Function Approximation

Given a problem domain with continuous states $s \in \mathrm{S} = \mathrm{R}^n$, we wish to find a way to represent the value function $v_\pi(s)$ (for prediction) or $q_\pi(s, a)$ (for control).

We can do this by choosing a parameterized function that *approximates* the true value function:

$\hat{v}(s, \mathbf{w}) \approx v_\pi(s)$
$\hat{q}(s, a, \mathbf{w}) \approx q_\pi(s, a)$

Our goal then reduces to finding a set of parameters $\mathbf{w}$ that yield an optimal value function. We can use the general reinforcement learning framework, with a Monte-Carlo or Temporal-Difference approach, and modify the update mechanism according to the chosen function.

Feature Vectors

A common intermediate step is to compute a feature vector that is representative of the state: $\mathbf{x}(s)$