They have latent variables $\boldsymbol{w}$ and

$$\boldsymbol{w} = \boldsymbol{X\beta} + \sum_{k=1}^{q} \boldsymbol{Z}_k \boldsymbol{r}_k + \boldsymbol{\epsilon}$$

Assuming something about the errors this means

$$\boldsymbol{w}|\boldsymbol{X}, \boldsymbol{\beta}, \boldsymbol{\Sigma}_\theta \sim N\left(\boldsymbol{X\beta}, \boldsymbol{\Sigma}_\theta\right)$$

or equivalently

$$f\left(\boldsymbol{w}|\boldsymbol{X}, \boldsymbol{\beta}, \boldsymbol{\Sigma}_\theta\right) = (2\pi)^{-n/2} \left|\boldsymbol{\Sigma}_\theta\right|^{-1/2} \exp\left\{-\frac{1}{2}\left(\boldsymbol{w} - \boldsymbol{X\beta}\right)^T \boldsymbol{\Sigma}_\theta^{-1}\left(\boldsymbol{w} - \boldsymbol{X\beta}\right)\right\}$$

where

$$\boldsymbol{\Sigma}_\theta = \sum_{k=1}^{q} \boldsymbol{Z}_k \boldsymbol{V}_k \boldsymbol{Z}_k^T + \sigma_0^2 \boldsymbol{I}$$

The joint density

$$\left[\boldsymbol{y}, \boldsymbol{w}|\boldsymbol{\phi}, \boldsymbol{X}, \boldsymbol{\beta}, \boldsymbol{\Sigma}_\theta\right] = \left[\boldsymbol{y}|\boldsymbol{g}^{-1}\left(\boldsymbol{w}\right), \boldsymbol{\phi}\right]\left[\boldsymbol{w}|\boldsymbol{X}, \boldsymbol{\beta}, \boldsymbol{\Sigma}_\theta\right]$$

where $\left[\boldsymbol{y}|\boldsymbol{g}^{-1}\left(\boldsymbol{w}\right), \boldsymbol{\phi}\right]$ is the data model and $\left[\boldsymbol{w}|\boldsymbol{X}, \boldsymbol{\beta}, \boldsymbol{\Sigma}_\theta\right]$ is the process model. We have

$$\begin{aligned}
\left[\boldsymbol{y}|\boldsymbol{\phi}, \boldsymbol{\theta}\right] &= \int_{\boldsymbol{w}} \int_{\boldsymbol{\beta}} \left[\boldsymbol{y}, \boldsymbol{w}|\boldsymbol{\phi}, \boldsymbol{X}, \boldsymbol{\beta}, \boldsymbol{\Sigma}_\theta\right] d\boldsymbol{\beta} d\boldsymbol{w} \\
&= \int_{\boldsymbol{w}} \left[\boldsymbol{y}|\boldsymbol{g}^{-1}\left(\boldsymbol{w}\right), \boldsymbol{\phi}\right] \left\{\int_{\boldsymbol{\beta}} \left[\boldsymbol{w}|\boldsymbol{X}, \boldsymbol{\beta}, \boldsymbol{\Sigma}_\theta\right] d\boldsymbol{\beta}\right\} d\boldsymbol{w}
\end{aligned}$$

We have

$$\int_{\boldsymbol{\beta}} \left[\boldsymbol{w}|\boldsymbol{X}, \boldsymbol{\beta}, \boldsymbol{\Sigma}_\theta\right] d\boldsymbol{\beta} = \int_{\boldsymbol{\beta}} (2\pi)^{-n/2} \left|\boldsymbol{\Sigma}_\theta\right|^{-1/2} \exp\left\{-\frac{1}{2} \underbrace{\left(\boldsymbol{w} - \boldsymbol{X\beta}\right)^T \boldsymbol{\Sigma}_\theta^{-1}\left(\boldsymbol{w} - \boldsymbol{X\beta}\right)}_{A_1}\right\} d\boldsymbol{\beta}$$

