Counting Perfect Matchings

1 Sampling Matchings from a Dense Bipartite Graph

Let G be a bipartite graph. Define $\mathcal{M}_k$ as the set of all matchings of G with exactly k edges. We want to compute $|\mathcal{M}_n|$.

In dense graphs, each $|\mathcal{M}_i|/|\mathcal{M}_{i-1}|$ is bounded between n^{-2} and n^2. Last time we showed we can approximate this quantity by sampling from $\mathcal{M}_n \cup \mathcal{M}_{n-1}$ with random walks. Suppose we are at a matching M. We choose a uniformly random edge $e = (u, v) \in G$.

- If M is perfect and $e \in M$, remove e from M.
- If M is not perfect and both u and v are unmatched, add e to M.
- If M is not perfect and u is matched to some w and v is unmatched, replace (u, w) with (u, v) – symmetrically if v is matched and u is unmatched.
- Otherwise, do nothing.

The Markov graph H is connected, undirected, and m-regular, so the stationary distribution is uniform (as desired). We will show it mixes rapidly using canonical paths. We will route one unit of flow from each $M \in \mathcal{M}_n \cup \mathcal{M}_{n-1}$ to each $M' \in \mathcal{M}_n \cup \mathcal{M}_{n-1}$ while minimizing the maximum number of paths using any edge (L, L').

Lemma 1. *In H, each matching $M \in \mathcal{M}_{n-1}$ is within at most two hops of some $M' \in \mathcal{M}_n$.*

Proof. Let u, v be the unmatched vertices in M. If $(u, v) \in E(G)$, then by adding (u, v) we reach M' in one hop. Otherwise, we will show that we can perform one rotation and one addition to reach M'. Let V' be the neighbors of u in G and U' be the neighbors of v in G. By density, $|U'|, |V'| \geq {}^{n}\!/_{2}$.

Because all of U' is matched in M and non are matched to v, if there were no edge between U' and V' then U' would be matched to V'', which has size $\leq {}^{n}\!/_{2} - 1$. So there is an edge $(u', v') \in M$ with $u' \in U'$ and $v' \in V'$. The two-hop path is to rotate (u', v') to (u, v') and add (u', v). $\square$

It is not difficult to show using the lemma that for each $M \in \mathcal{M}_n$, at most n^2 total $M' \in \mathcal{M}_{n-1}$ are within two hops. All such M' must take at most two hops to an $M \in \mathcal{M}_n$, so we can use the path from M to an $\hat{M} \in \mathcal{M}_n$ near the destination, then at most two hops to $M' \in \mathcal{M}_{n-1}$. The takeaway is that at a blowup of n^4 in congestion, we can focus only on canonical paths between perfect matchings, ignoring near-perfect matchings.

We will define a canonical path in H from $M \in \mathcal{M}_n$ to $M' \in \mathcal{M}_n$. Define $\Gamma = M \oplus M'$ to be the symmetric difference. Γ is a union of disjoint even length cycles, alternating edges from M and M'. For each cycle C_i, let v_i be the node with smallest index. Sort the cycles by increasing v_i.

For each cycle in this order,

1. Remove M edge incident on v_i.

2. Rotate edge to an edge M' that leaves v_i still unmatched.

3. After all rotations, add M' edge incident on v_i.

What is the maximum congestion on any edge (L, L') using these paths? We want to show it is at most $N \cdot poly(n)$ where $N = |\mathcal{M}_n \cup \mathcal{M}_{n-1}|$.

Claim 2. If we focus only on $M, M' \in \mathcal{M}_n$, then the congestion of (L, L') is at most N. Then, the congestion overall is at most $\mathcal{O}(N \cdot n^4)$.

The key proof idea for the claim is to show that source/sink pairs (M, M') using (L, L') can be "encoded" in one matching $R \in \mathcal{M}_n \cup \mathcal{M}_{n-1}$ so that M, M' can be uniquely recovered from L, L', and R. Thus, at most N pairs use (L, L').

Suppose in the path from M to M', the edge (L, L') is used as part of fixing the cycle C_i. Then every C_j for $j < i$ has already been fixed, so we need to "store" the edges of M for those C_j. And, every C_j for $j > i$ has not been fixed yet, so we need to store the edges of M' for those C_j. The rough idea that accomplishes both is to define $R = \Gamma \oplus (L \cup L')$. From L, L', and R, we can reconstruct all cycles and figure out where in the "unwinding sequence" we are.

