CSS.201.1 Algorithms

Instructor: Umang Bhaskar

TIFR 2024, Aug-Dec

Scribe: Soham Chatterjee

sohamchatterjee999@gmail.com
Website: sohamch08.github.io

CONTENTS

Matching Algorithms

1.1 Bipartite Matching

1.1.1 Using Max Flow

1.1.2 Using Augmenting Paths

1.1.3 Using Matrix Scaling

Here we will show a new algorithm for bipartite matching using matrix scaling. The paper which we will follow is [Linial-Samorodnitsky-Wigderson, STOC'1998]

Suppose $G = (L \cup R, E)$ a bipartite graph. If bipartite adjacency matrix of the graph G is A then the permanent of the matrix A,

$$per(A) = \sum_{\sigma \in S_n} \prod_{i=1}^{n} x_{i,\sigma(i)}$$

counts the number of perfect matchings in G. So we want to check if for a given bipartite graph $(L \cup R, E)$, $per(A) > 0$ or not where A is the bipartite adjacency matrix. Now there is a necessary and sufficient condition for existence of perfect matching in a bipartite graph which is called Hall's condition.

> **Theorem 1.1.1** Hall's Condition
>
> A bipartite graph $G = (L \cup R, E)$ has an L-perfect matching if and only if $\forall\, S \subseteq L$, $|S| \leq |N(S)|$ where $N(S) = \{v \in R \colon \exists\, u \in L,\ (u,v) \in E\}$

Proof: Now if G has a L-perfect matching then for every $S \subseteq L$, S is matched with some $T \subseteq R$ such that $|S| = |T|$. Therefore $T \subseteq N(S) \implies |S| = |T| \leq |N(S)|$.

Now we will prove the opposite direction. Suppose for all $S \subseteq L$ we have $|S| \leq |N(S)|$. Assume there is no L-perfect matching in G. Let M be a maximum L-matching in G. Let $u \in L$ is unmatched. Now consider the following sets:

$$X = \{x \in L \colon \exists\, M\text{-alternating path from } u \text{ to } x\}, \qquad Y = \{y \in R \colon \exists\, M\text{-alternating path from } u \text{ to } y\}$$

Now notice that $N(X) \subseteq Y$. Since in a M-alternating path from u whenever the odd edges are not matching edges and the even edges are matching edges. So in the odd edges we can pick any neighbor except the one it is matched with and the immediate even edge before that connects that vertex with the vertex in R it is matched with. Hence we have $N(X) \subseteq Y$.

Now it suffices to prove that $|X| > |Y|$. Now let $y \in Y$. Suppose $u \rightsquigarrow x' \to y$ be the M-alternating path. If y is not matched then we could increase the matching by taking the odd edges of the path and thus obtain a matching with larger size than M. But M is maximum matching. Hence y is matched. Therefore we can extend the path by taking the matching edge incident on y and go the the vertex $x'' \in L$ i.e. the new M-alternating path becomes $u \rightsquigarrow x' \to y \to x''$ to have a M-alternating path $u \rightsquigarrow x''$. So $|X| > |Y|$.

Therefore we obtained a set of vertices $X \subseteq Y$ such that $|X| > |Y| \geq N(X)|$. This contradicts the assumption. Hence contradiction. Therefore G has a L-perfect matching. ∎

We will use hall's condition on the adjacency matrix to check if $per(A)$ is positive or not. Now multiplying a row or a column of a matrix by some constant c also multiplies the permanent of the matrix by c as well. In fact if $d_1, d_2 \in \mathbb{R}_+^n$ and $D_1 = diag(d_1)$ and $D_2 = diag(D_2)$ then $per(D_1 A D_2) = \left(\prod_{i=1}^{n} d_{1_i} \right) \left(\prod_{i=1}^{n} d_{2_i} \right) per(A)$. So we can scale our original matrix A to obtain a different matrix B and from B we can approximate $per(A)$ by approximating $per(B)$. A natural strategy is to seek an efficient algorithm for scaling A to a doubly stochastic B.

> ### Definition 1.1.1: Doubly Stochastic
>
> A matrix $M \in \mathbb{R}^{m \times m}$ is doubly stochastic if entries are non-negative and each row and column sum to 1.

First we will show that Hall's Condition holds for doubly stochastic matrix. First let's see what it means for a matrix to satisfy hall's condition. A matrix with all entries non-negative holds Hall's Condition if for all $S \subseteq [n]$ if $T = \{ i \in [n] : \exists j \in S, A(i, j) \neq 0 \}$ then $|T| \geq |S|$. This also corresponds to the bipartite adjacency matrix satisfying the hall's condition since for any set of rows S the number of columns for which in the S rows at least one entry is non zero should be greater than or equal to $|S|$.

> ### Lemma 1.1.2
>
> Hall's Condition holds for doubly stochastic matrix.

