15780: Graduate AI (Spring 2017)

Practice Midterm Exam (Solutions)

February 23, 2017

Topic	Total Score	Score
Heuristic Search	25	
VC Dimension	25	
Integer Programming	25	
Convex Optimization	25	
Total	100	

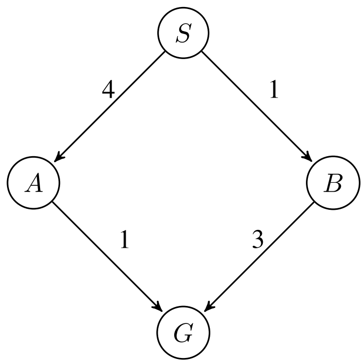
1 Heuristic Search [25 points]

Consider the problem of informed search with a heuristic. For each state x, let $h^*(x)$ be the length of the cheapest path from x to a goal.

Prove or disprove the following statements:

1.1 [15 points] If $h(x) = 2h^*(x)$ for all states x, then A^* tree search with the heuristic h is optimal.

Solution: This is false. We give a counterexample.



Note that $f(A) = 4 + 2(1) = 6$, $f(B) = 1 + 2(3) = 7$ so A^* will expand node A first. Then from node A, we have $f(G) = 5$, so we will expand node G and return the path S-A-G as the optimal path. However this is not the real optimal path. The real optimal path is S-B-G with a cost of 4.

1.2 [10 points] If h is a consistent heuristic, A^* graph search with the heuristic $h'(x) = h(x)/2$ is optimal.

Solution: This is true. Since h is consistent, this means for any node x and its successor x' we know that $h(x) \leq c(x, x') + h(x')$. This implies $h(x)/2 \leq c(x, x')/2 + h(x')/2$. Since costs are nonnegative, this also implies that $h(x)/2 \leq c(x, x') + h(x')/2$. Thus h' is also consistent and we know that A^* graph search with a consistent heuristic is optimal.

2 Learning Theory [25 points]

Determine the VC dimension of the following function classes.

2.1 [15 points] Define F to be the set of strings of length 3 composed of the symbols 0, 1, and $*$. Each $f \in F$ acts as a pattern matcher; i.e., when applied to a binary string s, it either accepts or rejects s. For example, when we apply the schema $f = 1 * *$ to the string $s = 101$, it accepts, and when we apply f to $s' = 010$, it rejects. What is the VC dimension of F?

Solution: The VC dimension is 3. The set $\{001, 010, 100\}$ can be shattered. For any set of size 4, note that if there are any two strings that differ at all three positions (call them s and s'), then the set $+\{s, s'\}$ can only be labeled with three wildcard characters, which also matches the rest of the strings not labeled +. Further, this means that there must be at least two pairs of strings at distance two from each other. In order to see this, put the strings on the vertices of a cube connected by edges between strings that differ from one another at exactly one position. Now, note that we can't realize this split. A pattern that matches one pair of strings must necessarily also match one string in the other pair. Concretely, this is because a pattern that matches strings s_1 and s_2 that differ in two positions must have two wildcards, and the third position, which is shared by s_1 and s_2, must differ between s_3 and s_4, meaning that one of s_3 and s_4 must match the pattern as well.

2.2 [10 points] The union of n intervals on the real line.

Solution: The VC dimension is $2n$. It's pretty clear that we can shatter $2n$ points, as this is equivalent to essentially using one interval for every consecutive pair of adjacent points. It's also not possible to shatter $2n + 1$ points because the assignment that alternates between +1 and -1 needs $n + 1$ intervals.

3 Integer Programming [25 points]

Consider an undirected graph $G = (V, E)$. A *minimum dominating set* is a smallest subset S of V such that every node not in S is adjacent to at least one node in S. A *minimum independent dominating set* is a smallest subset S of V such that (1) every node not in S is adjacent to at least one node in S and (2) no pair of nodes in S are adjacent. In your answer, you can use $N(i)$ to denote the set of neighbors of node i (i.e., $N(i)$ is a set of nodes adjacent to i) for each node $i \in V$. Note that $i \notin N(i)$. You also can use $(i, j) \in E$ to denote the edge between node $i \in V$ and node $j \in V$.

