_mean	-20.1	
rollout/actions_mean	0.00816	
rollout/actions_std	0.714	
rollout/episode_steps	200	
rollout/episodes	170	
rollout/return	-833	
rollout/return_history	-599	
total/duration	220	
total/episodes	170	
total/epochs	17	
total/steps	34000	
total/steps_per_second	155	
train/loss_actor	-82.4	
train/loss_critic	185	
train/param_noise_di...	0.14	

obs_rms_mean	0.262	
obs_rms_std	1.76	
param_noise_stddev	0.00915	
reference_Q_mean	-110	
reference_Q_std	115	
reference_action_mean	0.734	
reference_action_std	0.557	
reference_actor_Q_mean	-105	
reference_actor_Q_std	117	
reference_perturbed_...	0.565	
rollout/Q_mean	-10.7	
rollout/actions_mean	0.0105	
rollout/actions_std	0.708	
rollout/episode_steps	200	
rollout/episodes	180	
rollout/return	-804	
rollout/return_history	-559	
total/duration	233	
total/episodes	180	
total/epochs	18	
total/steps	36000	
total/steps_per_second	155	
train/loss_actor	-99.3	

train/loss_critic	164	
train/param_noise_di...	0.29	

obs_rms_mean	0.262	
obs_rms_std	1.72	
param_noise_stddev	0.00828	
reference_Q_mean	-107	
reference_Q_std	122	
reference_action_mean	0.583	
reference_action_std	0.703	
reference_actor_Q_mean	-101	
reference_actor_Q_std	123	
reference_perturbed_...	0.627	
rollout/Q_mean	0.606	
rollout/actions_mean	0.00981	
rollout/actions_std	0.694	
rollout/episode_steps	200	
rollout/episodes	190	
rollout/return	-770	
rollout/return_history	-509	
total/duration	246	
total/episodes	190	
total/epochs	19	
total/steps	38000	
total/steps_per_second	155	
train/loss_actor	-120	
train/loss_critic	184	
train/param_noise_di...	0.508	

obs_rms_mean	0.272	
obs_rms_std	1.7	
param_noise_stddev	0.00812	
reference_Q_mean	-92	
reference_Q_std	121	
reference_action_mean	0.647	
reference_action_std	0.634	
reference_actor_Q_mean	-84.6	
reference_actor_Q_std	120	
reference_perturbed_...	0.288	
rollout/Q_mean	10.4	
rollout/actions_mean	0.0103	
rollout/actions_std	0.686	
rollout/episode_steps	200	
rollout/episodes	200	

rollout/return	-748	
rollout/return_history	-480	
total/duration	259	
total/episodes	200	
total/epochs	20	
total/steps	40000	
total/steps_per_second	155	
train/loss_actor	-137	
train/loss_critic	158	
train/param_noise_di...	0.252	

obs_rms_mean	0.272	
obs_rms_std	1.68	
param_noise_stddev	0.00796	
reference_Q_mean	-80.2	
reference_Q_std	123	
reference_action_mean	0.663	
reference_action_std	0.637	
reference_actor_Q_mean	-71.1	
reference_actor_Q_std	123	
reference_perturbed_...	0.629	
rollout/Q_mean	19.1	
rollout/actions_mean	0.0109	
rollout/actions_std	0.682	
rollout/episode_steps	200	
rollout/episodes	210	
rollout/return	-727	
rollout/return_history	-463	
total/duration	272	
total/episodes	210	
total/epochs	21	
total/steps	42000	
total/steps_per_second	155	
train/loss_actor	-147	
train/loss_critic	350	
train/param_noise_di...	0.415	

obs_rms_mean	0.286	
obs_rms_std	1.67	
param_noise_stddev	0.00735	
reference_Q_mean	-74.1	
reference_Q_std	123	
reference_action_mean	0.657	
reference_action_std	0.663	

reference_actor_Q_mean	-66	
reference_actor_Q_std	124	
reference_perturbed_...	0.736	
rollout/Q_mean	26.9	
rollout/actions_mean	0.0124	
rollout/actions_std	0.68	
rollout/episode_steps	200	
rollout/episodes	220	
rollout/return	-716	
rollout/return_history	-451	
total/duration	285	
total/episodes	220	
total/epochs	22	
total/steps	44000	
total/steps_per_second	155	
train/loss_actor	-164	
train/loss_critic	257	
train/param_noise_di...	0.381	

obs_rms_mean	0.288	
obs_rms_std	1.66	
param_noise_stddev	0.00721	
reference_Q_mean	-53.9	
reference_Q_std	126	
reference_action_mean	0.599	
reference_action_std	0.723	
reference_actor_Q_mean	-45.2	
reference_actor_Q_std	128	
reference_perturbed_...	0.787	
rollout/Q_mean	34.4	
rollout/actions_mean	0.00895	
rollout/actions_std	0.68	
rollout/episode_steps	200	
rollout/episodes	230	
rollout/return	-705	
rollout/return_history	-395	
total/duration	298	
total/episodes	230	
total/epochs	23	
total/steps	46000	
total/steps_per_second	154	
train/loss_actor	-185	
train/loss_critic	386	
train/param_noise_di...	0.373	

obs_rms_mean	0.286	
obs_rms_std	1.65	
param_noise_stddev	0.00652	
reference_Q_mean	-52.4	
reference_Q_std	130	
reference_action_mean	0.55	
reference_action_std	0.726	
reference_actor_Q_mean	-44.8	
reference_actor_Q_std	131	
reference_perturbed_...	0.714	
rollout/Q_mean	43.7	
rollout/actions_mean	0.00866	
rollout/actions_std	0.677	
rollout/episode_steps	200	
rollout/episodes	240	
rollout/return	-690	
rollout/return_history	-366	
total/duration	311	
total/episodes	240	
total/epochs	24	
total/steps	48000	
total/steps_per_second	154	
train/loss_actor	-212	
train/loss_critic	586	
train/param_noise_di...	0.26	

obs_rms_mean	0.285	
obs_rms_std	1.64	
param_noise_stddev	0.00602	
reference_Q_mean	-31.8	
reference_Q_std	140	
reference_action_mean	0.404	
reference_action_std	0.86	
reference_actor_Q_mean	-19.9	
reference_actor_Q_std	140	
reference_perturbed_...	0.838	
rollout/Q_mean	52.2	
rollout/actions_mean	0.00446	
rollout/actions_std	0.676	
rollout/episode_steps	200	
rollout/episodes	250	
rollout/return	-680	
rollout/return_history	-370	
total/duration	324	
total/episodes	250	

total/epochs	25	
total/steps	50000	
total/steps_per_second	154	
train/loss_actor	-242	
train/loss_critic	573	
train/param_noise_di...	0.234	

obs_rms_mean	0.284	
obs_rms_std	1.62	
param_noise_stddev	0.00591	
reference_Q_mean	0.633	
reference_Q_std	153	
reference_action_mean	0.445	
reference_action_std	0.828	
reference_actor_Q_mean	10.2	
reference_actor_Q_std	153	
reference_perturbed_...	0.785	
rollout/Q_mean	63.4	
rollout/actions_mean	-0.00375	
rollout/actions_std	0.671	
rollout/episode_steps	200	
rollout/episodes	260	
rollout/return	-663	
rollout/return_history	-338	
total/duration	337	
total/episodes	260	
total/epochs	26	
total/steps	52000	
total/steps_per_second	154	
train/loss_actor	-275	
train/loss_critic	727	
train/param_noise_di...	0.14	

obs_rms_mean	0.285	
obs_rms_std	1.61	
param_noise_stddev	0.00568	
reference_Q_mean	27.6	
reference_Q_std	139	
reference_action_mean	0.363	
reference_action_std	0.854	
reference_actor_Q_mean	36.9	
reference_actor_Q_std	137	
reference_perturbed_...	0.805	
rollout/Q_mean	72.7	

rollout/actions_mean	-0.00731	
rollout/actions_std	0.67	
rollout/episode_steps	200	
rollout/episodes	270	
rollout/return	-653	
rollout/return_history	-348	
total/duration	350	
total/episodes	270	
total/epochs	27	
total/steps	54000	
total/steps_per_second	154	
train/loss_actor	-297	
train/loss_critic	1.19e+03	
train/param_noise_di...	0.386	

obs_rms_mean	0.284	
obs_rms_std	1.6	
param_noise_stddev	0.00535	
reference_Q_mean	51.9	
reference_Q_std	129	
reference_action_mean	0.262	
reference_action_std	0.891	
reference_actor_Q_mean	62.9	
reference_actor_Q_std	128	
reference_perturbed_...	0.875	
rollout/Q_mean	82.4	
rollout/actions_mean	-0.0068	
rollout/actions_std	0.667	
rollout/episode_steps	200	
rollout/episodes	280	
rollout/return	-645	
rollout/return_history	-359	
total/duration	363	
total/episodes	280	
total/epochs	28	
total/steps	56000	
total/steps_per_second	154	
train/loss_actor	-302	
train/loss_critic	747	
train/param_noise_di...	0.322	

