1. Consider a p-dim Gaussian random variable $X \sim N_p(\mu, \Sigma)$. Partition X, μ, Σ, and the precision matrix $\Theta = \Sigma^{-1}$ as the following (suppose X_a, X_b, X_c are all multidimensional):

$$X = \begin{pmatrix} X_a \\ X_b \\ X_c \end{pmatrix}, \mu = \begin{pmatrix} \mu_a \\ \mu_b \\ \mu_c \end{pmatrix}, \Sigma = \begin{pmatrix} \Sigma_{aa} & \Sigma_{ab} & \Sigma_{ac} \\ \Sigma_{ba} & \Sigma_{bb} & \Sigma_{bc} \\ \Sigma_{ca} & \Sigma_{cb} & \Sigma_{cc} \end{pmatrix}, \Theta = \begin{pmatrix} \Theta_{aa} & \Theta_{ab} & \Theta_{ac} \\ \Theta_{ba} & \Theta_{bb} & \Theta_{bc} \\ \Theta_{ca} & \Theta_{cb} & \Theta_{cc} \end{pmatrix}.$$

(a) Show that $X_a \perp X_b$ if and only if $\Sigma_{ab} = \Sigma_{ba} = 0$.

(b) Show that $X_a \perp X_b \mid X_c$ if and only if $\Theta_{ab} = \Theta_{ba} = 0$.

(c) Ex 17.2 in “Elements of Statistical Learning”.

Ex. 17.2 Consider random variables X_1, X_2, X_3, X_4. In each of the following cases draw a graph that has the given independence relations:

(a) $X_1 \perp X_3 | X_2$ and $X_2 \perp X_4 | X_3$.

(b) $X_1 \perp X_4 | X_2, X_3$ and $X_2 \perp X_4 | X_1, X_3$.

(c) $X_1 \perp X_4 | X_2, X_3$, $X_1 \perp X_3 | X_2, X_4$ and $X_3 \perp X_4 | X_1, X_2$.

(a) For Gaussian r.v. X and Y, $Z = [X, Y]$, then $f_Z(z) \propto \exp(-\frac{1}{2}(z-\mu)^T\Sigma^{-1}(z-\mu))$

then $Cov(X,Y) = 0 \Leftrightarrow \Sigma = \begin{pmatrix} \Sigma_{XX} & 0 \\ 0 & \Sigma_{YY} \end{pmatrix} \Leftrightarrow \Sigma^{-1} = \begin{pmatrix} \Sigma_{XX}^{-1} & 0 \\ 0 & \Sigma_{YY}^{-1} \end{pmatrix} \Leftrightarrow (z-\mu)^T\Sigma^{-1}(z-\mu) = \underbrace{(x-\mu_x)^T\Sigma_{XX}^{-1}(x-\mu_x)}_{A} + \underbrace{(y-\mu_y)^T\Sigma_{YY}^{-1}(y-\mu_y)}_{B}$

$\Leftrightarrow f_Z(z) \propto \exp(A+B) = e^A \cdot e^B = f_X(x) \cdot f_Y(y) \Leftrightarrow X \perp Y$

To prove $X_a \perp X_b \Rightarrow \Sigma_{ab} = \Sigma_{ba} = 0$: $X_a \perp X_b \Rightarrow Cov(X_a, X_b) = 0 \Rightarrow \Sigma_{ab} = \Sigma_{ba} = 0$

To prove $\Sigma_{ab} = \Sigma_{ba} = 0 \Rightarrow X_a \perp X_b$: $\Sigma_{ab} = \Sigma_{ba} = 0 \Rightarrow Cov(X_a, X_b) = 0 \Rightarrow X_a \perp X_b$

(b) $f_X(x) \propto \exp(-\frac{1}{2}(x-\mu)^T\Theta(x-\mu))$, let $Y := X - \mu = (y_a, y_b, y_c)^T$, $f_Y(y) \propto \exp(-\frac{1}{2}y^T\Theta y)$

then $y^T\Theta y = y_a^T\Theta_{aa}y_a + y_b^T\Theta_{bb}y_b + y_c^T\Theta_{cc}y_c + 2(y_a^T\Theta_{ab}y_b + y_b^T\Theta_{bc}y_c + y_c^T\Theta_{ca}y_a)$

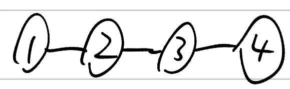
while X_c given, then Y_c given, then $f_Y(y) \propto \exp[-\frac{1}{2}(y_a^T\Theta_{aa}y_a + y_b^T\Theta_{bb}y_b + 2y_a^T\Theta_{ab}y_b + 2y_b^T\Theta_{bc}y_c + 2y_c^T\Theta_{ca}y_a)]$

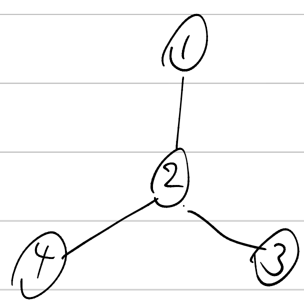
here $y_c^T\Theta_{cc}y_c$ disappear for Y_c given, so only $2y_a^T\Theta_{ab}y_b$ contains both y_a and y_b

so $y_a^T\Theta_{ab}y_b = 0 \Leftrightarrow f_Y(y)$ can be factorized into $g(y_a) \cdot h(y_b) \Leftrightarrow Y_a \perp Y_b | Y_c \Leftrightarrow X_a \perp X_b | X_c$

so $\Theta_{ab} = \Theta_{ba} \Leftrightarrow X_a \perp X_b | X_c$

(c)

(a) 

(c) 

(b)