1 Kalman Filtering and Smoothing Calculations

Using the preferred timing for the state space model,

$$\begin{aligned} y_t &= Z_t\alpha_t + d_t + \varepsilon_t \qquad && \varepsilon_t \sim N(0, H_t) \\ \alpha_t &= T_t\alpha_{t-1} + c_t + R_t\eta_t \qquad && \eta_t \sim N(0, Q_t) \end{aligned}$$

the filtered estimates are given by

$$\begin{aligned} a_{t+1} &= T_{t+1}a_t + c_{t+1} + K_tv_t \qquad && P_{t+1} = T_{t+1}P_tL_t' + R_{t+1}Q_{t+1}R_{t+1}' \\ v_t &= y_t - Z_ta_t - d_t \qquad && K_t = T_{t+1}P_tZ_t'F_t^{-1} \\ F_t &= Z_tP_tZ_t' + H_t \qquad && L_t = T_{t+1} - K_tZ_t \end{aligned}$$

and the smoothed estimates are given by

$$\begin{aligned} \hat{\alpha}_t &= a_t + P_tr_t \qquad && V_t = P_t - P_tN_tP_t \\ r_t &= Z_t'F_t^{-1}v_t + L_t'r_{t+1} \qquad && N_t = Z_t'F_t^{-1}Z_t + L_t'N_{t+1}L_t \end{aligned}$$

where $r_{n+1} = 0$ and $N_{n+1} = 0$.

Univariate Filter

When any H_t is non-diagonal, the observation equation is transformed by taking the LDL factorization of the H_t matrices:

$$y_t^* = Z_t^*\alpha_t + d_t^* + \varepsilon_t^* \qquad \varepsilon_t \sim N(0, H_t^*)$$
$$y_t^* = C_t^{-1}y_t \qquad Z_t^* = C_t^{-1}Z_t \qquad d_t^* = C_t^{-1}d_t \qquad \varepsilon_t^* = C_t^{-1}\varepsilon_t \qquad H_t = C_tH_t^*C_t'$$

To avoid taking the inverse of F_t in the above recursion, several quantities are computed using the univariate filter that can be used on this transformed system:

$$a_{t,i+1} = a_{t,i} + K_{t,i}v_{t,i}^* \qquad P_{t,i+1} = P_{t,i} - K_{t,i}F_{t,i}K_{t,i}'$$
$$v_{t,i}^* = y_{t,i}^* - Z_{t,i}^*a_{t,i} - d_{t,i}^* \qquad F_{t,i} = Z_{t,i}^*P_{t,i}Z_{t,i}^{*\prime} + H_{t,i}^* \qquad K_{t,i} = P_{t,i}Z_{t,i}^{*\prime}F_{t,i}^{-1}$$
$$a_{t+1,1} = T_{t+1}a_{t,p+1} + c_{t+1} \qquad P_{t+1,1} = T_{t+1}P_{t,p+1}T_{t+1}' + R_{t+1}Q_{t+1}R_{t+1}'$$

where $Z_{t,i}^*$ is the ith row of Z_t^*, $d_{t,i}^*$ is the ith element of d_t^*, and $H_{t,i}^*$ is the ith diagonal element of H_t^*. The filtered estimates of the state $a_t = a_{t,1}$ and $P_t = P_{t,1}$ are equivalent to those computed above. Similarly, the univariate smoother provides $r_t = r_{t,0}$ and $N_t = N_{t,0}$:

$$r_{t,i-1} = Z_{t,i}^{*\prime}F_{t,i}^{-1}v_{t,i}^* + L_{t,i}'r_{t,i} \qquad N_{t,i-1} = Z_{t,i}^{*\prime}F_{t,i}^{-1}Z_{t,i}^* + L_{t,i}'N_{t,i}L_{t,i} \qquad L_{t,i} = I_m - K_{t,i}Z_{t,i}^*F_{t,i}^{-1}$$
$$r_{t-1,p} = T_t'r_{t,0} \qquad N_{t-1,p} = T_{t-1}'N_{t,0}T_{t-1}$$

where $r_{t+1,p} = 0$ and $N_{t+1,p} = 0$.

Exact Initial Kalman Filter

When the state α_t is stationary, the initial values a_1 and P_1 can be computed as the unconditional mean and variance of the state given the system parameters by inverting the transition equation.

To handle cases where some states are nonstationary, the state is separated into states with known variance (those that are stationary) and those that are considered diffuse (nonstationary states). Let $\tilde{R}$ be a selection matrix with columns from the identity such that the initial shock η_1 is applied to the states with known variance. The selection matrix A is composed of the columns of the identity matrix associated with the diffuse states such that taking the limit as $\kappa \to \infty$ allows the diffuse states to have infinite initial variance where the initial values are given by

$$\begin{aligned} a_1 &= a + \tilde{R}\eta_1 + A\delta \qquad & \eta_1 &\sim N(0, Q_1) \qquad & \delta &\sim N(0, \kappa I) \\ P_1 &= P_{*,1} + \kappa P_{\infty,1} \qquad & P_{*,1} &= \tilde{R}\tilde{Q}\tilde{R}' \qquad & P_{\infty,1} &= AA' \end{aligned}$$

The unconditional mean, a, and variance, $P_{*,t,i}$, of the state are computed by inverting the stationary portion of the system. Any elements of a associated with nonstationary or diffuse states are set to 0.

$$\begin{aligned} \tilde{R}'a &= (I_m - T_1)^{-1} c_1 \\ \text{vec}(\tilde{R}\tilde{Q}\tilde{R}') &= (I_{m^2} - T_1 \otimes T_1)^{-1} \text{vec}(R_1 Q_1 R_1') \end{aligned}$$

When all states are stationary, this initialization collapses down to the simple case where $\alpha_1 \sim N(a_1, P_1)$, where a_1 and P_1 are determined by inverting the full system.