obs_rms_mean	0.281	
obs_rms_std	1.59	
param_noise_stddev	0.00494	

reference_Q_mean	56.1	
reference_Q_std	134	
reference_action_mean	0.294	
reference_action_std	0.914	
reference_actor_Q_mean	66.5	
reference_actor_Q_std	134	
reference_perturbed_...	0.898	
rollout/Q_mean	92.7	
rollout/actions_mean	-0.00738	
rollout/actions_std	0.662	
rollout/episode_steps	200	
rollout/episodes	290	
rollout/return	-633	
rollout/return_history	-373	
total/duration	376	
total/episodes	290	
total/epochs	29	
total/steps	58000	
total/steps_per_second	154	
train/loss_actor	-309	
train/loss_critic	893	
train/param_noise_di...	0.599	

obs_rms_mean	0.285	
obs_rms_std	1.57	
param_noise_stddev	0.00465	
reference_Q_mean	88.8	
reference_Q_std	120	
reference_action_mean	0.301	
reference_action_std	0.914	
reference_actor_Q_mean	100	
reference_actor_Q_std	118	
reference_perturbed_...	0.896	
rollout/Q_mean	102	
rollout/actions_mean	-0.0106	
rollout/actions_std	0.656	
rollout/episode_steps	200	
rollout/episodes	300	
rollout/return	-620	
rollout/return_history	-364	
total/duration	389	
total/episodes	300	
total/epochs	30	
total/steps	60000	
total/steps_per_second	154	
train/loss_actor	-317	

train/loss_critic	622	
train/param_noise_di...	0.193	

obs_rms_mean	0.284	
obs_rms_std	1.56	
param_noise_stddev	0.00456	
reference_Q_mean	119	
reference_Q_std	110	
reference_action_mean	0.382	
reference_action_std	0.871	
reference_actor_Q_mean	129	
reference_actor_Q_std	110	
reference_perturbed_...	0.892	
rollout/Q_mean	111	
rollout/actions_mean	-0.0125	
rollout/actions_std	0.651	
rollout/episode_steps	200	
rollout/episodes	310	
rollout/return	-608	
rollout/return_history	-356	
total/duration	402	
total/episodes	310	
total/epochs	31	
total/steps	62000	
total/steps_per_second	154	
train/loss_actor	-319	
train/loss_critic	811	
train/param_noise_di...	0.171	

obs_rms_mean	0.282	
obs_rms_std	1.54	
param_noise_stddev	0.00465	
reference_Q_mean	125	
reference_Q_std	104	
reference_action_mean	0.365	
reference_action_std	0.844	
reference_actor_Q_mean	134	
reference_actor_Q_std	102	
reference_perturbed_...	0.821	
rollout/Q_mean	120	
rollout/actions_mean	-0.0146	
rollout/actions_std	0.647	
rollout/episode_steps	200	
rollout/episodes	320	

rollout/return	-595	
rollout/return_history	-328	
total/duration	415	
total/episodes	320	
total/epochs	32	
total/steps	64000	
total/steps_per_second	154	
train/loss_actor	-317	
train/loss_critic	551	
train/param_noise_di...	0.347	

obs_rms_mean	0.279	
obs_rms_std	1.53	
param_noise_stddev	0.00465	
reference_Q_mean	129	
reference_Q_std	105	
reference_action_mean	0.63	
reference_action_std	0.596	
reference_actor_Q_mean	134	
reference_actor_Q_std	103	
reference_perturbed_...	0.662	
rollout/Q_mean	128	
rollout/actions_mean	-0.0166	
rollout/actions_std	0.641	
rollout/episode_steps	200	
rollout/episodes	330	
rollout/return	-582	
rollout/return_history	-299	
total/duration	428	
total/episodes	330	
total/epochs	33	
total/steps	66000	
total/steps_per_second	154	
train/loss_actor	-304	
train/loss_critic	1.23e+03	
train/param_noise_di...	0.101	

obs_rms_mean	0.276	
obs_rms_std	1.51	
param_noise_stddev	0.00474	
reference_Q_mean	114	
reference_Q_std	93.6	
reference_action_mean	0.469	
reference_action_std	0.753	

reference_actor_Q_mean	120	
reference_actor_Q_std	95.1	
reference_perturbed_...	0.699	
rollout/Q_mean	135	
rollout/actions_mean	-0.0158	
rollout/actions_std	0.636	
rollout/episode_steps	200	
rollout/episodes	340	
rollout/return	-571	
rollout/return_history	-283	
total/duration	441	
total/episodes	340	
total/epochs	34	
total/steps	68000	
total/steps_per_second	154	
train/loss_actor	-285	
train/loss_critic	493	
train/param_noise_di...	0.102	

obs_rms_mean	0.268	
obs_rms_std	1.51	
param_noise_stddev	0.00494	
reference_Q_mean	82.4	
reference_Q_std	105	
reference_action_mean	0.651	
reference_action_std	0.658	
reference_actor_Q_mean	89.4	
reference_actor_Q_std	106	
reference_perturbed_...	0.659	
rollout/Q_mean	140	
rollout/actions_mean	-0.0136	
rollout/actions_std	0.634	
rollout/episode_steps	200	
rollout/episodes	350	
rollout/return	-563	
rollout/return_history	-273	
total/duration	454	
total/episodes	350	
total/epochs	35	
total/steps	70000	
total/steps_per_second	154	
train/loss_actor	-266	
train/loss_critic	512	
train/param_noise_di...	0.281	

obs_rms_mean	0.263	
obs_rms_std	1.5	
param_noise_stddev	0.00535	
reference_Q_mean	72.4	
reference_Q_std	103	
reference_action_mean	0.589	
reference_action_std	0.725	
reference_actor_Q_mean	78.3	
reference_actor_Q_std	104	
reference_perturbed_...	0.636	
rollout/Q_mean	145	
rollout/actions_mean	-0.0111	
rollout/actions_std	0.631	
rollout/episode_steps	200	
rollout/episodes	360	
rollout/return	-556	
rollout/return_history	-277	
total/duration	467	
total/episodes	360	
total/epochs	36	
total/steps	72000	
total/steps_per_second	154	
train/loss_actor	-256	
train/loss_critic	696	
train/param_noise_di...	0.192	

obs_rms_mean	0.262	
obs_rms_std	1.48	
param_noise_stddev	0.00545	
reference_Q_mean	54.8	
reference_Q_std	100	
reference_action_mean	0.525	
reference_action_std	0.743	
reference_actor_Q_mean	61.3	
reference_actor_Q_std	99	
reference_perturbed_...	0.716	
rollout/Q_mean	149	
rollout/actions_mean	-0.0125	
rollout/actions_std	0.625	
rollout/episode_steps	200	
rollout/episodes	370	
rollout/return	-545	
rollout/return_history	-252	
total/duration	480	
total/episodes	370	

total/epochs	37	
total/steps	74000	
total/steps_per_second	154	
train/loss_actor	-246	
train/loss_critic	439	
train/param_noise_di...	0.301	

obs_rms_mean	0.26	
obs_rms_std	1.47	
param_noise_stddev	0.00504	
reference_Q_mean	30.8	
reference_Q_std	97.7	
reference_action_mean	0.281	
reference_action_std	0.898	
reference_actor_Q_mean	36.6	
reference_actor_Q_std	99	
reference_perturbed_...	0.895	
rollout/Q_mean	152	
rollout/actions_mean	-0.0145	
rollout/actions_std	0.621	
rollout/episode_steps	200	
rollout/episodes	380	
rollout/return	-537	
rollout/return_history	-234	
total/duration	493	
total/episodes	380	
total/epochs	38	
total/steps	76000	
total/steps_per_second	154	
train/loss_actor	-241	
train/loss_critic	377	
train/param_noise_di...	0.183	

obs_rms_mean	0.254	
obs_rms_std	1.46	
param_noise_stddev	0.00474	
reference_Q_mean	22.8	
reference_Q_std	108	
reference_action_mean	0.242	
reference_action_std	0.907	
reference_actor_Q_mean	29.4	
reference_actor_Q_std	109	
reference_perturbed_...	0.934	
rollout/Q_mean	156	

rollout/actions_mean	-0.0158	
rollout/actions_std	0.617	
rollout/episode_steps	200	
rollout/episodes	390	
rollout/return	-530	
rollout/return_history	-233	
total/duration	506	
total/episodes	390	
total/epochs	39	
total/steps	78000	
total/steps_per_second	154	
train/loss_actor	-235	
train/loss_critic	522	
train/param_noise_di...	0.177	

obs_rms_mean	0.252	
obs_rms_std	1.45	
param_noise_stddev	0.00465	
reference_Q_mean	28.6	
reference_Q_std	110	
reference_action_mean	0.247	
reference_action_std	0.908	
reference_actor_Q_mean	36	
reference_actor_Q_std	111	
reference_perturbed_...	0.915	
rollout/Q_mean	159	
rollout/actions_mean	-0.0194	
rollout/actions_std	0.613	
rollout/episode_steps	200	
rollout/episodes	400	
rollout/return	-523	
rollout/return_history	-232	
total/duration	519	
total/episodes	400	
total/epochs	40	
total/steps	80000	
total/steps_per_second	154	
train/loss_actor	-238	
train/loss_critic	439	
train/param_noise_di...	0.0642	

