Fit a Hessian and mode of a log likelihood or log posterior to the Hessian and mode of a Wishart distribution

The inverse Wishart distribution with (hyper)parameters $\boldsymbol{\Upsilon}$ and ν and dimension $p \times p$. Random variable $\boldsymbol{\Phi}$.

$$f\left(\boldsymbol{\Phi};\boldsymbol{\Upsilon},\nu\right)=\frac{\left|\boldsymbol{\Upsilon}\right|^{\nu/2}}{2^{\nu p/2}\Gamma_p\left(\frac{\nu}{2}\right)}\left|\boldsymbol{\Phi}\right|^{-(\nu+p+1)/2}\exp\left\{-\frac{1}{2}\mathrm{Tr}\left(\boldsymbol{\Upsilon}\boldsymbol{\Phi}^{-1}\right)\right\}$$

with log

$$\log f\left(\boldsymbol{\Phi};\boldsymbol{\Upsilon},\nu\right)=\frac{\nu}{2}\log\left|\boldsymbol{\Upsilon}\right|-\frac{\nu p}{2}\log 2-\frac{\nu+p+1}{2}\log\left|\boldsymbol{\Phi}\right|-\frac{1}{2}\mathrm{Tr}\left(\boldsymbol{\Upsilon}\boldsymbol{\Phi}^{-1}\right)$$

Derivative with respect to $\Phi_{j,k}$

$$\begin{aligned}
\frac{\partial}{\partial\Phi_{j,k}}\log f\left(\boldsymbol{\Phi};\boldsymbol{\Upsilon},\nu\right) &= -\frac{\nu+p+1}{2}\frac{\partial}{\partial\Phi_{j,k}}\log\left|\boldsymbol{\Phi}\right|-\frac{1}{2}\frac{\partial}{\partial\Phi_{j,k}}\mathrm{Tr}\left(\boldsymbol{\Upsilon}\boldsymbol{\Phi}^{-1}\right)\\
&= -\frac{\nu+p+1}{2}\frac{\frac{\partial}{\partial\Phi_{j,k}}\left|\boldsymbol{\Phi}\right|}{\left|\boldsymbol{\Phi}\right|}-\frac{1}{2}\mathrm{Tr}\left(\boldsymbol{\Upsilon}\frac{\partial\boldsymbol{\Phi}^{-1}}{\partial\Phi_{j,k}}\right)\\
&= -\frac{\nu+p+1}{2}\frac{\left|\boldsymbol{\Phi}\right|\mathrm{Tr}\left[\boldsymbol{\Phi}^{-1}\frac{\partial\boldsymbol{\Phi}}{\partial\Phi_{j,k}}\right]}{\left|\boldsymbol{\Phi}\right|}-\frac{1}{2}\mathrm{Tr}\left(\boldsymbol{\Upsilon}\left[-\boldsymbol{\Phi}^{-1}\frac{\partial\boldsymbol{\Phi}}{\partial\Phi_{j,k}}\boldsymbol{\Phi}^{-1}\right]\right)\\
&= \mathrm{Tr}\left[-\frac{\nu+p+1}{2}\boldsymbol{\Phi}^{-1}\frac{\partial\boldsymbol{\Phi}}{\partial\Phi_{j,k}}\right]+\mathrm{Tr}\left[\frac{1}{2}\boldsymbol{\Upsilon}\boldsymbol{\Phi}^{-1}\frac{\partial\boldsymbol{\Phi}}{\partial\Phi_{j,k}}\boldsymbol{\Phi}^{-1}\right]\\
&= \mathrm{Tr}\left[-\frac{\nu+p+1}{2}\boldsymbol{\Phi}^{-1}\frac{\partial\boldsymbol{\Phi}}{\partial\Phi_{j,k}}\right]+\mathrm{Tr}\left[\frac{1}{2}\boldsymbol{\Phi}^{-1}\boldsymbol{\Upsilon}\boldsymbol{\Phi}^{-1}\frac{\partial\boldsymbol{\Phi}}{\partial\Phi_{j,k}}\right]\\
&= \mathrm{Tr}\left[-\frac{\nu+p+1}{2}\boldsymbol{\Phi}^{-1}\frac{\partial\boldsymbol{\Phi}}{\partial\Phi_{j,k}}+\frac{1}{2}\boldsymbol{\Phi}^{-1}\boldsymbol{\Upsilon}\boldsymbol{\Phi}^{-1}\frac{\partial\boldsymbol{\Phi}}{\partial\Phi_{j,k}}\right]\\
&= \frac{1}{2}\mathrm{Tr}\left[\left(-\left(\nu+p+1\right)\boldsymbol{I}+\boldsymbol{\Phi}^{-1}\boldsymbol{\Upsilon}\right)\boldsymbol{\Phi}^{-1}\frac{\partial\boldsymbol{\Phi}}{\partial\Phi_{j,k}}\right]
\end{aligned}$$