Given this initialization, the univariate filter recursions must be altered to separate the states with finite and infinite variances:

$$\begin{aligned} F_{*,t,i} &= Z^*_{t,i} P_{*,t,i} Z^{*\prime}_{t,i} + H^*_{t,i} \qquad & F_{\infty,t,i} &= Z^*_{t,i} P_{\infty,t,i} Z^{*\prime}_{t,i} \\ K_{*,t,i} &= P_{*,t,i} Z^{*\prime}_{t,i} \qquad & K_{\infty,t,i} &= P_{\infty,t,i} Z^{*\prime}_{t,i} \end{aligned}$$

$$a_{t,i+1} = \begin{cases} a_{t,i} + K_{*,t,i} F^{-1}_{*,t,i} v^*_{t,i} & F_{\infty,t,i} = 0 \\ a_{t,i} + K_{\infty,t,i} F^{-1}_{\infty,t,i} v^*_{t,i} & F_{\infty,t,i} \neq 0 \end{cases}$$

$$P_{*,t,i+1} = \begin{cases} P_{*,t,i} - K_{*,t,i} K'_{*,t,i} F^{-1}_{*,t,i} & F_{\infty,t,i} = 0 \\ P_{*,t,i} + K_{\infty,t,i} K'_{\infty,t,i} F_{*,t,i} F^{-2}_{\infty,t,i} - (K_{*,t,i} K'_{\infty,t,i} + K_{\infty,t,i} K'_{*,t,i}) F^{-1}_{\infty,t,i} & F_{\infty,t,i} \neq 0 \end{cases}$$

$$P_{\infty,t,i+1} = \begin{cases} P_{\infty,t,i} & F_{\infty,t,i} = 0 \\ P_{\infty,t,i} - K_{\infty,t,i} K'_{\infty,t,i} F^{-1}_{\infty,t,i} & F_{\infty,t,i} \neq 0 \end{cases}$$

$$a_{t+1,1} = T_{t+1} a_{t,p+1} + c_{t+1}$$
$$P_{\infty,t+1,1} = T_{t+1} P_{\infty,t,p+1} T'_{t+1} \qquad P_{*,t+1,1} = T_{t+1} P_{*,t,p+1} T'_{t+1} + R_{t+1} Q_{t+1} R'_{t+1}$$

For any set of system parameters where the state can be identified, there exists some time d such that $F_{\infty,d,i} = 0$ for all i. For time $t > d$, the simpler Kalman filter recursion above can be employed.

The smoother must be similarly altered so that beginning at $t = d$, the computation of $r_{t,i}$ is expanded to account for the initialization:

$$L_{\infty,t,i} = I_m - K_{\infty,t,i} Z^*_{t,i} F^{-1}_{\infty,t,i} \qquad L_{*,t,i} = I_m - K_{*,t,i} Z^*_{t,i} F^{-1}_{*,t,i}$$

$$L^{(0)}_{t,i} = \left(K_{\infty,t,i} F_{*,t,i} F^{-1}_{\infty,t,i} - K_{*,t,i}\right) Z^*_{t,i} F^{-1}_{\infty,t,i}$$

$$r^{(0)}_{t,i-1} = \begin{cases} Z^{*\prime}_{t,i} F^{-1}_{*,t,i} v^*_{t,i} + L'_{*,t,i} r^{(0)}_{t,i} & F_{\infty,t,i} = 0 \\ L'_{\infty,t,i} r^{(0)}_{t,i} & F_{\infty,t,i} \neq 0 \end{cases}$$

$$r^{(1)}_{t,i-1} = \begin{cases} r^{(1)}_{t,i} & F_{\infty,t,i} = 0 \\ Z^{*\prime}_{t,i} F^{-1}_{\infty,t,i} v^*_{t,i} + L^{(0)\prime}_{t,i} r^{(0)}_{t,i} + L'_{\infty,t,i} r^{(1)}_{t,i} & F_{\infty,t,i} \neq 0 \end{cases}$$

$$r^{(0)}_{t-1,p} = T'_t r^{(0)}_{t,0} \qquad r^{(1)}_{t-1,p} = T'_t r^{(1)}_{t,0}$$

$$\hat{\alpha}_t = a_t + P_{*,t,1} r^{(0)}_{t,0} + P_{\infty,t,1} r^{(1)}_{t,0}$$

where $r^{(0)}_{d,p} = r_{d,p}$ and $r^{(1)}_{d,p} = 0$. Note that $L_{\infty,t,i}$ and $L^{(0)}_{t,i}$ only need to be computed when $F_{\infty,t,i} \neq 0$ and $L_{*,t,i}$ only needs to be computed when $F_{\infty,t,i} = 0$.