where

$$\begin{aligned}
A_1 &= \left(\boldsymbol{w} - \boldsymbol{X\beta}\right)^T \boldsymbol{\Sigma}_\theta^{-1}\left(\boldsymbol{w} - \boldsymbol{X\beta}\right) \\
&= \boldsymbol{w}^T \boldsymbol{\Sigma}_\theta^{-1} \boldsymbol{w} + \boldsymbol{\beta}^T \boldsymbol{X}^T \boldsymbol{\Sigma}_\theta^{-1} \boldsymbol{X\beta} - 2\boldsymbol{\beta}^T \boldsymbol{X}^T \boldsymbol{\Sigma}_\theta^{-1} \boldsymbol{w} \\
&= \left(\boldsymbol{\beta} - \left[\boldsymbol{X}^T \boldsymbol{\Sigma}_\theta^{-1} \boldsymbol{X}\right]^{-1} \boldsymbol{X}^T \boldsymbol{\Sigma}_\theta^{-1} \boldsymbol{w}\right)^T \boldsymbol{X}^T \boldsymbol{\Sigma}_\theta^{-1} \boldsymbol{X} \left(\boldsymbol{\beta} - \left[\boldsymbol{X}^T \boldsymbol{\Sigma}_\theta^{-1} \boldsymbol{X}\right]^{-1} \boldsymbol{X}^T \boldsymbol{\Sigma}_\theta^{-1} \boldsymbol{w}\right) \\
&- \boldsymbol{w}^T \boldsymbol{\Sigma}_\theta^{-1} \boldsymbol{X} \left[\boldsymbol{X}^T \boldsymbol{\Sigma}_\theta^{-1} \boldsymbol{X}\right]^{-1} \boldsymbol{X}^T \boldsymbol{\Sigma}_\theta^{-1} \boldsymbol{w} + \boldsymbol{w}^T \boldsymbol{\Sigma}_\theta^{-1} \boldsymbol{w}
\end{aligned}$$

