Technical Note to Accompany Markov Regime-Switching Tests: Asymptotic Critical Values

Douglas G. Steigerwald and Andrew V. Carter

July 22, 2011

In this technical note we present detailed derivations to accompany the covariance calculations for the Gaussian process ($\mathbb{E}\left[\mathcal{G}\left(\theta_0\right)\mathcal{G}\left(\theta_0'\right)\right]$) that appears in the limit distribution of Carter and Steigerwald (2011). We begin with the single equation linear process with homoskedastic Gaussian errors and show that $\mathbb{E}\left[\mathcal{G}\left(\theta_0\right)\mathcal{G}\left(\theta_0'\right)\right]$ is not a function of covariates (regressors). We then derive the covariance for a multiple equation linear process, first with heteroskedastic errors and second with homoskedastic errors.

For the single equation process we parallel the construction in the paper and the calculations (in order) are: 1) construct the quasi-log-likelihood function L_n, 2) calculate the gradient, 3) calculate the information matrix $\mathcal{I}\left(\theta_0\right)$, 4) calculate the asymptotic variance of the score $\mathcal{V}\left(\theta_0\right)$ and 5) calculate the covariance of the score $\mathbb{E}\left[\mathcal{S}\left(\theta_0\right)\mathcal{S}\left(\theta_0'\right)\right]$.

1 Single Equation

The process is

$$Y_t = \theta_0 + \delta S_t + Z_t'\beta + U_t, \tag{1}$$

with $U_t \sim i.i.d.N\left(0,\nu\right)$. In what follows, we take Z_t to be a scalar for convenience.

1.1 Calculation of L_n

The quasi-log-likelihood for observation t is

$$\ln\left[(1-\pi)\tfrac{1}{\sqrt{c\nu}}\exp\tfrac{(y_t-\beta z_t-\theta_0)^2}{2\nu} + \pi\tfrac{1}{\sqrt{c\nu}}\exp\tfrac{(y_t-\beta z_t-\theta_1)^2}{2\nu}\right],$$

where $c = 2\cdot pi$. The resulting quasi-log-likelihood function, $L_n\left(\pi,\gamma,\theta_0,\theta_1\right)$ where $\gamma = (\nu,\beta)$, is

$$\sum_{t=1}^{n}\log\left[(1-\pi)\exp\left(\frac{2\theta_0\left(y_t-\beta z_t\right)-\theta_0^2}{2\nu}\right)+\pi\exp\left(\frac{2\theta_1\left(y_t-\beta z_t\right)-\theta_1^2}{2\nu}\right)\right]-\frac{n}{2}\log\left(c\nu\right)-\frac{1}{2\nu}\sum_{t=1}^{n}\left(y_t-\beta z_t\right)^2.$$

1.2 Calculation of the Gradient of L_n

The gradient for the quasi-log-likelihood is[1]

$$\frac{\partial}{\partial\pi}L_n\left(\pi,\gamma,\theta_0,\theta_1\right)=\sum_{t=1}^{n}\frac{\exp\left(\frac{2\theta_1(y_t-\beta z_t)-\theta_1^2}{2\nu}\right)-\exp\left(\frac{2\theta_0(y_t-\beta z_t)-\theta_0^2}{2\nu}\right)}{(1-\pi)\exp\left(\frac{2\theta_0(y_t-\beta z_t)-\theta_0^2}{2\nu}\right)+\pi\exp\left(\frac{2\theta_1(y_t-\beta z_t)-\theta_1^2}{2\nu}\right)}$$

$$\begin{aligned}&\frac{\partial}{\partial\nu}L_n\left(\pi,\gamma,\theta_0,\theta_1\right)\\=\quad&\sum_{t=1}^{n}\frac{-\frac{\pi\left(2(y_t-\beta z_t)\theta_1-\theta_1^2\right)}{2\nu^2}\exp\left(\frac{2\theta_1(y_t-\beta z_t)-\theta_1^2}{2\nu}\right)-\frac{(1-\pi)\left(2(y_t-\beta z_t)\theta_0-\theta_0^2\right)}{2\nu^2}\exp\left(\frac{2\theta_0(y_t-\beta z_t)-\theta_0^2}{2\nu}\right)}{(1-\pi)\exp\left(\frac{2\theta_0(y_t-\beta z_t)-\theta_0^2}{2\nu}\right)+\pi\exp\left(\frac{2\theta_1(y_t-\beta z_t)-\theta_1^2}{2\nu}\right)}\\&-\frac{n}{2\nu}+\frac{1}{2\nu^2}\sum_{t=1}^{n}\left(y_t-\beta z_t\right)^2.\end{aligned}$$

$$\begin{aligned}\frac{\partial}{\partial\beta}L_n\left(\pi,\gamma,\theta_0,\theta_1\right)\quad=\quad&\sum_{t=1}^{n}\frac{-\frac{\pi(\theta_1 z_t)}{\nu}\exp\left(\frac{2\theta_1(y_t-\beta z_t)-\theta_1^2}{2\nu}\right)-\frac{(1-\pi)(\theta_0 z_t)}{\nu}\exp\left(\frac{2\theta_0(y_t-\beta z_t)-\theta_0^2}{2\nu}\right)}{(1-\pi)\exp\left(\frac{2\theta_0(y_t-\beta z_t)-\theta_0^2}{2\nu}\right)+\pi\exp\left(\frac{2\theta_1(y_t-\beta z_t)-\theta_1^2}{2\nu}\right)}\\&+\frac{1}{\nu}\sum_{t=1}^{n}w_t\left(y_t-\beta z_t\right).\end{aligned}$$

