Random Graph Simulation Using Maximum Entropy Probabilities

CENTER FOR FINANCIAL & BUSINESS ANALYTICS
Data Science Institute

Paul Glasserman and Enrique Lelo de Larrea
DRO, Columbia Business School
IEOR, Columbia Engineering

Data Science Institute
COLUMBIA UNIVERSITY

Financial Networks with Partial Information

We are interested in asset-firm networks [3] (bipartite graphs) and inter-bank networks [2] (directed graphs). In a simplified setting, partial information amounts to knowing only the degree sequence of the network. The enumeration of all graphs consistent with a degree sequence becomes impossible as the sequence increases. Instead, we want to simulate (uniformly) from the set of such graphs.

Observe that both bipartite and directed graphs can be represented with 0—1 matrices.

$$c_1 = 1 \quad c_2 = 1 \quad c_3 = 2$$
$$\begin{pmatrix} \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \end{pmatrix} \begin{matrix} r_1 = 1 \\ r_2 = 1 \\ r_3 = 2 \end{matrix}$$

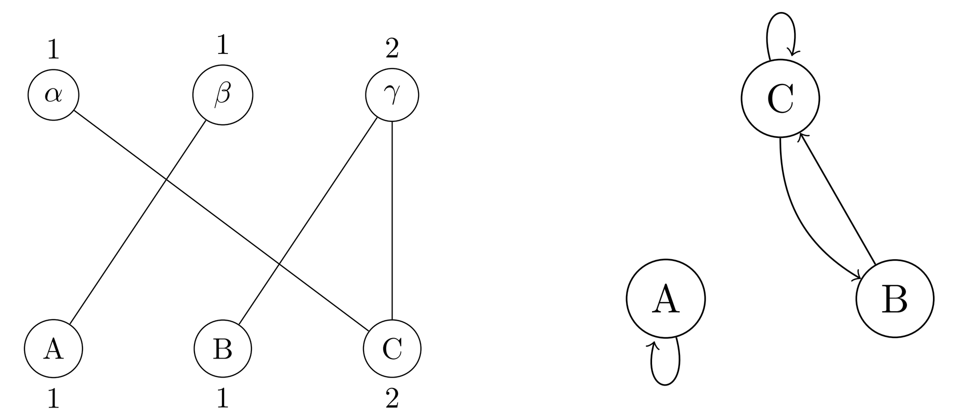

Figure 1. A particular degree sequence and two graphs (one bipartite and the other directed) that satisfy said sequence.

Notation

Let r and c be the degree sequences. Let W be a pattern matrix. If $w_{ij} = 0$ then that entry is excluded. Let $\Sigma(r, c; W)$ be our 0—1 matrix target space. Finally, let $P(r, c; W)$ be the convex hull of $\Sigma(r, c; W)$. X in $P(r, c; W)$ can be seen as a probability matrix and if we simulate the edges independently we get $P_X(M) = \prod_{ij} x_{ij}^{m_{ij}} (1 - x_{ij})^{1 - m_{ij}}$, **for any target matrix M in $\Sigma(r, c; W)$. We define the entropy as**

$$H(X) = -\sum_{ij} \left(x_{ij} \log(x_{ij}) + (1 - x_{ij}) \log(1 - x_{ij}) \right).$$

Maximum Entropy Probabilities

Denoted as $Z = Z(r, c; W)$, it is the unique solution of both the maximum entropy problem and the max-min probability problem:

$$\begin{aligned} &\underset{X}{\text{maximize}} && H(X) \\ &\text{subject to} && X \in \mathcal{P}(r, c; W). \end{aligned} \qquad \begin{aligned} &\underset{X}{\text{maximize}} && \min_{M \in \Sigma(r,c;W)} P_X(M) \\ &\text{subject to} && X \in \mathcal{P}(r, c; W). \end{aligned}$$

For instance:

- $r = c = (1, 1)$ and $w_{ij} = 1$.
- $\Sigma(r, c; W)$ has only two elements:

$$M_1 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \; M_2 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}.$$

- $\mathcal{P}(r, c; W)$ can be expressed as the matrices

$$X_p = \begin{pmatrix} p & 1-p \\ 1-p & p \end{pmatrix},$$

with p in $[0, 1]$.

- $Z = X_{1/2}$.

