1 Derandomization by Conditional Expectation

In the last lecture, we saw randomized algorithms for MAX CUT and MAX 3-SAT. In each of these algorithms, there is an underlying random process involving some random variables $X_0, X_1, \ldots, X_{n-1}$, and we have computed the expectation of some objective value $Y$, which is a function on those random variables.

We show that under some very general conditions, the randomized algorithm can be derandomized, i.e., there is some deterministic algorithm that finds values $X_0 := x_0, X_1 := x_1, \ldots, X_{n-1} := x_{n-1}$ such that the objective value $Y$ is at least its expectation.

1.1 Principle of Conditional Expectation

The success of the derandomization method depends on the following conditions.

**Sufficient Conditions**

1. The objective value $Y$ is a function of the random variables $X_0, X_1, X_2, \ldots, X_{n-1}$, i.e., if the values of the random variables are known, then the value of $Y$ is uniquely determined.

2. Given a partial assignment $X_{[i]} := x_{[i]}$ for $1 \leq i \leq n$, the conditional expectation $E[Y|X_{[i]} = x_{[i]}]$ can be computed efficiently.

We show that when such conditions are met, it is possible to find values $x_{[n]} := (x_0, x_1, \ldots, x_{n-1})$ for the random variables $X_{[n]} := (X_0, X_1, \ldots, X_{n-1})$ such that the objective value $Y$ is at least its expectation $\mu := E[Y]$.

1. **Initialization.** We begin when none of the random variables has been assigned any values, i.e. $i := 0$. We have the invariant: $E[Y|X_{[i]} = x_{[i]}] \geq \mu$. The left hand side is simply $E[Y]$.

2. **Assigning Value to One More Random Variable $X_i$.** Suppose for some $0 \leq i < n$, we already have the assignment $X_{[i]} := x_{[i]}$ and the invariant $E[Y|X_{[i]} = x_{[i]}] \geq \mu$. We show the following claim.

   **Claim 1.1** There exists some assignment $X_i := x_i$ such that $E[Y|X_{[i+1]} = x_{[i+1]}] \geq \mu$.

   **Proof:** Conditioning on the value of $X_i$, we have

---

1Recall we denote $[i] := \{0, 1, 2, \ldots, i-1\}$, and $[0] := \emptyset$. 

---
\[ E[Y|X[i] = x[i]] = \sum_x Pr(X_i = x|X[i] = x[i]) E[Y|X[i] = x[i] \land X_i = x], \]

where the summation is over the values \( x \) that \( X_i \) can take. Observe that by the invariant, that the left hand side is at least \( \mu \). Hence, it follows that there exists some \( x \) such that \( E[Y|X[i] = x[i] \land X_i = x[i]] \geq \mu \). \( \blacksquare \)

For each \( x \), we test if \( E[Y|X[i] = x[i] \land X_i = x[i]] \geq \mu \) and find such an \( x \).

We assign \( X_i := x_i \).

3. This continues until \( i = n \), when all random variables have received their values. In this case, we have \( E[Y|X[n] = x[n]] \geq \mu \), which means that under those values, the objective value \( Y \) is at least \( \mu \).

### 1.2 Derandomization for MAX CUT

Given a graph \( G = (V, E) \), where \( V := [n] \) is labeled by \( n \) integers. The randomized algorithm essentially assigns independently for each vertex \( i \), a random variable \( X_i \) taking values uniformly in \( \{0, 1\} \) (each with probability \( \frac{1}{2} \)). The cut can be defined by \( C := \{i: X_i = 1\} \).

The number \( Y \) of edges in the cut is a function of the random variables \( X_0, X_1, \ldots, X_{n-1} \). For each edge \( e = \{i, j\} \in E \), define \( Y_e \) to be 1 if \( X_i \neq X_j \) and 0 if \( X_i = X_j \). Then, \( Y := \sum_{e \in E} Y_e \).

It suffices to show that given a partial assignment \( X[i] := x[i] \), the conditional expectation \( E[Y|X[i] = x[i]] \) can be computed efficiently. By linearity of expectation, it is enough to consider, for each edge \( e = \{u, v\} \in E \), the quantity \( E[Y_e|X[i] = x[i]] \). There are 3 cases to consider.