$$\frac{\partial}{\partial\theta_1}L_n\left(\pi,\gamma,\theta_0,\theta_1\right)=\sum_{t=1}^{n}\frac{\pi\frac{(y_t-\beta z_t-\theta_1)}{\nu}\exp\left(\frac{2\theta_1(y_t-\beta z_t)-\theta_1^2}{2\nu}\right)}{(1-\pi)\exp\left(\frac{2\theta_0(y_t-\beta z_t)-\theta_0^2}{2\nu}\right)+\pi\exp\left(\frac{2\theta_1(y_t-\beta z_t)-\theta_1^2}{2\nu}\right)}$$

Evaluating the gradient at $(1,\gamma,\theta_0,\theta_*)$

The evaluated gradient, which has expectation zero, contains

$$\frac{\partial}{\partial\pi}L_n\left(1,\gamma,\theta_0,\theta_*\right)=\sum_{t=1}^{n}\left(1-e^{b_t}\right)\quad\text{where } b_t=(y_t-\beta z_t)\cdot\left(\frac{\theta_0-\theta_*}{\nu}\right)-\left(\frac{\theta_0^2-\theta_*^2}{2\nu}\right),$$

$$\frac{\partial}{\partial\nu}L_n\left(1,\gamma,\theta_0,\theta_*\right)=\sum_{t=1}^{n}\frac{\left(y_t-\beta z_t-\theta_*\right)^2}{2\nu^2}-\frac{n}{2\nu},$$

$$\frac{\partial}{\partial\beta}L_n\left(1,\gamma,\theta_0,\theta_*\right)=\sum_{t=1}^{n}\frac{z_t\left(y_t-\beta z_t-\theta_*\right)}{\nu},$$

$$\frac{\partial}{\partial\theta_1}L_n\left(1,\gamma,\theta_0,\theta_*\right)=\sum_{t=1}^{n}\frac{y_t-\beta z_t-\theta_*}{\nu}.$$

[1] If the derivatives are taken at $(0,\gamma,\theta_*,\theta_1)$ then the derivative with respect to π changes sign and the other derivatives are unchanged.

Evaluating the Behavior of e^b

In what follows, we drop the subscript t on (y, z) and use capital letters for the random variables (Y, Z). For the information matrix calculations, we need to know the behavior of four moments involving e^b. We first calculate

$$\mathbb{E}\left[e^b\right] = \mathbb{E}\left[e^{(Y-\beta Z)\cdot t - \frac{1}{2\nu}\left(\theta_0^2 - \theta_*^2\right)}\right] = \mathbb{E}\left[e^{(Y-\beta Z)\cdot t}\right] \cdot e^{-\frac{1}{2\nu}\left(\theta_0^2-\theta_*^2\right)}$$

where $t = \left(\frac{\theta_0 - \theta_*}{\nu}\right)$. Under H_0, $(Y - \beta Z) \sim N(\theta_*, \nu)$. From the definition of the moment generating function for a Gaussian random variable, for an arbitrary real number t

$$\begin{aligned}
\mathbb{E}\left[e^{(Y-\beta Z)\cdot t}\right] &= e^{\theta_* t + \frac{1}{2}\nu t^2} \\
&= \exp\left[\theta_*\left(\frac{\theta_0-\theta_*}{\nu}\right) + \frac{1}{2}\nu\left(\frac{\theta_0-\theta_*}{\nu}\right)^2\right] \\
&= \exp\left[\frac{1}{2\nu}\left(\theta_0^2 - \theta_*^2\right)\right].
\end{aligned}$$

Thus

$$\mathbb{E}\left[e^b\right] = 1.$$

We next calculate $\mathbb{E}\left[e^{2b}\right]$. Note

$$2b = (Y - \beta Z) \cdot 2t - \frac{\theta_0^2 - \theta_*^2}{\nu},$$

and

$$\begin{aligned}
\mathbb{E}\left[e^{2b}\right] &= \exp\left\{\theta_* \cdot 2t + \frac{1}{2}\nu\,(2t)^2 - \frac{\theta_0^2-\theta_*^2}{\nu}\right\} \\
&= \exp\left\{\frac{1}{\nu}\left[2\theta_0\left(\theta_0-\theta_*\right) - \theta_0^2 + \theta_*^2\right]\right\} \\
&= \exp\left[\frac{1}{\nu}\left(\theta_0-\theta_*\right)^2\right].
\end{aligned}$$

We next analyze

$$\mathbb{E}\left[e^b\left(Y - \beta Z - \theta_*\right)^2\right] = \int \left(y - \beta z - \theta_*\right)^2 e^b c e^{-\frac{1}{2\nu}\left(y-\beta z-\theta_*\right)^2} dy,$$

where $c = (2pi \cdot \nu)^{-\frac{1}{2}}$. Note $e^b \cdot e^{-\frac{1}{2\nu}(y-\beta z-\theta_*)^2}$ has exponent

$$\begin{aligned}
&\left(y-\beta z\right)\left(\frac{\theta_0-\theta_*}{\nu}\right) - \frac{1}{2}\frac{\theta_0^2-\theta_*^2}{\nu} - \frac{1}{2\nu}\left(y-\beta z-\theta_*\right)^2 \\
=\ & -\frac{1}{2\nu}\left(y-\beta z-\theta_0\right)^2.
\end{aligned}$$