Then

$$
\begin{aligned}
[\boldsymbol{w}|\boldsymbol{X},\boldsymbol{\Sigma}_\theta] &= \int_{\boldsymbol{\beta}} [\boldsymbol{w}|\boldsymbol{X},\boldsymbol{\beta},\boldsymbol{\Sigma}_\theta]\, d\boldsymbol{\beta} \\
&= (2\pi)^{-n/2} \left|\boldsymbol{\Sigma}_\theta\right|^{-1/2} \\
&\int_{\boldsymbol{\beta}} \exp\left\{-\frac{1}{2}\left(\boldsymbol{\beta} - \left[\boldsymbol{X}^T\boldsymbol{\Sigma}_\theta^{-1}\boldsymbol{X}\right]^{-1}\boldsymbol{X}^T\boldsymbol{\Sigma}_\theta^{-1}\boldsymbol{w}\right)^T \boldsymbol{X}^T\boldsymbol{\Sigma}_\theta^{-1}\boldsymbol{X}\left(\boldsymbol{\beta} - \left[\boldsymbol{X}^T\boldsymbol{\Sigma}_\theta^{-1}\boldsymbol{X}\right]^{-1}\boldsymbol{X}^T\boldsymbol{\Sigma}_\theta^{-1}\boldsymbol{w}\right)\right\} d\boldsymbol{\beta} \\
&\times \exp\left\{-\frac{1}{2}\boldsymbol{w}^T\left[\boldsymbol{\Sigma}_\theta^{-1} - \boldsymbol{\Sigma}_\theta^{-1}\boldsymbol{X}\left[\boldsymbol{X}^T\boldsymbol{\Sigma}_\theta^{-1}\boldsymbol{X}\right]^{-1}\boldsymbol{X}^T\boldsymbol{\Sigma}_\theta^{-1}\right]\boldsymbol{w}\right\} \\
&= (2\pi)^{-(n-p)/2} \left|\boldsymbol{\Sigma}_\theta\right|^{-1/2} \left|\boldsymbol{X}^T\boldsymbol{\Sigma}_\theta^{-1}\boldsymbol{X}\right|^{-1/2} \\
&\underbrace{\int_{\boldsymbol{\beta}} (2\pi)^{-p/2}\left|\boldsymbol{X}^T\boldsymbol{\Sigma}_\theta^{-1}\boldsymbol{X}\right|^{1/2} \exp\left\{-\frac{1}{2}\left(\boldsymbol{\beta} - \left[\boldsymbol{X}^T\boldsymbol{\Sigma}_\theta^{-1}\boldsymbol{X}\right]^{-1}\boldsymbol{X}^T\boldsymbol{\Sigma}_\theta^{-1}\boldsymbol{w}\right)^T \left[\boldsymbol{X}^T\boldsymbol{\Sigma}_\theta^{-1}\boldsymbol{X}\right]\left(\boldsymbol{\beta} - \left[\boldsymbol{X}^T\boldsymbol{\Sigma}\right.\right.\right.}_{=1} \\
&\times \exp\left\{-\frac{1}{2}\boldsymbol{w}^T\left[\boldsymbol{\Sigma}_\theta^{-1} - \boldsymbol{\Sigma}_\theta^{-1}\boldsymbol{X}\left[\boldsymbol{X}^T\boldsymbol{\Sigma}_\theta^{-1}\boldsymbol{X}\right]^{-1}\boldsymbol{X}^T\boldsymbol{\Sigma}_\theta^{-1}\right]\boldsymbol{w}\right\} \\
&= (2\pi)^{-(n-p)/2} \left|\boldsymbol{\Sigma}_\theta\right|^{-1/2} \left|\boldsymbol{X}^T\boldsymbol{\Sigma}_\theta^{-1}\boldsymbol{X}\right|^{-1/2} \\
&\times \exp\left\{-\frac{1}{2}\boldsymbol{w}^T\left[\boldsymbol{\Sigma}_\theta^{-1} - \boldsymbol{\Sigma}_\theta^{-1}\boldsymbol{X}\left[\boldsymbol{X}^T\boldsymbol{\Sigma}_\theta^{-1}\boldsymbol{X}\right]^{-1}\boldsymbol{X}^T\boldsymbol{\Sigma}_\theta^{-1}\right]\boldsymbol{w}\right\}
\end{aligned}
$$

Then the marginal

$$
\begin{aligned}
[\boldsymbol{y}|\boldsymbol{\phi},\boldsymbol{\theta}] &= \int_{\boldsymbol{w}} \left[\boldsymbol{y}|\boldsymbol{g}^{-1}(\boldsymbol{w}),\boldsymbol{\phi}\right][\boldsymbol{w}|\boldsymbol{X},\boldsymbol{\Sigma}_\theta]\, d\boldsymbol{w} \\
&\int_{\boldsymbol{w}} \left[\boldsymbol{y}|\boldsymbol{g}^{-1}(\boldsymbol{w}),\boldsymbol{\phi}\right]\left\{\int_{\boldsymbol{\beta}} [\boldsymbol{w}|\boldsymbol{X},\boldsymbol{\beta},\boldsymbol{\Sigma}_\theta]\, d\boldsymbol{\beta}\right\} d\boldsymbol{w} \\
&= (2\pi)^{-(n-p)/2} \left|\boldsymbol{\Sigma}_\theta\right|^{-1/2} \left|\boldsymbol{X}^T\boldsymbol{\Sigma}_\theta^{-1}\boldsymbol{X}\right|^{-1/2} \\
&\int_{\boldsymbol{w}} \left[\boldsymbol{y}|\boldsymbol{g}^{-1}(\boldsymbol{w}),\boldsymbol{\phi}\right] \exp\left\{-\frac{1}{2}\boldsymbol{w}^T\left[\boldsymbol{\Sigma}_\theta^{-1} - \boldsymbol{\Sigma}_\theta^{-1}\boldsymbol{X}\left[\boldsymbol{X}^T\boldsymbol{\Sigma}_\theta^{-1}\boldsymbol{X}\right]^{-1}\boldsymbol{X}^T\boldsymbol{\Sigma}_\theta^{-1}\right]\boldsymbol{w}\right\} d\boldsymbol{w}
\end{aligned}
$$

