CS5120: Probability & Computing	Jan-May 2020

Lecture 17-18
The Method of Conditional Expectations

Lecturer: N.R.Aravind | *Scribe: N.R.Aravind*

In this class, we will see two more examples of the method of conditional expecation: finding a large cut in graphs, and a generalized tic-toe game.

1 Cuts in graphs

Recall from class one, the definition of a cut: For an undirected graph $G = (V, E)$, a cut is a partition $(A, V \setminus A)$ of the vertices, and the set of edges $\{\{x, y\} : x \in A, y \in V \setminus A\}$ is called a cut-set, and sometimes the edge-set is also referred to as the cut, as we shall do.

In the max-cut problem, we are given a graph $G = (V, E)$, and our goal is to find a cut with the largest number of edges possible. Unlike the problem of finding a minimum cut, the problem of finding a max cut is NP-hard. However, we shall now see that we can efficiently find a cut with at least $|E|/2$ edges.

Consider a random cut; that is: place each vertex independently in A with probability $1/2$ (and in $B = V \setminus A$ with probability $1/2$). Then the probability that an edge uv is part of the cut is equal to the probability that u, v are placed in different sets, that is $1/2$. Thus, the expected number of edges in the cut is equal to $m/2$, where $|E| = m$.

Our goal is to devise a deterministic algorithm that finds a cut of size at least $m/2$. To this end, we shall follow the method of conditional expectation; we order the vertices arbitarilly as $v_1, \ldots, v_n$, and place the vertices one by one into either A or B.

Suppose that vertices $v_1, v_2, \ldots, v_{i-1}$ have already been placed, and consider the vertex v_i. Let $E_A(i)$ be the expected number of edges in the cut (over a random placement of vertices $v_{i+1}, \ldots, v_n$) when v_i is placed in A, and let $E_B(i)$ be the expected number of edges in the cut (over a random placement of vertices $v_{i+1}, \ldots, v_n$) when v is placed in B.

We then compute and compare $E_A(i)$ and $E_B(i)$. If $E_A(i)$ is larger, v_i is placed in A, else it is placed in B. Thus, it now suffices to compute (or just compare) the two expectations.

Let m_0 denote the number of edges currently in the cut, d_A be the number of neighbors of v_i currently in A, and let d_B be the number of neighbors of v_i currently in B, and let d_C be the unplaced neighbors of v_i.

Then we have $E_A(i) = m_0 + d_B + \frac{d_C}{2}$, and $E_B(i) = m_0 + d_A + \frac{d_C}{2}$. We have $E_A(i) \geq E_B(i)$ if and only if $d_B \geq d_A$. Thus, if D_B is larger, we will place v_i in A and otherwise, we will place v_i in B.

The algorithm is thus:

- Set $A = \emptyset, B = \emptyset$.
- for $i = 1$ to n, do:
- Find the number of current neighbors of v_i in A, B respectively: call them d_A, d_B respectively.
- If $d_B > d_A$, set $A = A \cup \{v_i\}$, else set $B = B \cup \{v_i\}$.
- end for. Output A, B.

If we denote by $\mu(i)$, the expected number of edges in the cut after placing the first i vertices, then we have $\mu(0) = m/2$, and also $\mu(i) \geq \mu(i-1)$ for every i, so that $\mu(n) \geq m/2$.

The argument that $\mu(i) \geq \mu(i-1)$ is similar to that for 3-SAT: we have $\mu(i-1) = \frac{1}{2}\left(E_A(i) + E_B(i)\right)$, and $\mu(i) = max(E_A(i), E_B(i))$.

2 The Erdős-Selfridge game

We now see an example, which is one of the earliest applications of the method of conditional expectation.

Consider the game of tic-tac-toe. The game is played on a 3 by 3 grid, and two players, whom we shall call A and B, take turns to label an unlabeled cell with their respective symbol: we will use A,B, instead of the traditional crosses and circles. Further, we will assume that player A is the first player. The goal of the game for A is to have three cells in a row (or column, or diagonal) all labeled A. The goal for B is to stop A from achieving this (B cannot achieve the configuration before A, with perfect play from both players).

