Gradient-based TD methods

Weiwei Zhang

260684686

Mountain Car

State Variables

Two dimensional continuous state space.

Velocity = (-0.07, 0.07)

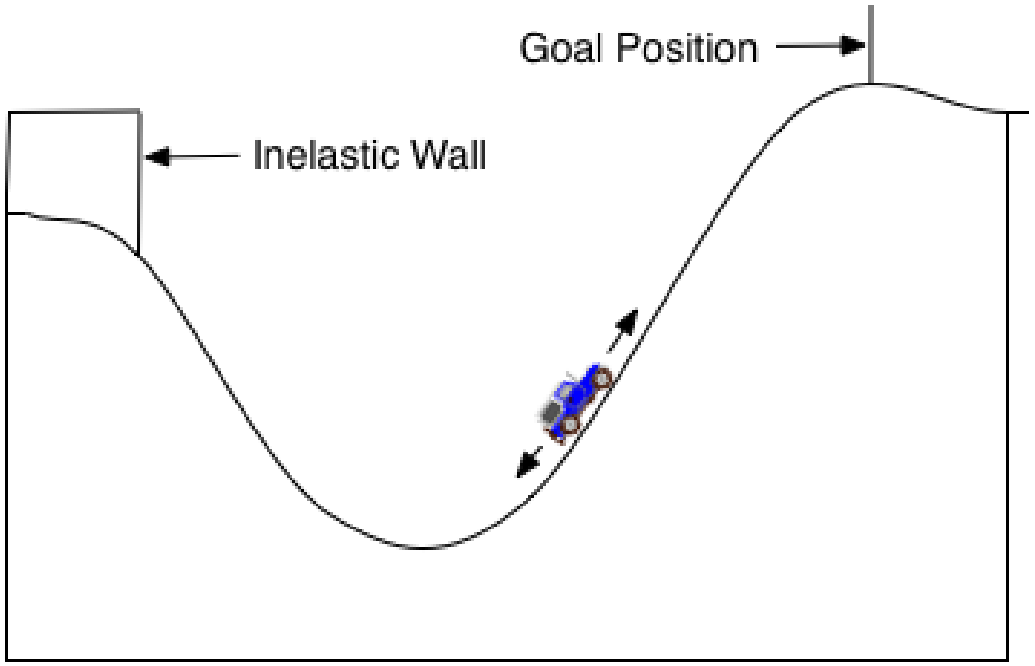
Position = (-1.2, 0.6)

Actions

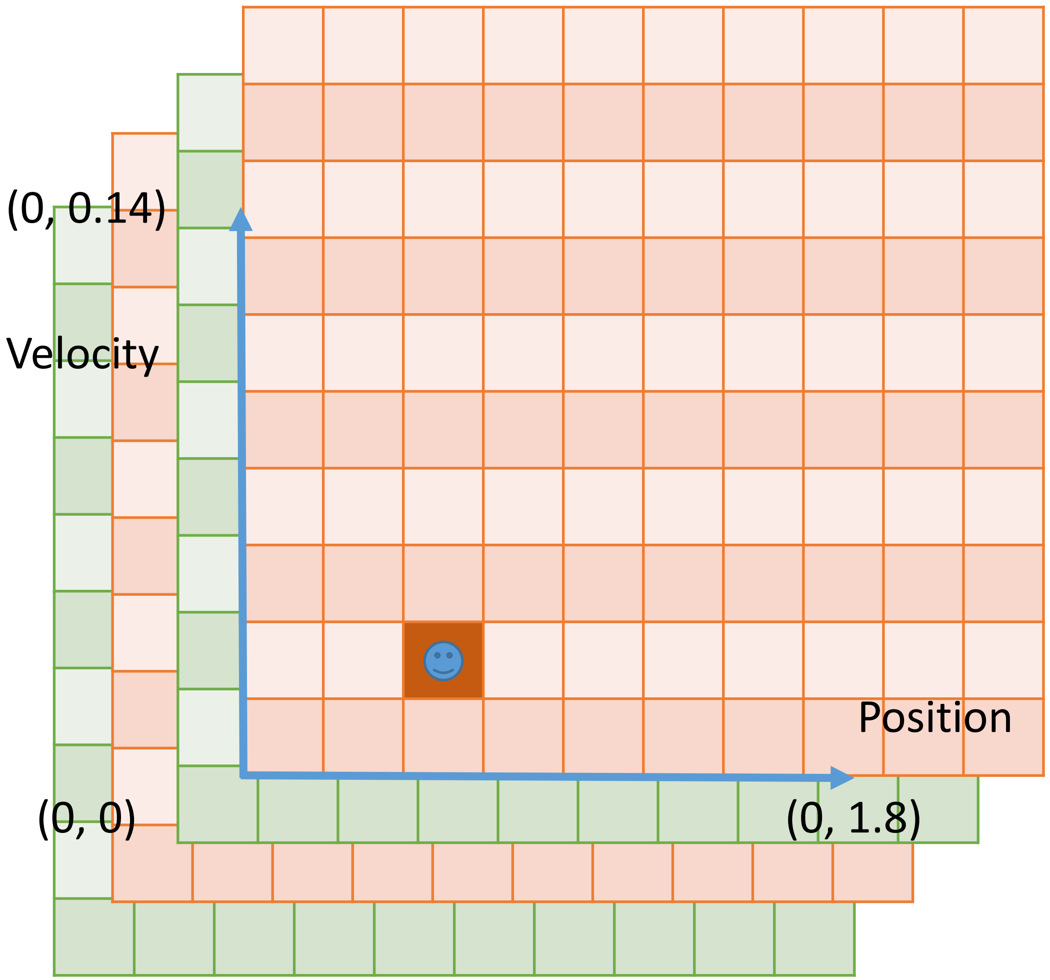
(reverse, coast, forward)