Further, $(y-\beta z-\theta_*)^2 = (y-\beta z-\theta_0)^2 + 2(\theta_0-\theta_*)(y-\beta z-\theta_0) + (\theta_0-\theta_*)^2$. Hence

$$\begin{aligned}
\mathbb{E}\left[e^b (Y-\beta Z-\theta_*)^2\right] &= \int (y-\beta z-\theta_0)^2 ce^{-\frac{1}{2\nu}(y-\beta z-\theta_0)^2} dy \\
&\quad + 2(\theta_0-\theta_*) \int (y-\beta z-\theta_0) ce^{-\frac{1}{2\nu}(y-\beta z-\theta_0)^2} dy \\
&\quad + (\theta_0-\theta_*)^2 \int ce^{-\frac{1}{2\nu}(y-\beta z-\theta_0)^2} dy \\
&= \nu + (\theta_0-\theta_*)^2 .
\end{aligned}$$

Finally, we analyze

$$\begin{aligned}
\mathbb{E}\left[e^b (Y-\beta Z)\right] &= \int (y-\beta z) ce^{-\frac{1}{2\nu}(y-\beta z-\theta_0)^2} dy \\
&= \theta_0.
\end{aligned}$$

1.3 Calculation of the Information Matrix $\mathcal{I}(\theta_0)$

The information matrix, $\mathcal{I}(\theta_0)$ is[2]

$$\begin{bmatrix}
e^{\frac{1}{\nu}(\theta_0-\theta_*)^2} - 1 & -\frac{1}{2\nu^2}(\theta_0-\theta_*)^2 & -\frac{\mathbb{E}Z}{\nu}(\theta_0-\theta_*) & -\frac{1}{\nu}(\theta_0-\theta_*) \\
. & \frac{1}{2\nu^2} & 0 & 0 \\
. & . & \frac{\mathbb{E}[Z^2]}{\nu} & \frac{\mathbb{E}[Z]}{\nu} \\
. & . & . & \frac{1}{\nu}
\end{bmatrix}$$

(1,1) Element

The (1,1) element of $\mathcal{I}(\theta_0)$ corresponds to $\mathbb{E}\left[\left(\frac{\partial}{\partial\pi} l_t\right)^2\right]$. We have

$$\left(\frac{\partial}{\partial\pi} l_t\right)^2 = 1 - 2e^b + e^{2b}.$$

Because $\mathbb{E}\left[e^b\right] = 1$ and $\mathbb{E}\left[e^{2b}\right] = \exp\left[\frac{1}{\nu}(\theta_0-\theta_*)^2\right]$:

$$\mathbb{E}\left[\left(\frac{\partial}{\partial\pi} l_t\right)^2\right] = \exp\left[\frac{1}{\nu}(\theta_0-\theta_*)^2\right] - 1.$$

(1,2) Element

The (1,2) element of $\mathcal{I}(\theta_0)$ corresponds to $\mathbb{E}\left[\left(\frac{\partial}{\partial\pi} l_t\right)\left(\frac{\partial}{\partial\nu} l_t\right)\right]$. We have

$$\left(\frac{\partial}{\partial\pi} l_t\right)\left(\frac{\partial}{\partial\nu} l_t\right) = \left(1-e^b\right) \cdot \frac{1}{2\nu^2}\left[(Y-\beta Z-\theta_*)^2 - \nu\right].$$

[2] In Drew Carter's notes, the (1,2) and (1,3) elements have a positive sign. This follows from the fact the he evaluates derivatives at the point $(0, \gamma, \theta_*, \theta_1)$, which (as noted above) changes the sign of the derivative with respect to π.

Because $\mathbb{E}\left[(Y-\beta Z-\theta_*)^2\right]=\nu$,

$$\mathbb{E}\left[\left(\frac{\partial}{\partial \pi}l_t\right)\left(\frac{\partial}{\partial \nu}l_t\right)\right]=-\frac{1}{2\nu^2}\mathbb{E}\left[e^b\left[(Y-\beta Z-\theta_*)^2-\nu\right]\right].$$

Because $\mathbb{E}\left[e^b\right]=1$ and $\mathbb{E}\left[e^b\,(Y-\beta Z-\theta_*)^2\right]=\nu+(\theta_0-\theta_*)^2$,

$$\mathbb{E}\left[\left(\frac{\partial}{\partial \pi}l_t\right)\left(\frac{\partial}{\partial \nu}l_t\right)\right]=-\frac{1}{2\nu^2}\left(\theta_0-\theta_*\right)^2.$$

(1,3) Element

The (1,3) element of $\mathcal{I}(\theta_0)$ corresponds to $\mathbb{E}\left[\left(\frac{\partial}{\partial\pi}l_t\right)\left(\frac{\partial}{\partial\beta}l_t\right)\right]$. We have

$$\left(\frac{\partial}{\partial\pi}l_t\right)\left(\frac{\partial}{\partial\beta}l_t\right) = \left(1-e^b\right)\cdot\frac{Z\left(Y-\beta Z-\theta_*\right)}{\nu}.$$

