Gradient vector & Hessian matrix

For now let's consider the case without right-censoring.

$$\mu(H_{u_i}, \theta_t) = \theta_0 + \sum_{j=1}^{n-1} \theta_j u_{i-j+1} \tag{1}$$

$$p(u_i|k, \theta_t) = \left[\frac{k}{2\pi \cdot {u_i}^3}\right]^{1/2} e^{-\frac{1}{2}\frac{k\left(u_i - \mu(H_{u_{i-1}}, \theta_t)\right)^2}{\mu(H_{u_{i-1}}, \theta_t)^2 \cdot u_i}} \tag{2}$$

$$\mathcal{L}(u_{t-m:t}|k, \eta_t, \theta_t) = -\sum_{i=0}^{m-1} \eta_i \log p(u_i|k, \theta_t) \tag{20}$$

Gradient `Vector`

$$\nabla \mathcal{L}(u_{t-m:t}|k, \eta_t, \theta_t) = \begin{bmatrix} \frac{\partial}{\partial k}\mathcal{L}(u_{t-m:t}|k, \eta_t, \theta_t) \\ \frac{\partial}{\partial \eta_0}\mathcal{L}(u_{t-m:t}|k, \eta_t, \theta_t) \\ \vdots \\ \frac{\partial}{\partial \eta_{m-1}}\mathcal{L}(u_{t-m:t}|k, \eta_t, \theta_t) \\ \frac{\partial}{\partial \theta_0}\mathcal{L}(u_{t-m:t}|k, \eta_t, \theta_t) \\ \vdots \\ \frac{\partial}{\partial \theta_{n-1}}\mathcal{L}(u_{t-m:t}|k, \eta_t, \theta_t) \end{bmatrix} \tag{4}$$

For k we have:

$$\begin{aligned} &\frac{\partial}{\partial k}\mathcal{L}(u_{t-m:t}|k, \eta_t, \theta_t) = \\ &\frac{\partial}{\partial k}\left(-\sum_{i=0}^{m-1} \eta_i \log p(u_i|k, \theta_t)\right) = \\ &\frac{\partial}{\partial k}\left(-\sum_{i=0}^{m-1} \eta_i \log\left(\left[\frac{k}{2\pi \cdot {u_i}^3}\right]^{1/2} e^{-\frac{1}{2}\frac{k\left(u_i - \mu(H_{u_{i-1}}, \theta_t)\right)^2}{\mu(H_{u_{i-1}}, \theta_t)^2 \cdot u_i}}\right)\right) = \\ &\frac{\partial}{\partial k}\left(\sum_{i=0}^{m-1} -\frac{1}{2}\eta_i \log\left(\left[\frac{k}{2\pi \cdot {u_i}^3}\right]\right) + \frac{1}{2}\eta_i \frac{k(u_i - \mu(H_{u_{i-1}}, \theta_t))^2}{\mu(H_{u_{i-1}}, \theta_t)^2 \cdot u_i}\right) = \\ &\frac{\partial}{\partial k}\left(\sum_{i=0}^{m-1} -\frac{1}{2}\eta_i \log(k) + \eta_i \log\left(2\pi \cdot {u_i}^3\right) + \frac{1}{2}\eta_i \frac{k(u_i - \mu(H_{u_{i-1}}, \theta_t))^2}{\mu(H_{u_{i-1}}, \theta_t)^2 \cdot u_i}\right) = \\ &\sum_{i=0}^{m-1} \frac{\eta_i}{2}\left(-\frac{1}{k} + \frac{(u_i - \mu(H_{u_{i-1}}, \theta_t))^2}{\mu(H_{u_{i-1}}, \theta_t)^2 \cdot u_i}\right) \end{aligned} \tag{5}$$

For η_j we have:

$$
\begin{aligned}
&\frac{\partial}{\partial \eta_j}\mathcal{L}(u_{t-m:t}|k,\eta_t,\theta_t) = \\
&\frac{\partial}{\partial \eta_j}\left(-\sum_{i=0}^{m-1}\eta_i \log p(u_i|k,\theta_t)\right) = \\
&\color{blue}{-\log p(u_j|k,\theta_t)}
\end{aligned} \tag{6}
$$

