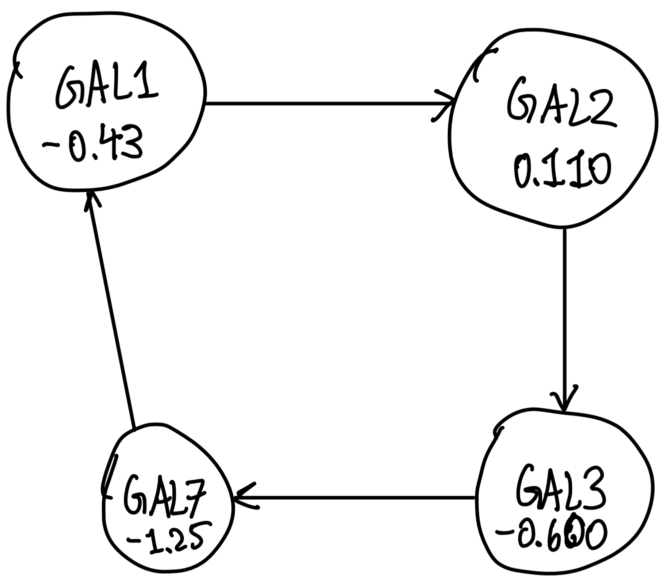
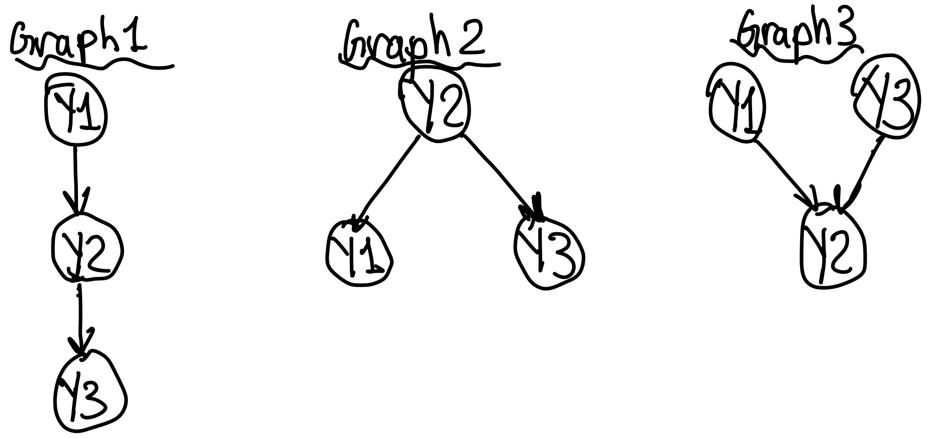
GAL1	Gal2	GAL3	GAL7
-0.43	0.110	-0.600	-1.25
-0.12	-0.510	-0.220	-0.71



1. This problem considers the 3 DAGs talked about during lecture.



a). Write down the regression representation of the Gaussian graphical model for each of the given graphs

In the lecture, we used the factorized joint distribution:

$$p(x) = \prod_{k=1}^{K} p(x_k \mid pa_k)$$

The conditional distribution of x_k given its parents pa_k:

$$x_k \mid pa_k \sim N\Big(\sum_{j \in pa_k} w_{kj} x_j + b_k, v_k\Big), \quad 1 \le k \le K$$

And we rewrote as:

$$x_k = \sum_{j \in pa_k} w_{kj} x_j + b_k + \sqrt{v_k}\,\epsilon_k$$

where $\epsilon_1, \dots, \epsilon_K$ are i.i.d $N(0,1)$

Thus, for Graph 1:

$$Y_1 = b_1 + \sqrt{V_1}\,\epsilon_1$$
$$Y_2 = W_{21} Y_1 + b_2 + \sqrt{V_2}\,\epsilon_2$$
$$Y_3 = W_{32} Y_2 + b_3 + \sqrt{V_3}\,\epsilon_3$$