################Try what they say, i.e., replaying $\boldsymbol{\beta}$ in $[\boldsymbol{w}|\boldsymbol{X},\boldsymbol{\beta},\boldsymbol{\Sigma}_\theta]$ by the maximum likelihood estimator $\hat{\boldsymbol{\beta}} = \left[\boldsymbol{X}^T\boldsymbol{\Sigma}_\theta^{-1}\boldsymbol{X}\right]^{-1}\boldsymbol{X}^T\boldsymbol{\Sigma}_\theta^{-1}\boldsymbol{w}$. Then

$$
\begin{aligned}
[\boldsymbol{w}|\boldsymbol{X},\boldsymbol{\Sigma}_\theta] &= \frac{1}{C_n}\exp\left\{\left(\boldsymbol{w} - \boldsymbol{X}\hat{\boldsymbol{\beta}}\right)^T \boldsymbol{\Sigma}_\theta^{-1}\left(\boldsymbol{w} - \boldsymbol{X}\hat{\boldsymbol{\beta}}\right)\right\} \\
&= \frac{1}{C_n}\exp\left\{\underbrace{\left(\boldsymbol{w} - \left[\boldsymbol{X}^T\boldsymbol{\Sigma}_\theta^{-1}\boldsymbol{X}\right]^{-1}\boldsymbol{X}^T\boldsymbol{\Sigma}_\theta^{-1}\boldsymbol{w}\right)^T \boldsymbol{\Sigma}_\theta^{-1}\left(\boldsymbol{w} - \left[\boldsymbol{X}^T\boldsymbol{\Sigma}_\theta^{-1}\boldsymbol{X}\right]^{-1}\boldsymbol{X}^T\boldsymbol{\Sigma}_\theta^{-1}\boldsymbol{w}\right)}_{B_1}\right\}
\end{aligned}
$$

where

$$\begin{aligned}
B_1 &= \left(\boldsymbol{I}\boldsymbol{w} - \left[\boldsymbol{X}^T\boldsymbol{\Sigma}_\theta^{-1}\boldsymbol{X}\right]^{-1}\boldsymbol{X}^T\boldsymbol{\Sigma}_\theta^{-1}\boldsymbol{w}\right)^T\boldsymbol{\Sigma}_\theta^{-1}\left(\boldsymbol{I}\boldsymbol{w} - \left[\boldsymbol{X}^T\boldsymbol{\Sigma}_\theta^{-1}\boldsymbol{X}\right]^{-1}\boldsymbol{X}^T\boldsymbol{\Sigma}_\theta^{-1}\boldsymbol{w}\right) \\
&= \left(\left\{\boldsymbol{I} - \left[\boldsymbol{X}^T\boldsymbol{\Sigma}_\theta^{-1}\boldsymbol{X}\right]^{-1}\right\}\boldsymbol{w}\right)^T\boldsymbol{\Sigma}_\theta^{-1}\left(\left\{\boldsymbol{I} - \left[\boldsymbol{X}^T\boldsymbol{\Sigma}_\theta^{-1}\boldsymbol{X}\right]^{-1}\right\}\boldsymbol{w}\right) \\
&= \boldsymbol{w}^T\left\{\boldsymbol{I} - \left[\boldsymbol{X}^T\boldsymbol{\Sigma}_\theta^{-1}\boldsymbol{X}\right]^{-1}\right\}\boldsymbol{\Sigma}_\theta^{-1}\left\{\boldsymbol{I} - \left[\boldsymbol{X}^T\boldsymbol{\Sigma}_\theta^{-1}\boldsymbol{X}\right]^{-1}\right\}\boldsymbol{w} \\
&= \boldsymbol{w}^T\left\{\boldsymbol{\Sigma}_\theta^{-1} + \left[\boldsymbol{X}^T\boldsymbol{\Sigma}_\theta^{-1}\boldsymbol{X}\right]^{-1}\boldsymbol{\Sigma}_\theta^{-1}\left[\boldsymbol{X}^T\boldsymbol{\Sigma}_\theta^{-1}\boldsymbol{X}\right]^{-1} - 2\left[\boldsymbol{X}^T\boldsymbol{\Sigma}_\theta^{-1}\boldsymbol{X}\right]^{-1}\right\}\boldsymbol{w}
\end{aligned}$$

