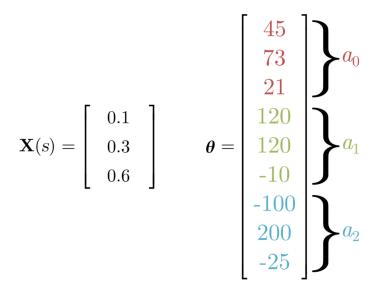
**10.** Consider the following state features and parameters  $\theta$  for three different actions (red, green, and blue):

1/1 point



Compute the action preferences for each of the three different actions using linear function approximation and stacked features for the action preferences.

What is the action preference of  $a_0$  (red)?

- 39
- O 35
- O 37
- O 33
- Correct
  Correct.

11. Which of the following statements are true about the Actor-Critic algorithm with softmax policies? (Choose all that apply)

1/1 point

- The actor and the critic share the same set of parameters.
- ☐ The preferences must be approximated using linear function approximation.
- ✓ The learning rate parameter of the actor and the critic can be different.
- ✓ Correct

Correct! In practice, it is preferable to have a slower learning rate for the actor so that the critic can accurately critique the policy.

Since the policy is written as a function of the current state, it is like having a different softmax distribution for each state.

Correct!

12. Which one is a reasonable parameterization for a Gaussian policy?

0 / 1 point

- $\bigcirc \;\; \mu$ : a linear function of parameters,  $\sigma$ : the exponential of a linear function of parameters.
- $\bigcirc$   $\mu$ : a linear function of parameters,  $\sigma$ : a linear function of parameters
- $\bullet$   $\mu$ : the exponential of a linear function of parameters,  $\sigma$ : a linear function of parameters.

⊗ Incorrect

Incorrect. Remember that the parameter sigma must be positive. A linear function does not guarantee that this constraint will be met.