

Computing Lab

Song Liu (song.liu@Bristol.ac.uk)

Requirement

- To do today's computing lab, you first need a working binary **logistic regression classifier**.

Factorization to Dependency

- Recall a **graph encoding the factorization** of a probability distribution, **encodes the dependency information** between R.V.s in such a probability distribution.
- Today, we try to discover the dependency between random variable by learning the factorization from a dataset.

US Senate

- Senate is a legislative branch of US government.
- Senators are (usually) affiliated to **Democratic** or **Republican** party.
- There are **100** senators, **2** per state.
- Our dataset contains rollcall votes of the 109th US senate.
 - Rollcall: senators say “yay” or “nay” when names were called.
- **Target:**

Analyze **dependencies between senators** using rollcall data.

Graphical Modelling

- Let $x^{(i)}$ be a vote that a senator i casts on a bill.
- $x^{(i)}$ can take two values: $\{1, -1\}$.
 - “Yay”: 1, “Nay”: -1
- Let us model $p(x^{(1)} \dots x^{(100)})$ as a Markov Net:
- $p(x^{(1)} \dots x^{(100)}) \propto \prod_{c \in \mathcal{C}} g_c(x^{(c)})$
- Assume p is a **pairwise Markov network**
 - Which means, all g_c can be further factorized over edges in clique c : $g_c = \prod_{(u,v) \in c} g_{u,v}(x^{(u)}, x^{(v)})$.

Graphical Modelling

- It means $p(x^{(1)} \dots x^{(100)}) \propto \prod_{u,v \in E} g'_{u,v}(x^{(u)}, x^{(v)})$
- For some unknown $g'_{u,v}$ function.
- Model $g'_{u,v} := \exp(w_{u,v} x^{(u)} x^{(v)})$ for all u, v , introducing an unknown parameter $w_{u,v}$. If $w_{u,v} = 0$, there is no edge $(X^{(u)}, X^{(v)})$ in the factorization graph G .
- **Target: Estimate $w_{u,v}$ from data**

Intractable Normalization

- Probability mass function parameterized by $w_{u,v}$ can be written as:
- $p(x^{(1)} \dots x^{(100)} | \mathbf{W}) = \frac{1}{Z(w_{u,v, \forall u,v})} \exp(\sum_{u,v, \forall u,v} w_{u,v} x^{(u)} x^{(v)})$
- $Z(w_{u,v, \forall u,v}) = \sum_{x \in \{-1,1\}^{100}} \exp(\sum_{u,v, \forall u,v} w_{u,v} x^{(u)} x^{(v)})$
- $Z(w_{u,v, \forall u,v})$ is **not** computationally tractable (why?)

Conditional Graphical Modelling

- Luckily, we can use conditional graphical model to reduce the complexity of the probability mass function model:
- 1. Write down $p(x^{(1)} | x^{(2)} \dots x^{(100)}, w_{1,v, \forall v})$ given the joint probability model $p(x^{(1)} \dots x^{(100)} | \mathbf{W})$.
- 2. Write down the likelihood function given dataset:
 $D := \{\mathbf{x}_1, \mathbf{x}_2 \dots \mathbf{x}_n\}$, where $\mathbf{x}_i \in R^{100}$ is the votes of 100 senators on the bill i . Assume D contains IID observations.
- 3. Maximize the above likelihood w.r.t. $w_{1,v, \forall v}$
 - Hint: use your logistic regressor.

Partisanship vs. Bipartisanship

- Try to find the edges in G connected to the following vertices:
- $x^{(v)} = \text{"Barack H. Obama"}$
- $x^{(v)} = \text{"Hillary R. Clinton"}$
- $x^{(v)} = \text{"John S. McCain"}$
- Who has the most bipartisan link?