However, a small problem is that rotations can cause R to not be a matching. Instead we define $R = (\Gamma \oplus (L \cup L') \setminus \{e\}$ for rotations involving $e \in G$. Thus, $R \in \mathcal{M}_{n-1}$.

From L and L' we can compute $L \cup L'$. From $L \cup L'$ and R we can compute $(L \cup L') \oplus R = \Gamma \setminus \{e\}$. From this, we can infer e as the unique edge completing a cycle. (This is for rotations only; for addition and deletion, there is no edge e removed). From Γ, we can compute the ordering of cycles $C_1, \ldots, C_k$. We can reconstruct from L, L' wich cycle C_i is being rotated. From this, we can recover all of M, M' by computing which cycles have been fixed. Thus the mapping from (M, M') is one-to-one.

So from L, L', and R, we can uniquely reconstruct an (M, M') using (L, L'). So the number of pairs (M, M') using (L, L') is at most the number of R. This is at most N, so at most N pairs use (L, L'). This shows our random walk mixes rapidly, and in summary, gives an approximation algorithm to count the perfect matchings in a dense bipartite graph.

2 Lovasz Local Lemma

The Lovasz Local Lemma is a tool for the probabilistic method (showing existence of an object via nonzero probability that it is generated by a random procedure). In the typical setting, we have a large number of "bad" events $\mathcal{E}_i$ and we would like none of them to happen. In extreme 1, the $\mathcal{E}_i$ are independent, so

$$\Pr[\bigcap \bar{\mathcal{E}}_i] = \prod (1 - \Pr[\mathcal{E}_i]) > 0$$

so the "good" thing happens with positive probability. In extreme 2, we know nothing about the dependence of the $\mathcal{E}_i$. In this case the union bound gives

$$\Pr[\bigcap \bar{\mathcal{E}}_i] \geq 1 - \sum_i \Pr[\mathcal{E}_i].$$

Unless all $\Pr[\mathcal{E}_i]$ are very small, this bound can be negative and thus trivial. The Lovasz Local Lemma is a tool for proving that $\Pr[\bigcap \bar{\mathcal{E}}_i] > 0$ when the dependence among the $\mathcal{E}_i$ is limited.

We say that $\mathcal{E}_j$ is mutually independent of a set T of events iff for all disjoint $S, S' \subseteq T$ (but not necessarily a partition) we have

$$\Pr[\mathcal{E}_j \mid \bigcap_{i \in S} \mathcal{E}_i \cap \bigcap_{i \in S'} \bar{\mathcal{E}}_i] = \Pr[\mathcal{E}_j].$$

Define a **dependency graph** G on these $(\mathcal{E}_i)$ as follows: Let

$$T_j := \{\mathcal{E}_i \mid (\mathcal{E}_j, \mathcal{E}_i) \notin G\}.$$

Then $\mathcal{E}_j$ is mutually independent of T_j. In other words, non-edges guarantee independence and edges capture possible dependence. Notice that the complete graph is a dependency graph which gives trivial bounds; we want *sparse* dependency graphs.

Theorem 3. *[Lovasz Local Lemma].* Let G be a dependency graph on $(\mathcal{E}_i)$. Assume that there exists x_i with $i \in [n]$ with

$$\Pr[\mathcal{E}_i] \leq x_i \prod_{(i,j)\in E} (1 - x_j)$$

for all i. Then

$$\Pr[\bigcap_i \bar{\mathcal{E}}_i] \geq \prod (1 - x_i).$$

Conceptually, the x_i are upper bounds on the probability of $\mathcal{E}_i$ subject to some conditioning on limited independence.

Corollary 4. Let $(\mathcal{E}_i)$ be events with $\Pr[\mathcal{E}_i] \leq p$ for all i, and each $\mathcal{E}_i$ be mutually independent of all except at most d other events. If

$$p \leq \frac{1}{e(d+1)}$$

then $\Pr[\bigcap_i \bar{\mathcal{E}}_i] > 0$. In contrast, if we were using the union bound, we would need $p \leq 1/n$. So we get a factor d instead of a factor n.