For θ_j we have:

$$
\begin{aligned}
&\frac{\partial}{\partial \theta_j}\mathcal{L}(u_{t-m:t}|k,\eta_t,\theta_t) = \\
&\frac{\partial}{\partial \theta_j}\left(-\sum_{i=0}^{m-1}\eta_i \log p(u_i|k,\theta_t)\right) = \\
&\frac{\partial}{\partial \theta_j}\left(-\sum_{i=0}^{m-1}\eta_i \log\left(\left[\frac{k}{2\pi\cdot {u_i}^3}\right]^{1/2} e^{-\frac{1}{2}\frac{k\left(u_i-\mu(H_{u_{i-1}},\theta_t)\right)^2}{\mu(H_{u_{i-1}},\theta_t)^2\cdot u_i}}\right)\right) = \\
&\frac{\partial}{\partial \theta_j}\left(\sum_{i=0}^{m-1}-\frac{1}{2}\eta_i \log\left(\left[\frac{k}{2\pi\cdot {u_i}^3}\right]\right)+\frac{1}{2}\eta_i\frac{k(u_i-\mu(H_{u_{i-1}},\theta_t))^2}{\mu(H_{u_{i-1}},\theta_t)^2\cdot u_i}\right) = \\
&\frac{\partial}{\partial \theta_j}\left(\sum_{i=0}^{m-1}-\frac{1}{2}\eta_i \log(k)+\eta_i\log\left(2\pi\cdot {u_i}^3\right)+\frac{1}{2}\eta_i\frac{k(u_i-\mu(H_{u_{i-1}},\theta_t))^2}{\mu(H_{u_{i-1}},\theta_t)^2\cdot u_i}\right) = \\
&\frac{\partial}{\partial \theta_j}\left(\sum_{i=0}^{m-1}\frac{\eta_i k}{2u_i}\frac{(u_i-\mu(H_{u_{i-1}},\theta_t))^2}{\mu(H_{u_{i-1}},\theta_t)^2}\right) = \\
&\sum_{i=0}^{m-1}\frac{\eta_i k}{2u_i}\cdot\frac{2\left(u_i-\mu(H_{u_{i-1}},\theta_t)\right)(-u_{i-j})\cdot\mu(H_{u_{i-1}},\theta_t)^2-2\mu(H_{u_{i-1}},\theta_t)(u_{i-j})(u_i-\mu(H_{u_{i-1}},\theta_t))^2}{\mu(H_{u_{i-1}},\theta_t)^4} = \\
&\sum_{i=0}^{m-1}\frac{\eta_i k}{u_i}\cdot\frac{-\left(u_i-\mu(H_{u_{i-1}},\theta_t)\right)(u_{i-j})\cdot\mu(H_{u_{i-1}},\theta_t)-(u_{i-j})(u_i-\mu(H_{u_{i-1}},\theta_t))^2}{\mu(H_{u_{i-1}},\theta_t)^3} = \\
&\sum_{i=0}^{m-1}\frac{\eta_i k}{u_i}\cdot\frac{\left(u_i-\mu(H_{u_{i-1}},\theta_t)\right)(u_{i-j})\left(-\mu(H_{u_{i-1}},\theta_t)-u_i+\mu(H_{u_{i-1}},\theta_t)\right)}{\mu(H_{u_{i-1}},\theta_t)^3} = \\
&\sum_{i=0}^{m-1}\frac{\eta_i k}{u_i}\cdot\frac{\left(u_i-\mu(H_{u_{i-1}},\theta_t)\right)(u_{i-j})\left(-\mu(H_{u_{i-1}},\theta_t)-u_i+\mu(H_{u_{i-1}},\theta_t)\right)}{\mu(H_{u_{i-1}},\theta_t)^3} = \\
&-\sum_{i=0}^{m-1}\frac{\eta_i k}{u_i}\cdot\frac{\left(u_i-\mu(H_{u_{i-1}},\theta_t)\right)(u_{i-j})(u_i)}{\mu(H_{u_{i-1}},\theta_t)^3} = \\
&\color{blue}{-\sum_{i=0}^{m-1}\eta_i k\cdot\frac{\left(u_i-\mu(H_{u_{i-1}},\theta_t)\right)(u_{i-j})}{\mu(H_{u_{i-1}},\theta_t)^3}}
\end{aligned} \tag{7}
$$

Since $\mu(H_{u_{i-1}},\theta_t)=\theta_0+\sum_{j=1}^{n-1}\theta_j u_{i-j}$ we note that in case $j=0$, we must have that $u_{i-j}=u_i=1$

(`Note: In code this case is handled by the fact the u[i] = 1`)