$$\begin{aligned}
B_1 &= \left(\boldsymbol{I}\boldsymbol{w} - \left[\boldsymbol{X}^T\boldsymbol{\Sigma}_\theta^{-1}\boldsymbol{X}\right]^{-1}\boldsymbol{X}^T\boldsymbol{\Sigma}_\theta^{-1}\boldsymbol{w}\right)^T\boldsymbol{\Sigma}_\theta^{-1}\left(\boldsymbol{I}\boldsymbol{w} - \left[\boldsymbol{X}^T\boldsymbol{\Sigma}_\theta^{-1}\boldsymbol{X}\right]^{-1}\boldsymbol{X}^T\boldsymbol{\Sigma}_\theta^{-1}\boldsymbol{w}\right) \\
&= \boldsymbol{w}^T\left\{\boldsymbol{I} - \left(\boldsymbol{\Sigma}_\theta^{-1}\boldsymbol{X}\right)\left[\boldsymbol{X}^T\boldsymbol{\Sigma}_\theta^{-1}\boldsymbol{X}\right]^{-1}\right\}\boldsymbol{\Sigma}_\theta^{-1}\left\{\boldsymbol{I} - \left(\boldsymbol{\Sigma}_\theta^{-1}\boldsymbol{X}\right)\left[\boldsymbol{X}^T\boldsymbol{\Sigma}_\theta^{-1}\boldsymbol{X}\right]^{-1}\right\}\boldsymbol{w} \\
&= \boldsymbol{w}^T\left\{\boldsymbol{\Sigma}_\theta^{-1} + \boldsymbol{\Sigma}_\theta^{-1}\boldsymbol{X}\left[\boldsymbol{X}^T\boldsymbol{\Sigma}_\theta^{-1}\boldsymbol{X}\right]^{-1}\boldsymbol{\Sigma}_\theta^{-1}\left[\boldsymbol{X}^T\boldsymbol{\Sigma}_\theta^{-1}\boldsymbol{X}\right]^{-1}\boldsymbol{X}^T\boldsymbol{\Sigma}_\theta^{-1} - 2\boldsymbol{\Sigma}_\theta^{-1}\boldsymbol{X}\left[\boldsymbol{X}^T\boldsymbol{\Sigma}_\theta^{-1}\boldsymbol{X}\right]^{-1}\boldsymbol{\Sigma}_\theta^{-1}\right\}\boldsymbol{w} \\
&= \boldsymbol{w}^T\left\{\boldsymbol{\Sigma}_\theta^{-1} + \boldsymbol{\Sigma}_\theta^{-1}\boldsymbol{X}\left[\boldsymbol{X}^T\boldsymbol{\Sigma}_\theta^{-1}\boldsymbol{X}\right]^{-1}\left(\boldsymbol{\Sigma}_\theta^{-1}\left[\boldsymbol{X}^T\boldsymbol{\Sigma}_\theta^{-1}\boldsymbol{X}\right]^{-1} - 2\boldsymbol{\Sigma}_\theta^{-1}\right)\boldsymbol{X}^T\boldsymbol{\Sigma}_\theta^{-1}\right\}\boldsymbol{w}
\end{aligned}$$

