STATISTICAL MACHINE LEARNING (WS2016)
SEMINAR 7

Assignment 1. Prove that:

$$v_\pi(s) = \sum_{a \in \mathcal{A}} \pi(a|s) q_\pi(s,a), \tag{1}$$

$$q_\pi(s,a) = \mathrm{R}_s^a + \gamma \sum_{s' \in \mathcal{S}} \mathcal{P}_{ss'}^a v_\pi(s'). \tag{2}$$

Use the definition of MDP and the definitions of the value functions $v_\pi(s)$ and $q_\pi(s,a)$. Both equations are used in derivation of *Bellman expectation equations.*

Assignment 2. Consider a non-episodic (continuing) MDP and a policy π. How would the state-value function $v_\pi(s)$ change when all rewards get increased by a constant c?

Assignment 3. The ϵ-greedy policy selects a random action with a probability ϵ, otherwise it selects a maximum valued (greedy) action. Show that ϵ-greedy policy can be described by the following expression:

$$\pi(a|s) = \begin{cases} \frac{\epsilon}{|\mathcal{A}(s)|} & \text{for non-greedy actions,} \\ 1 - \epsilon + \frac{\epsilon}{|\mathcal{A}(s)|} & \text{for the greedy action,} \end{cases} \tag{3}$$

where $a \in \mathcal{A}(s)$ and $s \in \mathcal{S}$.

Assignment 4. The output of a regression tree is defined as:

$$h(\boldsymbol{x}) = \sum_{r=1}^{M} c_r \mathbb{I}\{\boldsymbol{x} \in R_r\}, \tag{4}$$

where R_r is an input space region defined by the r-th tree leaf and $c_r \in \mathbb{R}$ is the corresponding region's response. The tree is trained using set $\mathcal{T} = \{(\boldsymbol{x}_i, y_i) \mid i = 1, \ldots, m\}$. Show that the sum of squares loss function $\sum_{i=1}^{m}(y_i - h(\boldsymbol{x}_i))^2$ is minimized by choosing the following region responses:

$$c_r = \frac{1}{|S_r|} \sum_{(\boldsymbol{x}_i, y_i) \in S_r} y_i, \tag{5}$$

where $S_r = \{(\boldsymbol{x}_i, y_i) \mid (\boldsymbol{x}_i, y_i) \in \mathcal{T} \wedge \boldsymbol{x}_i \in R_r\}$.

Assignment 5. Prove that:

$$I(Y;X) = H(X) - H(X|Y) = H(Y) - H(Y|X) \tag{6}$$

where $I(Y;X)$ is the mutual information (information gain) defined as:

$$I(Y;X) = \sum_y \sum_x \mathbb{P}(X = x, Y = y) \log \left(\frac{\mathbb{P}(X = x, Y = y)}{\mathbb{P}(X = x)\mathbb{P}(Y = y)} \right).$$

Assignment 6. Bootstrapping is a method which produces M datasets $\mathcal{T}_i$ for $i = 1, \dots M$ by uniformly sampling the original dataset $\mathcal{T}$ with replacement. Bootstrap datasets have the same size as the original dataset $|\mathcal{T}_i| = |\mathcal{T}| = m$. Show that as $m \to \infty$ the fraction of unique samples in $\mathcal{T}_i$ approaches $1 - \frac{1}{e} \approx 63.2\%$.

Hint: apply exponential of a logarithm to a limit which emerges in a last step in order to solve it.