An optimal algorithm for Vertex Cover and Maximum Matching on Bipartite graphs

Linear independence

A collection of row vectors $\{v_i^T\}$ are *independent* if there are no constants $\{c_i\}$ so that $\sum_i c_i v_i^T = 0$.

For an $n \times n$ matrix the rows are independent if and only if the determinant is not 0.

The rank of a matrix the maximum subset of rows that are independent. The rank of the rows and the rank of the columns is the same.

This can be shown by Gaus eliminations.

The rank of a matrix

If we have Maximize $c \cdot v$ such that $Ax \geq b$, and A is an $m \times n$ type matrix. Since for every variable we have and $x \geq 0$, these rows induce the identity matrix of dimension $n \times n$. Thus the rank of the columns is n so this is also the rank of the rows.

Recall the a BFS is obtained by taking n independent *rows* and put equality and solve this $n \times n$, system of equalities. It looks as $A' \cdot x = b'$. (Note that x does not change since x had size n to begin with). Since A' has independent rows the inverse matrix A'^{-1} exists since this is equivalent to the determinant is not 0..

Thus $x = {A'}^{-1} \cdot b$. A unique solution exists. Which is a corner (a basic feasible solution).

Summary

For minimize $c \cdot x$ under $Ax \geq b$, $x \geq 0$,

The main property we use is:

Theorem 1 *The number of independent rows is the number of variables.*

All corners or basic feasible solution are derived by taking n independent *rows and putting $A' \cdot x = b'$. The BFS is $(A')^{-1} \cdot b'$.*

The Vertex Cover problem

Given a graph $G(V,E)$, a subset $U \subseteq V$ is a *Vertex Cover* if for every edge $e = (u,v)$, either $u \in U$, or $v \in U$ or both of the above hold.

The vertex Cover problem

Input: An undirected weighted graph $G(V,E)$ with a cost function $c : V \mapsto \mathcal{Q}_+$ $c(v)$ for every v.

Required: A minimum cost subgraph U that is a Vertex Cover.

An example of a minimum Vertex Cover

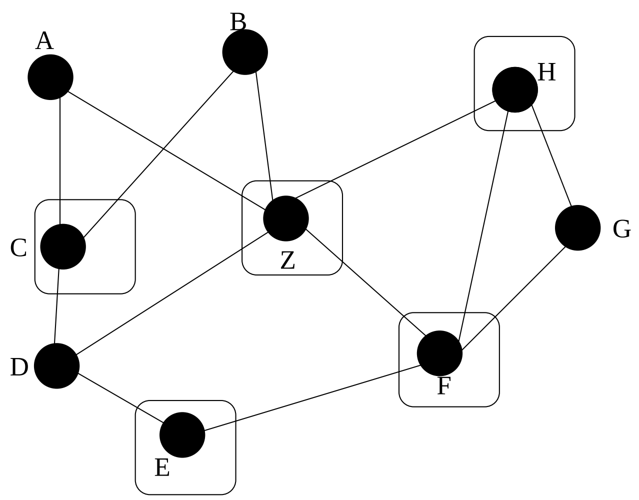

Figure 1: The squared vertices are a vertex cover. Every edge is touched by at least one squared vertex and some times EF by two of the chosen

A fractional LP for VC

$$\begin{aligned} &\text{Minimize} \quad opt_f = c_v \cdot x_v \\ &\qquad\qquad \text{subject to} \\ &\qquad x_v + x_u \geq 1 \qquad \text{for every } e = (u, v) \\ &\qquad x_v \geq 0 \end{aligned}$$

The x_v vertices in a Linear program are 1 if the vertex that is in the solution. And its zero of its not in the solution.

The main inequality $x_v + x_u \geq 1$ indicates that either v or u (or both of them, and so we can not have equality) belongs to the vertex cover.

We need to relax this to a fractional program.

Bipartite graphs

Definition 2 *A bipartite graph $G(V_1, V_2, E)$, is a graph so that $V = V_1 \cup V_2$, and there are no edges inside V_1 and no edges inside V_2, thus all edges go from V_1 to V_2.*

Such graphs are also called 2-colorable and in an equivalent definition its the collection of graphs that do not contains simple odd cycles.

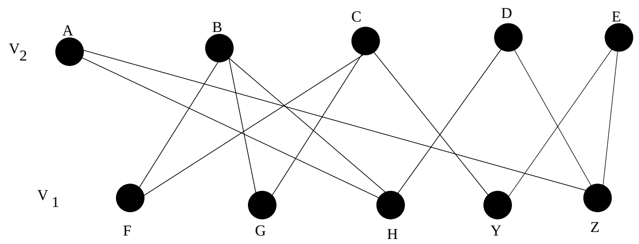

Figure 2: An example of a bipartite graph

The rows correspond to edges

Note that every row has two entries. For an edge $e = uv$ the row of e will have all 0 but the columns of u and v that have 1.

The number of variables is n. Which is the row rank of the matrix.

Consider a collection of edges of a cycle

The cycle is even.

It can also decomposed into two matching.

Every edge in the matching corresponds to a row.

Summing the rows of the first matching M_1 will give the vector with 1 in every column of a vertex in the cycle.

Summing the rows of M_2 gives the same thing.

Hence an independent set of rows (edges) has no cycles and thus induces a forest.

An optimal iterative rounding algorithm

Theorem 3 *Say that $x_v > 0$ for every v. Then there exists a subset $F \subseteq E$ of the edges so that*

1. $x_v + x_u = 1$ *for* $e = vu$.
2. *The rows of the edges are linearly independent*
3. $|F| = |V|$

This follows from the characterization of a BFS.