Same as Lukas.
##
Let

$$\begin{aligned}
\ell\left(\boldsymbol{w},\cdot\right) &= \log\left\{\left[\boldsymbol{y}|\boldsymbol{g}^{-1}\left(\boldsymbol{w}\right),\boldsymbol{\phi}\right]\left[\boldsymbol{w}|\boldsymbol{X},\boldsymbol{\Sigma}_\theta\right]\right\} \\
&= \log\left\{\left[\boldsymbol{y}|\boldsymbol{g}^{-1}\left(\boldsymbol{w}\right),\boldsymbol{\phi}\right]\right\} - \frac{1}{2}\boldsymbol{w}^T\left[\boldsymbol{\Sigma}_\theta^{-1} - \boldsymbol{\Sigma}_\theta^{-1}\boldsymbol{X}\left[\boldsymbol{X}^T\boldsymbol{\Sigma}_\theta^{-1}\boldsymbol{X}\right]^{-1}\boldsymbol{X}^T\boldsymbol{\Sigma}_\theta^{-1}\right]\boldsymbol{w}
\end{aligned}$$

And (again)

$$[\boldsymbol{y}|\boldsymbol{\phi},\boldsymbol{\theta}] = \int_{\boldsymbol{w}}\exp\left\{\ell\left(\boldsymbol{w},\cdot\right)\right\}d\boldsymbol{w}$$

Using $\hat{\boldsymbol{w}} = \underset{\boldsymbol{w}}{\arg\max}\,\ell\left(\boldsymbol{w},\cdot\right)$ and approximate

$$\begin{aligned}
\ell\left(\boldsymbol{w},\cdot\right) &\approx \ell\left(\hat{\boldsymbol{w}}\right) + \frac{1}{2}\left(\boldsymbol{w}-\hat{\boldsymbol{w}}\right)^T\underbrace{\left.\frac{\partial^2\ell\left(\boldsymbol{w},\cdot\right)}{\partial\boldsymbol{w}\partial\boldsymbol{w}^T}\right|_{\boldsymbol{w}=\hat{\boldsymbol{w}}}}_{\boldsymbol{H}}\left(\boldsymbol{w}-\hat{\boldsymbol{w}}\right) \\
&= \ell\left(\hat{\boldsymbol{w}}\right) - \frac{1}{2}\left(\boldsymbol{w}-\hat{\boldsymbol{w}}\right)^T\left[-\boldsymbol{H}\right]\left(\boldsymbol{w}-\hat{\boldsymbol{w}}\right)
\end{aligned}$$

Then

$$
\begin{aligned}
[\boldsymbol{y}|\boldsymbol{\phi},\boldsymbol{\theta}] &= \int_{\boldsymbol{w}} \exp\{\ell(\boldsymbol{w},\cdot)\}\, d\boldsymbol{w} \\
&\approx \int_{\boldsymbol{w}} \exp\left\{\ell(\hat{\boldsymbol{w}}) - \frac{1}{2}(\boldsymbol{w}-\hat{\boldsymbol{w}})^T [-\boldsymbol{H}] (\boldsymbol{w}-\hat{\boldsymbol{w}})\right\} d\boldsymbol{w} \\
&= (2\pi)^{n/2} |-\boldsymbol{H}|^{-1/2} \exp\{\ell(\hat{\boldsymbol{w}})\} \underbrace{\int_{\boldsymbol{w}} (2\pi)^{-n/2} [-\boldsymbol{H}]^{1/2} \exp\left\{-\frac{1}{2}(\boldsymbol{w}-\hat{\boldsymbol{w}})^T [-\boldsymbol{H}] (\boldsymbol{w}-\hat{\boldsymbol{w}})\right\} d\boldsymbol{w}}_{1} \\
&= (2\pi)^{n/2} |-\boldsymbol{H}|^{-1/2} \exp\{\ell(\hat{\boldsymbol{w}})\}
\end{aligned}
$$