Because $\mathbb{E}\left(Y-\beta Z|Z\right)=\theta_*$

$$\begin{aligned}\mathbb{E}\left[\left(\frac{\partial}{\partial\pi}l_t\right)\left(\frac{\partial}{\partial\theta_1}l_t\right)\right] &= -\mathbb{E}\left[e^b\cdot\frac{Z\left(Y-\beta Z-\theta_*\right)}{\nu}\right]\\ &= -\frac{\mathbb{E}Z}{\nu}\mathbb{E}\left[e^b\left(Y-\beta Z\right)-e^b\theta_*\right].\end{aligned}$$

Because $\mathbb{E}\left[e^b\right]=1$ and $\mathbb{E}\left[e^b\left(Y-\beta Z\right)\right]=\theta_0$,

$$\mathbb{E}\left[\left(\frac{\partial}{\partial\pi}l_t\right)\left(\frac{\partial}{\partial\beta}l_t\right)\right] = -\frac{\mathbb{E}Z}{\nu}\left(\theta_0-\theta_*\right).$$

(1,4) Element

The (1,4) element of $\mathcal{I}(\theta_0)$ corresponds to $\mathbb{E}\left[\left(\frac{\partial}{\partial\pi}l_t\right)\left(\frac{\partial}{\partial\theta_1}l_t\right)\right]$. We have

$$\left(\frac{\partial}{\partial\pi}l_t\right)\left(\frac{\partial}{\partial\theta_1}l_t\right) = \left(1-e^b\right)\cdot\frac{\left(Y-\beta Z-\theta_*\right)}{\nu}.$$

Because $\mathbb{E}\left(Y-\beta Z\right)=\theta_*$

$$\begin{aligned}\mathbb{E}\left[\left(\frac{\partial}{\partial\pi}l_t\right)\left(\frac{\partial}{\partial\theta_1}l_t\right)\right] &= -\mathbb{E}\left[e^b\cdot\frac{\left(Y-\beta Z-\theta_*\right)}{\nu}\right]\\ &= -\frac{1}{\nu}\mathbb{E}\left[e^b\left(Y-\beta Z\right)-e^b\theta_*\right].\end{aligned}$$

Because $\mathbb{E}\left[e^b\right]=1$ and $\mathbb{E}\left[e^b\left(Y-\beta Z\right)\right]=\theta_0$,

$$\mathbb{E}\left[\left(\frac{\partial}{\partial\pi}l_t\right)\left(\frac{\partial}{\partial\theta_1}l_t\right)\right] = -\frac{1}{\nu}\left(\theta_0-\theta_*\right).$$

(2,2) Element

The (2,2) element of $\mathcal{I}(\theta_0)$ corresponds to $\mathbb{E}\left[\left(\frac{\partial}{\partial\nu}l_t\right)^2\right]$. We have

$$\left(\frac{\partial}{\partial\nu}l_t\right)^2 = \frac{1}{4\nu^4}\left\{\left(Y-\beta Z-\theta_*\right)^4+\nu^2-2\left(Y-\beta Z-\theta_*\right)^2\nu\right\}.$$

Because $\mathbb{E}\left[\left(Y-\beta Z-\theta_*\right)^2\right]=\nu$,

$$\begin{aligned}\mathbb{E}\left[\left(\frac{\partial}{\partial\nu}l_t\right)^2\right] &= \frac{1}{4\nu^4}\left\{\mathbb{E}\left[\left(Y-\beta Z-\theta_*\right)^4\right]-\nu^2\right\}\\ &= \frac{1}{4\nu^4}\left\{3\nu^2-\nu^2\right\}=\frac{1}{2\nu^2}.\end{aligned}$$

(2,3) Element

The (2,3) element of $\mathcal{I}(\theta_0)$ corresponds to $\mathbb{E}\left[\left(\frac{\partial}{\partial\nu}l_t\right)\left(\frac{\partial}{\partial\beta}l_t\right)\right]$. We have

$$\begin{aligned}\left(\frac{\partial}{\partial\nu}l_t\right)\left(\frac{\partial}{\partial\beta}l_t\right) &= \frac{1}{2\nu^2}\left\{(Y-\beta Z-\theta_*)^2-\nu\right\}\cdot\frac{Z}{\nu}(Y-\beta Z-\theta_*)\\ &= \frac{1}{2\nu^3}\left\{Z(Y-\beta Z-\theta_*)^3-Z(Y-\beta Z-\theta_*)\nu\right\}.\end{aligned}$$

Because $\mathbb{E}\left[(Y-\beta Z-\theta_*)^3\right]=0$,

$$\mathbb{E}\left[\left(\frac{\partial}{\partial\nu}l_t\right)\left(\frac{\partial}{\partial\beta}l_t\right)\right]=0.$$

(2,4) Element

The (2,4) element of $\mathcal{I}(\theta_0)$ corresponds to $\mathbb{E}\left[\left(\frac{\partial}{\partial\nu}l_t\right)\left(\frac{\partial}{\partial\theta_1}l_t\right)\right]$. We have

$$\begin{aligned}\left(\frac{\partial}{\partial\nu}l_t\right)\left(\frac{\partial}{\partial\theta_1}l_t\right) &= \frac{1}{2\nu^2}\left\{(Y-\beta Z-\theta_*)^2-\nu\right\}\cdot\frac{1}{\nu}(Y-\beta Z-\theta_*)\\ &= \frac{1}{2\nu^3}\left\{(Y-\beta Z-\theta_*)^3-(Y-\beta Z-\theta_*)\nu\right\}.\end{aligned}$$

