1 Matchings in graphs

This week we will be talking about finding matchings in graphs: a set of edges that do not share endpoints.

Definition 1.1 (Maximum Matching)**.** *Given an undirected graph $G = (V, E)$, find a subset of edges $M \subseteq E$ of maximum size such that every pair of edges $e, e' \in M$ do not share endpoints $e \cap e' = \emptyset$.*

Definition 1.2 (Perfect Matching)**.** *Given an undirected graph $G = (V, E)$ where $|V| = n$ is even, find a subset of edges $M \subseteq E$ of size $n/2$ such that every pair of edges $e, e' \in M$ do not share endpoints $e \cap e' = \emptyset$. That is, every node must be covered by the matching M.*

Obviously, any algorithm for Maximum Matching gives an algorithm for Perfect Matching. First, we will show that these problems are roughly equivalent. In particular we will prove the following claim:

Claim. *If one can solve Perfect Matching in $T(n)$ time, then one can solve Maximum Matching in time $\tilde{O}(T(2n))$.*

Proof. We will binary search for the maximum k for which there is a matching M with $|M| \geq k$. For our current value of k, to check whether such M exists, we can add a clique on $n - 2k$ nodes to the graph G and connect it to the original graph with all possible edges. (An independent set instead of a clique also works.) We will now show that the new graph H has perfect matching if and only if G has matching with k edges.

Suppose G has a matching of size k. Then we can create a perfect matching in H by matching all of the unmatched vertices in G to vertices in the clique.

Suppose H has a matching a perfect matching. At most $n - 2k$ vertices in the clique are matched to a vertex in G. Thus, at least $2k$ vertices in G are matched to each other. This gives a matching of size k in G.

The size of H is at most $2n$ so performing a binary search for k on H yields running time $\tilde{O}(T(2n))$. □

We will focus on Perfect Matching and give algebraic algorithms for it. Because of the above reduction, this will also imply algorithms for Maximum Matching.

First we will give an $\tilde{O}(n^\omega)$ time algorithm that detects whether a graph has a pefect matching. Then we will give an $\tilde{O}(n^{\omega+1})$ time algorithm for finding a perfect matching if one exists. In the next lecture, we will improve this to $\tilde{O}(n^\omega)$.

The idea of these algorithms will be to define some matrix such that the determinant of this matrix is non-zero if and only if the graph has a perfect matching.

1.1 The Tutte Matrix

The Tutte matrix is a *symbolic* matrix i.e. each entry is a variable.

Definition 1.3. *For a graph $G = (V, E)$ with $|V| = n$, the following $n \times n$ matrix T is the Tutte matrix of G:*

$$T[i,j] = \begin{cases} 0 & \text{if } i = j \text{ or if } (i,j) \notin E \\ x_{i,j} & \text{if } (i,j) \in E \text{ and } i < j \\ -x_{i,j} & \text{if } (i,j) \in E \text{ and } i > j \end{cases}$$

Example 1. Consider the path on 2 edges where the vertices are labeled 1, 2, and 3 in order along the path. The Tutte matrix of this graph is:

$$\begin{bmatrix} 0 & x_{1,2} & 0 \\ -x_{1,2} & 0 & x_{2,3} \\ 0 & -x_{2,3} & 0 \end{bmatrix}$$

The Tutte matrix is a *skew symmetric* matrix i.e. $T = -T^t$, that is, it is symmetric about the diagonal except the entries below the diagonal are negated.

The following theorem is at the core of all the algorithms for Perfect Matching that we will discuss.

Theorem 1.1 (Tutte)**.** *For any graph $G = (V, E)$, the determinant of the Tutte matrix T is non-zero if and only if G contains a perfect matching.*

$$\det(T) \neq 0 \iff G \textit{contains a perfect matching.}$$

Note that the determinant of a symbolic matrix is a polynomial and we say that a polynomial is zero if it is identically the zero-polynomial, and non-zero otherwise. That is, it is possible for a polynomial to be non-zero but still have a setting of the variables that makes it evaluate to zero.

Before proving the theorem we will recall the definition of the determinant of a matrix.