And its log

$$
\begin{aligned}
\log[\boldsymbol{y}|\boldsymbol{\phi},\boldsymbol{\theta}] &= \frac{n}{2}\log(2\pi) + \ell(\hat{\boldsymbol{w}}) - \frac{1}{2}\log|-\boldsymbol{H}| \\
&= \frac{n}{2}\log(2\pi) + \log\left[\boldsymbol{y}|\boldsymbol{g}^{-1}(\hat{\boldsymbol{w}}), \boldsymbol{\phi}\right] + \log[\hat{\boldsymbol{w}}|\boldsymbol{X},\boldsymbol{\Sigma}_\theta] - \frac{1}{2}\log|-\boldsymbol{H}|
\end{aligned}
$$

where

$$
\begin{aligned}
\log[\hat{\boldsymbol{w}}|\boldsymbol{X},\boldsymbol{\Sigma}_\theta] = &-\frac{n-p}{2}\log(2\pi) - \frac{1}{2}|\boldsymbol{\Sigma}_\theta| - \frac{1}{2}\log\left|\boldsymbol{X}^T\boldsymbol{\Sigma}_\theta^{-1}\boldsymbol{X}\right| \\
&- \frac{1}{2}\underbrace{\hat{\boldsymbol{w}}^T\left[\boldsymbol{\Sigma}_\theta^{-1} - \boldsymbol{\Sigma}_\theta^{-1}\boldsymbol{X}\left[\boldsymbol{X}^T\boldsymbol{\Sigma}_\theta^{-1}\boldsymbol{X}\right]^{-1}\boldsymbol{X}^T\boldsymbol{\Sigma}_\theta^{-1}\right]}_{C_1}\hat{\boldsymbol{w}}
\end{aligned}
$$

########
Compare to them with replacing C_1 above to be (with $\hat{\boldsymbol{\beta}} = \left[\boldsymbol{X}^T\boldsymbol{\Sigma}_\theta^{-1}\boldsymbol{X}\right]\boldsymbol{X}^T\boldsymbol{\Sigma}_\theta^{-1}\hat{\boldsymbol{w}}$)

$$
\begin{aligned}
C_1 &= \left(\hat{\boldsymbol{w}} - \boldsymbol{X}\hat{\boldsymbol{\beta}}\right)^T \boldsymbol{\Sigma}_\theta^{-1} \left(\hat{\boldsymbol{w}} - \boldsymbol{X}\hat{\boldsymbol{\beta}}\right) \\
&= \left(\hat{\boldsymbol{w}} - \left[\boldsymbol{X}^T\boldsymbol{\Sigma}_\theta^{-1}\boldsymbol{X}\right]\boldsymbol{X}^T\boldsymbol{\Sigma}_\theta^{-1}\hat{\boldsymbol{w}}\right)^T \boldsymbol{\Sigma}_\theta^{-1} \left(\hat{\boldsymbol{w}} - \left[\boldsymbol{X}^T\boldsymbol{\Sigma}_\theta^{-1}\boldsymbol{X}\right]\boldsymbol{X}^T\boldsymbol{\Sigma}_\theta^{-1}\hat{\boldsymbol{w}}\right) \\
&= \hat{\boldsymbol{w}}^T\left\{\boldsymbol{I} - \boldsymbol{\Sigma}_\theta^{-1}\boldsymbol{X}\boldsymbol{X}^T\boldsymbol{\Sigma}_\theta^{-1}\boldsymbol{X}\right\}\boldsymbol{\Sigma}_\theta^{-1}\left\{\boldsymbol{I} - \boldsymbol{X}^T\boldsymbol{\Sigma}_\theta^{-1}\boldsymbol{X}\boldsymbol{X}^T\boldsymbol{\Sigma}_\theta^{-1}\right\}\hat{\boldsymbol{w}} \\
&= \hat{\boldsymbol{w}}^T\left\{\boldsymbol{\Sigma}_\theta^{-1} - 2\boldsymbol{\Sigma}_\theta^{-1}\boldsymbol{X}\boldsymbol{X}^T\boldsymbol{\Sigma}_\theta^{-1} + \boldsymbol{\Sigma}_\theta^{-1}\boldsymbol{X}\boldsymbol{X}^T\boldsymbol{\Sigma}_\theta^{-1}\boldsymbol{X}\boldsymbol{\Sigma}_\theta^{-1}\boldsymbol{X}^T\boldsymbol{\Sigma}_\theta^{-1}\boldsymbol{X}\boldsymbol{X}^T\boldsymbol{\Sigma}_\theta^{-1}\right\}\hat{\boldsymbol{w}}
\end{aligned}
$$

