Matching in bipartite graphs

Pramesh Kumar

IIT Delhi

April 7, 2024

Introduction

- Matching has numerous applications including but not limited to
  - Ridesharing/ridehailing (riders to drivers)
  - Jobs to employees
  - Roommate matching
  - and so on

Definitions

Definition (Bipartite graph). A graph $G(N, A)$ is bipartite if we can partition the set of nodes $N = L \cup R$ into two subsets L and R such that $L \cap R = \phi$ and every link in A is incident on one node in L and other node in R.

Definition (Matching). Given an undirected graph $G(N, A)$, a matching is a subset of edges $M \subseteq A$ such that $\forall i \in N$, at most one edge of M is incident on i.

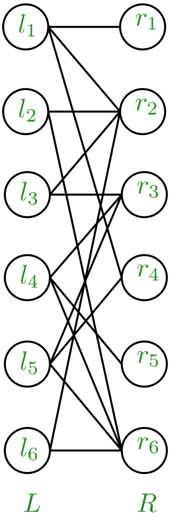
Definition (). A node i is said to be matched by matching M if some edge in M is incident on i, and otherwise, i is unmatched.

Definition (Perfect matching). A matching M is perfect if all nodes are matched i.e., $|M| = \frac{|N|}{2}$.

Definition (Vertex cover). A vertex cover is a set $C \subseteq N$ such that all edges in E are incident on at least one node in C.

Maximum cardinality matching

Definition (Maximum matching). A matching M is called maximum cardinality matching if $|M| \geq |M^{'}|, \forall$ matching $|M^{'}|$ of bipartite G.



LP formulation

Let $x_{ij} = \begin{cases} 1 & \text{if } i \in L \text{ is matched with } j \in R \\ 0 & \text{otherwise} \end{cases}$

Primal (Max matching in bipartite graphs)

$$\begin{aligned} \underset{\mathbf{x}}{\text{maximize}} \quad & \sum_{j \in R} \sum_{i \in L} x_{ij} \\ \text{subject to} \quad & \sum_{j \in R} x_{ij} = 1, \forall i \in L \\ & \sum_{i \in L} x_{ij} = 1, \forall i \in R \\ & x_{ij} \geq 0, \forall i \in L, \forall j \in R \end{aligned}$$

Dual (Vertex cover in bipartite graphs)

$$\begin{aligned} \underset{\mathbf{u}}{\text{minimize}} \quad & \sum_{i \in N} u_i \\ \text{subject to} \quad & u_i + u_j \geq 1, \forall (i,j) \in A \\ & u_i \geq 0, \forall i \in N \end{aligned}$$

Theorem (König)

For any bipartite graph, the maximum size of matching is equal to the minimum size of vertex cover.

Proof.

Follows from strong duality. □

Converting max matching into max flow problem

Construct $G^{'}(N^{'}, A^{'})$ from $G(N, A)$ as follows:

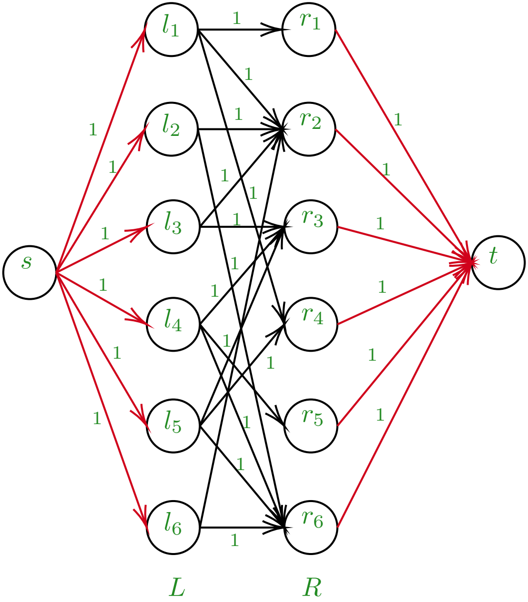
- Let s be source and t be sink and $N^{'} = N \cup \{s, t\}$.
- $A^{'} = \{(s, i) \ : \ i \in L\} \cup \{(i, j) \ : \ i \in L, j \in R, \text{ and } (i, j) \in A\} \cup \{(j, t) \ : \ j \in R\}$.
- Assign capacity $u_{ij} = 1$ to each edge in $A^{'}$.

Lemma

Let $G(N, A)$ be a bipartite graph with nodes partition $N = L \cup R$ and let $G^{'}(N^{'}, A^{'})$ be its corresponding flow network. If M is a matching in G, then there is an integer-valued flow f in $G^{'}$ with value $|f| = |M|$. Conversely, if f is an integer-valued flow in $G^{'}$, then there is a matching M in G with cardinality $|M| = |f|$ consisting of edges $(u, v) \in A$ such that $f(u, v) > 0$.

Remark. The max matching problem can be solved by solving a max flow problem on a modified flow network (e.g., using Ford-Fulkerson labeling algorithm).

Converting max matching into max flow problem



Definitions

Definition (M-alternating path). Given a matching M in undirected graph $G(N, A)$, an $M-$alternating path is a simple path whose edges alternate between being in M and being in $A\backslash M$.

Definition (M-augmenting path). An $M-$alternating path is an $M-$augmenting path if its first and last edges belong to $A\backslash M$. It must contain an odd number of edges.

Remark. By removing the edges in M on an $M-$augmenting path and including the edges not in M on that path will increase the matching by exactly one more edge than M.

Definition (Symmetric difference). For any two sets X and Y, the symmetric difference of X and Y is $X \bigoplus Y = (X \cup Y)\backslash(Y \cap X)$.

Remark. The operator $\bigoplus$ is commutative and associative.

