Outline

Introduction of Q-Learning

Tips of Q-Learning

Q-Learning for Continuous Actions

Continuous Actions

- Action a is a *continuous vector*

$$a = arg \max_{a} Q(s,a)$$

Solution 1

Sample a set of actions: $\{a_1, a_2, \cdots, a_N\}$
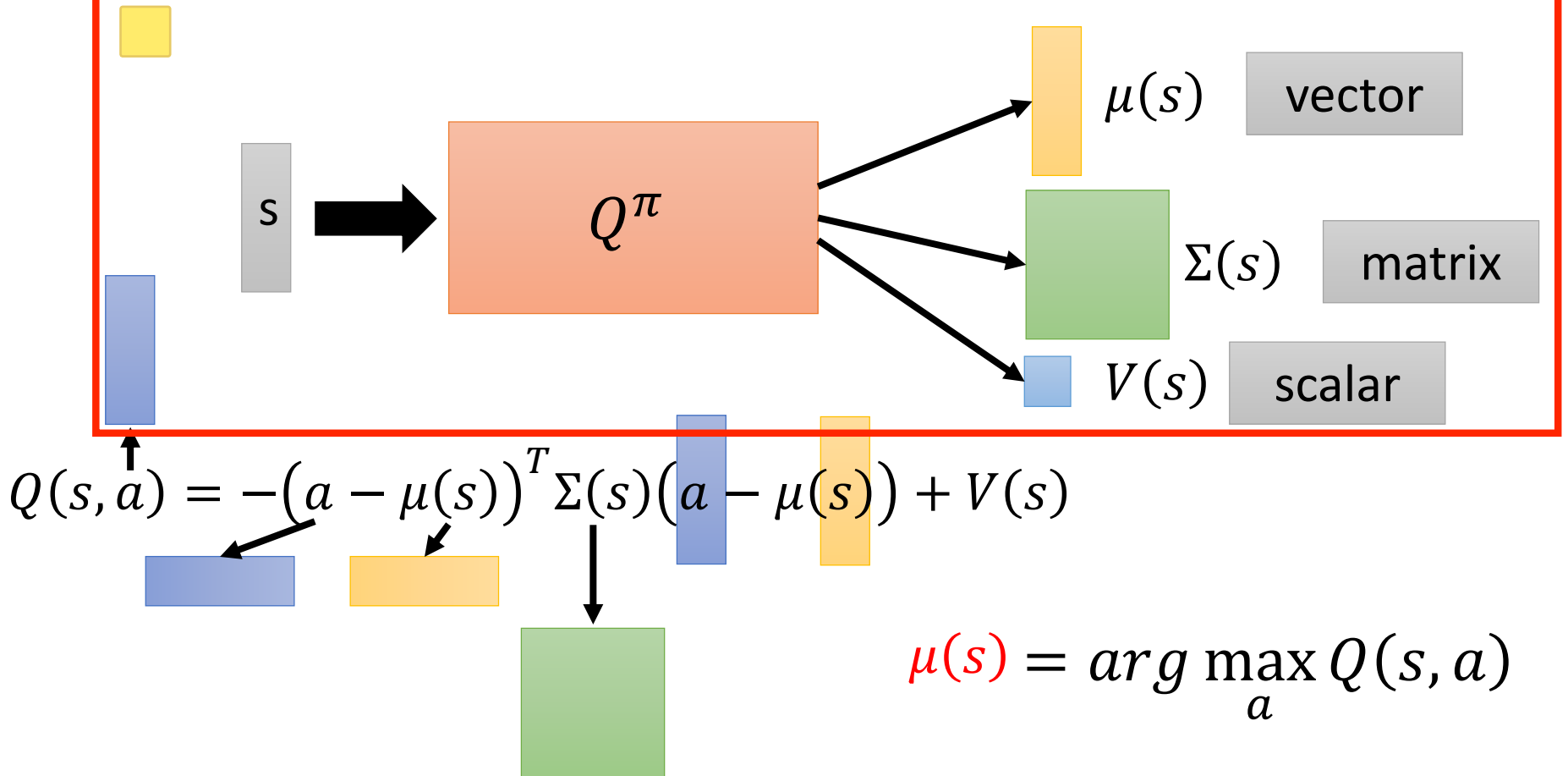
See which action can obtain the largest Q value

Solution 2

Using gradient ascent to solve the optimization problem.

Continuous Actions

s → Q^π → $\mu(s)$ **vector**

$\Sigma(s)$ **matrix**

$V(s)$ **scalar**

$$Q(s,a) = -\big(a - \mu(s)\big)^T \Sigma(s)\big(a - \mu(s)\big) + V(s)$$

$$\mu(s) = arg \max_a Q(s,a)$$

Continuous Actions

Solution 4   Don't use Q-learning

Policy-based   Value-based

Learning an Actor   Actor + Critic   Learning a Critic
(Next Lecture)