Find the zero

$$
\begin{aligned}
\frac{\partial}{\partial \Phi_{j,k}} \log f\left(\boldsymbol{\Phi}; \boldsymbol{\Upsilon}, \nu\right) &= 0 \\
\frac{1}{2}\mathrm{Tr}\left[\left(-\left(\nu+p+1\right)\boldsymbol{I} + \boldsymbol{\Phi}^{-1}\boldsymbol{\Upsilon}\right)\boldsymbol{\Phi}^{-1}\frac{\partial \boldsymbol{\Phi}}{\partial \Phi_{j,k}}\right] &= 0 \\
\left(-\left(\nu+p+1\right)\boldsymbol{I} + \boldsymbol{\Phi}^{-1}\boldsymbol{\Upsilon}\right) &= 0 \\
\boldsymbol{\Phi}^{-1}\boldsymbol{\Upsilon} &= \left(\nu+p+1\right)\boldsymbol{I} \\
\boldsymbol{\Upsilon} &= \left(\nu+p+1\right)\boldsymbol{\Phi} \\
\frac{\boldsymbol{\Upsilon}}{\nu+p+1} &= \boldsymbol{\Phi}
\end{aligned}
$$

Also found the same on Wikipedia. Hence, we have the mode

$$
\hat{\boldsymbol{\Phi}} = \frac{\boldsymbol{\Upsilon}}{\nu+p+1}
$$

Notice that, e.g.,

$$
\begin{aligned}
\frac{\partial \boldsymbol{\Phi}}{\partial \Phi_{2,2}} &= \begin{pmatrix} 0 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & 0 \\ 0 & 0 & \cdots & 0 \end{pmatrix} \\
\frac{\partial \boldsymbol{\Phi}}{\partial \Phi_{2,1}} &= \begin{pmatrix} 0 & 1 & \cdots & 0 \\ 1 & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & 0 \\ 0 & 0 & \cdots & 0 \end{pmatrix}
\end{aligned}
$$

etc. The second derivative is

$$
\frac{\partial^2}{\partial \Phi_{l,m}\partial \Phi_{j,k}} \log f\left(\boldsymbol{\Phi}; \boldsymbol{\Upsilon}, \nu\right) = -\frac{\nu+p+1}{2}\mathrm{Tr}\left[\underbrace{\frac{\partial}{\partial \Phi_{l,m}}\left(\boldsymbol{\Phi}^{-1}\frac{\partial \boldsymbol{\Phi}}{\partial \Phi_{j,k}}\right)}_{A_1}\right] + \frac{1}{2}\mathrm{Tr}\left[\boldsymbol{\Upsilon}\underbrace{\frac{\partial}{\partial \Phi_{l,m}}\left(\boldsymbol{\Phi}^{-1}\frac{\partial \boldsymbol{\Phi}}{\partial \Phi_{j,k}}\boldsymbol{\Phi}^{-1}\right)}_{A_2}\right]
$$

where

$$
\begin{aligned}
A_1 &= \left(\frac{\partial}{\partial \Phi_{l,m}}\boldsymbol{\Phi}^{-1}\right)\frac{\partial \boldsymbol{\Phi}}{\partial \Phi_{j,k}} \\
&= -\boldsymbol{\Phi}^{-1}\frac{\partial \boldsymbol{\Phi}}{\partial \Phi_{l,m}}\boldsymbol{\Phi}^{-1}\frac{\partial \boldsymbol{\Phi}}{\partial \Phi_{j,k}}
\end{aligned}
$$