###########Don't get it
#########Try myself

$$
\ell(\hat{\boldsymbol{w}}) = \log\left\{\left[\boldsymbol{y}|\boldsymbol{g}^{-1}(\hat{\boldsymbol{w}}), \boldsymbol{\phi}\right]\right\} - \frac{1}{2}\hat{\boldsymbol{w}}^T\left[\boldsymbol{\Sigma}_\theta^{-1} - \boldsymbol{\Sigma}_\theta^{-1}\boldsymbol{X}\left[\boldsymbol{X}^T\boldsymbol{\Sigma}_\theta^{-1}\boldsymbol{X}\right]^{-1}\boldsymbol{X}^T\boldsymbol{\Sigma}_\theta^{-1}\right]\hat{\boldsymbol{w}}
$$

and

$$
\frac{\partial \ell(\boldsymbol{w},\cdot)}{\partial \boldsymbol{w}} = \frac{\partial}{\partial \boldsymbol{w}}\log\left\{\left[\boldsymbol{y}|\boldsymbol{g}^{-1}(\boldsymbol{w}), \boldsymbol{\phi}\right]\right\} - \left[\boldsymbol{\Sigma}_\theta^{-1} - \boldsymbol{\Sigma}_\theta^{-1}\boldsymbol{X}\left[\boldsymbol{X}^T\boldsymbol{\Sigma}_\theta^{-1}\boldsymbol{X}\right]^{-1}\boldsymbol{X}^T\boldsymbol{\Sigma}_\theta^{-1}\right]\boldsymbol{w}
$$

and

$$\frac{\partial^2 \ell(\boldsymbol{w}, \cdot)}{\partial \boldsymbol{w} \partial \boldsymbol{w}^T} = \frac{\partial^2}{\partial \boldsymbol{w} \partial \boldsymbol{w}^T} \log \left\{ \left[\boldsymbol{y} | \boldsymbol{g}^{-1}(\boldsymbol{w}), \boldsymbol{\phi} \right] \right\} - \left[\boldsymbol{\Sigma}_\theta^{-1} - \boldsymbol{\Sigma}_\theta^{-1} \boldsymbol{X} \left[\boldsymbol{X}^T \boldsymbol{\Sigma}_\theta^{-1} \boldsymbol{X} \right]^{-1} \boldsymbol{X}^T \boldsymbol{\Sigma}_\theta^{-1} \right]$$

Hence,

$$\left. \frac{\partial^2 \ell(\boldsymbol{w}, \cdot)}{\partial \boldsymbol{w} \partial \boldsymbol{w}^T} \right|_{\boldsymbol{w}=\hat{\boldsymbol{w}}} = \underbrace{\left. \frac{\partial^2 \log \left\{ \left[\boldsymbol{y} | \boldsymbol{g}^{-1}(\boldsymbol{w}), \boldsymbol{\phi} \right] \right\}}{\partial \boldsymbol{w} \partial \boldsymbol{w}^T} \right|_{\boldsymbol{w}=\hat{\boldsymbol{w}}}}_{\boldsymbol{D}_\phi} - \underbrace{\left[\boldsymbol{\Sigma}_\theta^{-1} - \boldsymbol{\Sigma}_\theta^{-1} \boldsymbol{X} \left[\boldsymbol{X}^T \boldsymbol{\Sigma}_\theta^{-1} \boldsymbol{X} \right]^{-1} \boldsymbol{X}^T \boldsymbol{\Sigma}_\theta^{-1} \right]}_{\boldsymbol{P}_\theta}$$

Hm... I get the same gradient and Hessian as they...