obs_rms_mean	0.252	
obs_rms_std	1.44	
param_noise_stddev	0.00494	

reference_Q_mean	19.2	
reference_Q_std	121	
reference_action_mean	0.343	
reference_action_std	0.855	
reference_actor_Q_mean	25.1	
reference_actor_Q_std	121	
reference_perturbed_...	0.903	
rollout/Q_mean	162	
rollout/actions_mean	-0.0224	
rollout/actions_std	0.609	
rollout/episode_steps	200	
rollout/episodes	410	
rollout/return	-515	
rollout/return_history	-226	
total/duration	533	
total/episodes	410	
total/epochs	41	
total/steps	82000	
total/steps_per_second	154	
train/loss_actor	-255	
train/loss_critic	736	
train/param_noise_di...	0.161	

obs_rms_mean	0.25	
obs_rms_std	1.43	
param_noise_stddev	0.00484	
reference_Q_mean	17	
reference_Q_std	126	
reference_action_mean	0.352	
reference_action_std	0.823	
reference_actor_Q_mean	22.1	
reference_actor_Q_std	128	
reference_perturbed_...	0.809	
rollout/Q_mean	165	
rollout/actions_mean	-0.0313	
rollout/actions_std	0.607	
rollout/episode_steps	200	
rollout/episodes	420	
rollout/return	-508	
rollout/return_history	-228	
total/duration	546	
total/episodes	420	
total/epochs	42	
total/steps	84000	
total/steps_per_second	154	
train/loss_actor	-294	

train/loss_critic	1.02e+03	
train/param_noise_di...	0.259	

obs_rms_mean	0.246	
obs_rms_std	1.43	
param_noise_stddev	0.00484	
reference_Q_mean	5.24	
reference_Q_std	146	
reference_action_mean	0.625	
reference_action_std	0.639	
reference_actor_Q_mean	11.7	
reference_actor_Q_std	148	
reference_perturbed_...	0.644	
rollout/Q_mean	167	
rollout/actions_mean	-0.036	
rollout/actions_std	0.607	
rollout/episode_steps	200	
rollout/episodes	430	
rollout/return	-502	
rollout/return_history	-238	
total/duration	559	
total/episodes	430	
total/epochs	43	
total/steps	86000	
total/steps_per_second	154	
train/loss_actor	-316	
train/loss_critic	1.06e+03	
train/param_noise_di...	0.243	

obs_rms_mean	0.243	
obs_rms_std	1.42	
param_noise_stddev	0.00535	
reference_Q_mean	14.8	
reference_Q_std	164	
reference_action_mean	0.57	
reference_action_std	0.706	
reference_actor_Q_mean	21	
reference_actor_Q_std	167	
reference_perturbed_...	0.755	
rollout/Q_mean	172	
rollout/actions_mean	-0.0324	
rollout/actions_std	0.607	
rollout/episode_steps	200	
rollout/episodes	440	

rollout/return	-497	
rollout/return_history	-247	
total/duration	572	
total/episodes	440	
total/epochs	44	
total/steps	88000	
total/steps_per_second	154	
train/loss_actor	-346	
train/loss_critic	889	
train/param_noise_di...	0.14	

obs_rms_mean	0.238	
obs_rms_std	1.41	
param_noise_stddev	0.00545	
reference_Q_mean	34.2	
reference_Q_std	171	
reference_action_mean	0.457	
reference_action_std	0.793	
reference_actor_Q_mean	42	
reference_actor_Q_std	174	
reference_perturbed_...	0.744	
rollout/Q_mean	176	
rollout/actions_mean	-0.0294	
rollout/actions_std	0.607	
rollout/episode_steps	200	
rollout/episodes	450	
rollout/return	-491	
rollout/return_history	-240	
total/duration	585	
total/episodes	450	
total/epochs	45	
total/steps	90000	
total/steps_per_second	154	
train/loss_actor	-366	
train/loss_critic	981	
train/param_noise_di...	0.166	

obs_rms_mean	0.236	
obs_rms_std	1.4	
param_noise_stddev	0.00545	
reference_Q_mean	56.1	
reference_Q_std	177	
reference_action_mean	0.229	
reference_action_std	0.921	

reference_actor_Q_mean	66.8	
reference_actor_Q_std	177	
reference_perturbed_...	0.878	
rollout/Q_mean	181	
rollout/actions_mean	-0.0278	
rollout/actions_std	0.606	
rollout/episode_steps	200	
rollout/episodes	460	
rollout/return	-485	
rollout/return_history	-232	
total/duration	598	
total/episodes	460	
total/epochs	46	
total/steps	92000	
total/steps_per_second	154	
train/loss_actor	-387	
train/loss_critic	974	
train/param_noise_di...	0.238	

obs_rms_mean	0.23	
obs_rms_std	1.41	
param_noise_stddev	0.00524	
reference_Q_mean	111	
reference_Q_std	151	
reference_action_mean	0.271	
reference_action_std	0.874	
reference_actor_Q_mean	123	
reference_actor_Q_std	151	
reference_perturbed_...	0.889	
rollout/Q_mean	185	
rollout/actions_mean	-0.0297	
rollout/actions_std	0.606	
rollout/episode_steps	200	
rollout/episodes	470	
rollout/return	-484	
rollout/return_history	-260	
total/duration	611	
total/episodes	470	
total/epochs	47	
total/steps	94000	
total/steps_per_second	154	
train/loss_actor	-386	
train/loss_critic	538	
train/param_noise_di...	0.247	

obs_rms_mean	0.223	
obs_rms_std	1.42	
param_noise_stddev	0.00545	
reference_Q_mean	119	
reference_Q_std	134	
reference_action_mean	0.197	
reference_action_std	0.87	
reference_actor_Q_mean	127	
reference_actor_Q_std	134	
reference_perturbed_...	0.854	
rollout/Q_mean	189	
rollout/actions_mean	-0.0272	
rollout/actions_std	0.608	
rollout/episode_steps	200	
rollout/episodes	480	
rollout/return	-493	
rollout/return_history	-325	
total/duration	624	
total/episodes	480	
total/epochs	48	
total/steps	96000	
total/steps_per_second	154	
train/loss_actor	-372	
train/loss_critic	503	
train/param_noise_di...	0.2	

obs_rms_mean	0.213	
obs_rms_std	1.43	
param_noise_stddev	0.00494	
reference_Q_mean	110	
reference_Q_std	119	
reference_action_mean	0.287	
reference_action_std	0.857	
reference_actor_Q_mean	118	
reference_actor_Q_std	117	
reference_perturbed_...	0.87	
rollout/Q_mean	192	
rollout/actions_mean	-0.0256	
rollout/actions_std	0.608	
rollout/episode_steps	200	
rollout/episodes	490	
rollout/return	-501	
rollout/return_history	-385	
total/duration	637	
total/episodes	490	

total/epochs	49	
total/steps	98000	
total/steps_per_second	154	
train/loss_actor	-334	
train/loss_critic	761	
train/param_noise_di...	0.208	

obs_rms_mean	0.195	
obs_rms_std	1.44	
param_noise_stddev	0.00465	
reference_Q_mean	105	
reference_Q_std	94.4	
reference_action_mean	0.254	
reference_action_std	0.86	
reference_actor_Q_mean	114	
reference_actor_Q_std	93	
reference_perturbed_...	0.856	
rollout/Q_mean	194	
rollout/actions_mean	-0.0244	
rollout/actions_std	0.608	
rollout/episode_steps	200	
rollout/episodes	500	
rollout/return	-506	
rollout/return_history	-440	
total/duration	650	
total/episodes	500	
total/epochs	50	
total/steps	100000	
total/steps_per_second	154	
train/loss_actor	-287	
train/loss_critic	478	
train/param_noise_di...	0.231	

[-7.9141372303988362, -4.3881828725086862, -7.5796743246328981, -7.6300507806751012, -6.96479391

total runtime: 653.5052380561829s

WARN: gym.spaces.Box autodetected dtype as <class 'numpy.float32'>. Please provide explicit dtype

WARN: gym.spaces.Box autodetected dtype as <class 'numpy.float32'>. Please provide explicit dtype

rank 0: seed=0, logdir=C:\Users\LINDAI~1\AppData\Local\Temp\openai-2018-07-02-23-26-41-735687

