Improving some Combinatorial Tools

Angel Garrido
Facultad de Ciencias de la UNED, Madrid, Spain
agarrido@mat.uned.es

Abstract

We analyze here some very interesting and useful tools of Combinatorics, to develop the new and quickly evolving theory so-called as Probabilistic Graphical Models, with many current and quickly expanding applications.

Keywords: Combinatorics, Graph Theory, Artificial Intelligence.

Mathematics Subject Classification: 68727, 03E72, 26E50.

1 Introduction

Let G be a graph. Suppose that we denote by *V(G)* their set of nodes, and by *E(G)* their set of edges.

A graph, *G,* is said to be *node-transitive* (or *vertex-transitive*), if for any two of its nodes, n_i and n_j, there is an automorphism which maps n_i to n_j.

A simple graph, *G,* is said to be *edge-transitive* (or *link-transitive*), if for any two of its edges, e and $e\,'$, there is an automorphism which maps e into $e\,'$.

A simple graph, *G,* is said to be *symmetric,* when it is both, node-transitive and edge-transitive.

But a simple graph, *G,* which is edge-transitive, but not node-transitive, is said *semi-symmetric.* Obviously, such a graph will be necessarily a bipartite graph.

Let G be an undirected graph (UG). We says that G is *chordal,* if every cicle of length strictly greater than three possesses a "chord". This name ("chord") means an edge joining two non-consecutive nodes of the cycle. Therefore, an UG will be chordal, if it does not contain an induced subgraph isomorphic to the C_n, when $n > 3$.

The *chordality* result a *hereditary property,* because all the induced subgraphs of a chordal graph will be also chordal. For instance, the interval graphs are chordal.

[1]AMO - Advanced Modeling and Optimization. ISSN: 1841-4311

2 The Incidence Problem

We may to define the *Incidence Matrix* of an Incidence Structure as a *(pxq)*-matrix, where p and q are the number of points and the number of lines, respectively, in such a way that

$$b_{ij} = \begin{cases} +1, \; if\ the\ point\ p_i\ and\ the\ line\ L_j\ are\ incident \\ 0, \; otherwise \end{cases}$$

In this case, the Incidence Matrix will be also a biadjacency matrix of the Levi graph of the structure. And because there is a Hypergraph for every Levi graph, and a Levi graph for every Hypergraph (or vice versa), it is possible to conclude that *the incidence matrix of an incidence structure describes a hypergraph.*

We can also to introduce the matrix of *valencies,* also called *degrees* (deg), of the nodes in the graph. Such *degree matrix* will be denoted by $D(G)$. And there you are three fundamental matrices which may appears associated with a graph, as the *Incidence Matrix,* that encapsulates node-edge relationships, the *Adjacency Matrix,* that encapsulates node-node relationships, and the *Degree Matrix,* that encapsulates information about the degrees. But also there is one last and many times interesting matrix, which will be called the *Laplacian Matrix.*

Let G be a graph. We can define *the Laplacian Matrix of G,* denoted $L\left(G\right)$, by

$$L\left(G\right) \equiv D\left(G\right) - A\left(G\right)$$

where $D\left(G\right)$ is the degree matrix of *G,* and $A\left(G\right)$ is the adjacency matrix of G.

3 Enumerating Graphs

Among the different type of graphs, Bayesian Networks are the most sucessful class of models to represent uncertain knowledge. But the representation of conditional independencies (CIs, in acronym) does not have uniqueness. The reason is that probabilistically equivalent models may have different representations.

And this problem is overcome by the introduction of the concept of *Essential Graph (EGs),* as unique representant of each equivalence class.

Knowing the ratio of EGs to DAGs (Directed Acyclic Graphs) is a valuable tool, because through this information we may decide in which space to search. If the ratio is low, we may prefer to search the space of DAG models, rather than the space of DAGs directly, as it was usual until now.

Recall that a DAG, *G,* is *essential,* if every directed edge of G is protected. So, an *Essential Graph (EG)* is a graphical representation of a Markov equivalence class.

In relation with the Essential Graph, each directed edge would have the same direction in all the graphs that belongs to its equivalence class.

There is a bijective correspondence (one-to-one) among the set of Markov equivalence classes and the set of essential graphs, its representatives.

The labeled or unlabeled character of the graph means whether its nodes or edges are distinguishable or not.

For this, we will say that it is *vertex-labeled, vertex-unlabeled, edge-labeled, or edge-unlabeled.*

The labeling will be considered as a mathematical function, referred to a value or name (label), assigned to its elements, either nodes, edges, or both, which makes them distinguishable.

