Copenhagen interview questions

January 13, 2022

Exercise 1 - (Answer: $\beta = (\frac{1}{3}, \frac{4}{3})$)

a) Consider $\boldsymbol{X} = (X_1, \ldots X_d)^T \in \mathbb{R}^d$ vector of random variables and $Y \in \mathbb{R}$ (the response) a scalar random variable. Assume that all these random variables have mean zero and finite variance.
Let the covariance matrices be

$$\mathrm{Cov}(\boldsymbol{X}, Y) = \begin{bmatrix} \mathrm{Cov}(X_1, Y) \\ \vdots \\ \mathrm{Cov}(X_d, Y) \end{bmatrix} \in \mathbb{R}^d \tag{1}$$

$$\mathrm{Cov}(\boldsymbol{X}) = \begin{bmatrix} \mathrm{Var}(X_1) & \ldots & \mathrm{Cov}(X_1, X_d) \\ \vdots & & \vdots \\ \mathrm{Cov}(X_d, X_1) & \ldots & \mathrm{Var}(X_d) \end{bmatrix} \in \mathbb{R}^{d \times d} \ . \tag{2}$$

Given the zero mean assumption, covariance can be rewritten as follow:

$$\mathrm{Cov}(X_i, X_j) = \mathbb{E}(X_i X_j), \ \ \forall i, j \in [1, d] \tag{3}$$

$$\mathrm{Cov}(X_i, Y) = \mathbb{E}(X_i Y), \ \ \forall i \in [1, d] \ . \tag{4}$$

Finally, let us define the joint probability distribution as $f(Y, \boldsymbol{X})$ and assume variables changing on a continuous range.
In order to find β minimizing $\mathbb{E}([Y - \beta^t \boldsymbol{X}]^2)$, we derive the expectation over β and set the result equals to zero (the critical point is a minimum since we are dealing with a parabola).

$$\frac{\partial \mathbb{E}([Y-\beta^t \boldsymbol{X}]^2)}{\partial \beta} = -2\int_{\boldsymbol{x}}\int_y (y-\beta^T\boldsymbol{x})^T(y-\beta^T\boldsymbol{x})\ f(\boldsymbol{x},y)\ dy\ \boldsymbol{dx} = 0$$

$$\int_{\boldsymbol{x}}\int_y (y-\beta^T\boldsymbol{x})\boldsymbol{x}\ f(\boldsymbol{x},y)\ dy\ \boldsymbol{dx} = 0$$

$$\int_{\boldsymbol{x}}\int_y y\boldsymbol{x}\ f(\boldsymbol{x},y)\ dy\ \boldsymbol{dx} = \beta^T\int_{\boldsymbol{x}} \boldsymbol{x}^T\boldsymbol{x} f(\boldsymbol{x})\ \boldsymbol{dx}$$

$$\mathbb{E}(Y\boldsymbol{X}) = \beta^T\mathbb{E}(\boldsymbol{X}^T\boldsymbol{X})$$

$$\mathrm{Cov}(Y,\boldsymbol{X}) = \beta^T\,\mathrm{Cov}(\boldsymbol{X})$$

$$\beta^T = \mathrm{Cov}(\boldsymbol{X})^{-1}\,\mathrm{Cov}(Y,\boldsymbol{X})\ \ , \tag{5}$$

where we supposed that $\mathrm{Cov}(\boldsymbol{X})$ is an invertible matrix, and we used Equations (4) and (3) to get covariance from expectations.

b) We use the formula found in part a) in order to solve this exercise. From the provided definition of X, Y, Z random variables, we have the following distributions:

$$\begin{aligned} X &\sim N(0,1)\ \ ,\\ Y &\sim N(0,2)\ \ ,\\ Z &\sim N(0,4)\ \ . \end{aligned}$$

Given $\boldsymbol{X} = (X, Z)$ the covariance matrices are:

$$\mathrm{Cov}(\boldsymbol{X}) = \begin{pmatrix} \mathrm{Var}(X) & \mathrm{Cov}(X,Z) \\ \mathrm{Cov}(X,Z) & \mathrm{Var}(Z) \end{pmatrix}\ \ , \tag{6}$$

$$\mathrm{Cov}(\boldsymbol{X},Y) = \begin{pmatrix} \mathrm{Cov}(X,Y) \\ \mathrm{Cov}(Z,Y) \end{pmatrix}\ \ . \tag{7}$$

Let's do the math and find the required covariances using the same intuitions of Equations (3), (4) and using our knowledge on the variances.

$$\mathrm{Cov}(X,Z) = \mathbb{E}(XZ) = -1\ \ ,$$

$$\mathrm{Cov}(X,Y) = \mathbb{E}(XY) = -1 \ ,$$
$$\mathrm{Cov}(Z,Y) = \mathbb{E}(ZY) = 5 \ ,$$

which must be substituted in (6) and (7) leading to

$$\mathrm{Cov}(\boldsymbol{X}) = \begin{pmatrix} 1 & -1 \\ -1 & 4 \end{pmatrix} \ , \tag{8}$$

$$\mathrm{Cov}(\boldsymbol{X}, Y) = \begin{pmatrix} -1 \\ 5 \end{pmatrix} \ . \tag{9}$$

Using Gauss-Jordan method we start from $\big(\mathrm{Cov}(\boldsymbol{X}) \mid \mathbf{I}\big)$ in order to find the inverse matrix

$$\mathrm{Cov}(\boldsymbol{X})^{-1} = \begin{pmatrix} \frac{4}{3} & \frac{1}{3} \\ \frac{1}{3} & \frac{1}{3} \end{pmatrix} \tag{10}$$

Finally we get the value of β minimizing the expectation of the squared error by solving Equation (5):

$$\begin{aligned} \beta &= \mathrm{Cov}(\boldsymbol{X})^{-1}\,\mathrm{Cov}(Y, \boldsymbol{X}) = \\ &= \begin{pmatrix} \frac{4}{3} & \frac{1}{3} \\ \frac{1}{3} & \frac{1}{3} \end{pmatrix} \begin{pmatrix} -1 \\ 5 \end{pmatrix} = \begin{pmatrix} \frac{1}{3} \\ \frac{4}{3} \end{pmatrix} \end{aligned}$$

Exercise 2 - (Answer: 9)

Given the set of vertices $V = \{A, B, C, D\}$, the DAGs satisfying the problem's requirements are constrained as follow:

- No cycles in the graph
- The only d-separated nodes are A and C.

Since we are not conditioning on any variable, then $A \perp\!\!\!\perp C|\emptyset$ in the following cases:

- Node B is a collider relative to the path from A to C and every path is blocked by B
- Node D is a collider relative to the path from A to C and every path is blocked by D

- Both nodes B and D are colliders relative to the path from A to C.

Note the simmetry of the problem with respect to B and D: once we find the DAGs where $A \perp\!\!\!\perp C|\emptyset$ because every path is blocked by B, by replacing B with D we will simply double the number of DAGs satisfying the condition. There are also 3 additional graphs where both B and D are colliders and must be considered in the count.

Following the above reasoning, we can find 9 DAGs where $A \perp\!\!\!\perp C|\emptyset$ and no other group of vertices is d-separated.

Exercise 3

Example 1

$$A = \begin{pmatrix} 0 & 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 1 & 0 \end{pmatrix}$$

```
1: 2, 3, 4
2: 4
3: 4
4:
5: 3, 4
```

Example 2

$$A = \begin{pmatrix} 0 & 1 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

```
1: 2, 3, 4, 5
2:
3: 5
4: 5
5:
```