Definition 1.4. *The* determinant $\det(T)$ *of a matrix T is defined as*

$$\det(T) = \sum_{\sigma \in S_n} (-1)^{sign(\sigma)} \cdot \prod_{i=1}^{n} T(i, \sigma(i)) \tag{1}$$

where S_n is the set of permutations of $[n]$ and $sign(\sigma)$ is the parity of inversions for σ, i.e. the number of pairs $x < y$ for which $\sigma(x) > \sigma(y)$. Another interpretation of $sign(\sigma)$ is the parity of the number of even cycles in a product representing σ.

Example 2. Consider the permutation $\sigma = (1, 2)(5, 7, 3)(4, 6)$. It has two even cycles ($(1, 2)$ and $(4, 6)$) so $sign(\sigma) = 0$.

Proof of Theorem 1.1.

Claim. *If G has a perfect matching then* $\det(T) \neq 0$.

Proof. Say $\{(i_1, i_2), (i_3, i_4), \ldots, (i_{n-1}, i_n)\}$ is a perfect matching. Suppose the vertices are ordered so that $i_1 < i_2 < \cdots < i_n$. Consider the permutation $\sigma_M = (i_1, i_2)(i_3, i_4) \ldots (i_{n-1}, i_n)$. Now consider the term in the determinant that corresponds to σ_M. We will show that this term cannot cancel out, which means that the determinant is non-zero.

$$\begin{aligned} \prod_{i=1}^{n} T(i, \sigma_M(i)) &= T(x_{i_1,i_2}) \cdot T(x_{i_2,i_1}) \cdot T(x_{i_3,i_4}) \cdot T(x_{i_4,i_3}) \cdot \cdots \cdot T(x_{i_{n-1},i_n}) \cdot T(x_{i_n,i_{n-1}}) \\ &= x_{i_1,i_2}(-x_{i_1,i_2})x_{i_3,i_4}(-x_{i_3,i_4}) \ldots (x_{i_{n-1},i_n})(-x_{i_{n-1},i_n}) \\ &= (-1)^{n/2}(x_{i_1,i_2})^2 \ldots (x_{i_{n-1},i_n})^2 \end{aligned}$$

If you think about it, you can convince yourself that no other permutation yields a term that is exactly the product of the square of every variable, so this term can't cancel out. It follows that $det(T) \neq 0$. □

Claim. *If $det(T) \neq 0$ then G has a perfect matching.*

Proof. First we will show that all terms of $\det(T)$ corresponding to permutations with at least one odd cycle cancel out.

Let P be the set of permutations in S_n that contain at least 1 odd cycle. For each $\sigma \in P$, let C_σ be the odd cycle in σ with *minimum* element, and let σ' be σ with C_σ reversed.

For example, if $\sigma = (1,5)(2,3,4)(6,7,8)$, then $\sigma' = (1,5)(4,3,2)(6,7,8)$.

Note that this choice of pairing of permutations (σ, σ') is a perfect bijection of the permutations with odd cycles.

Now we will show that the term in the determinant corresponding to σ cancels with the term corresponding to σ'. Since σ and σ' have the same number of even cycles, $sign(\sigma) = sign(\sigma')$. Consider the sum of the terms corresponding to σ and σ':

$$\prod_{i=1}^{n} T(i, \sigma(i)) + \prod_{i=1}^{n} T(i, \sigma'(i))$$

Since σ and σ' only differ in C_σ, this is equal to

$$\prod_{i \notin C_\sigma} T(i, \sigma(i)) \cdot \Big(\prod_{i \in C_\sigma} T(i, \sigma(i) + \prod_{i \in C_\sigma} T(i, \sigma'(i)))\Big). \tag{2}$$

Say that $C_\sigma = (i_1 i_2 \dots i_t)$ for some odd t. Then $\prod_{i \in C_\sigma} T(i, \sigma(i) = \prod_{j=1}^{t} T(i_j, i_{j+1})$ Since σ' reverses C_σ, $\prod_{i \in C_\sigma} T(i, \sigma'(i) = \prod_{j=1}^{t} T(i_{j+1}, i_j) = (-1)^t \prod_{j=1}^{t} T(i_j, i_{j+1})$ since T is skew-symmetric.

Thus

$$\Big(\prod_{i \in C_\sigma} T(i, \sigma(i) + \prod_{i \in C_\sigma} T(i, \sigma'(i)))\Big) = 0.$$

So the above expression (2) is the zero-polynomial. Thus, we have shown that all terms of $\det(T)$ corresponding to permutations with at least one odd cycle cancel out.

