Why this matters (intuition). In multivariate Gaussians, edges are about *direct* relationships after controlling for all other variables. The conditional law of a single coordinate X_i given the rest X_{-i} reveals which variables directly "enter" the model for X_i. Algebraically, this local behavior is encoded by the ith row/column of the precision matrix $\Omega = \Sigma^{-1}$.

Problem 1 (Neighborhood regression in the Gaussian case). *Let $X = (X_1, \ldots, X_p)^\top \sim \mathcal{N}_p(0, \Sigma)$, and write $\Omega = \Sigma^{-1}$. For $i \in \{1, \ldots, p\}$, denote $-i = \{1, \ldots, p\} \setminus \{i\}$.*

(a) ***Conditional law of one node.*** *Show that $X_i \mid X_{-i}$ is Gaussian with a linear mean in X_{-i}. Express the regression coefficients $\beta_{i\leftarrow -i} \in \mathbb{R}^{p-1}$ and the conditional variance $\sigma^2_{i|-i}$ in terms of Ω. (You may solve this in any valid way; in the sample solution we provide two derivations: (i) by completing the square, and (ii) via Schur complements.)*

(b) ***One neighbor at a time.*** *Fix $j \neq i$. Prove that the coefficient of X_j in the mean of $X_i \mid X_{-i}$ equals zero if and only if $\Omega_{ij} = 0$. Then deduce (Gaussian case) that*

$$\Omega_{ij} = 0 \quad \Longleftrightarrow \quad X_i \perp X_j \;\big|\; X_{-(i,j)}.$$

(c) ***No neighbors $\Rightarrow$ factorization.*** *Suppose every coefficient in the mean of $X_i \mid X_{-i}$ is zero (i.e., $\beta_{i\leftarrow -i} = 0$). Show that the joint density factorizes as $p(x) = p(x_i)\, p(x_{-i})$. Briefly interpret the implication for X_i relative to the rest.*

Sample Solution. ***(a) Conditional law of one node. Two derivations.***

Version (i): Completing the square (no block inverses). *The joint density is $p(x) \propto \exp(-\frac{1}{2} x^\top \Omega x)$. Isolate the terms involving x_i:*

$$x^\top \Omega x \;=\; \Omega_{ii} x_i^2 \;+\; 2x_i \sum_{j\neq i} \Omega_{ij} x_j \;+\; C(x_{-i}),$$

where $C(\cdot)$ is independent of x_i. For scalars $A > 0, B$, recall the identity

$$Ax^2 + 2Bx \;=\; A\Big(x + \tfrac{B}{A}\Big)^2 \;-\; \tfrac{B^2}{A}.$$

With $A = \Omega_{ii}$ and $B = \sum_{j\neq i} \Omega_{ij} x_j$, we obtain

$$X_i \mid X_{-i} \;\sim\; \mathcal{N}\Big(-\tfrac{1}{\Omega_{ii}} \sum_{j\neq i} \Omega_{ij} X_j, \;\; \tfrac{1}{\Omega_{ii}} \Big).$$

Writing the mean as $\beta_{i\leftarrow -i}^\top X_{-i}$ gives

$$\boxed{\beta_{i\leftarrow -i} \;=\; -\frac{\Omega_{-i,i}}{\Omega_{ii}}, \qquad \sigma^2_{i|-i} \;=\; \frac{1}{\Omega_{ii}}}.$$

Version (ii): Schur complements and block inversion (fully defined). *Reorder variables (if needed) so that coordinate i comes first. Partition*

$$\Sigma \;=\; \begin{bmatrix} a & b^\top \\ b & C \end{bmatrix}, \quad \text{with } a = \Sigma_{ii} \in \mathbb{R},\; b = \Sigma_{-i,i} \in \mathbb{R}^{p-1},\; C = \Sigma_{-i,-i} \in \mathbb{R}^{(p-1)\times(p-1)}.$$

Definition (Schur complement). *If C is invertible, the Schur complement of C in Σ is*

$$S \;:=\; a - b^\top C^{-1} b \quad (\textit{a scalar here}).$$

Block inversion formula. *If C is invertible, then*

$$\Sigma^{-1} = \begin{bmatrix} S^{-1} & -\,S^{-1} b^\top C^{-1} \\ -\,C^{-1} b S^{-1} & C^{-1} + C^{-1} b S^{-1} b^\top C^{-1} \end{bmatrix}.$$

Therefore (identifying $\Omega = \Sigma^{-1}$),

$$\Omega_{ii} = S^{-1}, \qquad \Omega_{-i,i} = -\,C^{-1} b\, S^{-1}.$$

For a multivariate normal, the standard conditional mean/variance in covariance form *are*

$$\mathbb{E}[X_i \mid X_{-i}] = \underbrace{b^\top C^{-1}}_{=\ \Sigma_{i,-i}\Sigma_{-i,-i}^{-1}} X_{-i}, \qquad \mathrm{Var}(X_i \mid X_{-i}) = S.$$

Equivalently, in precision form*, by the identities above,*

$$\boxed{\beta_{i\leftarrow -i} = b^\top C^{-1} = \left(C^{-1} b\right)^\top = -\frac{\Omega_{-i,i}}{\Omega_{ii}}, \qquad \sigma^2_{i|-i} = S = \frac{1}{\Omega_{ii}}}.$$

(These expressions agree with Version (i).)

(b) One neighbor at a time; conditional independence. *From part (a),*

$$\beta_{i\leftarrow j} \;=\; -\frac{\Omega_{ij}}{\Omega_{ii}}.$$

Since $\Omega_{ii} > 0$, we have $\beta_{i\leftarrow j} = 0 \iff \Omega_{ij} = 0$. If $\Omega_{ij} = 0$, the conditional density $p(x_i \mid x_{-i})$ does not depend on x_j, hence $p(x_i \mid x_{-i}) = p(x_i \mid x_{-(i,j)})$, which is exactly $X_i \perp X_j \mid X_{-(i,j)}$. Conversely, in a Gaussian model the conditional mean is linear and unique; if $X_i \perp X_j \mid X_{-(i,j)}$, the coefficient of X_j must vanish, implying $\Omega_{ij} = 0$.

(c) No neighbors ⇒ factorization. *If $\beta_{i\leftarrow -i} = 0$, then by (a) $\Omega_{-i,i} = 0$. Thus $x^\top \Omega x = \Omega_{ii} x_i^2 + x_{-i}^\top \Omega_{-i,-i} x_{-i}$ has no cross-terms $x_i x_j$, and*

$$p(x) \;\propto\; \exp\!\Big(-\tfrac{1}{2}\Omega_{ii} x_i^2\Big) \;\cdot\; \exp\!\Big(-\tfrac{1}{2} x_{-i}^\top \Omega_{-i,-i} x_{-i}\Big) \;=\; p(x_i)\, p(x_{-i}).$$

Hence X_i is independent of X_{-i}.