scaling actions by [2.] before executing in env

setting up param noise

param_noise_actor/dense/kernel:0 <- actor/dense/kernel:0 + noise

param_noise_actor/dense/bias:0 <- actor/dense/bias:0 + noise

param_noise_actor/LayerNorm/beta:0 <- actor/LayerNorm/beta:0

param_noise_actor/LayerNorm/gamma:0 <- actor/LayerNorm/gamma:0

param_noise_actor/dense_1/kernel:0 <- actor/dense_1/kernel:0 + noise

```

param_noise_actor/dense_1/bias:0 <- actor/dense_1/bias:0 + noise
param_noise_actor/LayerNorm_1/beta:0 <- actor/LayerNorm_1/beta:0
param_noise_actor/LayerNorm_1/gamma:0 <- actor/LayerNorm_1/gamma:0
param_noise_actor/dense_2/kernel:0 <- actor/dense_2/kernel:0 + noise
param_noise_actor/dense_2/bias:0 <- actor/dense_2/bias:0 + noise
adaptive_param_noise_actor/dense/kernel:0 <- actor/dense/kernel:0 + noise
adaptive_param_noise_actor/dense/bias:0 <- actor/dense/bias:0 + noise
adaptive_param_noise_actor/LayerNorm/beta:0 <- actor/LayerNorm/beta:0
adaptive_param_noise_actor/LayerNorm/gamma:0 <- actor/LayerNorm/gamma:0
adaptive_param_noise_actor/dense_1/kernel:0 <- actor/dense_1/kernel:0 + noise
adaptive_param_noise_actor/dense_1/bias:0 <- actor/dense_1/bias:0 + noise
adaptive_param_noise_actor/LayerNorm_1/beta:0 <- actor/LayerNorm_1/beta:0
adaptive_param_noise_actor/LayerNorm_1/gamma:0 <- actor/LayerNorm_1/gamma:0
adaptive_param_noise_actor/dense_2/kernel:0 <- actor/dense_2/kernel:0 + noise
adaptive_param_noise_actor/dense_2/bias:0 <- actor/dense_2/bias:0 + noise
setting up actor optimizer
  actor shapes: [[3, 64], [64], [64], [64], [64, 64], [64], [64], [64], [64, 1], [1]]
  actor params: 4737
setting up critic optimizer
  regularizing: critic/dense/kernel:0
  regularizing: critic/dense_1/kernel:0
  regularizing: critic/dense_2/kernel:0
  applying l2 regularization with 0.01
  critic shapes: [[3, 64], [64], [64], [64], [65, 64], [64], [64], [64], [64, 1], [1]]
  critic params: 4801
setting up target updates ...
  target_actor/dense/kernel:0 <- actor/dense/kernel:0
  target_actor/dense/bias:0 <- actor/dense/bias:0
  target_actor/LayerNorm/beta:0 <- actor/LayerNorm/beta:0
  target_actor/LayerNorm/gamma:0 <- actor/LayerNorm/gamma:0
  target_actor/dense_1/kernel:0 <- actor/dense_1/kernel:0
  target_actor/dense_1/bias:0 <- actor/dense_1/bias:0
  target_actor/LayerNorm_1/beta:0 <- actor/LayerNorm_1/beta:0
  target_actor/LayerNorm_1/gamma:0 <- actor/LayerNorm_1/gamma:0
  target_actor/dense_2/kernel:0 <- actor/dense_2/kernel:0
  target_actor/dense_2/bias:0 <- actor/dense_2/bias:0
setting up target updates ...
  target_critic/dense/kernel:0 <- critic/dense/kernel:0
  target_critic/dense/bias:0 <- critic/dense/bias:0
  target_critic/LayerNorm/beta:0 <- critic/LayerNorm/beta:0
  target_critic/LayerNorm/gamma:0 <- critic/LayerNorm/gamma:0
  target_critic/dense_1/kernel:0 <- critic/dense_1/kernel:0
  target_critic/dense_1/bias:0 <- critic/dense_1/bias:0
  target_critic/LayerNorm_1/beta:0 <- critic/LayerNorm_1/beta:0
  target_critic/LayerNorm_1/gamma:0 <- critic/LayerNorm_1/gamma:0
  target_critic/dense_2/kernel:0 <- critic/dense_2/kernel:0
  target_critic/dense_2/bias:0 <- critic/dense_2/bias:0
Using agent with the following configuration:

```



```
dict_items([('obs0', <tf.Tensor 'obs0:0' shape=(?, 3) dtype=float32>), ('obs1', <tf.Tensor 'obs1
```

```
-----
| obs_rms_mean          | -0.0591 |
| obs_rms_std           | 1.72    |
| param_noise_stddev    | 0.188   |
| reference_Q_mean      | -58.3   |
| reference_Q_std       | 11.9    |
| reference_action_mean | 0.935   |
| reference_action_std  | 0.0133  |
| reference_actor_Q_mean | -58.2   |
| reference_actor_Q_std | 11.9    |
| reference_perturbed_... | 0.276   |
| rollout/Q_mean        | -25.1   |
| rollout/actions_mean  | 0.03    |
| rollout/actions_std   | 0.617   |
| rollout/episode_steps | 200     |
| rollout/episodes      | 10      |
| rollout/return        | -1.55e+03 |
| rollout/return_history | -1.55e+03 |
| total/duration        | 13      |
| total/episodes        | 10      |
| total/epochs          | 1       |
| total/steps           | 2000    |
| total/steps_per_second | 154     |
| train/loss_actor      | 55      |
| train/loss_critic     | 17      |
| train/param_noise_di... | 0.132   |
-----
```

```
-----
| obs_rms_mean          | -0.202  |
| obs_rms_std           | 1.96    |
| param_noise_stddev    | 0.158   |
| reference_Q_mean      | -114    |
| reference_Q_std       | 20.7    |
| reference_action_mean | -0.475  |
| reference_action_std  | 0.869   |
| reference_actor_Q_mean | -113    |
| reference_actor_Q_std | 21.8    |
| reference_perturbed_... | 0.0659  |
| rollout/Q_mean        | -53     |
| rollout/actions_mean  | -0.0864 |
| rollout/actions_std   | 0.721   |
| rollout/episode_steps | 200     |
| rollout/episodes      | 20      |
| rollout/return        | -1.5e+03 |
| rollout/return_history | -1.5e+03 |
| total/duration        | 25.6    |
-----
```

total/episodes	20	
total/epochs	2	
total/steps	4000	
total/steps_per_second	156	
train/loss_actor	108	
train/loss_critic	56.4	
train/param_noise_di...	0.644	

obs_rms_mean	-0.0779	
obs_rms_std	1.94	
param_noise_stddev	0.129	
reference_Q_mean	-161	
reference_Q_std	23.2	
reference_action_mean	0.494	
reference_action_std	0.836	
reference_actor_Q_mean	-159	
reference_actor_Q_std	24.3	
reference_perturbed_...	0.524	
rollout/Q_mean	-77.9	
rollout/actions_mean	-0.0793	
rollout/actions_std	0.743	
rollout/episode_steps	200	
rollout/episodes	30	
rollout/return	-1.48e+03	
rollout/return_history	-1.48e+03	
total/duration	38.2	
total/episodes	30	
total/epochs	3	
total/steps	6000	
total/steps_per_second	157	
train/loss_actor	149	
train/loss_critic	100	
train/param_noise_di...	0.621	

obs_rms_mean	-0.0434	
obs_rms_std	1.85	
param_noise_stddev	0.106	
reference_Q_mean	-189	
reference_Q_std	36	
reference_action_mean	0.449	
reference_action_std	0.836	
reference_actor_Q_mean	-186	
reference_actor_Q_std	37.2	
reference_perturbed_...	0.552	

rollout/Q_mean	-98.4	
rollout/actions_mean	-0.0262	
rollout/actions_std	0.737	
rollout/episode_steps	200	
rollout/episodes	40	
rollout/return	-1.44e+03	
rollout/return_history	-1.44e+03	
total/duration	50.8	
total/episodes	40	
total/epochs	4	
total/steps	8000	
total/steps_per_second	158	
train/loss_actor	175	
train/loss_critic	104	
train/param_noise_di...	0.88	

obs_rms_mean	-0.0486	
obs_rms_std	1.83	
param_noise_stddev	0.0867	
reference_Q_mean	-206	
reference_Q_std	41	
reference_action_mean	0.0805	
reference_action_std	0.937	
reference_actor_Q_mean	-203	
reference_actor_Q_std	42.2	
reference_perturbed_...	0.612	
rollout/Q_mean	-110	
rollout/actions_mean	0.00382	
rollout/actions_std	0.735	
rollout/episode_steps	200	
rollout/episodes	50	
rollout/return	-1.34e+03	
rollout/return_history	-1.34e+03	
total/duration	63.5	
total/episodes	50	
total/epochs	5	
total/steps	10000	
total/steps_per_second	158	
train/loss_actor	182	
train/loss_critic	128	
train/param_noise_di...	0.598	

obs_rms_mean	0.0786	
obs_rms_std	1.85	

param_noise_stddev	0.0711	
reference_Q_mean	-216	
reference_Q_std	46.9	
reference_action_mean	0.062	
reference_action_std	0.935	
reference_actor_Q_mean	-212	
reference_actor_Q_std	49.1	
reference_perturbed_...	0.58	
rollout/Q_mean	-115	
rollout/actions_mean	0.00566	
rollout/actions_std	0.723	
rollout/episode_steps	200	
rollout/episodes	60	
rollout/return	-1.27e+03	
rollout/return_history	-1.27e+03	
total/duration	76.1	
total/episodes	60	
total/epochs	6	
total/steps	12000	
total/steps_per_second	158	
train/loss_actor	182	
train/loss_critic	247	
train/param_noise_di...	0.55	

obs_rms_mean	0.193	
obs_rms_std	1.85	
param_noise_stddev	0.0582	
reference_Q_mean	-219	
reference_Q_std	56	
reference_action_mean	0.0381	
reference_action_std	0.942	
reference_actor_Q_mean	-214	
reference_actor_Q_std	59.2	
reference_perturbed_...	0.449	
rollout/Q_mean	-117	
rollout/actions_mean	0.019	
rollout/actions_std	0.714	
rollout/episode_steps	200	
rollout/episodes	70	
rollout/return	-1.2e+03	
rollout/return_history	-1.2e+03	
total/duration	88.7	
total/episodes	70	
total/epochs	7	
total/steps	14000	
total/steps_per_second	158	

