Gaussian conditional independence models

"Algebraic & Geometric Methods in Statistics"

Sonja Petrović
Created for Math/Stat 561

Jan 25, 2023.

Objective

Understand how to translate conditional independence statements into polynomials, and what these polynomials mean.

- Previous lecture: the discrete case.
- This lecture: the Gaussian case.

Recall conditional independence (CI) from Lectures 3&4:

Definition 4.1.2. Let $A, B, C \subseteq [m]$ be pairwise disjoint. The random vector X_A *is conditionally independent of* X_B *given* X_C if and only if

$$f_{A \cup B|C}(x_A, x_B \mid x_C) = f_{A|C}(x_A \mid x_C) \cdot f_{B|C}(x_B \mid x_C)$$

for all x_A, x_B, and x_C. The notation $X_A \perp\!\!\!\perp X_B | X_C$ is used to denote that the random vector X satisfies the conditional independence statement that X_A is conditionally independent of X_B given X_C. This is often further abbreviated to $A \perp\!\!\!\perp B | C$.

Figure 1: Source: Algebraic statistics, Seth Sullivant, AMS-GSM book

The Gaussian density

A random vector $X = (X_1, \ldots, X_m)$ has a Gaussian (normal) distribution if

$$f(x) = \frac{1}{(2\pi)^{m/2} \det \Sigma^{1/2}} \exp\left(-\frac{1}{2}(x-\mu)^T \Sigma^{-1} (x-\mu)\right)$$

for some $\mu \in \mathbb{R}^m$ and $\Sigma \in PD_m$ a positive definite matrix.

- $P(X \in A) = \int_A f(x)dx$ for any $A \subseteq \mathbb{R}^m$;
- μ is the mean;
- Σ^{-1} is the concentration matrix;
- Σ is the covariance matrix.

Marginals

The marginal X_A is also a Gaussian: $\mu_A = (\mu_i)_{i \in A}$ and $\Sigma_{A \times A} = (\Sigma_{ij})_{i,j \in A}$.

Independence [Proposition 2.4.4 @book]

For disjoint random variables $A, B \subset [m]$, $X_A \perp\!\!\!\perp X_B$ if and only if $\Sigma_{A \times B} = 0$.

Example (independence)

- $X_1 =$ delay of your flight to Chicago,
- $X_2 =$ delay of my flight to Chicago.

With no further information, a reasonable assumption: $X_1 \perp\!\!\!\perp X_2$.

Example (independence)

- X_1 = delay of your flight to Chicago,
- X_2 = delay of my flight to Chicago.

With no further information, a reasonable assumption: $X_1 \perp\!\!\!\perp X_2$.

or not?

Suppose X_3 = random variable on amount of rain of our common arrival day, taking high value.

- X_1 and X_2 are correlated (e.g. both more likely delayed)
- this correlation explained by X_3
- Conditionally on X_3 being large: X_1, X_2 still independent.
- Capture this by dividing by the marginal density if X_3 to get teh conditional

Review: definition of conditional density; Theorem 2.4.2 in the book.

CI for Gaussian random vectors

Proposition [4.1.9. @book]

The CI statement $X_A \perp\!\!\!\perp X_B | X_C$ holds for $X \sim N(\mu, \Sigma)$ if and only if the submatrix $\Sigma_{A\cup C, B\cup C}$ of the covariance matrix Σ has rank $\#C$.

Proof: see book, page 77!

- What is $\Sigma_{A\cup C, B\cup C}$?

$$\Sigma_{A\cup C, B\cup C} = \begin{bmatrix} \Sigma_{A,B} & \Sigma_{A,C} \\ \Sigma_{C,B} & \Sigma_{C,C} \end{bmatrix}.$$

- . . . . How else might we describe this **rank condition** on this submatrix?

Example [4.1.11 @book]

Take two statements: $\{1 \perp\!\!\!\perp 3\}$ and $\{1 \perp\!\!\!\perp 3|2\}$.
What are the explicit rank conditions?

- $\{1 \perp\!\!\!\perp 3\}$:

Example [4.1.11 @book]

Take two statements: $\{1 \perp\!\!\!\perp 3\}$ and $\{1 \perp\!\!\!\perp 3|2\}$.
What are the explicit rank conditions?

- $\{1 \perp\!\!\!\perp 3\}$:
  - $\{1 \perp\!\!\!\perp 3\} = \{1 \perp\!\!\!\perp 3|\emptyset\}$

Example [4.1.11 @book]

Take two statements: $\{1 \perp\!\!\!\perp 3\}$ and $\{1 \perp\!\!\!\perp 3|2\}$.

What are the explicit rank conditions?

- $\{1 \perp\!\!\!\perp 3\}$:
  - $\{1 \perp\!\!\!\perp 3\} = \{1 \perp\!\!\!\perp 3|\emptyset\}$
  - $\Sigma_{A\cup C,B\cup C} = \sigma_{13}$, yes?
  - So this 1×1 matrix has rank...

Example [4.1.11 @book]

Take two statements: $\{1 \perp\!\!\!\perp 3\}$ and $\{1 \perp\!\!\!\perp 3|2\}$.
What are the explicit rank conditions?

- $\{1 \perp\!\!\!\perp 3\}$:
  - $\{1 \perp\!\!\!\perp 3\} = \{1 \perp\!\!\!\perp 3|\emptyset\}$
  - $\Sigma_{A\cup C, B\cup C} = \sigma_{13}$, yes?
  - So this 1×1 matrix has rank... 0.
  - $\sigma_{13} = 0$.

Example [4.1.11 @book]

Take two statements: $\{1 \perp\!\!\!\perp 3\}$ and $\{1 \perp\!\!\!\perp 3|2\}$.
What are the explicit rank conditions?