Hessian `Matrix`

$$\nabla^2 \mathcal{L}(u_{t-m:t}|k,\eta_t,\theta_t) = \begin{bmatrix} \frac{\partial^2}{\partial^2 k}\mathcal{L}(\cdot) & \frac{\partial}{\partial k\partial\eta_0}\mathcal{L}(\cdot) & \cdots & \frac{\partial}{\partial k\partial\eta_{m-1}}\mathcal{L}(\cdot) & \frac{\partial}{\partial k\partial\theta_0}\mathcal{L}(\cdot) & \cdots & \frac{\partial}{\partial k\partial\theta_{n-1}}\mathcal{L}(\cdot) \\ \cdot & \frac{\partial^2}{\partial^2\eta_0}\mathcal{L}(\cdot) & \cdots & \frac{\partial}{\partial\eta_0\partial\eta_{m-1}}\mathcal{L}(\cdot) & \frac{\partial}{\partial\eta_0\partial\theta_0}\mathcal{L}(\cdot) & \cdots & \frac{\partial}{\partial\eta_0\partial\theta_{n-1}}\mathcal{L}(\cdot) \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \vdots \\ \cdot & \cdot & \cdots & \frac{\partial^2}{\partial^2\eta_{m-1}}\mathcal{L}(\cdot) & \frac{\partial}{\partial\eta_{m-1}\partial\theta_0}\mathcal{L}(\cdot) & \cdots & \frac{\partial}{\partial\eta_{m-1}\partial\theta_{n-1}}\mathcal{L}(\cdot) \\ \cdot & \cdot & \cdots & \cdot & \frac{\partial^2}{\partial^2\theta_0}\mathcal{L}(\cdot) & \cdots & \frac{\partial}{\partial\theta_0\theta_{n-1}}\mathcal{L}(\cdot) \\ \cdot & \cdot & \vdots & \cdot & \cdot & \ddots & \vdots \\ \cdot & \cdot & \cdots & \cdot & \cdot & \cdots & \frac{\partial^2}{\partial^2\theta_{n-1}}\mathcal{L}(\cdot) \end{bmatrix} \tag{8}$$

k

For k we have:

$$\begin{gathered} \frac{\partial^2}{\partial^2 k}\mathcal{L}(u_{t-m:t}|k,\eta_t,\theta_t) = \\ \frac{\partial}{\partial k}\left(\sum_{i=0}^{m-1}\frac{\eta_i}{2}\left(-\frac{1}{k}+\frac{(u_i-\mu(H_{u_{i-1}},\theta_t))^2}{\mu(H_{u_{i-1}},\theta_t)^2\cdot u_i}\right)\right) = \\ \frac{\partial}{\partial k}\left(-\sum_{i=0}^{m-1}\frac{\eta_i}{2}k^{-1}\right) = \\ \color{blue}{\sum_{i=0}^{m-1}\frac{\eta_i}{2}k^{-2}} \end{gathered} \tag{9}$$

$$\begin{gathered} \frac{\partial}{\partial k\partial\eta_j}\mathcal{L}(u_{t-m:t}|k,\eta_t,\theta_t) = \\ \frac{\partial}{\partial\eta_j}\left(\sum_{i=0}^{m-1}\frac{\eta_i}{2}\left(-\frac{1}{k}+\frac{(u_i-\mu(H_{u_{i-1}},\theta_t))^2}{\mu(H_{u_{i-1}},\theta_t)^2\cdot u_i}\right)\right) = \\ \frac{\partial}{\partial\eta_j}\left(\frac{\eta_j}{2}\left(-\frac{1}{k}+\frac{(u_j-\mu(H_{u_{j-1}},\theta_t))^2}{\mu(H_{u_{j-1}},\theta_t)^2\cdot u_j}\right)\right) = \\ \color{blue}{\frac{1}{2}\left(-\frac{1}{k}+\frac{(u_j-\mu(H_{u_{j-1}},\theta_t))^2}{\mu(H_{u_{j-1}},\theta_t)^2\cdot u_j}\right)} \end{gathered} \tag{10}$$

$$\begin{gathered} \frac{\partial}{\partial k\partial\theta_j}\mathcal{L}(u_{t-m:t}|k,\eta_t,\theta_t) = \\ \frac{\partial}{\partial\theta_j}\left(\sum_{i=0}^{m-1}\frac{\eta_i}{2}\left(-\frac{1}{k}+\frac{(u_i-\mu(H_{u_{i-1}},\theta_t))^2}{\mu(H_{u_{i-1}},\theta_t)^2\cdot u_i}\right)\right) = \\ \frac{\partial}{\partial\theta_j}\left(\sum_{i=0}^{m-1}\frac{\eta_i}{2u_i}\frac{(u_i-\mu(H_{u_{i-1}},\theta_t))^2}{\mu(H_{u_{i-1}},\theta_t)^2}\right) = \\ \cdots \\ \color{blue}{-\sum_{i=0}^{m-1}\eta_i\cdot\frac{(u_i-\mu(H_{u_{i-1}},\theta_t))\,(u_{i-j})}{\mu(H_{u_{i-1}},\theta_t)^3}} \end{gathered} \tag{11}$$

