Log-Concave Random Graphs

Alan Frieze and Santosh Vempala and Juan Vera

March 19, 2008

The geometric construction of a random graph

The geometric construction of a random graph

Let $K = [0,1]^{\binom{[n]}{2}}$.

Algorithm Generate(K, p):
Choose X uniformly from K and let

$$G_{K,p} = ([n], E_p)$$

where

$$E_p = \{e : X_e \leq p\}.$$

Here $G_{K,p} = G_{n,p}$.

The geometric construction of a random graph

Let K be any convex subset of the non-negative orthant.

Algorithm Generate(K, p):
Choose X uniformly from K and let

$$G_{K,p} = ([n], E_p)$$

where

$$E_p = \{e : X_e \leq p\}.$$

Here $G_{K,p}$ is a new model of a random graph.

Special Classes of Graph

Notice that $G_{K,p}$ is triangle free if we take $p < p_0$ and K to be

$$x_{ij} + x_{jk} + x_{ki} \geq 3p_0 \quad \forall i, j, k$$

$$0 \leq x_{ij} \leq 1 \quad \forall i, j$$

We can (almost) generate $G_{K,p}$ in polynomial time.

Special Classes of Graph

Notice that $G_{K,p}$ is triangle free if we take $p < p_0$ and K to be

$$x_{ij} + x_{jk} + x_{ki} \geq 3p_0 \quad \forall i,j,k$$

$$0 \leq x_{ij} \leq 1 \quad \forall i,j$$

We can (almost) generate $G_{K,p}$ in polynomial time.

We can exclude any fixed graph H in this way.

Special Classes of Graph

Notice that $G_{K,p}$ is triangle free if we take $p < p_0$ and K to be

$$x_{ij} + x_{jk} + x_{ki} \geq 3p_0 \quad \forall i, j, k$$

$$0 \leq x_{ij} \leq 1 \quad \forall i, j$$

We can (almost) generate $G_{K,p}$ in polynomial time.

We can exclude any fixed graph H in this way.
We can also generate graphs with a fixed degree sequence.

Special Classes of Graph

Notice that $G_{K,p}$ is triangle free if we take $p < p_0$ and K to be

$$x_{ij} + x_{jk} + x_{ki} \geq 3p_0 \quad \forall i,j,k$$

$$0 \leq x_{ij} \leq 1 \quad \forall i,j$$

We can (almost) generate $G_{K,p}$ in polynomial time.

We can exclude any fixed graph H in this way.
We can also generate graphs with a fixed degree sequence.
Unfortunately, we have not found a way to make this generation uniform.

More generally, let F be any integrable log-concave function on the positive orthant of $\mathbb{R}^N$.

Algorithm Generate(F, p):
Choose X uniformly from the distribution proportional to F and let

$$G_{F,p} = ([n], E_p)$$

where

$$E_p = \{e : X_e \leq p\}.$$

More generally, let F be any integrable log-concave function on the positive orthant of $\mathbb{R}^N$.

Algorithm Generate(F,p):
Choose X uniformly from the distribution proportional to F and let

$$G_{F,p} = ([n], E_p)$$

where

$$E_p = \{e: \ X_e \leq p\}.$$

We get a graph process by increasing p from 0 to ∞.

F is axis-symmetric if it is invariant under permutation of coordinates.

So,

$$G_{F,p} \textit{ given } |E_p| = m \textit{ is distributed as } G_{n,m}.$$

So, for this case, it is merely a question of analysing $|E_p|$.

Some results:

Theorem

Let F be distribution in the positive orthant with a down-monotone logconcave density and second moment σ^2 along every axis. There exist constants $A_1 < A_2$ such that

$$\lim_{n\to\infty} \mathbf{Pr}(G_{F,p} \text{ is connected}) = \begin{cases} 0 & p < \frac{A_1 \sigma \ln n}{n} \\ 1 & p > \frac{A_2 \sigma \ln n}{n} \end{cases}$$

By down-monotone we mean that if $x \geq y$ then $f(x) \leq f(y)$.

In the second moment condition $\sigma^2 = \mathbf{E}(X_e^2)$.

Theorem

Let F be distribution in the positive orthant with a down-monotone logconcave density and second moment σ^2 along every axis. There exist constants $A_3 < A_4$ such that

$$\lim_{\substack{n\to\infty \\ n \text{ even}}} \mathbf{Pr}(G_{F,p} \textit{ has a perfect matching}) = \begin{cases} 0 & p < \frac{A_3 \sigma \ln n}{n} \\ 1 & p > \frac{A_4 \sigma \ln n}{n} \end{cases}$$

Theorem

Let F be distribution in the positive orthant with a down-monotone logconcave density and second moment σ^2 along every axis. Then there exists an absolute constant A_5 such that if

$$p \geq A_5 \frac{\ln n}{n} \cdot \frac{\ln \ln \ln n}{\ln \ln \ln \ln n}$$

then $G_{F,p}$ is Hamiltonian **whp**.