Because $\mathbb{E}\left[(Y-\beta Z-\theta_*)^3\right]=0$,

$$\mathbb{E}\left[\left(\frac{\partial}{\partial\nu}l_t\right)\left(\frac{\partial}{\partial\theta_1}l_t\right)\right]=0.$$

(3,3) Element

The (3,3) element of $\mathcal{I}(\theta_0)$ corresponds to $\mathbb{E}\left[\left(\frac{\partial}{\partial\beta}l_t\right)^2\right]$. We have

$$\begin{aligned}\mathbb{E}\left[\left(\frac{\partial}{\partial\beta}l_t\right)^2\right] &= \mathbb{E}\left[Z^2\left(\frac{Y-\beta Z-\theta_*}{\nu}\right)^2\right]\\ &= \frac{\mathbb{E}\left[Z^2\right]}{\nu}.\end{aligned}$$

(3,4) Element

The (3,3) element of $\mathcal{I}(\theta_0)$ corresponds to $\mathbb{E}\left[\left(\frac{\partial}{\partial\beta}l_t\right)\left(\frac{\partial}{\partial\theta_1}l_t\right)\right]$. We have

$$\begin{aligned}\mathbb{E}\left[\left(\frac{\partial}{\partial\beta}l_t\right)\left(\frac{\partial}{\partial\theta_1}l_t\right)\right] &= \mathbb{E}\left[Z\left(\frac{Y-\beta Z-\theta_*}{\nu}\right)^2\right]\\ &= \frac{\mathbb{E}\left[Z\right]}{\nu}.\end{aligned}$$

(4,4) Element

The (4,4) element of $\mathcal{I}(\theta_0)$ corresponds to $\mathbb{E}\left[\left(\frac{\partial}{\partial\theta_1}l_t\right)^2\right]$. We have

$$\begin{aligned}\mathbb{E}\left[\left(\frac{\partial}{\partial\theta_1}l_t\right)^2\right] &= \mathbb{E}\left[\left(\frac{Y-\beta Z-\theta_*}{\nu}\right)^2\right]\\ &= \frac{1}{\nu}.\end{aligned}$$

1.4 Calculation of the Asymptotic Variance $\mathcal{V}(\theta_0)$

First

$$\mathcal{I}^2 = \begin{bmatrix} 2\nu^2 & 0 & 0 \\ \cdot & \frac{\nu}{Var(Z)} & -\frac{\nu\mathbb{E}[Z]}{Var(Z)} \\ \cdot & \cdot & \frac{\nu\mathbb{E}[Z^2]}{Var(Z)} \end{bmatrix}.$$

Second, $\mathcal{I}_1(\theta_0)\mathcal{I}^2(\theta_0)\mathcal{I}_1(\theta_0)^{\mathrm{T}}$ equals

$$\begin{aligned}&\begin{bmatrix} -\frac{(\theta_0-\theta_*)^2}{2\nu^2} & -\frac{\mathbb{E}[Z](\theta_0-\theta_*)}{\nu} & -\frac{\theta_0-\theta_*}{\nu}\end{bmatrix}\mathcal{I}^2\begin{bmatrix} -\frac{(\theta_0-\theta_*)^2}{2\nu^2} \\ -\frac{\mathbb{E}[Z](\theta_0-\theta_*)}{\nu} \\ -\frac{\theta_0-\theta_*}{\nu}\end{bmatrix}\\ =\; & \frac{(\theta_0-\theta_*)^2}{\nu}+\frac{(\theta_0-\theta_*)^4}{2\nu^2}.\end{aligned}$$

Hence

$$\begin{aligned}\mathcal{V}(\theta_0) &= \left(\mathcal{I}_{11}(\theta_0)-\mathcal{I}_1(\theta_0)\mathcal{I}^2(\theta_0)\mathcal{I}_1(\theta_0)^{\mathrm{T}}\right)^{-1}\\ &= \left(e^{\frac{1}{\nu}(\theta_0-\theta_*)^2}-1-\frac{(\theta_0-\theta_*)^2}{\nu}-\frac{(\theta_0-\theta_*)^4}{2\nu^2}\right)^{-1}.\end{aligned}$$

1.5 Calculation of $\mathbb{E}[\mathcal{S}(\theta_0)\mathcal{S}(\theta_0')]$

Because $\mathbb{E}\left[\mathcal{S}(\theta_0)\mathcal{S}(\theta_0')\right]$ is the (1,1) element of $\mathcal{I}(\theta_0')^{-1}\mathcal{I}(\theta_0,\theta_0')\mathcal{I}(\theta_0)^{-1}$, we begin with $\mathcal{I}(\theta_0)^{-1}$.