train/loss_actor	171	
train/loss_critic	229	
train/param_noise_di...	0.654	

obs_rms_mean	0.258	
obs_rms_std	1.83	
param_noise_stddev	0.0477	
reference_Q_mean	-221	
reference_Q_std	69.5	
reference_action_mean	-0.295	
reference_action_std	0.9	
reference_actor_Q_mean	-216	
reference_actor_Q_std	72.1	
reference_perturbed_...	0.769	
rollout/Q_mean	-116	
rollout/actions_mean	0.0181	
rollout/actions_std	0.708	
rollout/episode_steps	200	
rollout/episodes	80	
rollout/return	-1.14e+03	
rollout/return_history	-1.14e+03	
total/duration	101	
total/episodes	80	
total/epochs	8	
total/steps	16000	
total/steps_per_second	158	
train/loss_actor	151	
train/loss_critic	92.4	
train/param_noise_di...	0.484	

obs_rms_mean	0.264	
obs_rms_std	1.79	
param_noise_stddev	0.0391	
reference_Q_mean	-213	
reference_Q_std	79.3	
reference_action_mean	0.308	
reference_action_std	0.881	
reference_actor_Q_mean	-207	
reference_actor_Q_std	82.5	
reference_perturbed_...	0.783	
rollout/Q_mean	-110	
rollout/actions_mean	0.031	
rollout/actions_std	0.693	
rollout/episode_steps	200	

rollout/episodes	90	
rollout/return	-1.06e+03	
rollout/return_history	-1.06e+03	
total/duration	114	
total/episodes	90	
total/epochs	9	
total/steps	18000	
total/steps_per_second	158	
train/loss_actor	132	
train/loss_critic	126	
train/param_noise_di...	0.455	

obs_rms_mean	0.266	
obs_rms_std	1.74	
param_noise_stddev	0.0334	
reference_Q_mean	-207	
reference_Q_std	84.1	
reference_action_mean	-0.0385	
reference_action_std	0.924	
reference_actor_Q_mean	-200	
reference_actor_Q_std	86.9	
reference_perturbed_...	0.817	
rollout/Q_mean	-101	
rollout/actions_mean	0.0291	
rollout/actions_std	0.676	
rollout/episode_steps	200	
rollout/episodes	100	
rollout/return	-983	
rollout/return_history	-983	
total/duration	126	
total/episodes	100	
total/epochs	10	
total/steps	20000	
total/steps_per_second	158	
train/loss_actor	106	
train/loss_critic	136	
train/param_noise_di...	0.723	

obs_rms_mean	0.265	
obs_rms_std	1.7	
param_noise_stddev	0.0279	
reference_Q_mean	-196	
reference_Q_std	86.7	
reference_action_mean	-0.423	

reference_action_std	0.824	
reference_actor_Q_mean	-189	
reference_actor_Q_std	88.7	
reference_perturbed_...	0.82	
rollout/Q_mean	-94.2	
rollout/actions_mean	0.0364	
rollout/actions_std	0.662	
rollout/episode_steps	200	
rollout/episodes	110	
rollout/return	-927	
rollout/return_history	-864	
total/duration	139	
total/episodes	110	
total/epochs	11	
total/steps	22000	
total/steps_per_second	158	
train/loss_actor	88.1	
train/loss_critic	111	
train/param_noise_di...	0.388	

obs_rms_mean	0.248	
obs_rms_std	1.68	
param_noise_stddev	0.0233	
reference_Q_mean	-182	
reference_Q_std	89.8	
reference_action_mean	-0.484	
reference_action_std	0.826	
reference_actor_Q_mean	-174	
reference_actor_Q_std	92.8	
reference_perturbed_...	0.746	
rollout/Q_mean	-88.1	
rollout/actions_mean	0.0375	
rollout/actions_std	0.651	
rollout/episode_steps	200	
rollout/episodes	120	
rollout/return	-881	
rollout/return_history	-757	
total/duration	152	
total/episodes	120	
total/epochs	12	
total/steps	24000	
total/steps_per_second	158	
train/loss_actor	73.7	
train/loss_critic	109	
train/param_noise_di...	0.302	

obs_rms_mean	0.249	
obs_rms_std	1.64	
param_noise_stddev	0.0199	
reference_Q_mean	-179	
reference_Q_std	91.3	
reference_action_mean	0.0835	
reference_action_std	0.939	
reference_actor_Q_mean	-171	
reference_actor_Q_std	92.8	
reference_perturbed_...	0.877	
rollout/Q_mean	-81.1	
rollout/actions_mean	0.0387	
rollout/actions_std	0.64	
rollout/episode_steps	200	
rollout/episodes	130	
rollout/return	-830	
rollout/return_history	-634	
total/duration	164	
total/episodes	130	
total/epochs	13	
total/steps	26000	
total/steps_per_second	158	
train/loss_actor	62.8	
train/loss_critic	25.6	
train/param_noise_di...	0.265	

obs_rms_mean	0.25	
obs_rms_std	1.6	
param_noise_stddev	0.018	
reference_Q_mean	-163	
reference_Q_std	82.7	
reference_action_mean	0.0906	
reference_action_std	0.937	
reference_actor_Q_mean	-156	
reference_actor_Q_std	84.2	
reference_perturbed_...	0.87	
rollout/Q_mean	-75.1	
rollout/actions_mean	0.0282	
rollout/actions_std	0.628	
rollout/episode_steps	200	
rollout/episodes	140	
rollout/return	-783	
rollout/return_history	-520	
total/duration	177	

total/episodes	140	
total/epochs	14	
total/steps	28000	
total/steps_per_second	158	
train/loss_actor	52.9	
train/loss_critic	41.9	
train/param_noise_di...	0.319	

obs_rms_mean	0.253	
obs_rms_std	1.57	
param_noise_stddev	0.0163	
reference_Q_mean	-158	
reference_Q_std	82.4	
reference_action_mean	0.0745	
reference_action_std	0.956	
reference_actor_Q_mean	-152	
reference_actor_Q_std	84	
reference_perturbed_...	0.907	
rollout/Q_mean	-69.9	
rollout/actions_mean	0.025	
rollout/actions_std	0.616	
rollout/episode_steps	200	
rollout/episodes	150	
rollout/return	-743	
rollout/return_history	-443	
total/duration	190	
total/episodes	150	
total/epochs	15	
total/steps	30000	
total/steps_per_second	158	
train/loss_actor	47.3	
train/loss_critic	17.1	
train/param_noise_di...	0.206	

obs_rms_mean	0.255	
obs_rms_std	1.54	
param_noise_stddev	0.016	
reference_Q_mean	-156	
reference_Q_std	80.4	
reference_action_mean	-0.152	
reference_action_std	0.957	
reference_actor_Q_mean	-149	
reference_actor_Q_std	81	
reference_perturbed_...	0.938	

rollout/Q_mean	-65.2	
rollout/actions_mean	0.018	
rollout/actions_std	0.606	
rollout/episode_steps	200	
rollout/episodes	160	
rollout/return	-707	
rollout/return_history	-371	
total/duration	202	
total/episodes	160	
total/epochs	16	
total/steps	32000	
total/steps_per_second	158	
train/loss_actor	44.7	
train/loss_critic	27.2	
train/param_noise_di...	0.368	

obs_rms_mean	0.25	
obs_rms_std	1.51	
param_noise_stddev	0.0166	
reference_Q_mean	-153	
reference_Q_std	76.7	
reference_action_mean	-0.121	
reference_action_std	0.968	
reference_actor_Q_mean	-146	
reference_actor_Q_std	77.3	
reference_perturbed_...	0.898	
rollout/Q_mean	-61.2	
rollout/actions_mean	0.0302	
rollout/actions_std	0.599	
rollout/episode_steps	200	
rollout/episodes	170	
rollout/return	-676	
rollout/return_history	-308	
total/duration	215	
total/episodes	170	
total/epochs	17	
total/steps	34000	
total/steps_per_second	158	
train/loss_actor	39.2	
train/loss_critic	12.3	
train/param_noise_di...	0.17	

obs_rms_mean	0.248	
obs_rms_std	1.48	

param_noise_stddev	0.016	
reference_Q_mean	-152	
reference_Q_std	76.9	
reference_action_mean	0.146	
reference_action_std	0.958	
reference_actor_Q_mean	-146	
reference_actor_Q_std	78.3	
reference_perturbed_...	0.935	
rollout/Q_mean	-57.9	
rollout/actions_mean	0.034	
rollout/actions_std	0.591	
rollout/episode_steps	200	
rollout/episodes	180	
rollout/return	-650	
rollout/return_history	-259	
total/duration	228	
total/episodes	180	
total/epochs	18	
total/steps	36000	
total/steps_per_second	158	
train/loss_actor	38.6	
train/loss_critic	25.7	
train/param_noise_di...	0.172	

obs_rms_mean	0.245	
obs_rms_std	1.46	
param_noise_stddev	0.0148	
reference_Q_mean	-150	
reference_Q_std	74.6	
reference_action_mean	0.215	
reference_action_std	0.94	
reference_actor_Q_mean	-143	
reference_actor_Q_std	75.2	
reference_perturbed_...	0.912	
rollout/Q_mean	-54.8	
rollout/actions_mean	0.0342	
rollout/actions_std	0.583	
rollout/episode_steps	200	
rollout/episodes	190	
rollout/return	-625	
rollout/return_history	-231	
total/duration	240	
total/episodes	190	
total/epochs	19	
total/steps	38000	
total/steps_per_second	158	