Lemma

Let M be a matching in undirected graph $G(N, A)$ and let P be an $M-$augmenting path. Then the set of edges $M^{'} = M \bigoplus P$ is also a maching in G with $|M^{'}| = |M| + 1$.

Corollary

A matching M in graph $G(N, A)$ is a maximum matching if and only if G contains no $M-$augmenting path.

Remark.

1. Start with $M = \phi$
2. Find $M-$augmenting path P in G (one can use DFS for finding an augmenting path starting from an unmatched node)
3. Set $M = M \bigoplus P$
4. Continue steps 2-3 until there does not exist any $M-$augmenting path.
5. Finding the $M-$augmenting path can be done in $O(m)$ time. Therefore, the overall complexity is $O(nm)$. It can be improved to $O(\sqrt{n}m)$ by augmenting along several paths (Hopcroft-Karp algorithm).

Weighted bipartite matching problem

Definition (Assignment problem). Given a bipartite graph $G(L \cup R, A)$, and weights $w : A \mapsto \mathbb{R}_+$, find a perfect matching M with maximum total weight $\sum_{(i,j) \in M} w_{ij}$

LP formulation

Primal
(Max-weight matching in bipartite graphs)

$$\begin{aligned}\underset{\mathbf{x}}{\text{maximize}} \quad & \sum_{j \in R} \sum_{i \in L} w_{ij} x_{ij} \\ \text{subject to} \quad & \sum_{j \in R} x_{ij} = 1, \forall i \in L \\ & \sum_{i \in L} x_{ij} = 1, \forall i \in R \\ & x_{ij} \geq 0, \forall i \in L, \forall j \in R\end{aligned}$$

Dual

$$\begin{aligned}\underset{\mathbf{u}}{\text{minimize}} \quad & \sum_{i \in N} u_i \\ \text{subject to} \quad & u_i + u_j \geq w_{ij}, \forall (i,j) \in A \\ & u_i \geq 0, \forall i \in N\end{aligned}$$

Using complementary slackness, we know that when $x_{ij} = 1$ i.e., $(i, j) \in M$, then $w_{ij} - u_i - u_j = 0$.

Definition (Labeling). A labeling for a bipartite graph $G(N = L \cup R, A)$ is a function $u : N \mapsto \mathbb{R}$ such that $u_i + u_j \geq w_{ij}, \forall (i,j) \in A$. The labeling is a dual feasible solution.

Definition (Equality subgraph). An equality subgraph corresponding to labeling l is a subgraph $G_l(N, A_l)$ such that $A_l = \{(i,j) \in A \ : \ u_i + u_j = w_{ij}\}$.

Theorem (Kuhn-Munkres Theorem)

Let $G(N = L \cup R, A)$ be a bipartite graph where each edge $(i,j) \in A$ has weight w_{ij}. Let $\mathbf{u}$ be the labeling of G and G_l be the corresponding equality subgraph. If G_l contains a perfect matching M^*, then M^* is an optimal solution to the assignment problem on G.

Remark. Using above algorithm, the max-weight assignment problem reduces to finding right labeling and any perfect matching in the corresponding equality subgraph.

Hungarian algorithm

1. Start with some matching $M = \phi$ and labeling $\mathbf{u}$ defined as:
   $u_k = \begin{cases} 0 & \text{if } k \in R \\ \max_{j' \in R} w_{kj} & \text{if } k \in L \end{cases}$
2. Repeat until M is a perfect matching.
   (i) Find an augmenting path in G_l and augment the matching.
   (ii) If augmenting path does not exist, improve the labeling to $l^{'}$ and go to step (i).

Improving the labeling

- Let P be an alternating path. Let $S \subseteq L$ and $T \subseteq R$ be the respective subset of nodes on path P.
- Compute $\delta_l = \min_{i \in S, j \notin T}\{u_i + u_j - w_{ij}\}$.
- Then improve the labeling to $u^{'}$ as: $u_k^{'} = \begin{cases} u_k - \delta_l & \text{if } k \in S \\ u_k + \delta_l & \text{if } k \in T \\ u_k & \text{otherwise} \end{cases}$

Complexity

Overall complexity is $O(n^3)$.

Stable matching problem

Consider a set of women L and a set of men R. Each women in $i \in L$ provides the ranking of each men $j \in R$ and vice-versa.

Definition (Blocking pair). Given matching M, $i \in L$ and $j \in R$ forms a blocking pair if i prefers j over their current match and j also prefers i over their current match.

Definition (Stable matching). A stable matching is a perfect matching with no blocking pairs.

Gale-Shapley's Deferred-acceptance algorithm

procedure $\mathrm{DA}(G, ranking)$
    Assign each man and woman as free.
    while some woman i is free **do**
        Let j be the first man on i's ranked list to whom she has not proposed.
        if j is free **then**
            i and j get matched (not free)
        else if j ranks i higher than their current match $i^{'}$ **then**
            Replace current match $i^{'}$ of j by i and make $i^{'}$ free.
        else
            j rejects i and i remains free.
        end if
    end while
end procedure

Perfect matching

Theorem (Hall's marriage theorem (1935))

There exists a perfect matching in a bipartite graph $G(N, A)$ *if and only if* $|\delta(S)| \geq |S|, \forall S \subseteq N$, *where* $\delta(S)$ *represents the set of neighbors of nodes in* S.

Remark. A necessary and sufficient condition for a solution of the marriage problem is that each set of k women collectively knows at least k men, for $1 \leq k \leq |L|$

Final remarks

- Matching in non-bipartite graphs is more involved.
- Please read about Edmond's blossom algorithm from book if you are interested.

Suggested reading

- AMO Chapter 12
- https://www.eecs.harvard.edu/cs286r/courses/fall09/papers/galeshapley.pdf

Thank you!