1.5.1 Calculation of $\mathcal{I}(\theta_0)^{-1}$

The formula for $\mathcal{I}(\theta_0)^{-1}$ comes from Harville (1997, p. 99). Recall

$$\mathcal{I}(\theta_0) = \begin{bmatrix} \mathcal{I}_{11} & \mathcal{I}_1 \\ \mathcal{I}_1^{\mathrm{T}} & \mathcal{I}_2 \end{bmatrix}$$

$$\mathcal{I}(\theta_0)^{-1} = \begin{bmatrix} 0 & \cdots & 0 \\ \vdots & & \mathcal{I}^2 \end{bmatrix} + \begin{bmatrix} 1 \\ -\mathcal{I}^2\mathcal{I}_1^{\mathrm{T}} \end{bmatrix}\mathcal{V}(\theta_0)\begin{bmatrix} 1 \\ -\mathcal{I}^2\mathcal{I}_1^{\mathrm{T}} \end{bmatrix}^{\mathrm{T}}$$

$$= \begin{bmatrix} 0 & 0 & 0 & 0 \\ \cdot & 2\nu^2 & 0 & 0 \\ \cdot & \cdot & \frac{\nu}{Var(Z)} & -\frac{\nu\mathbb{E}[Z]}{Var(Z)} \\ \cdot & \cdot & \cdot & \frac{\nu\mathbb{E}\left[Z^2\right]}{Var(Z)} \end{bmatrix} + \mathcal{V}(\theta_0) \begin{bmatrix} 1 & (\theta_0-\theta_*)^2 & 0 & (\theta_0-\theta_*) \\ \cdot & (\theta_0-\theta_*)^4 & 0 & (\theta_0-\theta_*)^3 \\ \cdot & \cdot & 0 & 0 \\ \cdot & \cdot & \cdot & (\theta_0-\theta_*)^2 \end{bmatrix}$$

$$= \mathcal{V}(\theta_0) \begin{bmatrix} 1 & (\theta_0-\theta_*)^2 & 0 & (\theta_0-\theta_*) \\ \cdot & 2\nu^2+(\theta_0-\theta_*)^4 & 0 & (\theta_0-\theta_*)^3 \\ \cdot & \cdot & \frac{\nu}{Var(Z)} & -\frac{\nu\mathbb{E}[Z]}{Var(Z)} \\ \cdot & \cdot & \cdot & \frac{\nu\mathbb{E}\left[Z^2\right](\theta_0-\theta_*)^2}{Var(Z)} \end{bmatrix}.$$

1.5.2 Calculation of $\mathbb{E}\left[\mathcal{S}(\theta_0)\,\mathcal{S}(\theta_0')\right]$

We have

$$\mathcal{I}(\theta_0,\theta_0') = \begin{bmatrix} e^{\frac{1}{\nu}(\theta_0-\theta_*)(\theta_0'-\theta_*)}-1 & -\frac{1}{2\nu^2}(\theta_0'-\theta_*)^2 & -\frac{\mathbb{E}Z}{\nu}(\theta_0'-\theta_*) & -\frac{1}{\nu}(\theta_0'-\theta_*) \\ -\frac{(\theta_0-\theta_*)^2}{2\nu^2} & \frac{1}{2\nu^2} & 0 & 0 \\ -\frac{\mathbb{E}Z}{\nu}(\theta_0-\theta_*) & 0 & \frac{\mathbb{E}\left[Z^2\right]}{\nu} & \frac{\mathbb{E}[Z]}{\nu} \\ -\frac{\theta_0-\theta_*}{\nu} & 0 & \frac{\mathbb{E}[Z]}{\nu} & \frac{1}{\nu} \end{bmatrix}.$$

Hence $\mathbb{E}\left[\mathcal{S}(\theta_0)\,\mathcal{S}(\theta_0')\right]$ equals

$$\begin{aligned} &\mathcal{I}^{11}(\theta_0)\left(\mathcal{I}^{11}(\theta_0')\,\mathcal{I}_{11}(\theta_0,\theta_0')+\mathcal{I}^{12}(\theta_0')\,\mathcal{I}_{21}(\theta_0,\theta_0')+\mathcal{I}^{13}(\theta_0')\,\mathcal{I}_{31}(\theta_0,\theta_0')\right) \\ &+\mathcal{I}^{21}(\theta_0)\left(\mathcal{I}^{11}(\theta_0')\,\mathcal{I}_{12}(\theta_0,\theta_0')+\mathcal{I}^{12}(\theta_0')\,\mathcal{I}_{22}(\theta_0,\theta_0')+\mathcal{I}^{13}(\theta_0')\,\mathcal{I}_{32}(\theta_0,\theta_0')\right) \\ &+\mathcal{I}^{31}(\theta_0)\left(\mathcal{I}^{11}(\theta_0')\,\mathcal{I}_{13}(\theta_0,\theta_0')+\mathcal{I}^{12}(\theta_0')\,\mathcal{I}_{23}(\theta_0,\theta_0')+\mathcal{I}^{13}(\theta_0')\,\mathcal{I}_{33}(\theta_0,\theta_0')\right) \\ &+\mathcal{I}^{41}(\theta_0)\left(\mathcal{I}^{11}(\theta_0')\,\mathcal{I}_{14}(\theta_0,\theta_0')+\mathcal{I}^{12}(\theta_0')\,\mathcal{I}_{24}(\theta_0,\theta_0')+\mathcal{I}^{13}(\theta_0')\,\mathcal{I}_{34}(\theta_0,\theta_0')\right), \end{aligned}$$