train/loss_actor	36.7	
train/loss_critic	12.4	
train/param_noise_di...	0.158	

obs_rms_mean	0.245	
obs_rms_std	1.44	
param_noise_stddev	0.0145	
reference_Q_mean	-152	
reference_Q_std	74.7	
reference_action_mean	-0.162	
reference_action_std	0.953	
reference_actor_Q_mean	-146	
reference_actor_Q_std	75.4	
reference_perturbed_...	0.935	
rollout/Q_mean	-52.3	
rollout/actions_mean	0.0386	
rollout/actions_std	0.577	
rollout/episode_steps	200	
rollout/episodes	200	
rollout/return	-605	
rollout/return_history	-227	
total/duration	253	
total/episodes	200	
total/epochs	20	
total/steps	40000	
total/steps_per_second	158	
train/loss_actor	34.9	
train/loss_critic	21.2	
train/param_noise_di...	0.333	

obs_rms_mean	0.252	
obs_rms_std	1.41	
param_noise_stddev	0.0139	
reference_Q_mean	-150	
reference_Q_std	74.4	
reference_action_mean	-0.133	
reference_action_std	0.945	
reference_actor_Q_mean	-143	
reference_actor_Q_std	74.4	
reference_perturbed_...	0.897	
rollout/Q_mean	-49.1	
rollout/actions_mean	0.0242	
rollout/actions_std	0.571	
rollout/episode_steps	200	

rollout/episodes	210	
rollout/return	-582	
rollout/return_history	-203	
total/duration	266	
total/episodes	210	
total/epochs	21	
total/steps	42000	
total/steps_per_second	158	
train/loss_actor	34.1	
train/loss_critic	18.5	
train/param_noise_di...	0.245	

obs_rms_mean	0.256	
obs_rms_std	1.39	
param_noise_stddev	0.0139	
reference_Q_mean	-149	
reference_Q_std	74.3	
reference_action_mean	0.184	
reference_action_std	0.912	
reference_actor_Q_mean	-143	
reference_actor_Q_std	75.3	
reference_perturbed_...	0.909	
rollout/Q_mean	-46.1	
rollout/actions_mean	0.00381	
rollout/actions_std	0.571	
rollout/episode_steps	200	
rollout/episodes	220	
rollout/return	-562	
rollout/return_history	-180	
total/duration	279	
total/episodes	220	
total/epochs	22	
total/steps	44000	
total/steps_per_second	158	
train/loss_actor	28.4	
train/loss_critic	50.8	
train/param_noise_di...	0.216	

obs_rms_mean	0.263	
obs_rms_std	1.38	
param_noise_stddev	0.0142	
reference_Q_mean	-147	
reference_Q_std	72.5	
reference_action_mean	0.219	

reference_action_std	0.911	
reference_actor_Q_mean	-141	
reference_actor_Q_std	73.3	
reference_perturbed_...	0.905	
rollout/Q_mean	-43.6	
rollout/actions_mean	-0.017	
rollout/actions_std	0.575	
rollout/episode_steps	200	
rollout/episodes	230	
rollout/return	-547	
rollout/return_history	-179	
total/duration	291	
total/episodes	230	
total/epochs	23	
total/steps	46000	
total/steps_per_second	158	
train/loss_actor	21.3	
train/loss_critic	23.7	
train/param_noise_di...	0.284	

obs_rms_mean	0.27	
obs_rms_std	1.36	
param_noise_stddev	0.0154	
reference_Q_mean	-145	
reference_Q_std	73.4	
reference_action_mean	0.0563	
reference_action_std	0.936	
reference_actor_Q_mean	-139	
reference_actor_Q_std	74.6	
reference_perturbed_...	0.874	
rollout/Q_mean	-40.9	
rollout/actions_mean	-0.0368	
rollout/actions_std	0.58	
rollout/episode_steps	200	
rollout/episodes	240	
rollout/return	-532	
rollout/return_history	-179	
total/duration	304	
total/episodes	240	
total/epochs	24	
total/steps	48000	
total/steps_per_second	158	
train/loss_actor	16.8	
train/loss_critic	48.2	
train/param_noise_di...	0.186	

obs_rms_mean	0.276	
obs_rms_std	1.35	
param_noise_stddev	0.0163	
reference_Q_mean	-145	
reference_Q_std	79.2	
reference_action_mean	-0.198	
reference_action_std	0.924	
reference_actor_Q_mean	-139	
reference_actor_Q_std	80.5	
reference_perturbed_...	0.894	
rollout/Q_mean	-38.5	
rollout/actions_mean	-0.0553	
rollout/actions_std	0.583	
rollout/episode_steps	200	
rollout/episodes	250	
rollout/return	-521	
rollout/return_history	-189	
total/duration	317	
total/episodes	250	
total/epochs	25	
total/steps	50000	
total/steps_per_second	158	
train/loss_actor	13	
train/loss_critic	27.5	
train/param_noise_di...	0.0989	

obs_rms_mean	0.278	
obs_rms_std	1.34	
param_noise_stddev	0.016	
reference_Q_mean	-142	
reference_Q_std	81.5	
reference_action_mean	0.172	
reference_action_std	0.928	
reference_actor_Q_mean	-136	
reference_actor_Q_std	82	
reference_perturbed_...	0.907	
rollout/Q_mean	-36.6	
rollout/actions_mean	-0.0685	
rollout/actions_std	0.585	
rollout/episode_steps	200	
rollout/episodes	260	
rollout/return	-514	
rollout/return_history	-206	
total/duration	329	

total/episodes	260	
total/epochs	26	
total/steps	52000	
total/steps_per_second	158	
train/loss_actor	10.2	
train/loss_critic	28.6	
train/param_noise_di...	0.215	

obs_rms_mean	0.285	
obs_rms_std	1.33	
param_noise_stddev	0.016	
reference_Q_mean	-138	
reference_Q_std	80.9	
reference_action_mean	0.119	
reference_action_std	0.941	
reference_actor_Q_mean	-132	
reference_actor_Q_std	81.2	
reference_perturbed_...	0.906	
rollout/Q_mean	-34.4	
rollout/actions_mean	-0.0855	
rollout/actions_std	0.588	
rollout/episode_steps	200	
rollout/episodes	270	
rollout/return	-505	
rollout/return_history	-215	
total/duration	342	
total/episodes	270	
total/epochs	27	
total/steps	54000	
total/steps_per_second	158	
train/loss_actor	4.83	
train/loss_critic	31.8	
train/param_noise_di...	0.288	

obs_rms_mean	0.29	
obs_rms_std	1.32	
param_noise_stddev	0.015	
reference_Q_mean	-134	
reference_Q_std	79.9	
reference_action_mean	0.126	
reference_action_std	0.96	
reference_actor_Q_mean	-128	
reference_actor_Q_std	80	
reference_perturbed_...	0.928	

rollout/Q_mean	-32.1	
rollout/actions_mean	-0.102	
rollout/actions_std	0.59	
rollout/episode_steps	200	
rollout/episodes	280	
rollout/return	-496	
rollout/return_history	-219	
total/duration	355	
total/episodes	280	
total/epochs	28	
total/steps	56000	
total/steps_per_second	158	
train/loss_actor	2.51	
train/loss_critic	18.6	
train/param_noise_di...	0.265	

obs_rms_mean	0.293	
obs_rms_std	1.31	
param_noise_stddev	0.0148	
reference_Q_mean	-130	
reference_Q_std	76.3	
reference_action_mean	-0.093	
reference_action_std	0.958	
reference_actor_Q_mean	-125	
reference_actor_Q_std	76.2	
reference_perturbed_...	0.908	
rollout/Q_mean	-30	
rollout/actions_mean	-0.118	
rollout/actions_std	0.592	
rollout/episode_steps	200	
rollout/episodes	290	
rollout/return	-487	
rollout/return_history	-224	
total/duration	367	
total/episodes	290	
total/epochs	29	
total/steps	58000	
total/steps_per_second	158	
train/loss_actor	-1.8	
train/loss_critic	28.7	
train/param_noise_di...	0.267	

obs_rms_mean	0.298	
obs_rms_std	1.3	

param_noise_stddev	0.0139	
reference_Q_mean	-126	
reference_Q_std	74.8	
reference_action_mean	-0.147	
reference_action_std	0.95	
reference_actor_Q_mean	-120	
reference_actor_Q_std	74.9	
reference_perturbed_...	0.933	
rollout/Q_mean	-27.7	
rollout/actions_mean	-0.135	
rollout/actions_std	0.593	
rollout/episode_steps	200	
rollout/episodes	300	
rollout/return	-476	
rollout/return_history	-217	
total/duration	380	
total/episodes	300	
total/epochs	30	
total/steps	60000	
total/steps_per_second	158	
train/loss_actor	-5.55	
train/loss_critic	24.3	
train/param_noise_di...	0.274	

obs_rms_mean	0.302	
obs_rms_std	1.29	
param_noise_stddev	0.0142	
reference_Q_mean	-126	
reference_Q_std	77.2	
reference_action_mean	0.11	
reference_action_std	0.953	
reference_actor_Q_mean	-119	
reference_actor_Q_std	77.3	
reference_perturbed_...	0.943	
rollout/Q_mean	-25.7	
rollout/actions_mean	-0.151	
rollout/actions_std	0.594	
rollout/episode_steps	200	
rollout/episodes	310	
rollout/return	-466	
rollout/return_history	-222	
total/duration	393	
total/episodes	310	
total/epochs	31	
total/steps	62000	
total/steps_per_second	158	