Reward

-1



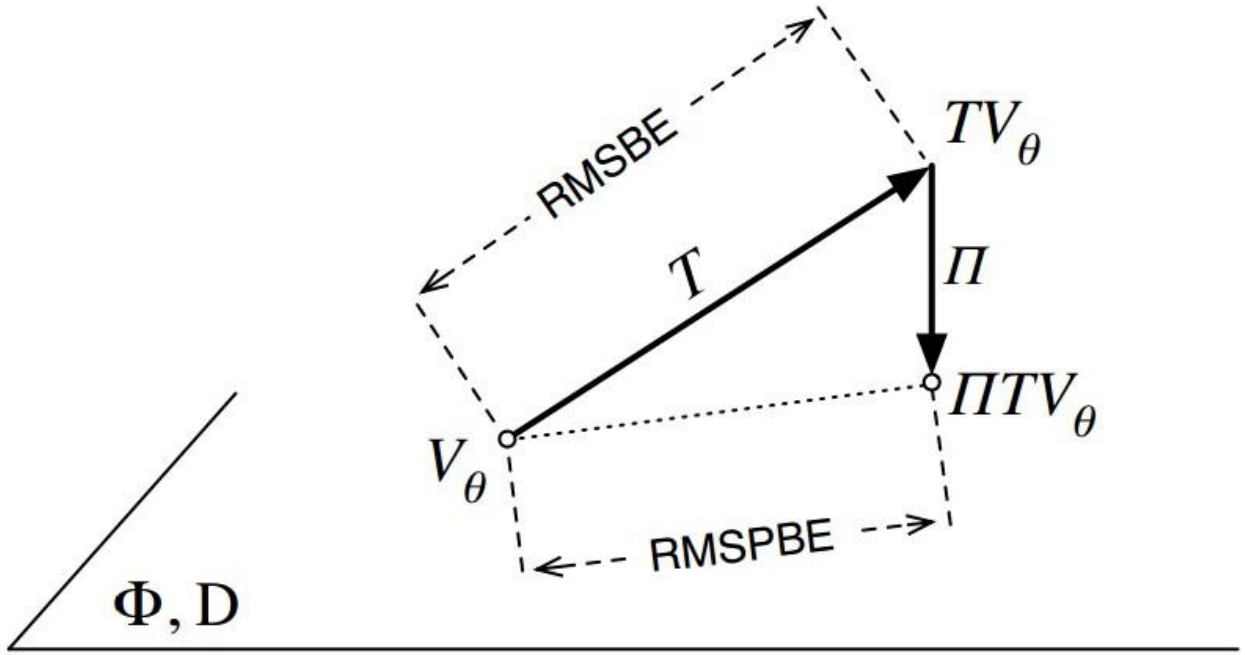
Tile Coding



[0, 0, 0, 0, 0, 0, **1**, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0,
0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0,
0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0,
0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, **1**, 0, 0, 0, 0, 0, 0, 0, 0, 0,
0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0,
0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0,
0, 0, 0, 0, 0, 0, 0, **1**, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0,
0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0,
0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0,
0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0,
0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, **1**, 0, 0, 0, 0, 0, 0, 0,
0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0,
0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0,
0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0,
0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0,
0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0,
0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0]

Mean Square Projected Bellman Error (MSPBE)