Let

$$F(s) \equiv \sum_{i \in N^*} \frac{c_i \, s^i}{i!}$$

Then, if we denotes as Δ the linear operation on exponential generating functions which divides by $\exp_2 C_{i,2}$, i.e.

$$\Delta F(s) \equiv \sum_{i \in N^*} \frac{c_i \, s^i}{i! \, \exp_2 C_{i,2}}$$

So, we can use the function $\Delta F(s)$ to count labeled DAGs. It will be called as the *Special Generating Function for F*.

Let a_n be the number of labeled n-DAGs. We can found the zeroes of the function

$$\Delta \exp(-s) = \sum_{i \in N^*} \frac{(-s)^i}{i! \, \exp_2 C_{i,2}}$$

by Mathematical Analysis, more exactly by *Theory of Residua.*

4 Searching for bounds

Let a_n be the number of essential (labeled) DAGs, and $a_{n'}$ be the number of (also labeled) DAGs. Then, a_n is given by the recurrence equation

$$a_n = \sum_{s=1}^{n} (-1)^{s+1} C_{n,s} \left(2^{n-s} - n + s\right)^{s} a_{n-s}, \; with \; a_0 = 1$$

Whereas we can obtain for the number of labeled n-DAGs,

$$a_n' = \sum_{s=1}^{n} (-1)^{s+1} C_{n,s} \left(2^{n-s}\right)^{s} a_{n-s}', \; with \; a_0' = 1$$

The basic idea is to count the number of *n-DAGs* considering each digraph as created by adding terminal nodes to a DAG with lesser number of nodes. After this addition, we obtain a new DAG. So, the new formula would be recursive, and it is a direct application of the *IEP.* From which, we can reach directly the equation.

We may rewrite the equation as

$$\sum_{s=0}^{n} (-1)^{n-s} C_{n,s} \left(2^{s} - s\right)^{n-s} a_{s} = 0,\ with\ n \geq 1$$

Another case of application of *Inclusion-Exclusion Principle* is to find the cardinal of the set of essential DAGs, *E,* with a set of nodes $\{1, 2, \ldots, n\}$. For this, we start with a family of sets, as the aforementioned $\{A_{k}\}_{k=1}^{n}$.

Therefore, to know the cardinal of *E,* first we compute the intersection that appears in the last summatory, for $j = 1,\ 2,\ \ldots,\ n.$ With the total allowed connection numbers, from a given node being

$$2^{n-s} - n + s$$

So, the number of possible ways of adding directed edges from the essential graph until all the s terminal nodes will be

$$[2^{n-s} - n + s]^{s}$$

If we denote a_{n} the number of *essential* $n - DAGs$, this would be

$$a_{n} = \sum_{s=1}^{n} (-1)^{s+1} C_{n,s} \left(2^{n-s} - (n-s)\right)^{s} a_{n-s},\ with\ a_{0} = 1$$

It is possible to obtain a very similar expression. In such case, the purpose was to obtain a number of labeled *n-DAGs.* It would be

$$a_{n}{'} = \sum_{s=1}^{n} (-1)^{s+1} C_{n,s} \left(2^{n-s}\right)^{s} a_{n-s}{'},\ with\ a_{0}{'} = 1$$

If we denote e_{n} the number of *essential n-graphs,* also labeled, it holds

$$a_{n} \leq e_{n} \leq a_{n}{'}$$

I.e. both precedent values, a_{n} and a_{n}', are the lower and upper bounds of e_{n}, for each selected order, n. So, it holds

$$\frac{1}{13.6} \leq \frac{a_n}{a_n'}$$

And also

$$\frac{1}{13.6} a_{n}{'} \leq e_{n} \leq a_{n}'$$

where we obtain lower and upper bounds for the cardinal of essential graphs.

5 Conclusion

We shall support our study on a more powerful analytical framework, improving theoretical foundations, being the different result coherent with the precedent known computations.

References

[1] E. A. Bender, L. B. Richmond, R. W. Robinson, and N. C. Wormald (1986). The asymptotic number of acyclic digraphs I. *Combinatorica* **6** (1): 15-22.

[2] E. A. Bender, and R. W. Robinson (1988). The asymptotic number of acyclic digraphs II. *J. Comb.Theory,* Serie B **44** (3): 363-369.

[3] F. Harary (1955). The number of linear, directed, rooted, and connected graphs. *Proc. Amer. Math. Soc.* (AMS) **78**: 445-463.

[4] F. Harary (1957). The number of oriented graphs. *Michigan Math. J.* Volume 4, Issue 3: 221-224.

[5] F. Jensen and T. Nielsen. 2007. *Bayesian Networks and Decision Graphs.* Springer-Verlag.

[6] Y. Liu. 2002. Computational Symmetry. In *Proceedings of Symmetry 2000.* Portland Press, London, 80(1):231–245.

[7] H. Weyl. 1952. *Symmetry.* Princeton University Press.