train/loss_actor	-7.83	
train/loss_critic	18.8	
train/param_noise_di...	0.181	

obs_rms_mean	0.303	
obs_rms_std	1.27	
param_noise_stddev	0.0131	
reference_Q_mean	-125	
reference_Q_std	77.5	
reference_action_mean	0.207	
reference_action_std	0.922	
reference_actor_Q_mean	-119	
reference_actor_Q_std	78.1	
reference_perturbed_...	0.877	
rollout/Q_mean	-23.7	
rollout/actions_mean	-0.164	
rollout/actions_std	0.594	
rollout/episode_steps	200	
rollout/episodes	320	
rollout/return	-456	
rollout/return_history	-224	
total/duration	406	
total/episodes	320	
total/epochs	32	
total/steps	64000	
total/steps_per_second	158	
train/loss_actor	-13.3	
train/loss_critic	22.3	
train/param_noise_di...	0.22	

obs_rms_mean	0.304	
obs_rms_std	1.26	
param_noise_stddev	0.0142	
reference_Q_mean	-122	
reference_Q_std	77.4	
reference_action_mean	-0.121	
reference_action_std	0.95	
reference_actor_Q_mean	-116	
reference_actor_Q_std	77.5	
reference_perturbed_...	0.911	
rollout/Q_mean	-21.7	
rollout/actions_mean	-0.176	
rollout/actions_std	0.592	
rollout/episode_steps	200	

rollout/episodes	330	
rollout/return	-447	
rollout/return_history	-216	
total/duration	419	
total/episodes	330	
total/epochs	33	
total/steps	66000	
total/steps_per_second	158	
train/loss_actor	-11.8	
train/loss_critic	46.4	
train/param_noise_di...	0.468	

obs_rms_mean	0.306	
obs_rms_std	1.25	
param_noise_stddev	0.0145	
reference_Q_mean	-124	
reference_Q_std	78.3	
reference_action_mean	-0.141	
reference_action_std	0.956	
reference_actor_Q_mean	-119	
reference_actor_Q_std	78	
reference_perturbed_...	0.877	
rollout/Q_mean	-19.8	
rollout/actions_mean	-0.189	
rollout/actions_std	0.59	
rollout/episode_steps	200	
rollout/episodes	340	
rollout/return	-438	
rollout/return_history	-213	
total/duration	432	
total/episodes	340	
total/epochs	34	
total/steps	68000	
total/steps_per_second	158	
train/loss_actor	-10.3	
train/loss_critic	17.3	
train/param_noise_di...	0.247	

obs_rms_mean	0.306	
obs_rms_std	1.24	
param_noise_stddev	0.0142	
reference_Q_mean	-120	
reference_Q_std	73.1	
reference_action_mean	-0.121	

reference_action_std	0.946	
reference_actor_Q_mean	-114	
reference_actor_Q_std	72.2	
reference_perturbed_...	0.911	
rollout/Q_mean	-18.3	
rollout/actions_mean	-0.2	
rollout/actions_std	0.591	
rollout/episode_steps	200	
rollout/episodes	350	
rollout/return	-431	
rollout/return_history	-205	
total/duration	444	
total/episodes	350	
total/epochs	35	
total/steps	70000	
total/steps_per_second	158	
train/loss_actor	-10	
train/loss_critic	28	
train/param_noise_di...	0.216	

obs_rms_mean	0.31	
obs_rms_std	1.23	
param_noise_stddev	0.0139	
reference_Q_mean	-124	
reference_Q_std	74	
reference_action_mean	-0.113	
reference_action_std	0.951	
reference_actor_Q_mean	-118	
reference_actor_Q_std	73.5	
reference_perturbed_...	0.852	
rollout/Q_mean	-16.8	
rollout/actions_mean	-0.209	
rollout/actions_std	0.589	
rollout/episode_steps	200	
rollout/episodes	360	
rollout/return	-424	
rollout/return_history	-189	
total/duration	457	
total/episodes	360	
total/epochs	36	
total/steps	72000	
total/steps_per_second	157	
train/loss_actor	-13	
train/loss_critic	22.8	
train/param_noise_di...	0.175	

obs_rms_mean	0.311	
obs_rms_std	1.22	
param_noise_stddev	0.0136	
reference_Q_mean	-126	
reference_Q_std	74.2	
reference_action_mean	-0.118	
reference_action_std	0.948	
reference_actor_Q_mean	-120	
reference_actor_Q_std	73.6	
reference_perturbed_...	0.847	
rollout/Q_mean	-15.2	
rollout/actions_mean	-0.219	
rollout/actions_std	0.587	
rollout/episode_steps	200	
rollout/episodes	370	
rollout/return	-416	
rollout/return_history	-174	
total/duration	470	
total/episodes	370	
total/epochs	37	
total/steps	74000	
total/steps_per_second	157	
train/loss_actor	-12.4	
train/loss_critic	16.7	
train/param_noise_di...	0.238	

obs_rms_mean	0.311	
obs_rms_std	1.21	
param_noise_stddev	0.0123	
reference_Q_mean	-125	
reference_Q_std	75	
reference_action_mean	-0.0533	
reference_action_std	0.956	
reference_actor_Q_mean	-120	
reference_actor_Q_std	74.6	
reference_perturbed_...	0.899	
rollout/Q_mean	-14	
rollout/actions_mean	-0.225	
rollout/actions_std	0.587	
rollout/episode_steps	200	
rollout/episodes	380	
rollout/return	-411	
rollout/return_history	-173	
total/duration	483	

total/episodes	380	
total/epochs	38	
total/steps	76000	
total/steps_per_second	157	
train/loss_actor	-14.3	
train/loss_critic	27	
train/param_noise_di...	0.247	

obs_rms_mean	0.311	
obs_rms_std	1.21	
param_noise_stddev	0.0131	
reference_Q_mean	-123	
reference_Q_std	74.8	
reference_action_mean	0.161	
reference_action_std	0.935	
reference_actor_Q_mean	-118	
reference_actor_Q_std	74.5	
reference_perturbed_...	0.896	
rollout/Q_mean	-12.8	
rollout/actions_mean	-0.235	
rollout/actions_std	0.586	
rollout/episode_steps	200	
rollout/episodes	390	
rollout/return	-406	
rollout/return_history	-170	
total/duration	495	
total/episodes	390	
total/epochs	39	
total/steps	78000	
total/steps_per_second	157	
train/loss_actor	-12.5	
train/loss_critic	40	
train/param_noise_di...	0.22	

obs_rms_mean	0.313	
obs_rms_std	1.2	
param_noise_stddev	0.0134	
reference_Q_mean	-126	
reference_Q_std	76.5	
reference_action_mean	-0.0707	
reference_action_std	0.928	
reference_actor_Q_mean	-120	
reference_actor_Q_std	75.7	
reference_perturbed_...	0.95	

rollout/Q_mean	-11.5	
rollout/actions_mean	-0.241	
rollout/actions_std	0.584	
rollout/episode_steps	200	
rollout/episodes	400	
rollout/return	-400	
rollout/return_history	-171	
total/duration	508	
total/episodes	400	
total/epochs	40	
total/steps	80000	
total/steps_per_second	157	
train/loss_actor	-13.6	
train/loss_critic	17.8	
train/param_noise_di...	0.105	

obs_rms_mean	0.316	
obs_rms_std	1.19	
param_noise_stddev	0.0134	
reference_Q_mean	-126	
reference_Q_std	74.2	
reference_action_mean	0.202	
reference_action_std	0.912	
reference_actor_Q_mean	-119	
reference_actor_Q_std	73.6	
reference_perturbed_...	0.877	
rollout/Q_mean	-10.2	
rollout/actions_mean	-0.247	
rollout/actions_std	0.583	
rollout/episode_steps	200	
rollout/episodes	410	
rollout/return	-393	
rollout/return_history	-168	
total/duration	521	
total/episodes	410	
total/epochs	41	
total/steps	82000	
total/steps_per_second	157	
train/loss_actor	-13.9	
train/loss_critic	31.2	
train/param_noise_di...	0.322	

obs_rms_mean	0.316	
obs_rms_std	1.18	

param_noise_stddev	0.015	
reference_Q_mean	-125	
reference_Q_std	74.7	
reference_action_mean	-0.106	
reference_action_std	0.936	
reference_actor_Q_mean	-119	
reference_actor_Q_std	74.5	
reference_perturbed_...	0.921	
rollout/Q_mean	-9.21	
rollout/actions_mean	-0.252	
rollout/actions_std	0.581	
rollout/episode_steps	200	
rollout/episodes	420	
rollout/return	-389	
rollout/return_history	-173	
total/duration	534	
total/episodes	420	
total/epochs	42	
total/steps	84000	
total/steps_per_second	157	
train/loss_actor	-16.8	
train/loss_critic	20.7	
train/param_noise_di...	0.264	

obs_rms_mean	0.319	
obs_rms_std	1.17	
param_noise_stddev	0.0157	
reference_Q_mean	-126	
reference_Q_std	75.7	
reference_action_mean	-0.362	
reference_action_std	0.872	
reference_actor_Q_mean	-120	
reference_actor_Q_std	75.9	
reference_perturbed_...	0.858	
rollout/Q_mean	-8.03	
rollout/actions_mean	-0.258	
rollout/actions_std	0.577	
rollout/episode_steps	200	
rollout/episodes	430	
rollout/return	-382	
rollout/return_history	-169	
total/duration	547	
total/episodes	430	
total/epochs	43	
total/steps	86000	
total/steps_per_second	157	