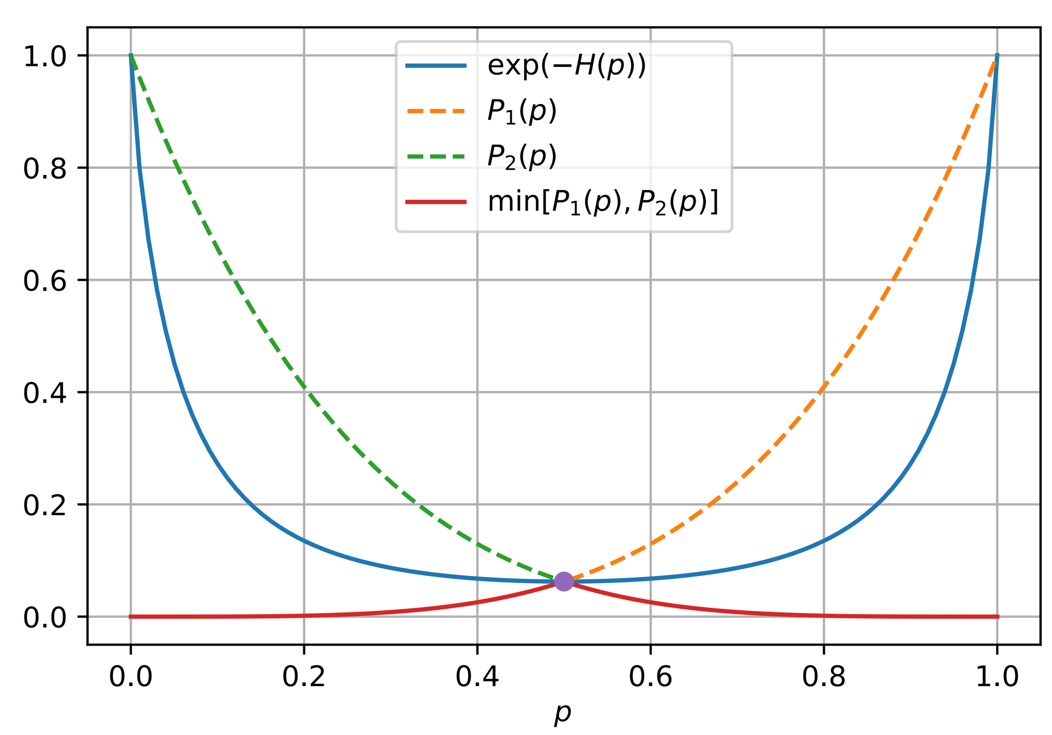

Figure 2. The relation between the maximum entropy matrix and the max-min probability for the simple case $r = c = (1, 1)$.

Sequential Algorithm [1]

1. **Compute $Z = Z(r, c; W)$.**
2. **Select and edge (i, j) and simulate $m_{ij} \sim \text{Ber}(z_{ij})$.**
3. **Update r, c, and W as follows:**
   - **If $m_{ij} = 0$ then $w_{ij} \leftarrow 0$.**
   - **If $m_{ij} = 1$ then $r_i \leftarrow r_i - 1$, $c_j \leftarrow c_j - 1$, and $w_{ij} \leftarrow 0$.**
4. **Repeat (1-3) for the rest of the edges.**

(1)

$$M = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \quad r = (1, 1, 2) \quad c = (1, 1, 2) \quad W = \begin{pmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{pmatrix}$$

$$Z = \begin{pmatrix} \mathbf{0.22} & 0.22 & 0.57 \\ 0.22 & 0.22 & 0.57 \\ 0.57 & 0.57 & 0.86 \end{pmatrix} \quad p_M = 1$$

(2)

$$M = \begin{pmatrix} \mathbf{0} & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \quad r = (1, 1, 2) \quad c = (1, 1, 2) \quad W = \begin{pmatrix} 0 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{pmatrix}$$

$$Z = \begin{pmatrix} 0 & \mathbf{0.32} & 0.68 \\ 0.32 & 0.18 & 0.5 \\ 0.68 & 0.5 & 0.82 \end{pmatrix} \quad p_M = 1 - 0.22$$

(3)

$$M = \begin{pmatrix} \mathbf{0} & \mathbf{1} & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \quad r = (0, 1, 2) \quad c = (1, 0, 2) \quad W = \begin{pmatrix} 0 & 0 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{pmatrix}$$

$$Z = \begin{pmatrix} 0 & 0 & \mathbf{0} \\ 0 & 0 & 1 \\ 1 & 0 & 1 \end{pmatrix} \quad p_M = (1 - 0.22) \times 0.32$$

(4)

$$M = \begin{pmatrix} \mathbf{0} & \mathbf{1} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{1} \\ \mathbf{1} & \mathbf{0} & \mathbf{1} \end{pmatrix} \quad r = (0, 0, 0) \quad c = (0, 0, 0) \quad W = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

$$Z = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \quad p_M = (1 - 0.22) \times 0.32 \approx 0.25$$

Figure 3. Illustration of the algorithm when $r = c = (1, 1, 2)$.

Properties of the Algorithm

- **If the target space is non-empty, the algorithm always terminates with a graph in the target space.**
- **Given a graph in the target space, the algorithm will generate it with positive probability.**
- **The algorithm does not sample uniformly. Importance sampling may be used instead.**

Current Research

- **Fast update of the maximum entropy matrix $Z = Z(r, c; W)$.**
- **Importance sampling weight distribution and its dependency on the edge order.**

References

[1] Glasserman P., Lelo de Larrea E., *Simulation of Bipartite or Directed Graphs with Prescribed Degree Sequences Using Maximum Entropy Probabilities* (2018).

[2] Glasserman P., Young H. P., *Contagion in Financial Networks* (2016).

[3] Squartini T., Almog A., Caldarelli G., van Lelyveld I., Garlaschelli D., and Cimini G., *Enhanced Capital-Asset Pricing Model for the Reconstruction of Bipartite Financial Networks* (2017).