1. If none of \( X_u \) or \( X_v \) is assigned a value yet, the conditional expectation of \( Y_e \) is \( \frac{1}{2} \), as before.
2. If exactly one of \( X_u \) and \( X_v \) is assigned a value, check that the conditional expectation \( Y_e \) is also \( \frac{1}{2} \).
3. If both of \( X_u \) and \( X_v \) already have been assigned values, then \( Y_e \) is 1 if they receive different values and 0 otherwise.

The running time of the derandomization algorithm is \( O(mn) \), where \( m \) is the number of edges.

### 1.3 Derandomization of MAX 3-SAT

The argument is similar. The important part is given a partial assignment of variables, what is the (conditional) probability that a clause is satisfied? There are some cases to consider:

1. If the partial assignment makes the clause satisfied, then it is 1;
2. if there are 3 unassigned variables in the clause, then it is \( \frac{7}{8} \);
3. if there are 2 unassigned variables in the clause, then it is \( \frac{3}{4} \);
4. if there is 1 unassigned variable in the clause, then it is \( \frac{1}{2} \);
5. if there is no more unassigned variables in the clause, then it is 0.

One can check that for \( m \) clauses in \( n \) variables, the derandomized procedure takes time \( O(mn) \).

## 2 Graphs with No Short Cycles: Method of Alteration

When we use the probabilistic method, after we run the experiment, sometimes we have to make minor alteration to the outcome in order to obtain a desirable solution. We demonstrate this method by considering the number of edges in a graph with no short cycles.

**Definition 2.1** An undirected graph \( G = (V,E) \) contains a cycle of length \( l \) if there are \( l \) vertices \( v_1, v_2, \ldots, v_l \) such that all \( l \) edges \( \{v_1, v_2\}, \{v_2, v_3\}, \ldots, \{v_{l-1}, v_l\}, \{v_l, v_1\} \in E \) are present. The minimum length of a cycle is 3; note that there is no cycle of length 2.

**Question.** Suppose a graph has no cycles of length \( l \) or less. What is the maximum number of edges that it can have?

Observe that we are trying to optimize two conflicting objectives: adding more edges means eventually creating short cycles. In the extreme case, in a complete graph, every 3 points form a 3-cycle.

**Theorem 2.2** There exists an \( n \)-vertex graph with no cycles of length \( l \) or less that has at least \( \Omega(n^{1 + \frac{1}{l-1}}) \) edges.

We proof the special case for \( l = 3 \). The general case will appear in a homework question.

**Definition 2.3 (Random Graph \( G_{n,p} \))** Consider the following experiment. Let \( V \) be a set of \( n \) vertices. We form a random graph \( (V,E) \) in the following way. For each unordered pair \( \{u,v\} \in \binom{V}{2} \), independently add an edge between \( u \) and \( v \) with probability \( p \), i.e., \( \Pr(\{u,v\} \in E) = p \). The resulting graph is known as \( G_{n,p} := (V,E) \).

Note that for any graph \( G \) with index set \( V \), \( \Pr(G_{n,p} = G) > 0 \). Our candidate graph could be generated by the process. If we want more edges, then \( p \) should be large; if we do not want short cycles, then \( p \) should be small. We will find the best value of \( p \) to balance between the two requirements.

### 2.1 Without alteration: How lucky can we be?

Our best hope is to prove that with non-zero probability, both of the following events \( A \) and \( B \) happen. Event \( A \) is the event that \( G_{n,p} \) has a large number \( X \) of edges. Event \( B \) is the event that there are no triangles in \( G_{n,p} \).

Consider event \( B \) first. Group the \( n \) vertices into \( \frac{n}{3} \) groups, each of size 3. (For the time being, assume \( n \) is divisible by 3.) Look at one such group. The probability that there is no triangle between 3 vertices is \( (1 - p^3) \). The probability that this holds for all \( \frac{n}{3} \) groups is \( (1 - p^3)^{n/3} \). Hence, \( \Pr(B) \leq (1 - p^3)^{n/3} \), which is quite small (exponentially small with respect to \( n \)).

Note that the events \( A \) and \( B \) are not independent. If we want to use the union bound \( \Pr(\overline{A} \cup \overline{B}) \leq \Pr(\overline{A}) + \Pr(\overline{B}) \) to give an upper bound on the failure probability, we would need to prove something
like the failure probability \( Pr(\overline{A}) \) is less than \((1 - p^3)^{n/3}\). Since this is small, it is difficult to show that the number of edges \( X \) is large using this approach.