In the example below, the cells were marked in the order $(2,2),(1,2),(1,1),(3,3),(2,1)$ at which point, irrespective of B's move, A will subsequently win the game.

	1	2	3
1	A	B	
2	A	A	
3			B

However, player B does have a strategy to stop A from winning, in tic-tac-toe. If the 3 grid were replaced by a larger grid, and the winning condition for A is to have a full row/column/diagonal filled with As, then we'd expect that B will still draw. The intuitive reason for this is that the number of winning configurations is small, namely $(2n+2)$, compared to the number of cells in such a configuration (n).

In 1973, Erdős and Selfridge generalized this game as follows. There's a universe U of elements, and a collection $\mathcal{F}$ of subsets $S_1,\ldots,S_m$, such that $|S_i| = n$. Players take turns to pick an element from U and give it their label (A or B). Player A wins if there is a set S_i all of whose elements are labeled A.

In the above example, the universe is the set of cells, and the subsets are the rows, columns and diagonals. The result of Erdős and Selfridge is the following.

Theorem 1 *If $m = |\mathcal{F}| < 2^{n-1}$, then player B has a strategy to draw the game.*

We now prove the above result by giving a strategy for B to draw the game, using the method of conditional expectation.

Proof of Theorem 1: For a set $S \in \mathcal{F}$, let $X_S = 1$ if all its elements are labeled as A, and $X_S = 0$ otherwise. Let $X = \sum_{S\in\mathcal{F}} X_S$, and let $t = \lceil U/2 \rceil$ be the number of elements that will be labeled by A.

If we initially labeled each element randomly as A or B, then we'll have:

$Pr[X_S = 1] = \dfrac{1}{2^n}$ for every set S, and $E[X] = \dfrac{|F|}{2^n} < \dfrac{1}{2}$.

For $i = 1, 2, \ldots, t$, we denote by $\mu_A(i)$, the expected value of X after A has labeled i elements (the expectation being over a random assignment of unpicked elements). Similarly, let $\mu_B(i)$ be the expected value of X after B has labeled i elements.

The aim of B will be to ensure that $\mu_A(i) < 1$ after A has played i moves, for every i. In particular, when the game ends with all the elements having been labeled, the value of $\mu_A(t)$ is less than one, that is, zero.

Let's calculate $\mu_A(1)$, that is $E[X]$ after the first move of A. Let u be the element labeled by A, and let S be a set containing u. Then $Pr[X_S = 1] = \dfrac{1}{2^{n-1}}$. For a set S that doesn't contain u, we have: $Pr[X_S = 1] = \dfrac{1}{2^n} < \dfrac{1}{2^{n-1}}$. Thus, $\mu_A(1) = \sum_{S \in \mathcal{F}} E[X_S] \leq \dfrac{|F|}{2^{n-1}} < 1$.

Now, it is B's move. We note that $\mu_A(i) > \mu_B(i) < \mu_A(i+1)$, because after B picks an element v, the value of $Pr[X_S]$ becomes zero for sets S containing v, and is unchanged for other sets. After that, when A picks an element w, the value of $Pr[X_S]$ doubles for sets containing w, and is unchanged for other sets. The statements are true irrespective of whether S already has an element labeled B or not.

Let $\mu(v) = \sum_{S: v \in S} E[X_S]$. By the argument in the preceding paragraph, if B labels v, then $E[X]$ reduces by $\mu(v)$, and if A labels v, then $E[X]$ increases by $\mu(v)$.

Thus, B will pick v such that $\mu(v)$ is maximum and label it as B. Note that if A subsequently picks w, then we have $\mu(w) \leq \mu(v)$. Thus $\mu_B(i) = \mu_A(i) - \mu(v)$ and $\mu_A(i+1) = \mu_B(i) + \mu(w)$, which gives: $\mu_A(i+1) = \mu_A(i) - \mu(v) + \mu(w) \leq \mu(A)$.

Thus, B ensures that $\mu_A(i)$ never increases with i, and since $\mu_A(1) < 1$, we have $\mu_A(t) < 1$, as desired. ■

Remark: The proof also works for cases where the sets do not necessarily have to be of the same size. The assumption needed is: $\sum_{S \in \mathcal{F}} \dfrac{1}{2^{|S|}} < \dfrac{1}{2}$.