Search spaces

When optimising MDPs, would we rather optimise within the set of potentially optimal policies, the $|A|^{|S|}$ discrete policies, or would we rather optimise within the set of possible value functions, $\mathbb{R}^{|S|}$, or …?

Q: How does the space we are searching within effect the search dynamics? Convergence, trajectories, …

Smaller search spaces are better. But added structure (for example, locally linear) can be exploited.

When transforming between two spaces, how does the optimisation space change? Does my abstraction make optimisation easier?

$$\max_{V} \mathbb{E}_{s \sim D} V(s) \tag{1}$$

$$\max_{\pi} \mathbb{E}_{s \sim D} V^{\pi}(s) \tag{2}$$

$$\max_{\theta} \mathbb{E}_{s \sim D} V_{\theta}(s) \tag{3}$$

$$\max_{\theta} \mathbb{E}_{s \sim D} V^{\pi_{\theta}}(s) \tag{4}$$

$$\max_{\phi} \mathbb{E}_{s \sim D} V^{\pi_{\theta_{\phi}}}(s) \tag{5}$$

$$\max_{\varphi} \mathbb{E}_{s \sim D} V^{\pi_{\theta_{\phi_{\varphi}}}}(s) \tag{6}$$

$$\tag{7}$$

We can pick the space we optimise in. Why would we want to pick one space over another?

- In which spaces can we go gradient descent?
- In which spaces can we do convex optimisation?
- In which spaces does momentum work well?
- …

Value iteration

```
value = init()
while not converged(value):
  value = update(value)
```

Policy iteration

```
policy = init()
while not converged(value):
  value = evaluate(policy)
  policy = update(value)
```

Parameter iteration

```
parameters = init()
while not converged(value):
  policy = fn(parameters)
  value = evaluate(policy)
  parameters = step(value)
```

Dynamics and connectivity

Ok, so if we parameterise our search space. We have now changed the topology of our search space. See these gradient flows for example;
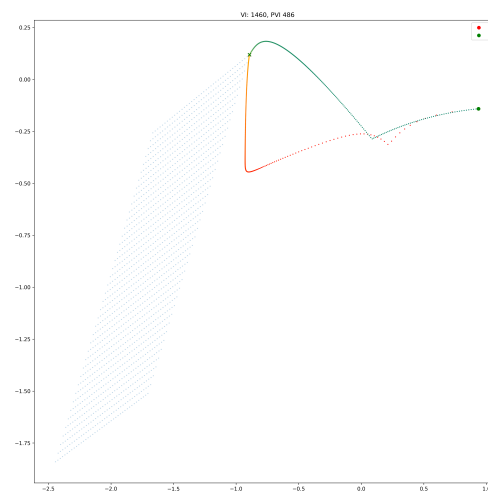

Figure 1: The optimisation dynamics of value iteration versus parameterised value iteration.
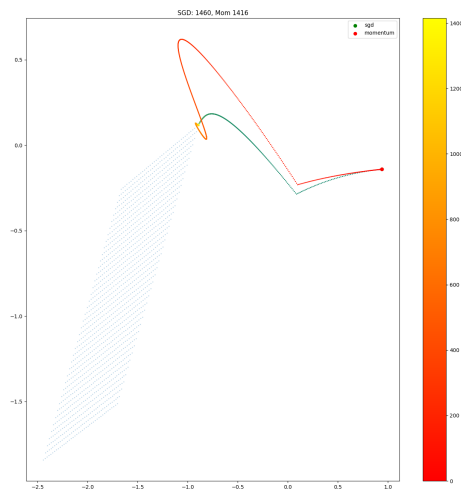
Figure 2: The optimisation dynamics of value iteration versus value iteration with momentum.

If we overparameterise the search space, then we can move between solutions in new ways. We can "tunnel" from A to B, without crossing C.

Q: What advantages does this provide?

Intuition: Every point is closer, under some measure of distance?!? But. Momentum seems like it might be a bad thing here?

Momentum in higher dimensions

Intuition. Something weird happens with momentum in overparameterised spaces.

It is necessary to consider the trajectory to study momentum. It depends on what has happened in the past. Can we construct a space of possible trajectories? What properties do trajectories have? They are connected by the update fn.

Continuous flow and its discretisation

A linear step of size, α, in parameter space, ie by gradient descent, is not necessrily a linear step in parameter space.
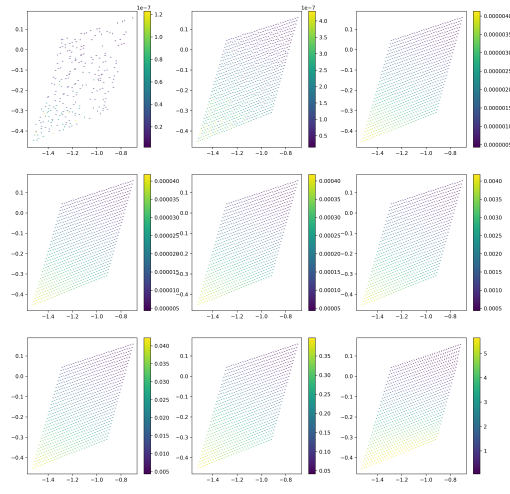
Figure 3: The gradient field dV as a function of different learning.