For Graph2:

$$Y_1 = W_{12} Y_2 + b_1 + \sqrt{V_1}\,\epsilon_1$$
$$Y_2 = b_2 + \sqrt{V_2}\,\epsilon_2$$
$$Y_3 = W_{32} Y_2 + b_3 + \sqrt{V_3}\,\epsilon_3$$

For Graph3:

$$Y_1 = b_1 + \sqrt{V_1}\,\epsilon_1$$
$$Y_2 = W_{21} Y_1 + W_{23} Y_3 + b_2 + \sqrt{V_2}\,\epsilon_2$$
$$Y_3 = b_3 + \sqrt{V_3}\,\epsilon_3$$

b) Write down the covariance matrix of the joint distribution corresponding to each of these three distributions

set Bias = 0:

For Graph 1:

$$Y_1 = b_1 + \sqrt{V_1}\,\epsilon_1$$
$$Y_2 = W_{21} Y_1 + b_2 + \sqrt{V_2}\,\epsilon_2$$
$$Y_3 = W_{32} Y_2 + b_3 + \sqrt{V_3}\,\epsilon_3$$
$$\Downarrow$$

$$Y_1 = \sqrt{V_1}\,\epsilon_1$$

$$-w_2 Y_1 + Y_2 = \sqrt{V_2}\,\epsilon_2$$

$$-w_{32} Y_2 + Y_3 = \sqrt{V_3}\,\epsilon_3$$

In Matrix Form:

$$\begin{bmatrix} 1 & 0 & 0 \\ -w_{21} & 1 & 0 \\ 0 & -w_{32}Y_2 & 1 \end{bmatrix} \begin{bmatrix} Y_1 \\ Y_2 \\ Y_3 \end{bmatrix} = \begin{bmatrix} \sqrt{V_1}\,\epsilon_1 \\ \sqrt{V_2}\,\epsilon_2 \\ \sqrt{V_3}\,\epsilon_3 \end{bmatrix}$$

$$\Downarrow \qquad \Downarrow \qquad \Downarrow$$

$$A * Y = \epsilon$$

$$Var(cX) = c^2 Var(X)$$

$$Y = A^{-1}\epsilon$$

$$cov(Y) = A^{-1} cov(\epsilon) (A^{-1})^T$$

$$cov(Y) = cov\begin{pmatrix} Y_1 \\ Y_2 \\ Y_3 \end{pmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ w_{21} & 1 & 0 \\ (w_{21}w_{32}) & w_{32} & 1 \end{bmatrix} \begin{bmatrix} V_1 & 0 & 0 \\ 0 & V_2 & 0 \\ 0 & 0 & V_3 \end{bmatrix} \begin{bmatrix} 1 & w_{21} & (w_{21}w_{32}) \\ 0 & 1 & w_{32} \\ 0 & 0 & 1 \end{bmatrix}$$

Multiply first two matrices:

$$= \begin{bmatrix} V_1 & 0 & 0 \\ w_{21}V_1 & V_2 & 0 \\ (w_{21}w_{32}V_1) & (w_{32}V_2) & V_3 \end{bmatrix} \begin{bmatrix} 1 & w_{21} & (w_{21}w_{32}) \\ 0 & 1 & w_{32} \\ 0 & 0 & 1 \end{bmatrix}$$

$$= \boxed{\begin{bmatrix} V_1 & V_1 w_{21} & V_1 w_{21} w_{32} \\ w_{21}V_1 & w_{21}^2 V_1 + V_2 & w_{21}^2 V_1 w_{32} + V_2 w_{32} \\ w_{21}w_{32}V_1 & w_{21}^2 w_{32} V_1 + w_{32}V_2 & w_{21}^2 w_{32}^2 V_1 + w_{32}^2 V_2 + V_3 \end{bmatrix}}$$

For Graph2:

