

The 3D space of all Value Functions over 3 states

$$BE = B^\pi \cdot V_w - V_w$$

$$PBE = \Pi_\phi \cdot B^\pi \cdot V_w - V_w$$

Path of Dynamic Programming Policy Evaluation

MC with linear function approximation gets here (eventually)

$$w_\pi = \arg \min_w d(V^\pi, V_w)$$

$$w_{TDE}$$

SGD with gradient as $\nabla_w (E_\pi [r + \gamma V_w(s) - V_w(s)]^2)$

Bellman Error Vector (BE)

$B^\pi \cdot V_w$

PBE

$\Pi_\phi \cdot B^\pi \cdot V_w$

V_w

Path of sequence of $\Pi_\phi \cdot B^\pi$ operations

$w_{PBE} = \arg \min_w d(\Pi_\phi \cdot B^\pi \cdot V_w, V_w)$

TD Fixed-Point

$$w_{BE} = \arg \min_w d(B^\pi \cdot V_w, V_w)$$

Bellman Error minimizing point

SGD with gradient as $\nabla_w (E_\pi [r + \gamma V_w(s) - V_w(s)]^2)$

The subspace of all Value Functions representable as V_w