Now we are left with permutations with only even cycles. Since $\det(T) \neq 0$, there must be at least one permutation σ with only even cycles whose corresponding term in $\det(T)$ is non-zero. Consider any cycle $C = (i_1, \dots, i_{2k})$ in σ. The term in $\det(T)$ that corresponds to C is $x_{i_1,i_2} \cdot x_{i_2,i_3} \cdot \dots \cdot x_{i_{2k},i_1}$. Since this term is non-zero, the edges $(i_1, i_2), (i_2, i_3), \dots, (i_{2k}, i_1)$ are all in G. Taking every other edge $(i_1, i_2), (i_3, i_4), \dots, (i_{2k-1}, i_{2k})$ forms a matching. Taking the union of these matchings over all cycles in σ yields a perfect matching.

□

□

Now we will show how to use Theorem 1.1 to detect whether G has a perfect matching. We would like to determine whether $\det(T)$ is non-zero, however because $\det(T)$ has n^2 variables and degree n, it can be computationally expensive to do this. Instead, we will use a useful tool, the Schwartz-Zippel lemma, which allows us to evaluate $\det(T)$ on a set of random values to determine, with high probability, whether $\det(T)$ is non-zero.

Lemma 1.1 (Schwartz-Zippel). *Let P be a non-zero polynomial over $\{x_1, \dots, x_N\}$ of degree d over a field $\mathbb{F}$. If we pick values $v_1, \dots, v_N$ randomly from a finite set $S \subseteq \mathbb{F}$ and let $P(\{v_i\})$ be the value obtained by setting $x_1 = v_1, \dots, x_N = v_N$ in P, then $P(\{v_i\}) \neq 0$ with probability at least $1 - \frac{d}{|S|}$.*

For $\det(T)$ we have $\deg(\det(T)) = n$ and therefore it is enough to pick $|S| = n^2$ to get probability $1 - 1/n$. We can repeat the algorithm to further boost the probability.

However, if we work over $\mathbb{Z}$ the entries of this determinant could be very large and we only get a running time of $O(n^{\omega+1})$. Instead, pick a prime $p = \tilde{\Theta}(n^2)$ and work over $\mathbb{Z}_p$, letting $S = \mathbb{Z}_p$. It is ok to work over $\mathbb{Z}_p$ for the following reason. If G has a perfect matching M then the term of the polynomial $\det(T) \mod p$ corresponding to the permutation representing M is non-zero and therefore $\det(T) \mod p$ is a non-zero polynomial.

To conclude, the algorithm is as follows. Construct the Tutte matrix T of G. Pick values v_{ij} for each x_{ij} uniformly at random from $\{1, \dots, p\}$ and let $T(\{v_{ij}\})$ be the matrix obtained from T by these substitutions. Then compute $\det(T(\{v_{ij}\}))$ over $\mathbb{F}_p$. Since each entry of $T(\{v_{ij}\})$ has $O(\log n)$ bits and we can compute determinants using $O(n^\omega)$ operations, computing $\det(T(\{v_{ij}\}))$ takes time $\tilde{O}(n^\omega)$. If $\det(T(\{v_{ij}\})) \neq 0$ we return that G has a perfect matching and if over all repetitions of this algorithm $\det(T(\{v_{ij}\}))$ is always 0, then return that G does not have a perfect matching.

2 Finding the matching

The above algorithm tells us in $\tilde{O}(n^\omega)$ time whether the graph contains a perfect matching. In the rest of this lecture (and the next one) we will discuss algorithms that can find a perfect matching.

There is a simple $\tilde{O}(n^{\omega+2})$ solution: for every edge $(x, y) \in E$, remove it and its endpoints from the graph and check if there is still a perfect matching in $\tilde{O}(n^\omega)$ time. If the graph does not contain a perfect matching, put the vertices x and y back in the graph, together with all their incident edges except for (x, y) and move on to the next edge. Otherwise include the edge (x, y) in the matching and remove the vertices x and y from the graph.

Today we will see an $\tilde{O}(n^{\omega+1})$ algorithm and next week we'll see an $\tilde{O}(n^\omega)$ one.

2.1 The Rabin-Vazirani Algorithm

We will prove this thorem.

Theorem 2.1 (Rabin-Vazirani)**.** *A perfect matching can be found in* $\tilde{O}(n^{\omega+1})$ *time.*

For any $n \times n$ matrix A and subsets $X, Y \subseteq [n]$, let $A[X, Y]$ denote A restricted to the rows indexed by X and columns indexed by Y. Let $A_{X,Y}$ denote the matrix obtained from T by removing the rows indexed by X and columns indexed by Y.