- $\{1 \perp\!\!\!\perp 3\}$:
  - $\{1 \perp\!\!\!\perp 3\} = \{1 \perp\!\!\!\perp 3|\emptyset\}$
  - $\Sigma_{A\cup C, B\cup C} = \sigma_{13}$, yes?
  - So this 1×1 matrix has rank... 0.
  - $\sigma_{13} = 0$.
- $\{1 \perp\!\!\!\perp 3|2\}$:

Example [4.1.11 @book]

Take two statements: $\{1 \perp\!\!\!\perp 3\}$ and $\{1 \perp\!\!\!\perp 3|2\}$.
What are the explicit rank conditions?

- $\{1 \perp\!\!\!\perp 3\}$:
  - $\{1 \perp\!\!\!\perp 3\} = \{1 \perp\!\!\!\perp 3|\emptyset\}$
  - $\Sigma_{A\cup C, B\cup C} = \sigma_{13}$, yes?
  - So this 1×1 matrix has rank... 0.
  - $\sigma_{13} = 0$.
- $\{1 \perp\!\!\!\perp 3|2\}$:
  - $\Sigma_{A\cup C, B\cup C} = \Sigma_{\{1,2\},\{2,3\}}$

Example [4.1.11 @book]

Take two statements: $\{1 \perp\!\!\!\perp 3\}$ and $\{1 \perp\!\!\!\perp 3|2\}$.

What are the explicit rank conditions?

- $\{1 \perp\!\!\!\perp 3\}$:
  - $\{1 \perp\!\!\!\perp 3\} = \{1 \perp\!\!\!\perp 3|\emptyset\}$
  - $\Sigma_{A\cup C, B\cup C} = \sigma_{13}$, yes?
  - So this 1×1 matrix has rank… 0.
  - $\sigma_{13} = 0$.
- $\{1 \perp\!\!\!\perp 3|2\}$:
  - $\Sigma_{A\cup C, B\cup C} = \Sigma_{\{1,2\},\{2,3\}}$
  - so the ? $\times$? matrix has rank ?. *[Write matrix on board.]*
  - And is there a polynomial equation for this?

The model $\{1 \perp\!\!\!\perp 3\}$ and $\{1 \perp\!\!\!\perp 3|2\}$ is defined by two polynomial equations in the entries of the 3×3 covariance matrix Σ. These two equations generate the conditional independence ideal for this model.

- Understanding this system of two equations:

$\sigma_{13} = 0$ and $\sigma_{13}\sigma_{22} - \sigma_{12}\sigma_{23} = 0$

is equivalent to this system:

$\sigma_{13} = 0$ **and** $\sigma_{12}\sigma_{23} = 0$

which is equivalent to the **union** ("or") of two linear spaces:

$L_1 = \{\Sigma : \sigma_{13} = \sigma_{12} = 0\}$, $L_2 = \{\Sigma : \sigma_{13} = \sigma_{23} = 0\}$.

- Therefore, the **variety** of the CI model defined by two statements $\{1 \perp\!\!\!\perp 3\}$ and $\{1 \perp\!\!\!\perp 3|2\}$ splits into a **union** of these two spaces.
  - Are there CI statements you can write corresponding to each of L_1 and L_2?
  - Check! L_1 characterizes $X_1 \perp\!\!\!\perp (X_2, X_3)$ and L_2 characterizes $X_3 \perp\!\!\!\perp (X_1, X_2)$.

When there is will, there's a way.

Gaussian CI ideal

Definition [4.1.10 @book]

$J_{A \perp\!\!\!\perp B|C}$ is the ideal geneated by the following polynomials in the indeterminates $\sigma_{ij}, 1 \leq i \leq j \leq m$:

$$J_{A \perp\!\!\!\perp B|C} = \left((\#C+1) \times (\#C+1) \text{ minors of } \Sigma_{A \cup C, B \cup C} \right).$$

- As before, a collection of CI statements $\leftrightarrow$ sum of ideals.
- What about the corresponding variety? The Gaussian CI model is: $M_{A \perp\!\!\!\perp B|C} = V(J_{A \perp\!\!\!\perp B|C}) \cap PD_m$.

```
Macaulay2, version 1.18
i1 : loadPackage "GraphicalModels";
i2 : R=gaussianRing 5
o2 = R
o2 : PolynomialRing
i3 : S={{{1},{2},{3,4}}, {{2,3},{1},{5}}}
o3 = {{{1}, {2}, {3, 4}}, {{2, 3}, {1}, {5}}}
o3 : List
i4 : conditionalIndependenceIdeal (R,S) / print;
```

and the output is:

$-s_{1,4}s_{2,4}s_{3,3} + s_{1,4}s_{2,3}s_{3,4} + s_{1,3}s_{2,4}s_{3,4} - s_{1,2}s_{3,4}^2 - s_{1,3}s_{2,3}s_{4,4} + s_{1,2}s_{3,3}s_{4,4}$,

$-s_{1,3}s_{2,5} + s_{1,2}s_{3,5}$, $\quad -s_{1,5}s_{2,5} + s_{1,2}s_{5,5}$,

$-s_{1,5}s_{3,5} + s_{1,3}s_{5,5}$. Task: verify these polynomials are correct!

Outlook

We will be using these CI statements and ideals when we study **causal discovery** algorithms and (discrete and Gaussian) graphical models.

License

This document is created for Math/Stat 561, Spring 2023, at Illinois Tech.

Main content of these slides is from Thomas Kahle's tutorial on Gaussian CI and graphical models. (We have not yet arrived to the part of the book about graphical models!)