The case of a Simplex

We now consider the case of $G_{K,p}$ where

$$K = \{X : \sum_e \alpha_e X_e \leq L\}$$

We usually assume $L = N = \binom{n}{2}$, which can be achieved by scaling.

We assume that α is $M = M(n)$-bounded in the sense that

$$\frac{1}{M} \leq \alpha_e \leq M \textit{ for all } e.$$

The case of a Simplex

We now consider the case of $G_{K,p}$ where

$$K = \{X : \sum_e \alpha_e X_e \leq L\}$$

We usually assume $L = N = \binom{n}{2}$, which can be achieved by scaling.

We assume that α is $M = M(n)$-bounded in the sense that

$$\frac{1}{M} \leq \alpha_e \leq M \text{ for all } e.$$

With no constraints on α, we can essentially generate random subgraphs of an arbitrary graph G.

Let

$$\alpha_v = \sum_{w \neq v} \alpha_{vw} \qquad \textit{for } v \in [n].$$

Theorem

Assume w.l.o.g. that $L = N$ (otherwise replace p by pN/L). Suppose that α is $M = o((\ln n)^{1/4})$-bounded.

Let p_0 be the solution to

$$\sum_{v \in [n]} \left(1 - \frac{\alpha_v p}{N}\right)^N = 1.$$

Then for any fixed $\epsilon > 0$,

$$\lim_{n \to \infty} \mathbf{Pr}(G_{K,p} \text{ is connected}) = \begin{cases} 0 & p \leq (1 - \epsilon)p \\ 1 & p \geq (1 + \epsilon)p \end{cases}.$$

Diameter

Theorem

Let $k \geq 2$ be a fixed integer. Suppose that α is M-bounded and for simplicity assume only that $M = n^{o(1)}$. Suppose that θ is fixed and satisfies $\frac{1}{k} < \theta < \frac{1}{k-1}$. Suppose that $p = \frac{1}{n^{1-\theta}}$. Then **whp** $\mathbf{diam}(S_{n,p,\alpha}) = k$.

Edge Weighted Problems

One can also use X_e as an edge weight and ask for the expected weight of various quantites.

One can do probabilistic analysis with edge weights generated in this model.

Asymmetric Traveling Salesman Problem (ATSP).

We can use a variant of an algorithm of Karp and Steele to find a tour within $1 + o(1)$ of optimum. Suppose that the edge weights of the complete digraph on n vertices are given by the X_e.
Suppose that $M \leq n^\delta$.

We need an extra assumption: f has *column symmetry*: for any permutation π

$$f(\mathbf{x}_{\pi(1)}, \mathbf{x}_{\pi(2)}, \ldots, \mathbf{x}_{\pi(n)}) = f(\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_n).$$

where $\mathbf{x}_i = (x_{1,i}, x_{2,i}, \ldots, x_{n,i})$.

Weight of Minimum Spanning Tree

Suppose we are in the simplex case and $\alpha_{vw} = d_v d_w$, where $1 \leq d_v \leq (\ln n)^{1/10}$. Suppose that the edge weights of the complete graph on n vertices are given by the X_e.

Let Z denote the length of the minimum spanning tree. Then,

$$\mathbf{E}(Z) \sim \sum_{k=1}^{\infty} \frac{(k-1)!}{D^k} \sum_{\substack{S \subseteq V \\ |S|=k}} \frac{\prod_{v \in S} d_v}{d_S^2}.$$

Here $d_S = \sum_{v \in S} d_v$ and $D = d_V$.
If $d_v = 1, \forall v$ then $\mathbf{E}(Z) \sim \sum_{k=1}^{\infty} \frac{(k-1)!}{n^k} \binom{n}{k} \frac{1}{k^2} \sim \zeta(3)$.

Proofs of theorems are based on modifying $G_{n,p}$ type proofs:

General Case:

Lemma

$$e^{-c_1 p|S|/\sigma} \leq \mathbf{Pr}(S \cap E_p = \emptyset) \leq e^{-c_2 p|S|/\sigma}$$

Lower bound requires $p/\sigma < 1/4$.

$$\left(\frac{c_3 p}{\sigma}\right)^{|S|} \leq \mathbf{Pr}(S \subseteq E_p) \leq \left(\frac{c_4 p}{\sigma}\right)^{|S|}.$$

Simplex Case

Lemma

(a) *If* $S \subseteq E_n$ *and* $E_p = E(G_{\Sigma_L,p})$,

$$\mathbf{Pr}(S \cap E_p = \emptyset) = \left(1 - \frac{\alpha(S)p}{L}\right)^N.$$