Set Bias=0

$$Y_1 = W_{12} Y_2 + b_1 + \sqrt{V_1}\,\epsilon_1 \qquad\qquad Y_1 - W_{12} Y_2 = \sqrt{V_1}\,\epsilon_1$$
$$Y_2 = b_2 + \sqrt{V_2}\,\epsilon_2 \qquad \Longrightarrow \qquad Y_2 = \sqrt{V_2}\,\epsilon_2$$
$$Y_3 = W_{32} Y_2 + b_3 + \sqrt{V_3}\,\epsilon_3 \qquad\qquad -W_{32} Y_2 + Y_3 = \sqrt{V_3}\,\epsilon_3$$

In Matrix Form:

$$\underbrace{\begin{bmatrix} 1 & -w_{12} & 0 \\ 0 & 1 & 0 \\ 0 & -w_{32} & 1 \end{bmatrix}}_{A} \underbrace{\begin{bmatrix} Y_1 \\ Y_2 \\ Y_3 \end{bmatrix}}_{Y} = \underbrace{\begin{bmatrix} \sqrt{V_1}\,\epsilon_1 \\ \sqrt{V_2}\,\epsilon_2 \\ \sqrt{V_3}\,\epsilon_3 \end{bmatrix}}_{\epsilon}$$

$$AY = \epsilon$$
$$Y = A^{-1}\epsilon$$
$$Cov(Y) = A^{-1} Cov(\epsilon) (A^{-1})^T$$

$$Cov(Y) = \begin{bmatrix} 1 & W_{12} & 0 \\ 0 & 1 & 0 \\ 0 & W_{32} & 1 \end{bmatrix} \begin{bmatrix} V_1 & 0 & 0 \\ 0 & V_2 & 0 \\ 0 & 0 & V_3 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ W_{12} & 1 & W_{32} \\ 0 & 0 & 1 \end{bmatrix}$$

$$= \begin{bmatrix} V_1 & W_{12}V_2 & 0 \\ 0 & V_2 & 0 \\ 0 & W_{32}V_2 & V_3 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ W_{12} & 1 & W_{32} \\ 0 & 0 & 1 \end{bmatrix}$$

$$= \boxed{\begin{bmatrix} V_1 + W_{12}^2 V_2 & W_{12} V_2^2 & W_{12} V_2 W_{32} \\ V_2 W_{12} & V_2 & V_2 W_{32} \\ W_{32} V_2 W_{12} & W_{32} V_2 & W_{32}^2 V_2 + V_3 \end{bmatrix}}$$

Cov

| by

For Graph 3:

$Y_1 = b_1 + \sqrt{V_1}\,\epsilon_1$

$Y_2 = W_{21} Y_1 + W_{23} Y_3 + b_2 + \sqrt{V_2}\,\epsilon_2$

$Y_3 = b_3 + \sqrt{V_3}\,\epsilon_3$

Set bias $= 0$

$$\begin{matrix} Y_1 & & & = \sqrt{V_1}\,\epsilon_1 \\ -W_{21} Y_1 & Y_2 & W_{23} Y_3 & = \sqrt{V_2}\,\epsilon_2 \\ & & Y_3 & = \sqrt{V_3}\,\epsilon_3 \end{matrix}$$

In Matrix Form:

$$\begin{bmatrix} 1 & 0 & 0 \\ -W_{21} & 1 & W_{23} \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} Y_1 \\ Y_2 \\ Y_3 \end{bmatrix} = \begin{bmatrix} \sqrt{V_1}\,\epsilon_1 \\ \sqrt{V_2}\,\epsilon_2 \\ \sqrt{V_3}\,\epsilon_3 \end{bmatrix}$$