- *T* takes you outside the space
- *Π* projects you back into it



GTD(0)

$$\delta = r + \gamma\theta^\top\phi' - \theta^\top\phi$$
$$w_{k+1} = w_k + \beta_k(\delta_k\phi_k - w_k)$$
$$\theta_{k+1} = \theta_k + \alpha_k(\phi_k - \gamma\phi'_k)(\phi_k^\top w_k)$$

very slow compared to conventional linear TD

GTD-2(0)

$$\delta = r + \gamma \theta^\top \phi' - \theta^\top \phi$$
$$w_{k+1} = w_k + \beta_k (\delta_k - \phi_k^\top w_k) \phi_k$$
$$\theta_{k+1} = \theta_k + \alpha_k (\phi_k - \gamma \phi'_k)(\phi_k^\top w_k)$$

TDC(0)

$$\delta = r + \gamma \theta^\top \phi' - \theta^\top \phi$$

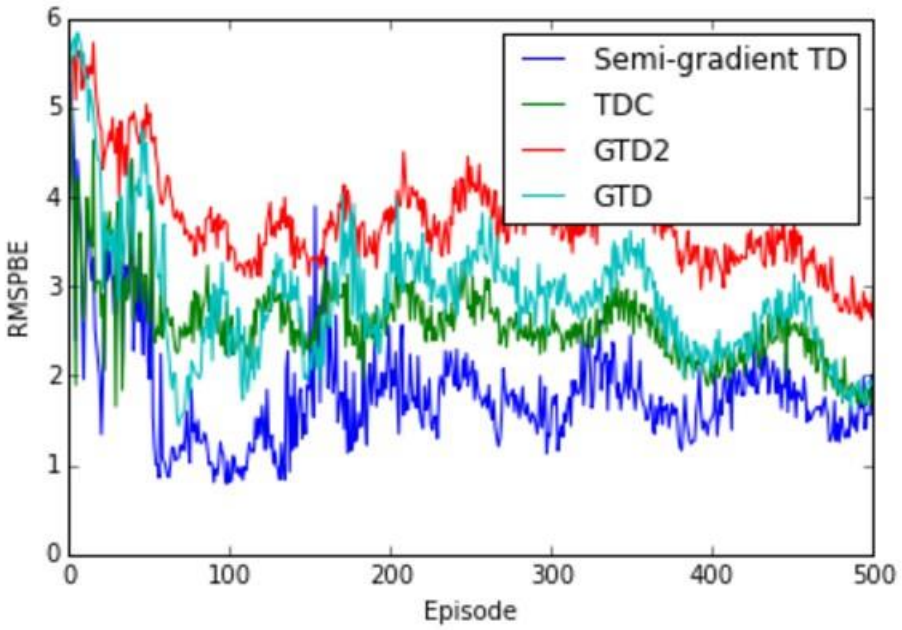
$$w_{k+1} = w_k + \beta_k (\delta_k - \phi_k^\top w_k) \phi_k$$

$$\theta_{k+1} = \theta_k + \alpha_k \delta_k \phi_k - \alpha \gamma \phi'_k (\phi_k^\top w_k)$$

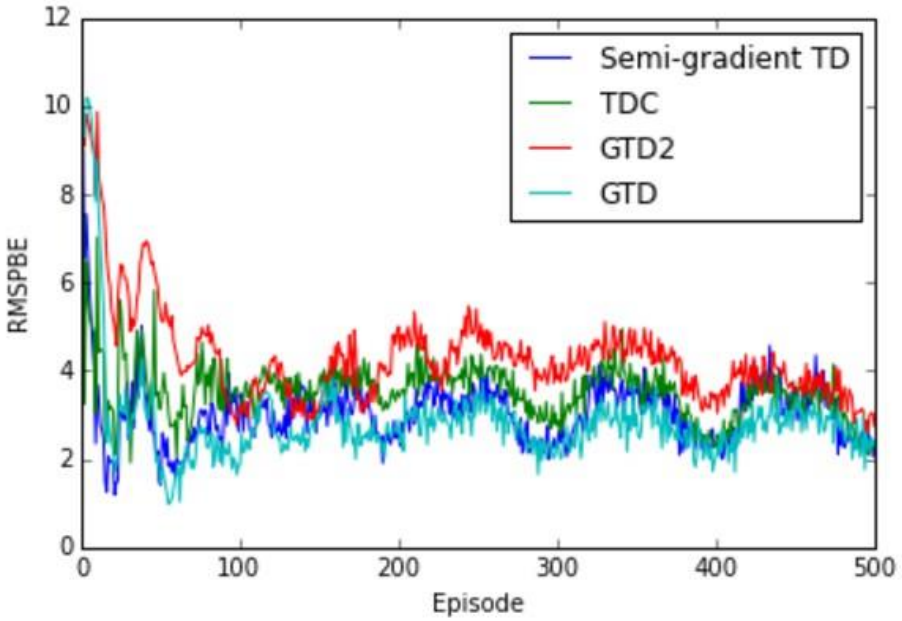
gradient correction

Results

4 tiles



8 tiles



Thanks