(b) *If* $S, T \subseteq E_n$ *and* $S \cap T = \emptyset$ *and* $|T| = o(n)$ *and* $\alpha(S)|T|p, \alpha(T)Np, MNp = o(L)$ *then*

$$\mathbf{Pr}(S \cap E_p = \emptyset,\ T \subseteq E_p) =$$
$$(1 + o(1)) \left(\prod_{e \in T} \alpha_e\right) \left(\frac{Np}{L}\right)^{|T|} \left(1 - \frac{\alpha(S)p}{L}\right)^N.$$

$p \geq \frac{A_1 \sigma \ln n}{n}$:

$$\begin{aligned}\mathbf{Pr}(G \text{ is not connected}) &\leq \sum_{k=1}^{\lfloor n/2 \rfloor} \binom{n}{k} e^{-c_2 pk(n-k)/\sigma} \\ &\leq \sum_{k=1}^{\lfloor n/2 \rfloor} \left(\frac{ne}{k} e^{-\frac{1}{2} A_1 c_2 \ln n} \right)^k \\ &= o(1).\end{aligned}$$

$p \leq \frac{C_1 \sigma \ln n}{n}$

$$\mathbf{Pr}(v \textit{ is isolated}) \geq e^{-c_1 p(n-1)/\sigma} \geq n^{-C_1 c_1}.$$

So if Z is the number of isolated vertices:

$$\mathbf{E}(Z) \geq n^{1-C_1 c_1}.$$

$$\begin{aligned}
&\mathbf{Pr}(v, w \textit{ isolated}) = \mathbf{Pr}(v \textit{ isolated and } w \textit{ has no edges to } V \setminus \{v\}) \\
&\leq \mathbf{Pr}(v \textit{ is isolated})\mathbf{Pr}(w \textit{ has no edges to } V \setminus \{v\}), \\
&\leq (1 + o(1))\mathbf{Pr}(v \textit{ is isolated})(\mathbf{Pr}(w \textit{ is isolated}) + \mathbf{Pr}(x_{vw} \leq p)) \\
&\leq (1 + o(1))\mathbf{Pr}(v \textit{ is isolated})(\mathbf{Pr}(w \textit{ is isolated}) + c_3 p/\sigma) \\
&\leq (1 + o(1))\mathbf{Pr}(v \textit{ is isolated})(\mathbf{Pr}(w \textit{ is isolated}) + O(\ln n/n)) \\
&= (1 + o(1))\mathbf{Pr}(v \textit{ is isolated})\mathbf{Pr}(w \textit{ is isolated}).
\end{aligned}$$

Chebyshev inequality implies that $Z \neq 0$ **whp**.

TSP Analysis

The matrix $X(i,j)$ can be viewed as weights of edges of complete digraph: Digraph View or as the weights of edges of a complete bipartite graph: Bipartite View.

Algorithm

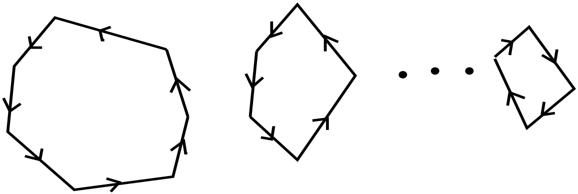
Step 1 Solve the assignment problem with cost matrix X i.e. find a minimum cost perfect matching in the bipartite view. The edges (i,j) of the optimal assignment form a set of vertex disjoint cycles $C_1, C_2, \ldots, C_k$ in the digraph view.

Step 2 Assume that $|C_1| \geq |C_2| \geq \cdots \geq |C_k|$.
For $i = k$ down to 2: $C_1 \leftarrow C_1 \oplus C_i$. (*Patch C_i into C_1*).

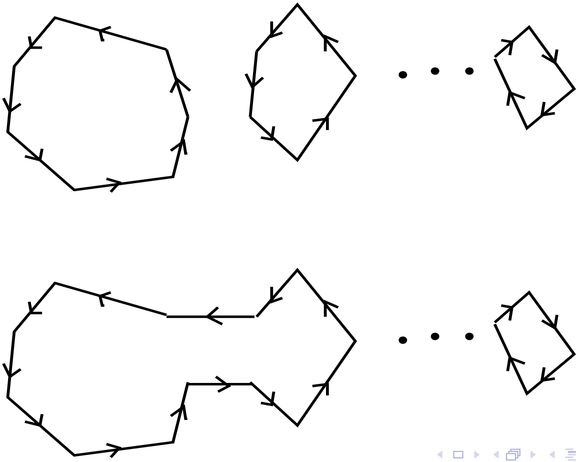
Here $C_1 \oplus C_i$ is obtained by removing an edge (a,b) from C_1 and an edge (c,d) from C_i and adding edges $(a,d),(c,b)$ to make one cycle. These two edges are chosen to minimise the cost $X_{ad} + X_{cb}$.

