

In short, choosing  $(i,j) = (1,2)$  w.l.o.g. and the covariance matrix  $\Sigma = \begin{pmatrix} \Sigma_{1,1} & 0 & \vdots \\ 0 & \Sigma_{2,2} & \vdots \\ \dots & \dots & \dots \end{pmatrix}$  means that  $\Sigma_{1,2} = \Sigma_{2,1} = 0$

Therefore using the equation from part b yields

$$\begin{aligned} p(x_1|x_2) &= \mathcal{N}(x_1; \mu_1 + \underbrace{\Sigma_{12}}_0 \cdot \Sigma_{2,2}^{-1} (x_2 - \mu_2), \Sigma_{11} - \Sigma_{12} \cdot \Sigma_{2,2}^{-1} \Sigma_{2,1}) \\ &= \mathcal{N}(x_1; \mu_1 + 0 \cdot \Sigma_{2,2}^{-1} (x_2 - \mu_2), \Sigma_{11} - 0 \cdot \Sigma_{2,2}^{-1} \cdot 0) \\ &= \mathcal{N}(x_1; \mu_1, \Sigma_{11}) \end{aligned}$$

$$p(x_1|x_2) = p(x_1) \rightarrow \text{marginal}$$

and  $x_1$  and  $x_2$  are marginally independent. Since this can be applied to any  $i \neq j$  pair (by permuting), we say that the statement in the first bullet point always holds.

In short, for a precision matrix  $\Lambda = \Sigma^{-1}$  and a zero entry on the  $(i,j) = (1,2)$ th index of the  $\Lambda$  means that  $x_i$  and  $x_j$  are conditionally independent where other variables are  $x_k$  with  $k \neq i, j$  needs to be proven.

Using the equation from part b,

$$\begin{aligned}
 p(x_1 | x_2, x_k \neq x_1, x_2) &= N\left(x_1; \mu_1 - \Lambda_{11}^{-1} \left( \underbrace{\Lambda_{1,2} \dots \Lambda_{1,k}}_0 \right) \begin{pmatrix} x_2 - \mu_2 \\ \vdots \\ x_k - \mu_k \end{pmatrix}, \right. \\
 &\quad \left. \Lambda_{11}^{-1} - \left( \underbrace{\Lambda_{1,2} \dots \Lambda_{1,k}}_0 \right) \begin{pmatrix} \Lambda_{2,2} \Lambda_{2,1} \\ \vdots \\ \Lambda_{k,2} \Lambda_{k,1} \end{pmatrix} \right) \\
 &\quad \swarrow \text{Setting } \Lambda_{1,2} = \Lambda_{2,1} = 0 \\
 &= N\left(x_1; \mu_1 - \Lambda_{11}^{-1} (-\Lambda_{1,k} \dots) \begin{pmatrix} x_k - \mu_k \\ \vdots \end{pmatrix}, \right. \\
 &\quad \left. \Lambda_{11}^{-1} - (\dots \Lambda_{1,k} \dots) \begin{pmatrix} \vdots \\ \Lambda_{k,k} \Lambda_{k,1} \end{pmatrix} \right) \\
 &= p(x_1 | x_k \neq 1, 2)
 \end{aligned}$$

Thus, we have proven that the last term does not involve  $x_2$  and  $x_1$  and  $x_2$  are conditionally independent given  $\{x_k\}$ .

By permuting the indices, this result can also be generalized for arbitrary  $(i,j)$  in the statement.