$\Downarrow$

$AY = \epsilon$

$\mathrm{Cov}(Y) = A^{-1}\,\mathrm{Cov}(\epsilon)\,(A^{-1})^T$

$$\mathrm{Cov}(Y) = \begin{bmatrix} 1 & 0 & 0 \\ W_{21} & 1 & W_{23} \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} V_1 & 0 & 0 \\ 0 & V_2 & 0 \\ 0 & 0 & V_3 \end{bmatrix} \begin{bmatrix} 1 & W_{21} & 0 \\ 0 & 1 & 0 \\ 0 & W_{23} & 1 \end{bmatrix}$$

$$= \begin{bmatrix} V_1 & 0 & 0 \\ W_{21}V_1 & V_2 & W_{23}V_3 \\ 0 & 0 & V_3 \end{bmatrix} \begin{bmatrix} 1 & W_{21} & 0 \\ 0 & 1 & 0 \\ 0 & W_{23} & 1 \end{bmatrix}$$

$$= \boxed{\begin{bmatrix} V_1 & V_1 W_{21} & 0 \\ W_{21}V_1 & W_{21}^2 V_1 + V_2 + W_{23}^2 V_3 & W_{23}V_3 \\ 0 & V_3 W_{23} & V_3 \end{bmatrix}}$$

?c). Are any of these graphical models equivalent? (Equivalent if any Matrix under one model can also be generated from the other model selecting suitable parameters and vice versa]. Prove your finding

Let's try Graph1 $\Leftrightarrow$ Graph2 :

- $V_{(G1)1} = V_{(G2)1} + V_{(G2)2} W_{(G2)12}^2$
- $V_{(G1)2} = V_{(G2)} - V_{(G1)1} W_{(G1)21}^2$
- $V_{(G1)3} = V_{(G2)3} + V_{(G2)2} W_{(G2)32}^2 - V_{(G1)1} W_{(G1)21}^2 W_{(G1)32}^2 - V_{(G1)2} W_{(G1)32}^2$
- $W_{(G1)32} = \dfrac{V_{(G2)2} W_{(G2)12} W_{(G2)32}}{V_{(G1)1} W_{(G1)21}}$
- $W_{(G1)21} = \dfrac{V_{(G2)2} W_{(G2)12}}{V_{(G1)1}}$

The number of equations is less than the number of unknowns so we cannot prove that Graph1 $\Leftrightarrow$ Graph2.

Lets try Graph1 $\Leftrightarrow$ Graph3 :

- $V_{(G1)1} = V_{(G3)1}$
- $V_{(G1)2} = V_{(G3)1} W_{(G3)21}^2 + V_{(G3)2} + V_{(G3)3} W^2_{(G3)} - V_{(G1)1} W^2_{(G1)}$
- $V_{(G1)3} = V_{(G3)3} - V_{(G1)1} W_{(G1)21}^2 W_{(G1)32}^2 - V_{(G1)2} W_{(G1)32}^2$
- $W_{(G1)21} = \dfrac{V_{(G3)2} W_{(G3)12}}{V_{(G1)1}}$
- $W_{(G1)32} = \dfrac{V_{(G3)2} W_{(G3)32} - V_{(G1)1} W_{(G1)21}^2 W_{(G1)32}}{V_{(G1)2}}$

The same problem emerges as when we tried to prove Graph1 $\Leftrightarrow$ Graph2. Thus Graph1 $\Leftrightarrow$ Graph3

Finally let's try Graph2 $\Longleftrightarrow$ Graph3

- $V_{(G2)1} = V_{(G3)1} - V_{(G2)2}\, W^2_{(G2)12}$
- $V_{(G2)2} = \dfrac{V_{(G3)1}\, W_{(G3)21}}{W_{(G2)12}}$
- $V_{(G2)3} = V_{(G3)3} - V_{(G2)2}\, W^2_{(G2)32}$
- $W_{(G2)12} = \dfrac{V_{(G3)1}\, W_{(G3)21}}{V_{(G2)2}}$
- $W_{(G2)32} = \dfrac{W_{(G3)23}\, V_{(G3)3}}{V_{(G2)2}}$

The number of equations is equal to the number of unknowns, so there exists a unique combination of the variables such that Graph 2 can be transformed into Graph 3 and vice versa.