which in turn equals

$$\mathcal{V}(\theta_0)\,\mathcal{V}(\theta_0')\cdot(A+B+C+D),$$

where

$$\begin{aligned} A &= \mathcal{I}_{11}(\theta_0,\theta_0')+(\theta_0'-\theta_*)^2\,\mathcal{I}_{21}(\theta_0,\theta_0')+0\cdot\mathcal{I}_{31}(\theta_0,\theta_0')+(\theta_0'-\theta_*)\,\mathcal{I}_{41}(\theta_0,\theta_0') \\ &= e^{\frac{1}{\upsilon}(\theta_0-\theta_*)(\theta_0'-\theta_*)}-1-(\theta_0'-\theta_*)^2\,\frac{(\theta_0-\theta_*)^2}{2\nu^2}-\frac{(\theta_0'-\theta_*)(\theta_0-\theta_*)}{\nu}, \end{aligned}$$

$$\begin{aligned} B &= (\theta_0-\theta_*)^2\left[\mathcal{I}_{12}(\theta_0,\theta_0')+(\theta_0'-\theta_*)^2\,\mathcal{I}_{22}(\theta_0,\theta_0')+(\theta_0'-\theta_*)\,\mathcal{I}_{42}(\theta_0,\theta_0')\right] \\ &= (\theta_0-\theta_*)^2\left[-\frac{(\theta_0'-\theta_*)^2}{2\nu^2}+\frac{(\theta_0'-\theta_*)^2}{2\nu^2}\right]=0, \end{aligned}$$

$$\begin{aligned} C &= 0\cdot\left[\mathcal{I}_{13}(\theta_0,\theta_0')+(\theta_0'-\theta_*)^2\,\mathcal{I}_{23}(\theta_0,\theta_0')+(\theta_0'-\theta_*)\,\mathcal{I}_{43}(\theta_0,\theta_0')\right] \\ &= 0, \end{aligned}$$

$$
\begin{aligned}
D &= (\theta_0 - \theta_*) \left[\mathcal{I}_{14}\left(\theta_0, \theta_0'\right) + \left(\theta_0' - \theta_*\right)^2 \mathcal{I}_{24}\left(\theta_0, \theta_0'\right) + \left(\theta_0' - \theta_*\right) \mathcal{I}_{34}\left(\theta_0, \theta_0'\right) \right] \\
&= (\theta_0 - \theta_*) \left[-\frac{\left(\theta_0' - \theta_*\right)}{\nu} + \frac{\left(\theta_0' - \theta_*\right)}{\nu} \right] = 0.
\end{aligned}
$$

1.5.3 Calculation of $\mathbb{E}\left[\mathcal{G}(\theta_0)\,\mathcal{G}\left(\theta_0'\right)\right]$

We have that $\mathbb{E}\left[\mathcal{G}(\theta_0)\,\mathcal{G}\left(\theta_0'\right)\right]$ equals

$$
\mathcal{V}(\theta_0)^{\frac{1}{2}}\,\mathcal{V}\left(\theta_0'\right)^{\frac{1}{2}} \cdot (A + B + C + D)\,.
$$

If we note that the covariance is indexed by $\eta = \frac{(\theta_0 - \theta_*)}{\sqrt{\nu}}$ and $\eta' = \frac{\left(\theta_0' - \theta_*\right)}{\sqrt{\nu}}$, we have

$$
\begin{aligned}
&\left[e^{\eta^2} - 1 - \eta^2 - \frac{\eta^4}{2} \right]^{-\frac{1}{2}} \left[e^{\left(\eta'\right)^2} - 1 - \left(\eta'\right)^2 - \left(\eta'\right)^4 \right]^{-\frac{1}{2}} \times \\
&\times e^{\eta\eta'} - 1 - \eta\eta' - \eta^2 \left(\eta'\right)^2 .
\end{aligned}
$$

2 Multiple Equations

The general structure we consider is for n independent observations on a multivariate normal $Y_t = \left(Y_{t1}, \ldots, Y_{td}\right)'$ with mean $\mu \in \mathbb{R}^d$ and covariance matrix Ω^{-1}. The density function is

$$
f\left(Y_t|\mu, \Omega\right) = (2pi)^{\frac{d}{2}} \left|\Omega\right|^{\frac{d}{2}} \exp\left[-\frac{1}{2} \left(Y_t - \mu\right)^{\mathrm{T}} \Omega \left(Y_t - \mu\right) \right],
$$

The mixture arises from a second model with mean $\mu + \delta$, so

$$
\frac{f\left(Y_t|\mu + \delta, \Omega\right)}{f\left(Y_t|\mu, \Omega\right)} = \exp\left[\delta^{\mathrm{T}} \Omega \left(Y_t - \mu\right) - \frac{1}{2} \delta^{\mathrm{T}} \Omega \delta \right].
$$

The covariance $\mathbb{E}\left[\mathcal{G}(\delta_1)\,\mathcal{G}(\delta_2)\right]$ is given by

$$
\begin{aligned}
&\left[\mathcal{I}_{\pi\pi}(\delta_1) - \mathcal{I}_{\pi\mu}(\delta_1)^{\mathrm{T}} \mathcal{I}_{\mu\mu}^{-1} \mathcal{I}_{\pi\mu}(\delta_1) \right]^{-\frac{1}{2}} \left[\mathcal{I}_{\pi\pi}(\delta_2) - \mathcal{I}_{\pi\mu}(\delta_2)^{\mathrm{T}} \mathcal{I}_{\mu\mu}^{-1} \mathcal{I}_{\pi\mu}(\delta_2) \right]^{-\frac{1}{2}} \\
&\times \left[\mathbb{E}\left(\frac{f\left(Y_t|\mu + \delta_1, \Omega\right) f\left(Y_t|\mu + \delta_2, \Omega\right)}{f\left(Y_t|\mu, \Omega\right)^2} \right) - 1 - \mathcal{I}_{\pi\mu}(\delta_1)^{\mathrm{T}} \mathcal{I}_{\mu\mu}^{-1} \mathcal{I}_{\pi\mu}(\delta_2) \right].
\end{aligned} \tag{2}
$$