An optimum algorithm

Let $\delta(v)$ be the edges touching v.

1. Set $U \leftarrow \emptyset$
2. **While** $V(G) \neq \emptyset$ **do**
   (a) Find an optimum solution for the above LP
   (b) If there is a vertex with $x_v = 0$it follows for the remaining edges that $deg_{E'}(v) = 0$ remove this vertex
   (c) If there is a vertex with $x_v = 1$, add v to u
   and set $U \leftarrow U \cup \{v\}$ and $E \leftarrow E \setminus \delta(v)$
   (d) Remove all edges (rows) of edges covered by v.
3. Return U

Correctness

We need to show that in any iteration there is a vertex v so that $x_v = 1$ or $x_v = 0$. Then optimality follows like in the *Assignment problem*

Let E' be the non covered edges.

Claim 1 *In every iteration either we find a vertex v with $x_v = 0$ and $deg_{E'}(v) = 0$ or we find a vertex v with $x_v = 1$.*

Proof

For the sake of contradiction, assume that Claim 1 is false.

Thus for every vertex $x_v < 1$ and if $x_v = 0$, $deg_{E'}(v) > 0$. Note that for no vertex touched by at least one edge, $x_v = 0$ since in this case all the neighbors if v have value 1

Let F be the set of edges whose rows are independent and $|F| = |V|$.

This gives a contradiction as any cycle implies that the rows of F are not independent. See Corollary **??** that implies that F is a forest, and so has at most $n - 1$ edges.

A polynomial algorithm for maximum matching in Bipartite graphs

The LP

Maximize $\sum_e y_e \cdot c_e$

Subject to:

$$\sum_{e\in E(v)} y_e \leq 1,$$

$$y_e \leq 1, y_e \geq 0.$$

The algorithm we will show that works

The algorithm:

1. $F \leftarrow \emptyset$
2. **While** $E(G) \neq \emptyset$ **do:**
   (a) Compute a solution to the above LP.
   (b) Remove every edge e with $y_e = 0$.
   (c) If there is an edge $e = uv$ so that $y_e = 1$ then set $F \leftarrow F \cup F \cup \{e\}$ and $E \leftarrow E \setminus \{uv\}$
   (d) Remove all edges with at least one endpoint in e.
3. Return F

When is a graph a collection of cycles?

Claim 2 *If $deg(v) = 2$ for every v, then there is a collection of vertex disjoint cycles that contains all of V*

Proof. Consider an edge $e = uv$. Since both u, v have degree 1 now, but degree 2 in the graph, we can make the path longer by two, adding an edge to u and to v. In general we get a path whose first and last vertices have degree 1. Since the graph is finite, these two paths must meet getting a cycle. The same argument implies that G is a collection of cycles. $\square$

Characterization of extreme points

There are more rows than columns so we need to choose a subset of the rows W so that $|W| = |E|$ because there are $|E|$ variables and so you need to choose $|E|$ rows. Note that we choose a set of $|E|$ **vertices**. WE should think of that as a set $W \subset V$.

Note that for the chosen rows the inequality also hold with **equality**. The vectors whose inequality is chosen to hold with equality Also, the vectors of the vertices must be **linearly independent**.

there are no cycles

Say that the graph restricted to the vertices in W with all the edges with both endpoints in W is a independent set. If the cycle is $x_1 - -y_1 - -x_2 - -y_2 - -x_3 - -y_3 - x_1$ for example, the vectors of x_1, x_2, x_3 have the same edges as y_1, y_2, y_3. This implies that the matrix is dependent. Hence the graph induced by W has no cycles.

This means that the graph $G(W)$ with W as vertices and edges with both endpoints in W is a **forest**.

Proof

Say that we have no edges that are 1 or 0. Namely for every edge $0 < y_e < 1$.

Let $deg_W(v)$ be the number of neighbors v has in W Its the degree of the vertex in $G(W)$ (we are assuming here that $v \in W$).

Claim 3 $deg_W(v) \geq 2$

Proof. We know by the characterization of an extreme point that that for every $v \in W$ $\sum_{uv,\ u \in W} y_u = 1$.

Since for $e = vu$, $y_{vu} = y_e < 1$ for every e, the degree $deg_W(v)$ of v in $G(W)$ is at least 2. □

Now we show its almost 2.

Claim 4 *Vertices in W have degree at most 2 in G. In particular vertices not in W have no edges to vertices in W.*

Proof. $2|W| = 2|E| = \sum_{v\in V} deg(v) \geq \sum_v deg_W(v) \geq 2|W|$

The last inequality follows because the degree of every vertex in W is at least 2.

This means that all inequalities are equalities:

$2|W| = 2|E| = \sum_{v\in V} deg(v) = \sum_v deg_{G(W)}(v) = 2|W|$

We know that $deg_W(v) \geq 2$. But if there exists a vertex that has a neighbor outside W then its degree is at least 3 as its degree inside W is 2. In this case we get contradiction if a degree 3 appears since we get $2|W| = 2|W| + 1$.

Thus vertices in $V - W$ have no edges to vertices in W. Also vertices have degree exactly 2 in $G(W)$. □

Proof continued

By the claim above the graph $G(W)$ is a collection of cycles. Because the degrees are exactly 2 in $G(W)$. But we can not have cycles as it leads to linear dependence.

Thus there is an e so that $y_e = 0$ or $y_e = 1$.

This ends the proof.

One thing we proved was that Minimum Vertex cover is polynomial in bipartite graphs. Thus so is Maximum Independent Set.