Lecture 5

Xuanxi Zhang

1 Picture 1 and 2: Graphical Models and the Hammersley–Clifford Theorem

Dependency Graph Definition

- We have a probability distribution $p(x_1, x_2, \ldots, x_n)$ over n variables in some set X^n.
- We draw a graph G on the vertex set $\{1, \ldots, n\}$. We place an edge between i and j if and only if x_i is *not* independent of x_j when conditioned on all remaining variables.

$$x_i \not\perp x_j \;\Big|\; \{x_k : k \neq i, j\} \quad \Longleftrightarrow \quad \text{edge } (i, j) \text{ in } G.$$

- The distribution p satisfies the (pairwise) Markov property w.r.t. this graph if whenever i and j are *not* connected by an edge, then

$$x_i \perp x_j \;\Big|\; \{x_k : k \neq i, j\}.$$

Hammersley–Clifford Theorem (Informal Statement)

- Assume $p(x) > 0$ for all configurations x.
- Then p satisfies the Markov property w.r.t. a graph G if and only if there exist *potential functions* $\Psi_C(x_C)$ (one for each *clique* C of G) such that

$$p(x) \;=\; \prod_C \Psi_C(x_C),$$

  where x_C is the sub-configuration $\{x_i : i \in C\}$.

2 Picture 3: Example of a Factorization

- Consider a small graph of five variables x_1, x_2, x_3, x_4, x_5.
- Suppose the edges yield cliques: $\{1, 2, 3\}, \{2, 3, 4\}, \{4, 5\}$.
- The factorization of p then is:

$$p(x_1, x_2, x_3, x_4, x_5) \;=\; \Psi_{123}(x_1, x_2, x_3) \;\times\; \Psi_{234}(x_2, x_3, x_4) \;\times\; \Psi_{45}(x_4, x_5).$$

- Example of a conditional independence implied by the absence of edges:

$$x_2 \perp x_5 \;\Big|\; (x_1, x_3, x_4).$$

3 Picture 4: Proof Sketch of Hammersley–Clifford

1. **Factorization $\implies$ Markov property:** If $p(x) = \prod_C \Psi_C(x_C)$, one can check directly that if i and j are not in the same clique, then x_i and x_j become conditionally independent given the rest.
2. **Markov property $\implies$ Factorization:**
   - One shows that

   $$\log p(x) \;=\; \sum_{A \subseteq \{1, \ldots, n\}} \Psi_A(x_A),$$

     for some set of functions Ψ_A.
   - Then by exploiting the conditional independence structure, one shows $\Psi_A = 0$ whenever A is not contained in a clique. So only clique-based terms remain.

4 Picture 5 and 6: Interacting Particles, Ising Model, and Boltzmann Distribution

Interacting Particles in Statistical Physics

- We have n particles (or sites), each in a discrete state space X.
- They interact via an *energy function* $E(x_1, \ldots, x_n)$, often written as a sum of pairwise or multi-wise potentials:
$$E(x) \;=\; \sum_{i,j} V(x_i, x_j) \quad \text{(for example)}.$$

Ising Model (Example)

- Each site i has a spin $x_i \in \{+1, -1\}$.
- A common Hamiltonian:
$$E(x) \;=\; -\sum_{\langle i,j \rangle} x_i\, x_j,$$
summing over neighboring pairs $\langle i, j \rangle$.
- Ground states ($T \to 0$) are all spins up or all spins down.

Boltzmann (Gibbs) Distribution

- At temperature T, the probability of configuration x is:
$$p(x) \;=\; \frac{1}{Z}\, e^{-\beta\, E(x)} \quad \text{where} \quad \beta = \frac{1}{T}.$$
- The *partition function* is $Z = \sum_x e^{-\beta E(x)}$.
- As $T \to 0$, p concentrates on the configurations that minimize E (the ground states).

5 Picture 7 and 8: Typical Configurations and Phase Transitions

Typical Energies

- For large n, the number of configurations of a given energy can be huge, and that entropy factor can outweigh the Boltzmann factor for certain energy levels.
- In the thermodynamic limit $n \to \infty$, the system's energy distribution becomes sharply peaked around an *average* (or typical) energy.

Phase Transitions

- As β (inverse temperature) crosses a critical value β_c, we may observe a discontinuous change in certain order parameters (e.g., magnetization, energy).
- Example: In the Ising model, below a critical temperature T_c, the system "chooses" a magnetized state (mostly + or mostly –). Above T_c, it becomes disordered.

6 Organized Summary

- **Graphical Models and Hammersley–Clifford:** Distributions with conditional independences factorize over cliques of a graph.
- **Interacting Particles / Ising:** Systems in statistical physics can be seen as Markov random fields with an energy function.
- **Boltzmann Distribution:** $p(x) \propto e^{-\beta E(x)}$. As temperature goes to zero, the distribution concentrates on ground states.
- **Phase Transitions:** In large systems, typical states concentrate in energy "bands." Sudden changes (discontinuities) in observables at critical temperatures are phase transitions.