train/loss_actor	-15.1	
train/loss_critic	21.9	
train/param_noise_di...	0.104	

obs_rms_mean	0.322	
obs_rms_std	1.17	
param_noise_stddev	0.0154	
reference_Q_mean	-122	
reference_Q_std	71.5	
reference_action_mean	0.0115	
reference_action_std	0.93	
reference_actor_Q_mean	-116	
reference_actor_Q_std	71.2	
reference_perturbed_...	0.873	
rollout/Q_mean	-7.04	
rollout/actions_mean	-0.263	
rollout/actions_std	0.575	
rollout/episode_steps	200	
rollout/episodes	440	
rollout/return	-377	
rollout/return_history	-172	
total/duration	559	
total/episodes	440	
total/epochs	44	
total/steps	88000	
total/steps_per_second	157	
train/loss_actor	-16.1	
train/loss_critic	25.5	
train/param_noise_di...	0.275	

obs_rms_mean	0.321	
obs_rms_std	1.16	
param_noise_stddev	0.0145	
reference_Q_mean	-123	
reference_Q_std	72.6	
reference_action_mean	-0.0899	
reference_action_std	0.933	
reference_actor_Q_mean	-117	
reference_actor_Q_std	72	
reference_perturbed_...	0.926	
rollout/Q_mean	-6.1	
rollout/actions_mean	-0.268	
rollout/actions_std	0.572	
rollout/episode_steps	200	

rollout/episodes	450	
rollout/return	-373	
rollout/return_history	-169	
total/duration	572	
total/episodes	450	
total/epochs	45	
total/steps	90000	
total/steps_per_second	157	
train/loss_actor	-15.5	
train/loss_critic	44	
train/param_noise_di...	0.211	

obs_rms_mean	0.32	
obs_rms_std	1.15	
param_noise_stddev	0.015	
reference_Q_mean	-126	
reference_Q_std	74.4	
reference_action_mean	0.168	
reference_action_std	0.909	
reference_actor_Q_mean	-121	
reference_actor_Q_std	74.8	
reference_perturbed_...	0.913	
rollout/Q_mean	-5.21	
rollout/actions_mean	-0.271	
rollout/actions_std	0.569	
rollout/episode_steps	200	
rollout/episodes	460	
rollout/return	-368	
rollout/return_history	-168	
total/duration	585	
total/episodes	460	
total/epochs	46	
total/steps	92000	
total/steps_per_second	157	
train/loss_actor	-15.1	
train/loss_critic	13.3	
train/param_noise_di...	0.0923	

obs_rms_mean	0.321	
obs_rms_std	1.15	
param_noise_stddev	0.0157	
reference_Q_mean	-127	
reference_Q_std	74.2	
reference_action_mean	0.0483	

reference_action_std	0.918	
reference_actor_Q_mean	-121	
reference_actor_Q_std	74.4	
reference_perturbed_...	0.888	
rollout/Q_mean	-4.41	
rollout/actions_mean	-0.275	
rollout/actions_std	0.567	
rollout/episode_steps	200	
rollout/episodes	470	
rollout/return	-363	
rollout/return_history	-169	
total/duration	598	
total/episodes	470	
total/epochs	47	
total/steps	94000	
total/steps_per_second	157	
train/loss_actor	-16.3	
train/loss_critic	11.4	
train/param_noise_di...	0.44	

obs_rms_mean	0.322	
obs_rms_std	1.14	
param_noise_stddev	0.0157	
reference_Q_mean	-129	
reference_Q_std	75.3	
reference_action_mean	0.0138	
reference_action_std	0.927	
reference_actor_Q_mean	-123	
reference_actor_Q_std	74.9	
reference_perturbed_...	0.873	
rollout/Q_mean	-3.64	
rollout/actions_mean	-0.278	
rollout/actions_std	0.565	
rollout/episode_steps	200	
rollout/episodes	480	
rollout/return	-359	
rollout/return_history	-163	
total/duration	610	
total/episodes	480	
total/epochs	48	
total/steps	96000	
total/steps_per_second	157	
train/loss_actor	-16.1	
train/loss_critic	11.6	
train/param_noise_di...	0.173	

obs_rms_mean	0.323	
obs_rms_std	1.13	
param_noise_stddev	0.016	
reference_Q_mean	-128	
reference_Q_std	75	
reference_action_mean	0.408	
reference_action_std	0.863	
reference_actor_Q_mean	-123	
reference_actor_Q_std	75.2	
reference_perturbed_...	0.88	
rollout/Q_mean	-2.88	
rollout/actions_mean	-0.28	
rollout/actions_std	0.563	
rollout/episode_steps	200	
rollout/episodes	490	
rollout/return	-355	
rollout/return_history	-160	
total/duration	623	
total/episodes	490	
total/epochs	49	
total/steps	98000	
total/steps_per_second	157	
train/loss_actor	-15.4	
train/loss_critic	15	
train/param_noise_di...	0.103	

obs_rms_mean	0.322	
obs_rms_std	1.13	
param_noise_stddev	0.0157	
reference_Q_mean	-123	
reference_Q_std	68.9	
reference_action_mean	-0.0272	
reference_action_std	0.941	
reference_actor_Q_mean	-117	
reference_actor_Q_std	69.5	
reference_perturbed_...	0.95	
rollout/Q_mean	-2.17	
rollout/actions_mean	-0.283	
rollout/actions_std	0.561	
rollout/episode_steps	200	
rollout/episodes	500	
rollout/return	-351	
rollout/return_history	-159	
total/duration	636	

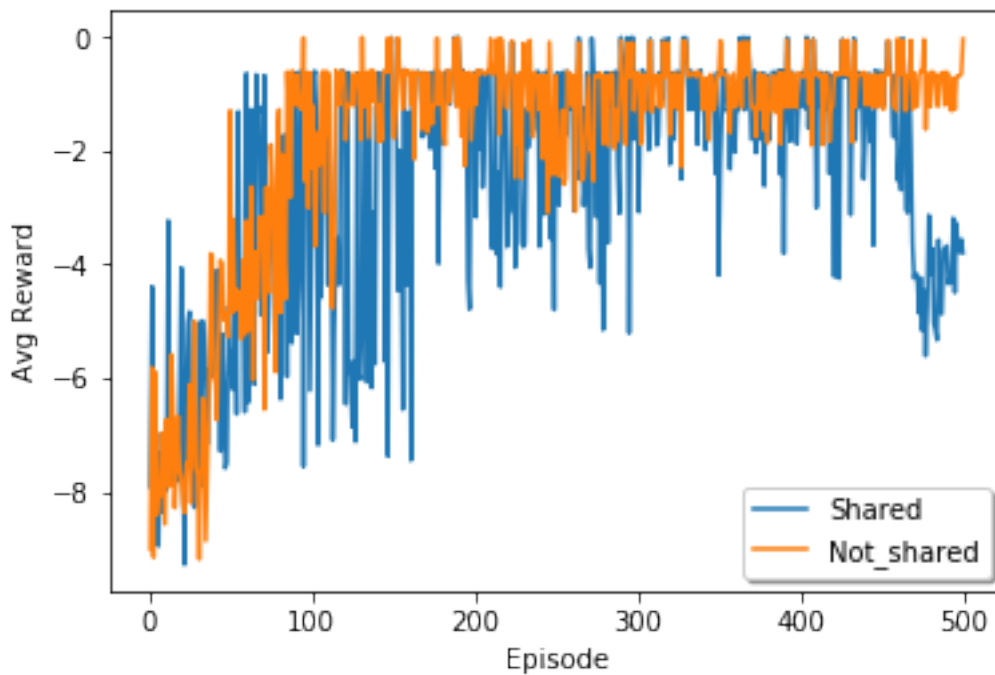
total/episodes	500	
total/epochs	50	
total/steps	100000	
total/steps_per_second	157	
train/loss_actor	-13.7	
train/loss_critic	40.9	
train/param_noise_di...	0.617	

[-9.0083088729298666, -5.8098595500575776, -9.1528298098407159, -5.8953800171774011, -8.40900707
total runtime: 639.1219322681427s

In [3]: %matplotlib inline

```
In [4]: import matplotlib.pyplot as plt
x = np.linspace(1,500,500)
fig, ax = plt.subplots()
ax.plot(x, r_history[0,:], label = "Shared")
ax.plot(x, r_history[1,:], label = "Not_shared")

legend = ax.legend(loc='lower right', shadow=True)
plt.xlabel("Episode")
plt.ylabel("Avg Reward")
plt.show()
```



From the graph above, we can see that the NN which has shared weights has a relatively faster convergence rate. However, variance of average rewards is larger when sharing weights between actor and critic network. This is because when sharing weights between critic and actor, the actor's network gets updates directly associated with rewards. This associate will accelerate training if the sample is a good move. However, as samples cannot always be good move, it will harm the learning process and increase the variance. Another possible cause is everytime the function calls `agent.train()` function, it updates the network parameters using `actor_optimizer` and `critic_optimizer`. These two will both update the first layer variables, meaning that the variables will be updated twice. This can cause faster learning, but will lead to higher variance and unstable performance.