4. Let $x=(x_1,\dots,x_p)'$ be a random vector with multivariate normal distribution $N(\mu,\Sigma)$. Partition x into two subvectors x_1 and x_2, and partition the mean vector and covariance matrix accordingly as:

$$x=\begin{pmatrix}x_1\\x_2\end{pmatrix},\quad \mu=\begin{pmatrix}\mu_1\\\mu_2\end{pmatrix},\quad \Sigma=\begin{pmatrix}\Sigma_{11} & \Sigma_{12}\\ \Sigma_{21} & \Sigma_{22}\end{pmatrix}$$

a). Define $z_2=x_2-[\mu_2+\Sigma_{21}\Sigma_{11}^{-1}(x_1-\mu_1)]$ Show that $E[z_2]=0$

$$E[z_2]=E[x_2-[\mu_2+\Sigma_{21}\Sigma_{11}^{-1}(x_1-\mu_1)]]$$

$$=E[x_2]-E[\mu_2]-\Sigma_{21}\Sigma_{11}^{-1}(E[x_1]-E[\mu_1])$$

$$=\underbrace{\cancel{\mu_2-\mu_2}}_{0}-\underbrace{\cancel{\Sigma_{21}\Sigma_{11}^{-1}(\mu_1-\mu_1)}}_{0}$$

$$=0$$

Thus, we have proven that $E[z_2]=0$

b). Show that $cov(x_1, z_2) = 0$ and $Var(z_2) = \Sigma_{22} - \Sigma_{21}\Sigma_{11}^{-1}\Sigma_{12}$

$$Cov(X_1, Z_2) = Cov(X_1, X_2 - [\mu_2 + \Sigma_{21}\Sigma_{11}^{-1}(x_1 - \mu_1)])$$

$$= Cov(X_1, X_2) - \cancelto{0}{Cov(X_1, \mu_2)} - Cov(X_1, \Sigma_{21}\Sigma_{11}^{-1}(X_1 - \mu_1))$$

(From Matrix) (Distrib Covariance)

$$= \Sigma_{12} - \Sigma_{21}\Sigma_{11}^{-1}(Cov(X_1, X_1) - \cancelto{0}{Cov(X_1, \mu_1)})$$

(Variance)

$$= \Sigma_{12} - \Sigma_{21}\Sigma_{11}^{-1}(\Sigma_{11})$$

$$= \Sigma_{12} - \Sigma_{21} I$$

$$= \Sigma_{12} - \Sigma_{21}$$

$$= 0$$

$$Var(Z_2) = Var[X_2 - [\mu_2 + \Sigma_{21}\Sigma_{11}^{-1}(x_1 - \mu_1)]]$$

$$= Var[X_2] - \cancelto{0}{Var[\mu_2]} - \Sigma_{21}\Sigma_{11}^{-1} Var(X_1 - \mu_1)(\Sigma_{21}\Sigma_{11}^{-1})^T$$

$$= Var[X_2] - \Sigma_{21}\Sigma_{11}^{-1} Var[X_1] (\Sigma_{21}\Sigma_{11}^{-1})^T$$

$$= \Sigma_{22} - \Sigma_{21}\Sigma_{11}^{-1}\Sigma_{11}(\Sigma_{21}\Sigma_{11}^{-1})^T$$

$$= \Sigma_{22} - \Sigma_{21} I (\Sigma_{21}\Sigma_{11}^{-1})^T$$

Thus, $Var(z_2) = \Sigma_{22} - \Sigma_{21}\Sigma_{11}^{-1}\Sigma_{12}$

c). Conclude from part (a) and (b) that

$$E(X_2|X_1) = \mu_2 + \Sigma_{21}\Sigma_{11}^{-1}(X_1 - \mu_1)$$

$$Var(X_2|X_1) = \Sigma_{22} - \Sigma_{21}\Sigma_{11}^{-1}\Sigma_{12}$$

$$Z_2 = X_2 - [\mu_2 + \Sigma_{21}\Sigma_{11}^{-1}(X_1 - \mu_1)]$$

$$X_2 = Z_2 + [\mu_2 + \Sigma_{21}\Sigma_{11}^{-1}(X_1 - \mu_1)]$$

$$E[X_2|X_1] = E(Z_2 + [\mu_2 + \Sigma_{21}\Sigma_{11}^{-1}(X_1 - \mu_1)] \,|\, X_1)$$