3.1 [15 points] Formulate an integer linear program to find a minimum dominating set.

Solution: Consider the following the integer program:

$$\begin{aligned} \text{minimize } & \sum_{i=1}^{n} x_i \\ \text{subject to } & \sum_{j \in N(i) \cup \{i\}} x_j \geq 1, \forall i \in V, \\ & \text{and } x_i = \{0, 1\}, \forall i \in V. \end{aligned}$$

If you solve this integer program, $S = \{i \in V : x_i = 1\}$ is a minimum dominating set.

3.2 [10 points] Formulate an integer linear program to find a minimum independent dominating set.

Solution: Consider the following the integer program:

$$\begin{aligned} \text{minimize } & \sum_{i=1}^{n} x_i \\ \text{subject to } & \sum_{j \in N(i) \cup \{i\}} x_j \geq 1, \forall i \in V, \\ & \text{and } x_i + x_j \leq 1, \forall (i, j) \in E, \\ & \text{and } x_i = \{0, 1\}, \forall i \in V. \end{aligned}$$

If you solve this integer program, $S = \{i \in V : x_i = 1\}$ is a minimum independent dominating set.

4 Convex Optimization [25 points]

Consider a linear program of the standard form: minimize $\mathbf{c}^T\mathbf{x}$ such that $\mathbf{A}\mathbf{x} \leq \mathbf{b}$. Here $\mathbf{x} \in \mathbb{R}^n$ is the vector of variables, and $\mathbf{c} \in \mathbb{R}^n$, $\mathbf{A} \in \mathbb{R}^{m\times n}$, and $\mathbf{b} \in \mathbb{R}^m$ are constants.

Prove from the definitions that this is a convex program.

Solution: First, we show that the objective function is linear, which we denote by f. Let $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$, and let $0 \leq \theta \leq 1$. We need to show $f(\theta\mathbf{x} + (1-\theta)\mathbf{y}) \leq \theta f(\mathbf{x}) + (1-\theta)f(\mathbf{y})$. We have

$$\begin{aligned} f(\theta\mathbf{x} + (1-\theta)\mathbf{y}) &= \mathbf{c}^T(\theta\mathbf{x} + (1-\theta)\mathbf{y}) = \mathbf{c}^T(\theta\mathbf{x}) + \mathbf{c}^T((1-\theta)\mathbf{y}) \\ &= \theta\mathbf{c}^T\mathbf{x} + (1-\theta)\mathbf{c}^T\mathbf{y} = \theta f(\mathbf{x}) + (1-\theta)f(\mathbf{y}). \end{aligned}$$

Thus, we conclude that the desired inequality holds (in fact, it holds with equality). Next, we show that the feasible region $\mathcal{F} = \{\mathbf{x} \in \mathbb{R}^n : \mathbf{A}\mathbf{x} \leq \mathbf{b}\}$ is convex. For this, let $\mathbf{x}, \mathbf{y} \in \mathcal{F}$ and let $0 \leq \theta \leq 1$. We need to show that $\theta\mathbf{x} + (1-\theta)\mathbf{y} \in \mathcal{F}$ as well, which amounts to showing $\mathbf{A}(\theta\mathbf{x} + (1-\theta)\mathbf{y}) \leq \mathbf{b}$. We have

$$\begin{aligned} \mathbf{A}(\theta\mathbf{x} + (1-\theta)\mathbf{y}) &= \mathbf{A}(\theta\mathbf{x}) + \mathbf{A}((1-\theta)\mathbf{y}) = \theta\mathbf{A}\mathbf{x} + (1-\theta)\mathbf{A}\mathbf{y} \\ &\leq \theta\mathbf{b} + (1-\theta)\mathbf{b} = (\theta + 1 - \theta)\mathbf{b} = \mathbf{b}. \end{aligned}$$

This completes the proof that $\mathcal{F}$ is convex, and hence the proof that a linear program is a convex program.