and

$$
\begin{aligned}
A_2 &= -\frac{\partial}{\partial \Phi_{l,m}} \left(\boldsymbol{\Phi}^{-1} \frac{\partial \boldsymbol{\Phi}}{\partial \Phi_{j,k}} \boldsymbol{\Phi}^{-1} \right) \\
&= -\left(\frac{\partial \boldsymbol{\Phi}^{-1}}{\partial \Phi_{l,m}} \right) \frac{\partial \boldsymbol{\Phi}}{\partial \Phi_{j,k}} \boldsymbol{\Phi}^{-1} - \boldsymbol{\Phi}^{-1} \frac{\partial \boldsymbol{\Phi}}{\partial \Phi_{j,k}} \frac{\partial \boldsymbol{\Phi}^{-1}}{\partial \Phi_{l,m}} \\
&= -\boldsymbol{\Phi}^{-1} \frac{\partial \boldsymbol{\Phi}}{\partial \Phi_{l,m}} \boldsymbol{\Phi}^{-1} \frac{\partial \boldsymbol{\Phi}}{\partial \Phi_{j,k}} \boldsymbol{\Phi}^{-1} - \boldsymbol{\Phi}^{-1} \frac{\partial \boldsymbol{\Phi}}{\partial \Phi_{j,k}} \boldsymbol{\Phi}^{-1} \frac{\partial \boldsymbol{\Phi}}{\partial \Phi_{l,m}} \boldsymbol{\Phi}^{-1}
\end{aligned}
$$

Since all matrices $\boldsymbol{\Phi}^{-1}$, $\frac{\partial \boldsymbol{\Phi}}{\partial \Phi_{l,m}}$ and $\frac{\partial \boldsymbol{\Phi}}{\partial \Phi_{j,k}}$ are symmetric, we can write

$$
\begin{aligned}
& A_2 - \boldsymbol{\Phi}^{-1} \frac{\partial \boldsymbol{\Phi}}{\partial \Phi_{l,m}} \boldsymbol{\Phi}^{-1} \frac{\partial \boldsymbol{\Phi}}{\partial \Phi_{j,k}} \boldsymbol{\Phi}^{-1} - \left[\boldsymbol{\Phi}^{-1} \frac{\partial \boldsymbol{\Phi}}{\partial \Phi_{j,k}} \boldsymbol{\Phi}^{-1} \frac{\partial \boldsymbol{\Phi}}{\partial \Phi_{l,m}} \boldsymbol{\Phi}^{-1} \right]^T \\
&= -\boldsymbol{\Phi}^{-1} \frac{\partial \boldsymbol{\Phi}}{\partial \Phi_{l,m}} \boldsymbol{\Phi}^{-1} \frac{\partial \boldsymbol{\Phi}}{\partial \Phi_{j,k}} \boldsymbol{\Phi}^{-1} - \boldsymbol{\Phi}^{-1} \frac{\partial \boldsymbol{\Phi}}{\partial \Phi_{l,m}} \boldsymbol{\Phi}^{-1} \frac{\partial \boldsymbol{\Phi}}{\partial \Phi_{j,k}} \boldsymbol{\Phi}^{-1} + \\
&= -2\boldsymbol{\Phi}^{-1} \frac{\partial \boldsymbol{\Phi}}{\partial \Phi_{l,m}} \boldsymbol{\Phi}^{-1} \frac{\partial \boldsymbol{\Phi}}{\partial \Phi_{j,k}} \boldsymbol{\Phi}^{-1}
\end{aligned}
$$