$$= \cancel{E(Z_2|X_1)}^{0} + E(\mu_2|X_1) + E(\Sigma_{21}\Sigma_{11}^{-1}(X_1 - \mu_1)|X_1)$$

$$= \mu_2 + \Sigma_{21}\Sigma_{11}^{-1}(E(X_1|X_1) - E(\mu_1|X_1))$$

$$= \mu_2 + \Sigma_{21}\Sigma_{11}^{-1}(X_1 - \mu_1)$$

$$Var(X_2|X_1) = Var(Z_2 + [\mu_2 + \Sigma_{21}\Sigma_{11}^{-1}(X_1 - \mu_1)] \,|\, X_1)$$

$$= Var(Z_2|X_1) + 0$$

$$= Var(Z_2)$$

$$= \Sigma_{22} - \Sigma_{21}\Sigma_{11}^{-1}\Sigma_{12} \quad \text{(From Part b)}$$

d). Let $\mu = 0$

From Part (b) we know:

$$\mathrm{cov}\begin{pmatrix} X_1 \\ Z_2 \end{pmatrix} = \begin{pmatrix} \Sigma_{11} & 0 \\ 0 & \Sigma_{22} - \Sigma_{21}\Sigma_{11}^{-1}\Sigma_{12} \end{pmatrix}$$

$$\begin{pmatrix} X_1 \\ X_2 \end{pmatrix} = \begin{pmatrix} I & 0 \\ \Sigma_{21}\Sigma_{11}^{-1} & I \end{pmatrix}\begin{pmatrix} X_1 \\ X_2 \end{pmatrix}$$

$$\Sigma = \underbrace{\begin{pmatrix} I & 0 \\ \Sigma_{21}\Sigma_{11}^{-1} & I \end{pmatrix}}_{M_1}\underbrace{\begin{pmatrix} \Sigma_{11} & 0 \\ 0 & \Sigma_{22} - \Sigma_{21}\Sigma_{11}^{-1}\Sigma_{12} \end{pmatrix}}_{M_2}\underbrace{\begin{pmatrix} I & \Sigma_{11}^{-1}\Sigma_{21} \\ 0 & I \end{pmatrix}}_{M_3}$$

$$(M_1 M_2 M_3)^{-1} = M_3^{-1} M_2^{-1} M_1^{-1}$$

$$\Sigma^{-1} = \begin{pmatrix} I & \Sigma_{11}^{-1}\Sigma_{21} \\ 0 & I \end{pmatrix}^{-1}\begin{pmatrix} \Sigma_{11} & 0 \\ 0 & \Sigma_{22} - \Sigma_{21}\Sigma_{11}^{-1}\Sigma_{12} \end{pmatrix}^{-1}\begin{pmatrix} I & 0 \\ \Sigma_{21}\Sigma_{11}^{-1} & I \end{pmatrix}^{-1}$$

$$= \begin{pmatrix} I & -\Sigma_{11}^{-1}\Sigma_{21} \\ 0 & I \end{pmatrix}\begin{pmatrix} \Sigma_{11}^{-1} & 0 \\ 0 & (\Sigma_{22} - \Sigma_{21}\Sigma_{11}^{-1}\Sigma_{12})^{-1} \end{pmatrix}\begin{pmatrix} I & 0 \\ -\Sigma_{21}\Sigma_{11}^{-1} & I \end{pmatrix}$$

e). Let $\theta = \Sigma^{-1}$ be the Precision Matrix, and denote its elements by θ_{jk}. Show that for any two components X_j and X_k of X, it holds that

$$\mathrm{Cor}(X_j, X_k \mid X \setminus \{X_j, X_k\}) = \frac{-\theta_{jk}}{\sqrt{\theta_{jj}\theta_{kk}}}$$

Also known as Partial Correlation of X_j and X_k.
Tells us whether X_j and X_k are conditional independent given the rest of X depends on whether θ_{jk} is 0.