Proof: Let M be the doubly stochastic matrix. Let $S \subseteq [n]$. So consider the $|S| \times n$ matrix which only consists of the rows in S. Call this matrix M_S^r. Now suppose T be the set of columns in M_S^r which has nonzero entries. Now consider the $n \times |T|$ matrix which only consists of the columns in T. Call this matrix M_T^c. Now since M is doubly stochastic we know sum of entries of M_S^r is $|S|$ and sum of entries of M_T^c is $|T|$. Our goal is to show $|S| \leq |T|$. Now since T is the only set of columns which have nonzero columns in M_S^r the elements which contributes to the sum of entries in M_S^r are in the T columns in M_S^r. Since these elements are also present in M_T^c we have $|T| \geq |S|$. ∎

Hence for doubly stochastic matrices the permanent is positive. Now not all matrices are doubly stochastic. And in fact matrices with permanent zero will not be doubly stochastic so no amount of scaling will make it doubly stochastic. So we will settle for approximately doubly stochastic matrix. In order to make a matrix doubly stochastic first for each row we will divide the row with their row some. Now it becomes row stochastic. Then if its not approximately doubly stochastic for each column we will divide the column entries with their column sum. But first what ϵ-approximate doubly stochastic matrix means.

> ### Definition 1.1.2: $\epsilon-$Approximate Doubly Stochastic Matrix
>
> A matrix is $\epsilon-$approximate doubly stochastic if for each column, the column sum is in $(1 - \epsilon, 1 + \epsilon)$ and for each row, the row sum is in $(1 - \epsilon, 1 + \epsilon)$

Now we will show that even for ϵ-approximate doubly stochastic matrix the hall's condition holds.

> ### Lemma 1.1.3
>
> Halls's Condition holds for ϵ-approximate doubly stochastic matrix for $\epsilon < \frac{1}{10n}$

Proof: Let M is $\epsilon-$approximate doubly stochastic matrix. Let $S \subseteq [n]$. So consider the $|S| \times n$ matrix which only consists of the rows in S. Call this matrix M_S^r. Now suppose T be the set of columns in M_S^r which has nonzero entries. Now consider the $n \times |T|$ matrix which only consists of the columns in T. Call this matrix M_T^c. Now the sum of entries in M_S^r is $\geq |S|(1 - \epsilon)$ and sum of entries in M_T^c is $\leq |T|(1 - \epsilon)$. Now since T is the only set of columns which have nonzero columns in M_S^r the elements which contributes to the sum of entries in M_S^r are in the T columns in M_S^r. Since these elements are also present in M_T^c we have $||T|(1 + \epsilon) \geq |S|(1 - \epsilon)$. Therefore we have

$$|T| \geq |S| \frac{1 - \epsilon}{1 + \epsilon} = |S| \left(1 - \frac{2\epsilon}{1 + \epsilon} \right) \geq |S|(1 - 2\epsilon) > |S| \left(1 - \frac{1}{5n} \right) \geq |S| \left(1 - \frac{1}{|S|} \right) > |S| - 1$$

Since T is an integer we have $|T| \geq |S|$. Hence the Hall's condition holds. ∎

Therefore permanent of ϵ-approximate doubly stochastic matrix is also positive. Hence our algorithm for bipartite perfect matching is:

Algorithm 1: BP-MATRIX-SCALING

Input: Bipartite adjacency matrix A of $G = (L \cup R, E)$
Output: Decide if G has a perfect matching.

1 **begin**
2   **while** *True* **do**
3    $A \longleftarrow$ Scale every rows of A to make it row stochastic.
4   **if** *All column-sums are in* $(1 - \epsilon, 1 + \epsilon)$ **then**
5   **return** *Yes*
6    $A \longleftarrow$ Scale every column of A to make it column stochastic.
7   **if** *All row-sums are in* $(1 - \epsilon, 1 + \epsilon)$ **then**
8   **return** *Yes*

In both if conditions we are checking if the matrix is $\epsilon-$approximate doubly stochastic matrix. The moment it becomes a $\epsilon-$approximate doubly stochastic matrix we are done.

Now if G doesn't have a perfect matching then we will never reach a ϵ-approximate doubly stochastic matrix since otherwise Hall's condition will hold and then we will have that the permanent is positive. So if G doesn't have a perfect matching the algorithm will run in an infinite loop. We only need to check if G has a perfect matching the algorithm returns Yes.

We will now define a potential function $\Phi \colon \mathbb{Z}_0 \to \mathbb{R}_+$. Let $\sigma \in S_n$ such that $a_{i,\sigma(i)} \neq 0$ for all $i \in [n]$. Now if an entry of the matrix is nonzero then it is always nonzero since all the entries are non-negative. Now since the scalings are symmetric we will define the potential function for i^{th} scaling (row/column) is $\Phi(i) = \prod\limits_{i=1}^{n} a_{i,\sigma(i)}$. So we have $\Phi(0) = 1$ since at first all the entries of the matrix are from $\{0, 1\}$. Also we know $\Phi(t) \leq 1$ for all t since every time we are scaling the matrix. Now $\Phi(1) \geq \frac{1}{n^n}$ since every row-sum can be at most n so it will be divided by n and therefore $a_{i,\sigma(i)} \geq \frac{1}{n}$ for all $i \in [n]$. Now to show the while loop stops if G has a perfect matching it suffices to show that $\Phi(t)$ increases by a multiplicative factor. So we have the following lemma.