Hence,

$$
\begin{aligned}
\frac{\partial^2}{\partial \Phi_{l,m} \partial \Phi_{j,k}} \log f\left(\boldsymbol{\Phi}; \boldsymbol{\Upsilon}, \nu\right) &= -\frac{\nu+p+1}{2} \mathrm{Tr} \left[-\boldsymbol{\Phi}^{-1} \frac{\partial \boldsymbol{\Phi}}{\partial \Phi_{l,m}} \boldsymbol{\Phi}^{-1} \frac{\partial \boldsymbol{\Phi}}{\partial \Phi_{j,k}} \right] - \frac{1}{2} \mathrm{Tr} \left[2\boldsymbol{\Upsilon} \boldsymbol{\Phi}^{-1} \frac{\partial \boldsymbol{\Phi}}{\partial \Phi_{l,m}} \boldsymbol{\Phi}^{-1} \frac{\partial \boldsymbol{\Phi}}{\partial \Phi_{j,k}} \boldsymbol{\Phi}^{-1} \right] \\
&= \frac{\nu+p+1}{2} \mathrm{Tr} \left[\boldsymbol{\Phi}^{-1} \frac{\partial \boldsymbol{\Phi}}{\partial \Phi_{l,m}} \boldsymbol{\Phi}^{-1} \frac{\partial \boldsymbol{\Phi}}{\partial \Phi_{j,k}} \right] - \mathrm{Tr} \left[\boldsymbol{\Upsilon} \boldsymbol{\Phi}^{-1} \frac{\partial \boldsymbol{\Phi}}{\partial \Phi_{l,m}} \boldsymbol{\Phi}^{-1} \frac{\partial \boldsymbol{\Phi}}{\partial \Phi_{j,k}} \boldsymbol{\Phi}^{-1} \right] \\
&= \frac{\nu+p+1}{2} \mathrm{Tr} \left[\boldsymbol{\Phi}^{-1} \frac{\partial \boldsymbol{\Phi}}{\partial \Phi_{l,m}} \boldsymbol{\Phi}^{-1} \frac{\partial \boldsymbol{\Phi}}{\partial \Phi_{j,k}} \right] - \mathrm{Tr} \left[\boldsymbol{\Phi}^{-1} \boldsymbol{\Upsilon} \boldsymbol{\Phi}^{-1} \frac{\partial \boldsymbol{\Phi}}{\partial \Phi_{l,m}} \boldsymbol{\Phi}^{-1} \frac{\partial \boldsymbol{\Phi}}{\partial \Phi_{j,k}} \right] \\
&= \mathrm{Tr} \left[\left(\frac{\nu+p+1}{2} \boldsymbol{I} - \boldsymbol{\Phi}^{-1} \boldsymbol{\Upsilon} \right) \boldsymbol{\Phi}^{-1} \frac{\partial \boldsymbol{\Phi}}{\partial \Phi_{l,m}} \boldsymbol{\Phi}^{-1} \frac{\partial \boldsymbol{\Phi}}{\partial \Phi_{j,k}} \right]
\end{aligned}
$$

The elements in $\frac{\partial^2}{\partial \Phi_{l,m} \partial \Phi_{j,k}} \log f\left(\boldsymbol{\Phi}; \boldsymbol{\Upsilon}, \nu\right)$ will be matched with the Hessian of the log likelihood or log posterior to identify the Wishart distribution approximating the posterior distribution of the dispersion parameters. Insert the mode

$$
\begin{aligned}
\hat{\boldsymbol{\Phi}} &= \frac{\boldsymbol{\Upsilon}}{\nu+p+1} \\
\hat{\boldsymbol{\Phi}}^{-1} &= \left(\nu+p+1\right) \boldsymbol{\Upsilon}^{-1}
\end{aligned}
$$

$$
\begin{aligned}
\frac{\partial^2}{\partial \Phi_{l,m} \partial \Phi_{j,k}} \log f \left(\boldsymbol{\Phi}; \boldsymbol{\Upsilon}, \nu \right) \bigg|_{\boldsymbol{\Phi} = \hat{\boldsymbol{\Phi}}} &= \operatorname{Tr} \left[\left(\frac{\nu + p + 1}{2} \boldsymbol{I} - \hat{\boldsymbol{\Phi}}^{-1} \boldsymbol{\Upsilon} \right) \hat{\boldsymbol{\Phi}}^{-1} \frac{\partial \boldsymbol{\Phi}}{\partial \Phi_{l,m}} \hat{\boldsymbol{\Phi}}^{-1} \frac{\partial \hat{\boldsymbol{\Phi}}}{\partial \Phi_{j,k}} \right] \\
&= \operatorname{Tr} \left[\left(\frac{\nu + p + 1}{2} \boldsymbol{I} - \left(\nu + p + 1 \right) \boldsymbol{\Upsilon}^{-1} \boldsymbol{\Upsilon} \right) \left(\nu + p + 1 \right) \boldsymbol{\Upsilon}^{-1} \frac{\partial \boldsymbol{\Phi}}{\partial \Phi_{l,m}} \left(\nu + p + \right. \right. \\
&= \operatorname{Tr} \left[\left(\frac{\nu + p + 1}{2} \boldsymbol{I} - \left(\nu + p + 1 \right) \boldsymbol{I} \right) \left(\nu + p + 1 \right) \boldsymbol{\Upsilon}^{-1} \frac{\partial \boldsymbol{\Phi}}{\partial \Phi_{l,m}} \left(\nu + p + 1 \right) \boldsymbol{\Upsilon}^{-} \right. \\
&= -\frac{1}{2} \operatorname{Tr} \left[\left(\nu + p + 1 \right)^3 \boldsymbol{\Upsilon}^{-1} \frac{\partial \boldsymbol{\Phi}}{\partial \Phi_{l,m}} \boldsymbol{\Upsilon}^{-1} \frac{\partial \boldsymbol{\Phi}}{\partial \Phi_{j,k}} \right]
\end{aligned}
$$