2.1 Shifted Likelihood

The first term in the covariance that we calculate is the expected value of the shifted likelihood

$$
\mathbb{E}\left(\frac{f\left(Y_t|\mu + \delta_1, \Omega\right) f\left(Y_t|\mu + \delta_2, \Omega\right)}{f\left(Y_t|\mu, \Omega\right)^2} \right) - 1.
$$

We have

$$\mathbb{E}\left(\frac{f\left(Y_t|\mu+\delta_1,\Omega\right)f\left(Y_t|\mu+\delta_2,\Omega\right)}{f\left(Y_t|\mu,\Omega\right)^2}\right) = \mathbb{E}\exp\left[\left(\delta_1+\delta_2\right)^{\mathrm{T}}\Omega\left(Y_t-\mu\right)-\frac{1}{2}\delta_1^{\mathrm{T}}\Omega\delta_1-\frac{1}{2}\delta_2^{\mathrm{T}}\Omega\delta_2\right]$$
$$= \exp\delta_1^{\mathrm{T}}\Omega\delta_2.$$

Further,

$$\mathcal{I}_{\pi\pi}\left(\delta\right)=\mathbb{E}\left(\frac{f\left(Y_t|\mu+\delta,\Omega\right)}{f\left(Y_t|\mu,\Omega\right)}\right)-1=\exp\delta^{\mathrm{T}}\Omega\delta-1.$$

2.2 Mean Information

The block diagonal structure of the information matrix leads to separation between information from the mean and information from the covariance. We have

$$\begin{aligned}\mathcal{I}_{\pi\mu}\left(\delta\right) &= \mathbb{E}\left(\frac{f\left(Y_t|\mu+\delta,\Omega\right)\nabla_\mu f\left(Y_t|\mu,\Omega\right)}{f\left(Y_t|\mu,\Omega\right)^2}\right)\\ &= \mathbb{E}_\delta\left(\frac{\nabla_\mu f\left(Y_t|\mu,\Omega\right)}{f\left(Y_t|\mu,\Omega\right)}\right)\\ &= \mathbb{E}_\delta\Omega\left(Y_t-\mu\right)=\Omega\delta.\end{aligned}$$

and $\mathcal{I}_{\mu\mu}^{-1}=\Omega$. Thus

$$\mathcal{I}_{\pi\mu}\left(\delta_1\right)^{\mathrm{T}}\mathcal{I}_{\mu\mu}^{-1}\mathcal{I}_{\pi\mu}\left(\delta_2\right)=\delta_1^{\mathrm{T}}\Omega\delta_2.$$

2.3 Covariance Information

We have $\mathcal{I}_{\pi\Omega}\left(\delta\right)$ consists of the matrix of components $\mathbb{E}_\delta\nabla_{\omega_{jk}}\log f\left(Y_t|\mu,\Omega\right)$. Matrix differentiation reveals

$$\mathcal{I}_{\pi\Omega}\left(\delta\right)=\frac{1}{2}\delta\delta^{\mathrm{T}}.$$

For $\mathcal{I}_{\Omega\Omega}$ we use the eigenvalue decomposition $\Omega=VDV^{\mathrm{T}}$, where D is a diagonal matrix with elements λ_i. Let Γ be a matrix with elements $\gamma_{ij}=\frac{\partial\lambda_i}{\partial\theta_j}$, then

$$\mathcal{I}_{\Omega\Omega}=\frac{1}{2}\Gamma D^{-2}\Gamma^{\mathrm{T}}.$$

The actual values of the calculation depend on the parameterization of Ω.

If

$$\Omega=\begin{pmatrix}\nu_1 & \rho\sqrt{\upsilon_1\nu_2}\\ \rho\sqrt{\upsilon_1\nu_2} & \upsilon_2\end{pmatrix},$$

then

$$\mathcal{I}_{\pi\Omega}\left(\delta_1\right)^{\mathrm{T}}\mathcal{I}_{\Omega\Omega}^{-1}\mathcal{I}_{\pi\Omega}\left(\delta_2\right)=\frac{1}{2}\left(\delta_1^{\mathrm{T}}\Omega\delta_2\right)^2.$$

If

$$\Omega = \nu \begin{pmatrix} 1 & \rho \\ \rho & 1 \end{pmatrix},$$

then

$$\mathcal{I}_{\pi\Omega}\left(\delta_1\right)^{\mathrm{T}} \mathcal{I}_{\Omega\Omega}^{-1} \mathcal{I}_{\pi\Omega}\left(\delta_2\right) = \frac{1}{2}\left(\delta_1^{\mathrm{T}} \Omega \delta_2\right)^2 + \frac{\left(\delta_{1,1}^2 - \delta_{1,2}^2\right)\left(\delta_{2,1}^2 - \delta_{2,2}^2\right)}{4\upsilon^2\left(1-\rho^2\right)},$$

where $\delta_{1,2}$ is the second element of δ_1.