η

For η_j we have:

$$
\begin{aligned}
&\frac{\partial}{\partial \eta_j \partial k}\mathcal{L}(u_{t-m:t}|k,\eta_t,\theta_t) = \\
&\frac{\partial}{\partial k}(-\log p(u_j|k,\theta_t)) = \\
&\frac{\partial}{\partial k}\left(-\log\left(\left[\frac{k}{2\pi \cdot {u_j}^3}\right]^{1/2} e^{-\frac{1}{2}\frac{k\left(u_j-\mu(H_{u_{j-1}},\theta_t)\right)^2}{\mu(H_{u_{j-1}},\theta_t)^2 \cdot u_j}}\right)\right) = \\
&\frac{\partial}{\partial k}\left(-\frac{1}{2}\log\left(\frac{k}{2\pi \cdot {u_j}^3}\right) + \frac{1}{2}\frac{k\left(u_j-\mu(H_{u_{j-1}},\theta_t)\right)^2}{\mu(H_{u_{j-1}},\theta_t)^2 \cdot u_j}\right) = \\
&\frac{\partial}{\partial k}\left(-\frac{1}{2}\log k + \frac{1}{2}\log(2\pi \cdot {u_j}^3) + \frac{1}{2}\frac{k\left(u_j-\mu(H_{u_{j-1}},\theta_t)\right)^2}{\mu(H_{u_{j-1}},\theta_t)^2 \cdot u_j}\right) = \\
&\color{purple}{\frac{1}{2}\left(-\frac{1}{k} + \frac{\left(u_j-\mu(H_{u_{j-1}},\theta_t)\right)^2}{\mu(H_{u_{j-1}},\theta_t)^2 \cdot u_j}\right)}
\end{aligned}
\tag{12}
$$

We have that $\frac{\partial}{\partial \eta_j \partial k}\mathcal{L}(\cdot) = \frac{\partial}{\partial k \partial \eta_j}\mathcal{L}(\cdot)$ as *expected*.

$$
\begin{aligned}
&\frac{\partial^2}{\partial^2 \eta_j}\mathcal{L}(u_{t-m:t}|k,\eta_t,\theta_t) = \\
&\frac{\partial}{\partial \eta_j}(-\log p(u_j|k,\theta_t)) = \color{blue}{0}
\end{aligned}
\tag{13}
$$

$$
\begin{aligned}
&\frac{\partial}{\partial \eta_j \eta_q}\mathcal{L}(u_{t-m:t}|k,\eta_t,\theta_t) = \\
&\frac{\partial}{\partial \eta_q}(-\log p(u_j|k,\theta_t)) = \color{blue}{0}
\end{aligned}
\tag{14}
$$

$$
\begin{aligned}
&\frac{\partial}{\partial \eta_j \theta_q}\mathcal{L}(u_{t-m:t}|k,\eta_t,\theta_t) = \\
&\frac{\partial}{\partial \theta_q}(-\log p(u_j|k,\theta_t)) = \\
&\frac{\partial}{\partial \theta_q}\left(-\frac{1}{2}\log k + \frac{1}{2}\log(2\pi \cdot {u_j}^3) + \frac{1}{2}\frac{k\left(u_j-\mu(H_{u_{j-1}},\theta_t)\right)^2}{\mu(H_{u_{j-1}},\theta_t)^2 \cdot u_j}\right) = \\
&\frac{\partial}{\partial \theta_q}\left(\frac{k}{2u_j}\frac{\left(u_j-\mu(H_{u_{j-1}},\theta_t)\right)^2}{\mu(H_{u_{j-1}},\theta_t)^2}\right) = \\
&\color{blue}{-k \cdot \frac{\left(u_j-\mu(H_{u_{j-1}},\theta_t)\right)(u_{j-q})}{\mu(H_{u_{j-1}},\theta_t)^3}}
\end{aligned}
\tag{15}
$$

θ

For θ_j we have:

$$\frac{\partial}{\partial\theta_j\partial k}\mathcal{L}(u_{t-m:t}|k,\eta_t,\theta_t) = \tag{16}$$
$$\frac{\partial}{\partial k}\left(-\sum_{i=0}^{m-1}\eta_i k\cdot\frac{(u_i-\mu(H_{u_{i-1}},\theta_t))\,(u_{i-j})}{\mu(H_{u_{i-1}},\theta_t)^3}\right) =$$
$$-\sum_{i=0}^{m-1}\eta_i\cdot\frac{(u_i-\mu(H_{u_{i-1}},\theta_t))\,(u_{i-j})}{\mu(H_{u_{i-1}},\theta_t)^3}$$