We know $X \sim N\left(\begin{pmatrix}\mu_1\\ \mu_2\end{pmatrix}, \begin{pmatrix}\Sigma_{11} & \Sigma_{12}\\ \Sigma_{21} & \Sigma_{22}\end{pmatrix}\right)$

$$\mu_2 = \begin{pmatrix}\mu_j\\ \mu_k\end{pmatrix} \quad \text{and} \quad \Sigma_{22}\begin{pmatrix}\mathrm{Var}(X_j) & \mathrm{Cov}(X_j, X_k)\\ \mathrm{Cov}(X_k, X_j) & \mathrm{Var}(X_k)\end{pmatrix}$$

We have shown $\mathrm{Var}(X_2 \mid X_1) = \Sigma_{22} - \Sigma_{21}\Sigma_{11}^{-1}\Sigma_{12} = \Sigma_{jk}^{-1}$

We have also shown

$$\Sigma^{-1} = \begin{pmatrix} I & -\Sigma_{11}^{-1}\Sigma_{12}\\ 0 & I\end{pmatrix}\begin{pmatrix}\Sigma_{11}^{-1} & 0\\ 0 & (\Sigma_{22} - \Sigma_{21}\Sigma_{11}^{-1}\Sigma_{12})^{-1}\end{pmatrix}\begin{pmatrix} I & 0\\ -\Sigma_{21}\Sigma_{11}^{-1} & I\end{pmatrix}$$

$$= \begin{pmatrix}\Sigma_{11}^{-1} & \left(-\Sigma_{11}^{-1}\Sigma_{12}(\Sigma_{22} - \Sigma_{21}\Sigma_{11}^{-1}\Sigma_{12})^{-1}\right)\\ 0 & (\Sigma_{22} - \Sigma_{21}\Sigma_{11}^{-1}\Sigma_{12})^{-1}\end{pmatrix}\begin{pmatrix} I & 0\\ -\Sigma_{21}\Sigma_{11}^{-1} & I\end{pmatrix}$$

$$\begin{pmatrix} \Sigma_{11}^{-1} + \Sigma_{11}^{-1}\Sigma_{12}\left(\Sigma_{22} - \Sigma_{21}\Sigma_{11}^{-1}\Sigma_{12}^{-1}\right)^{-1}\Sigma_{21}\Sigma_{11}^{-1} & -\Sigma_{11}^{-1}\Sigma_{12}\left(\Sigma_{22} - \Sigma_{21}\Sigma_{11}^{-1}\Sigma_{12}\right)^{-1} \\ \left(\Sigma_{22} - \Sigma_{21}\Sigma_{11}^{-1}\Sigma_{12}\right)^{-1}\Sigma_{21}\Sigma_{11}^{-1} & \Sigma_j \end{pmatrix}$$

$\theta = \Sigma^{-1}$ so

$$\begin{pmatrix} \theta_{jj} & \theta_{jk} \\ \theta_{kj} & \theta_{kk} \end{pmatrix} = \begin{pmatrix} Var(X_j) & Cov(X_j, X_k) \\ Cov(X_k, X_j) & Var(X_k) \end{pmatrix}$$

$$Cov(X_j, X_k) = -\frac{\theta_{jk}}{\theta_{kk}} Var(X_j) = Cov(X_k, X_j) = -\frac{\theta_{jk}}{\theta_{jj}} Var(X_k)$$

Thus, $Cor(X_2, X_1) = Cor(X_j, X_k \mid X \setminus \{X_j, X_k\}) = \dfrac{Cov(X_j, X_k)}{\sqrt{Var(X_j)Var(X_k)}}$