ECE59500RL HW2

Robert (Cars) Chandler — chandl71@purdue.edu

Problem 1

1.1

Any matrix with only nonzero eigenvectors is necessarily invertible. So, we seek to prove that $I - \gamma P^\pi$ has only nonzero eigenvectors.

Using the properties of eigenvalues, if λ is an eigenvalue of P^π, then $1 - \gamma\lambda$ is an eigenvalue of $I - \gamma P^\pi$. So, if we can show that $\gamma\lambda \neq 1$ for all λ, then we can prove the matrix is invertible. Since we are in the discounted setting, we know that $0 < \gamma < 1$. So, if we can show that $\lambda \leq 1$, then we will have achieved our goal.

We know that P^π is a row-stochastic matrix, which is to say that each element must be in the interval $[0, 1]$ and the sum of the elements in each row must be 1. Given the definition of the eigenvalue and eigenvector:

$$\mathbf{A}\mathbf{x} = \lambda\mathbf{x}$$

We begin by analyzing the RHS of the equation.

If we consider some $\lambda > 1$ to exist, then by definition:

$$\|\lambda\mathbf{x}\|_\infty = |\lambda|\|\mathbf{x}\|_\infty$$

and since $\lambda > 1$, then

$$\|\lambda\mathbf{x}\|_\infty = |\lambda|\|\mathbf{x}\|_\infty > \|\mathbf{x}\|_\infty$$

Now, looking at the LHS,

$$\|P^\pi\mathbf{x}\|_\infty \leq \|\mathbf{x}\|_\infty$$

This is true because each element in the vector resulting from $P^\pi \mathbf{x}$ will be some linear combination of the components of $\mathbf{x}$ where the scalars are at least 0 and at most 1 and sum to 1. So, the greatest magnitude possible from this operation occurs in the case where a row of P^π has a 1 that aligns with the maximum magnitude of x and is zero everywhere else.

Now, we have proven that the LHS is less than or equal to $\|\mathbf{x}\|_\infty$ and that the RHS is greater than $\|\mathbf{x}\|_\infty$. But if these two sides of the equation are to be equal, then this is impossible. Therefore, it is impossible that $\lambda > 1$ for a row-stochastic matrix.

So if $\lambda < 1$ for all eigenvalues of P^π, then $1 - \gamma\lambda > 0$, because $0 < \gamma < 1$. Therefore, all eigenvalues of $I - \gamma P^\pi$ are nonzero, which means that $I - \gamma P^\pi$ must also be invertible.

1.2

Beginning with the Bellman consistency equation:

$$\begin{aligned} v^\pi(s) &= \mathbb{E}_{a\sim\pi(s)}\left[R(s,a) + \gamma \mathbb{E}_{s'\sim P(\cdot|s,a)}\left[v^\pi(s')\right]\right] \\ &= \sum_{a\in\mathcal{A}} \mathbb{P}(a|s)\left[R(s,a) + \gamma \sum_{s'\in\mathcal{S}}\left[\mathbb{P}(s'|a,s)v^\pi(s')\right]\right] \end{aligned}$$

Which we can write in matrix form. Let $\pi_{|\mathcal{S}|\times|\mathcal{A}|}$ be a matrix such that each element $\pi_{i,j} = \mathbb{P}(a_j|s_i)$ is the probability of applying action $a_j \in \mathcal{A}$ given the current state $s_i \in \mathcal{S}$.

For each $s \in \mathcal{S}$,

Problem 2

2.1

We wish to prove that

$$\left|\max_{x\in X} g_1(x) - \max_{x\in X} g_2(x)\right| \leq \max_{x\in X} |g_1(x) - g_2(x)|$$

It is true that

$$g_1(x) \leq |g_1(x) - g_2(x)| + g_2(x)$$
$$\left|\max_{x\in X} g_1(x) - max_{x\in X} g_2(x)\right| \leq \max_{x\in X} |g_1(x) - g_2(x)|$$