

STA 772: Probabilistic Graphical Models (Factorization of Joint Probability Distribution (JPD))

Professor O. E. Olubusoye

2023-11-15

Assignment on Factorization of JPD

Consider a simplified economic model to analyze factors influencing a person's decision to invest in the stock market. The model involves four random variables:

- E : Economic conditions (e.g., recession, boom)
- I : Personal income level (e.g., high, low)
- Edu : Education level (e.g., college degree, no degree)
- S : Decision to invest in the stock market (e.g., yes, no)

The joint probability distribution $P(E, I, Edu, S)$ represents the relationships between these variables. Assume there are two levels for each variable.

1. Write down the factorization of the joint probability distribution $P(E, I, Edu, S)$ based on the following conditional dependencies:

- E influences S .
- I influences S .
- Edu influences S .

2. Calculate the total number of probability values required to specify the factorization you wrote in part (a).

Solution

I. The factorization of the joint probability distribution $P(E, I, Edu, S)$ based on the given conditional dependencies is as follows:

$$P(E, I, Edu, S) = P(E) \cdot P(I) \cdot P(Edu) \cdot P(S|E, I, Edu)$$

In this factorization:

- $P(E)$, $P(I)$, and $P(Edu)$ represent the probability distributions of economic conditions, personal income level, and education level, respectively.
- $P(S|E, I, Edu)$ represents the conditional probability of the decision to invest in the stock market (S) given the values of economic conditions (E), personal income level (I), and education level (Edu). This conditional probability captures the dependencies between these variables.

Solution cont'd

2. To calculate the total number of probability values required for the factorization, we consider the number of values needed for each probability distribution:

- $P(E)$, $P(I)$, and $P(Edu)$ each require 2 probability values (one for each level).
- $P(S|E, I, Edu)$ requires $2^3 = 8$ probability values (since there are two levels for each of the three influencing variables: E , I , and Edu).

So, the total number of probability values required for the factorization is:

$$2 + 2 + 2 + 8 = 14$$

Therefore, 14 probability values are needed to specify the factorization of the joint probability distribution $P(E, I, Edu, S)$ based on the given conditional dependencies in this economic model.