Consider another approach. Although with high probability, there would exist a triangle, the number \( Y \) of triangles is not too big. Here is an idea. We run the experiment and form \( G_{n,p} \). Let \( X \) be the number of edges and \( Y \) be the number of triangles. We pick one edge from each triangle and remove it. In the worst case, we remove \( Y \) edges from the graph. After this alteration, the graph would have no triangles, and the number of edges in the remaining graph is at least \( Z := X - Y \). By choosing the probability \( p \) carefully, we show that with non-zero probability, \( Z \) is large.

First observe the following quantities.

**Claim 2.4** We have the following.

1. \( E[X] = \binom{n}{2}p \)
2. \( \text{var}[X] = \binom{n}{2}p(1 - p) \)
3. \( E[Y] = \binom{n}{3}p^3 \).

**Proof:** The first two results follow from the fact that \( X \) has a binomial distribution with \( n \) objects and probability \( p \). The last result follows from the fact that there are \( \binom{n}{3} \) ways to form a cycle in a graph, and the probability that each of them is formed is \( p^3 \).

We can proceed in two ways: (1) using probability or (2) using expectation.

### 2.2 By Probability

Suppose we can show the following:

1. For some \( \alpha > 0 \), \( Pr(X < E[X] - \alpha) < \frac{1}{2} \).
2. For some \( \beta > 0 \), \( Pr(Y > \beta) < \frac{1}{2} \).

Using union bound, we know that with non-zero probability neither events happen, and in that case, we have \( Z := X - Y \geq E[X] - \alpha - \beta \).

We first choose \( p \) such that \( E[X] \geq 4E[Y] \). Observe that we can choose \( p := \sqrt{\frac{3}{8n}} \). Check that \( E[X] = \Theta(n^{1.5}) \).

For the first event, observe by Chebyshev’s inequality, \( Pr(X < E[X] - \alpha) \leq Pr(|X - E[X]| > \alpha) < \frac{\text{var}[X]}{\alpha^2} \leq \frac{(\binom{n}{2})p}{\alpha^2} \). The last quantity is at most \( \frac{1}{2} \), if we set \( \alpha := \sqrt{2 \binom{n}{2}p} = \Theta(n^{0.75}) \).

For the second event, by Markov’s inequality, if we set \( \beta := 2E[Y] \), then \( Pr[Y > \beta] < \frac{1}{2} \).

Hence, it follows that with non-zero probability, we have

\[
Z := X - Y \geq E[X] - \alpha - 2E[Y] \geq E[X] - \alpha - \frac{1}{2}E[X] = \frac{E[X]}{2} - \sqrt{2 \binom{n}{2}p} \geq \Omega(n^{1.5}).
\]
2.3 By Expectation

We choose $p$ such that $E[X] \geq 2E[Y]$. We can set $p := \frac{\sqrt{3}}{4n}$. Check that $E[X] = \Theta(n^{1.5})$.

Then, it follows that

$$E[Z] = E[X] - E[Y] \geq E[X] - \frac{E[X]}{2} = \frac{E[X]}{2} = \Omega(n^{1.5}).$$

**Remark 2.5** Note that this is not the best result for triangle-free graphs. Consider a complete bipartite graph with $\frac{n}{2}$ vertices on each side. Then, the graph has no triangles and has $\Omega(n^2)$ edges. However, we obtain a weaker result using the probabilistic method to illustrate how a similar result could be proved for general $l$.

2.4 General Case

The general case would appear as a homework problem. If you would like a head start to work on the next homework, here is a preview.

1. **Graphs with No Short-Cycles.** In this question, we show the following result. For each $l \geq 3$, and $n \geq 2^{l-1}$, there exists a graph, with $n$ vertices and no cycles of length $l$ or less, that has $\Omega(n^{1+\frac{1}{l-1}})$ edges.

   (a) Consider the random graph $G_{n,p}$, where $p \geq \frac{2}{n}$. For $3 \leq i \leq l$, let $Y_i$ be the number of length-$i$ cycles in $G_{n,p}$. Compute $E[Y_i]$.

   (b) Let $Y := \sum_{3 \leq i \leq 1} Y_i$. Show that $E[Y] \leq (np)^l$.

   (c) By choosing an appropriate value of $p$, prove that there exists an $n$-vertex graph, with no cycles of length $l$ or less, that has $\Omega(n^{1+\frac{1}{l-1}})$ edges.

   (d) Derandomize the above procedure, i.e., give a deterministic algorithm that returns a graph with the desired properties. Analyze the running time of your algorithm.