Each patch reduces the number of cycles by one and so the procedure ends with a tour.
Column symmetry implies that the set of cycles found in Step 1 is a random cycle cover and then **whp** it has $O(\ln n)$ cycles.

Each patch reduces the number of cycles by one and so the procedure ends with a tour.
Column symmetry implies that the set of cycles found in Step 1 is a random cycle cover and then **whp** it has $O(\ln n)$ cycles.



Each patch reduces the number of cycles by one and so the procedure ends with a tour.
Column symmetry implies that the set of cycles found in Step 1 is a random cycle cover and then **whp** it has $O(\ln n)$ cycles.



Expected weight of MST in simplex case

T is minimum spanning tree. K denotes simplex.

$$
\begin{aligned}
\ell(T) &= \sum_{e\in T} X_e \\
&= \sum_{e\in T}\int_{p=0}^{N} 1_{X_e\geq p}dp \\
&= \int_{p=0}^{N}\sum_{e\in T}|\{e:\ X_e\geq p\}|dp \\
&= \int_{p=0}^{N}(\kappa(G_{K,p})-1)dp
\end{aligned}
$$

where κ denotes the number of components.

$$\mathbf{E}(T)=\int_{p=0}^{N}(\mathbf{E}(\kappa(G_{K,p}))-1)dp.$$

$\tau_{k,p}$ denotes the number of components of $G_{K,p}$ that are isolated trees with k vertices For $X \subseteq V$ we let $A_k = \{a \in [1,k]^k : \sum_{j=1}^k a_j = 2k-2\}$. Then, where $q = e^{-Dp}$

$$
\begin{aligned}
\mathbf{E}[\tau_{k,p}] &\sim (k-2)!p^{k-1} \sum_{a \in A_k} \sum_{f:[k]\to V} \prod_{j=1}^{k} \frac{d_{f(j)}^{a_j} q^{d_{f(j)}}}{(a_j-1)!} \\
&\sim (k-2)!p^{k-1} \sum_{a \in A_k} \prod_{i=1}^{k} \sum_{v=1}^{n} \frac{d_v^{a_i} q^{d_v}}{(a_i-1)!} \\
&\sim (k-2)!p^{k-1} [x^{2k-2}] \left(\sum_{v=1}^{n} \sum_{r=1}^{\infty} \frac{q^{d_v} d_v^r}{(r-1)!} x^r \right)^k \\
&= (k-2)!p^{k-1} [x^k] \left(\sum_{v=1}^{n} q^{d_v} d_v e^{d_v x} \right)^k \\
&= (k-2)!p^{k-1} \sum_{S \subseteq V, |S|=k} q^{d_S} \frac{d_S^{k-2}}{(k-2)!} \prod_{v \in S} d_v
\end{aligned}
$$

So,

$$\sum_{k\geq 1}\int_{p\geq 0}\mathbf{E}[\tau_{k,p}]dp \sim \sum_{k\geq 1}\sum_{\substack{S\subseteq V\\|S|=k}} d_S^{k-2}\prod_{v\in S}d_v\int_{p\geq 0}p^{k-1}e^{-d_S Dp}dp$$

$$= \sum_{k\geq 1}\sum_{\substack{S\subseteq V\\|S|=k}}\frac{\prod_{v\in S}d_v}{d_S^2 D^k}\int_{x\geq 0}x^{k-1}e^{-x}dx$$

$$\sim \sum_{k=1}^{\infty}\frac{(k-1)!}{D^k}\sum_{\substack{S\subseteq V\\|S|=k}}\frac{\prod_{v\in S}d_v}{d_S^2}$$

Open Questions

Open Questions

1. Does every monotone property have a threshold in $G_{F,p}$ (in simplex case)?

Open Questions

1. Does every monotone property have a threshold in $G_{F,p}$ (in simplex case)?
2. When does $G_{F,p}$ have a giant component?

Open Questions

1. Does every monotone property have a threshold in $G_{F,p}$ (in simplex case)?
2. When does $G_{F,p}$ have a giant component?
3. Do we need to restrict ourselves to down-monotone functions?

Open Questions

1. Does every monotone property have a threshold in $G_{F,p}$ (in simplex case)?
2. When does $G_{F,p}$ have a giant component?
3. Do we need to restrict ourselves to down-monotone functions?
4. Is there a polytope K that provides uniform generation of H-free sub-graphs of a fixed graph G. ($H = P_2$ gives matchings).

Open Questions

1. Does every monotone property have a threshold in $G_{F,p}$ (in simplex case)?
2. When does $G_{F,p}$ have a giant component?
3. Do we need to restrict ourselves to down-monotone functions?
4. Is there a polytope K that provides uniform generation of H-free sub-graphs of a fixed graph G. ($H = P_2$ gives matchings).
5. Do the above models of a random graph have a use in Ramsey theory?