> **Lemma 1.1.4**
>
> For all t, $\Phi(t + 1) \geq \Phi(t)(1 + \alpha)$ for some $\alpha \in (0, 1)$.

Proof: Let A' denote the matrix at the tth scaling where the $(t - 1)^{th}$ scaling was column-scaling. Let A'' denote the matrix after row-scaling. Now since we went to the next iteration not all column-sums are in $(1 - \epsilon, 1 + \epsilon)$ after scaling the rows. Now the row sums of A'' are 1. Therefore we have

$$\frac{\Phi(t)}{\Phi(t+1)} = \prod_{i=1}^{n} Col\text{-}sum_i(A'') \leq \left(\frac{\sum\limits_{i=1}^{n} Col\text{-}sum_i(A'')}{n} \right)^n = \left(\frac{\sum\limits_{i=1}^{n} Row\text{-}sum_i(A'')}{n} \right)^n = 1 \implies \Phi(t) \leq \Phi(t+1)$$

Similarly we can say the same if $(t - 1)^{th}$ scaling was row-scaling. Since not all column-sums are in $(1 - \epsilon, 1 + \epsilon)$ we have $\sum\limits_{i=1}^{n} (Col\text{-}sum_i(A'') - 1)^2 \geq \epsilon^2$. Therefore using Lemma 1.1.5 we have

$$\frac{\Phi(t)}{\Phi(t+1)} \leq 1 - \frac{\epsilon^2}{2} \implies \Phi(t+1) \geq \left(1 + \frac{\epsilon^2}{2} \right) \Phi(t)$$

Therefore we have the lemma.   ■

We have $\epsilon < \frac{1}{10n}$. Therefore if $t \geq 200n^4$ then we have

$$1 \geq \Phi(t) \geq \frac{1}{n^n} \left(1 + \frac{1}{200n^2} \right)^t \geq \frac{1}{n^n} e^{n^2} > 1$$

Hence the while loop will iterate for at most $200n^4$ iterations. Hence this algorithm takes $O(n^4)$ time. Hence if G has a perfect matching the algorithm runs for at most $O(n^4)$ iterations. And if G doesn't have a perfect matching then the loop never stops. So we have the new modified algorithm to prevent infinite looping:

Algorithm 2: BP-Matrix-Scaling

    Input: Bipartite adjacency matrix A of $G = (L \cup R, E)$
    Output: Decide if G has a perfect matching.
1 **begin**
2      $\epsilon \longleftarrow \frac{1}{20n}$
3   bp-matrix-scaling-potential-increase **for** $i \in [200n^4]$ **do**
4          $A \longleftarrow$ Scale every rows of A to make it row stochastic.
5   **if** *All column-sums are in* $(1 - \epsilon, 1 + \epsilon)$ **then**
6   **return** *Yes*
7          $A \longleftarrow$ Scale every column of A to make it column stochastic.
8   **if** *All row-sums are in* $(1 - \epsilon, 1 + \epsilon)$ **then**
9   **return** *Yes*

We will prove the helping lemma needed to prove Lemma 1.1.4.

> **Lemma 1.1.5**
>
> Suppose $x_1, \ldots, x_n \geq 0$ and $\sum\limits_{i=1}^{n} x_i = n$ and $\sum\limits_{i=1}^{n} (1 - x_i)^2 \geq \delta$. Then
>
> $$\prod_{i=1}^{n} x_i \leq 1 - \frac{\delta}{2} + o(\delta)$$

Proof: Denote $\rho_i = x_i - 1$. So $\sum\limits_{i=1}^{n} \rho_i = 0$ and $\sum\limits_{i=1}^{n} \rho^2 \geq \delta$. Now

$$\log(1 + \rho_i) = \sum_{j=1}^{\infty} (-1)^{j-1} \frac{\rho_i^j}{j} \implies \log(1 + \rho_i) \leq \rho_i - \frac{\rho_i^2}{3} + \frac{\rho_i^3}{3} \implies 1 + \rho_i \leq e^{\rho_i - \frac{\rho_i^2}{3} + \frac{\rho_i^3}{3}}$$

Therefore we have

$$\prod_{i=1}^{n} x_i \leq \exp\left[\sum_{i=1}^{n} \rho_i - \sum_{i=1}^{n} \frac{\rho_i^2}{3} + \sum_{i=1}^{n} \frac{\rho_i^3}{3}\right] \leq \exp\left[0 - \frac{\delta}{2} + \frac{\left(\sum\limits_{i=1}^{n} \rho_i^2\right)^{\frac{3}{2}}}{3}\right] = \exp\left[-\frac{\delta}{2} + \frac{\delta^{\frac{3}{2}}}{3}\right] \leq 1 - \frac{\delta}{2} + o(\delta)$$

Therefore we have the lemma. ∎