We have that $\frac{\partial}{\partial\theta_j\partial k}\mathcal{L}(\cdot)=\frac{\partial}{\partial k\partial\theta_j}\mathcal{L}(\cdot)$ as *expected*.

$$\frac{\partial}{\partial\theta_j\partial\eta_q}\mathcal{L}(u_{t-m:t}|k,\eta_t,\theta_t) = \tag{17}$$
$$\frac{\partial}{\partial\eta_q}\left(-\sum_{i=0}^{m-1}\eta_i k\cdot\frac{(u_i-\mu(H_{u_{i-1}},\theta_t))\,(u_{i-j})}{\mu(H_{u_{i-1}},\theta_t)^3}\right) =$$
$$-k\cdot\frac{\left(u_q-\mu(H_{u_{q-1}},\theta_t)\right)\left(u_{q-j}\right)}{\mu(H_{u_{q-1}},\theta_t)^3}$$

We have that $\frac{\partial}{\partial\theta_j\partial\eta_q}\mathcal{L}(\cdot)=\frac{\partial}{\partial\eta_q\partial\theta_j}\mathcal{L}(\cdot)$ as *expected*.

$$\frac{\partial^2}{\partial^2\theta_j}\mathcal{L}(u_{t-m:t}|k,\eta_t,\theta_t) = \tag{18}$$
$$\frac{\partial}{\partial\theta_j}\left(-\sum_{i=0}^{m-1}\eta_i k\cdot\left(\frac{(u_i-\mu(H_{u_{i-1}},\theta_t))\cdot(u_{i-j})}{\mu(H_{u_{i-1}},\theta_t)^3}\right)\right) =$$
$$-\sum_{i=0}^{m-1}\eta_i k(u_{i-j})\cdot\frac{\partial}{\partial\theta_j}\left(\frac{(u_i-\mu(H_{u_{i-1}},\theta_t))}{\mu(H_{u_{i-1}},\theta_t)^3}\right) =$$
$$-\sum_{i=0}^{m-1}\eta_i k(u_{i-j})\cdot\left(\frac{(-u_{i-j})\cdot\mu(H_{u_{i-1}},\theta_t)^3-3\mu(H_{u_{i-1}},\theta_t)^2\cdot u_{i-j}\cdot(u_i-\mu(H_{u_{i-1}},\theta_t))}{\mu(H_{u_{i-1}},\theta_t)^6}\right)0$$
$$-\sum_{i=0}^{m-1}\eta_i k(u_{i-j})\cdot\left(\frac{(-u_{i-j})\cdot\mu(H_{u_{i-1}},\theta_t)-3\cdot u_{i-j}\cdot(u_i-\mu(H_{u_{i-1}},\theta_t))}{\mu(H_{u_{i-1}},\theta_t)^4}\right) =$$
$$-\sum_{i=0}^{m-1}\eta_i k(u_{i-j})\cdot\left(\frac{-3\cdot u_{i-j}\cdot u_i+2\cdot u_{i-j}\cdot\mu(H_{u_{i-1}},\theta_t)}{\mu(H_{u_{i-1}},\theta_t)^4}\right) =$$
$$\sum_{i=0}^{m-1}\eta_i k(u_{i-j})^2\cdot\left(\frac{3\cdot u_i-2\cdot\mu(H_{u_{i-1}},\theta_t)}{\mu(H_{u_{i-1}},\theta_t)^4}\right)$$

$$
\begin{aligned}
&\frac{\partial}{\partial\theta_j\partial\theta_q}\mathcal{L}(u_{t-m:t}|k,\eta_t,\theta_t) = \\
&\frac{\partial}{\partial\theta_q}\left(-\sum_{i=0}^{m-1}\eta_i k\cdot\left(\frac{(u_i-\mu(H_{u_{i-1}},\theta_t))\cdot(u_{i-j})}{\mu(H_{u_{i-1}},\theta_t)^3}\right)\right) = \\
&\quad -\sum_{i=0}^{m-1}\eta_i k(u_{i-j})\cdot\frac{\partial}{\partial\theta_q}\left(\frac{(u_i-\mu(H_{u_{i-1}},\theta_t))}{\mu(H_{u_{i-1}},\theta_t)^3}\right) = \\
&\quad \sum_{i=0}^{m-1}\eta_i k(u_{i-j})(u_{i-q})\cdot\left(\frac{3\cdot u_i-2\cdot\mu(H_{u_{i-1}},\theta_t)}{\mu(H_{u_{i-1}},\theta_t)^4}\right)
\end{aligned}
\tag{19}
$$