We will use a different definition of the determinant than above (this is probably the definition you saw when you first saw determinants):

$$\det(T) = \sum_{j=1}^{n} (-1)^{1+j} \cdot T[1, j] \cdot \det(T_{\{1\}\{j\}}),$$

Suppose G contains a perfect matching. Then $\det(T) \neq 0$ (by Theorem 1.1), so there exists $j \in [n]$ such that $T[1, j]$ and $\det(T_{\{1\}\{j\}})$ are both non-zero. $T[1, j] \neq 0$ means that $(1, j) \in E$. For $\det(T_{\{1\}\{j\}})$, we will prove the following claim (we will prove it at the end).

Claim 1. *If* $\det(T_{\{1\}\{j\}}) \neq 0$ *then* $\det(T_{\{1,j\}\{1,j\}}) \neq 0$.

Note that $T_{\{1,j\}\{1,j\}}$ is the Tutte matrix for the graph $G \setminus \{1, j\}$. Thus, by Claim 1, the graph $G \setminus \{1, j\}$ has a perfect matching. Combining this with the fact that $(1, j) \in E$, this means that G has a perfect matching that contains the edge $(1, j)$.

Since $\det(T) \neq 0$, T is invertible. It will be useful to study T^{-1}. Recall the adjoin formula:

$$T^{-1}(i, j) = (-1)^{i+j} \cdot \frac{\det(T_{\{i\},\{j\}})}{\det(T)}.$$

Therefore, $T^{-1}(1, j) \neq 0$ if and only if $\det(T_{\{1\},\{j\}}) \neq 0$. This suggests the following algorithm, which iteratively peels off edges from the perfect matching:

Algorithm 1: Perf-matching(G)

$T \leftarrow T(\{v_{ij}\})$: a random substitution of the Tutte matrix modulo a large enough prime;
if $\det(T) = 0$ **then**
  return no perfect matching;
Set $M = \emptyset$;
while $|M| < n/2$ **do**
  Compute $N = T^{-1}$;
  Find j such that $N(1,j) \neq 0$ and $(1,j) \in E$;
  $M \leftarrow M \cup \{(1,j)\}$;
  $T \leftarrow T_{\{1,j\},\{1,j\}}$ i.e. remove rows 1 and j and columns 1 and j from T;
return M;

We claim that the running time of this algorithm is $O(n^{\omega+1})$. This is because computing T^{-1} takes $\tilde{O}(n^{\omega})$ time, finding j such that $N(1,j) \neq 0$ and $(1,j) \in E$ takes time $\tilde{O}(n)$, and we do each of these $O(n)$ times.

Finally, we will prove Claim 1. The following properties of the Tutte matrix will be useful.

Proposition 1. *Let A be an $n \times n$ skew symmetric matrix, then:*

1. *A^{-1} is skew symmetric.*
2. *If n is odd, then $\det(A) = 0$.*
3. *(Frobenius) Let $Y \subseteq [n]$ s.t. $|Y| = rank(A)$ and the column rank of $A[[n], Y]$ is $rank(A)$, then $\det(A[Y,Y]) \neq 0$.*

Property 1 is straightforward. Property 2 is straightforward for the matrices that we care about since a graph with an odd number of vertices cannot have a perfect matching, but it is also true in general. We will use property 3 without proof.

Proof of Claim 1. Assume without loss of generality that $j = 2$. By property 2 we know that $\det(T_{\{1\},\{1\}}) = 0$, so the rank of $T_{\{1\},\{1\}}$ is at most $n-2$. By our assumption, $\det(T_{\{1\},\{2\}}) \neq 0$ so $\det(T_{\{1\},\{2\}})$ has rank $n-1$. Therefore the column rank of $T_{\{1\},\{1,2\}}$ is $n-2$ and the rank of $T_{\{1\},\{1\}}$ is exactly $n-2$. $T_{\{1\},\{1\}}$ is skew symmetric, so it follows from the Frobenius property that for $Y = \{3,\ldots,n\}$, $\det(T_{\{1\},\{1\}}[Y,Y]) \neq 0$. By definition, $T_{\{1\},\{1\}}[Y,Y] = T_{\{1,2\},\{1,2\}}$. This completes the proof. □