For the smoothed variance of the state,

$$V_t = P_{*,t,1} - P_{*,t,1} N^{(0)}_{t,0} P_{*,t,1} - \left(P_{\infty,t,1} N^{(1)}_{t,0} P_{*,t,1}\right)' - P_{\infty,t,1} N^{(1)}_{t,0} P_{*,t,1} - P_{\infty,t,1} N^{(2)}_{t,0} P_{\infty,t,1}$$

$$N^{(0)}_{t-1,p} = T'_t N^{(0)}_{t,0} T_t \qquad N^{(1)}_{t-1,p} = T'_t N^{(1)}_{t,0} T_t \qquad N^{(2)}_{t-1,p} = T'_t N^{(1)}_{t,0} T_t$$

where when $F_{\infty,t,i} = 0$,

$$N^{(0)}_{t,i-1} = Z^{*\prime}_{t,i} F^{-1}_{*,t,i} Z^*_{t,i} + L'_{*,t,i} N^{(0)}_{t,i} L_{*,t,i} \qquad N^{(1)}_{t,i-1} = N^{(1)}_{t,i} \qquad N^{(2)}_{t,i-1} = N^{(2)}_{t,i}$$

and when $F_{\infty,t,i} \neq 0$,

$$N^{(0)}_{t,i-1} = L'_{\infty,t,i} N^{(0)}_{t,i} L_{\infty,t,i}$$

$$N^{(1)}_{t,i-1} = Z^{*\prime}_{t,i} F^{-1}_{\infty,t,i} Z^*_{t,i} + L'_{\infty,t,i} N^{(0)}_{t,i} L^{(0)}_{t,i} + L'_{\infty,t,i} N^{(1)}_{t,i} L_{\infty,t,i}$$

$$N^{(2)}_{t,i-1} = Z^{*\prime}_{t,i} F^{-2}_{\infty,t,i} Z^*_{t,i} F_{*,t,i} + L^{(0)\prime}_{t,i} N^{(1)}_{t,i} L^{(0)}_{t,i} + L'_{\infty,t,i} N^{(1)}_{t,i} L^{(0)}_{t,i} + L^{(0)\prime}_{t,i} N^{(1)}_{t,i} L_{\infty,t,i} + L'_{\infty,t,i} N^{(2)}_{t,i} L_{\infty,t,i}$$

where $N^{(0)}_{d,p} = N_{d,p}$ and $N^{(1)}_{d,p} = N^{(2)}_{d,p} = 0$.

2 Likelihood and Gradient Calculation

The likelihood of of data $y_1, \ldots, y_n$ in the standard filter as shown in [1] (§7.2) is given by the prediction error decomposition:

$$\log L(Y_n) = -\frac{np}{2}\log 2\pi - \frac{1}{2}\sum_{t=1}^{n}\left(\log|F_t| + v_t' F_{t,i}^{-1} v_t\right)$$

For the univariate filter, the same decomposition works in the univariate context:

$$\log L(Y_n) = -\frac{np}{2}\log 2\pi - \frac{1}{2}\sum_{t=1}^{n}\sum_{i=1}^{p}\log F_{t,i} + v_{t,i}^2/F_{t,i} \tag{1}$$

The likelihood for the exact initial filter allows for some simplifications in $F_{\infty,t,i}$ as shown in [1]:

$$\log L_d(Y_n) = -\frac{1}{2}\sum_{t=1}^{n}\sum_{i=1}^{p}\iota_{t,i}\log 2\pi - \frac{1}{2}\sum_{t=1}^{d}\sum_{i=1}^{p} w_{t,i} - \frac{1}{2}\sum_{t=d}^{n}\sum_{i=1}^{p}\iota_{t,i}\left(\log F_{t,i} + v_{t,i}^2/F_{t,i}\right) \tag{2}$$

where $\iota_{t,i} = 1$ if $F_{*,t,i} \neq 0$ or $t > d$, and

$$w_{t,i} = \begin{cases} \iota_{t,i}(\log(F_{*,t,i}) + v_{t,i}^{(0)2}/F_{*,t,i}) & F_{\infty,t,i} = 0 \\ \log F_{\infty,t,i} & F_{\infty,t,i} \neq 0 \end{cases}$$

Since these quantities are naturally produced by the Kalman filter, this is the preferred method to calculate the likelihood.

Univariate Gradient

Following the notation of [2], we denote $\partial\left[\text{vec}(A)'\right]/\partial\theta$ by $G_\theta(A)$, where for an $m \times n$ matrix A and a θ of length n_θ, $G_\theta(A)$ will be a $n_\theta \times mn$ matrix of partial derivatives. See Appendix A for details and properties of $G_\theta(\cdot)$.

Given the univariate likelihood function (1), the identities for the gradient in Appendix A produce the expression for the univariate gradient:

$$\begin{aligned} G_\theta(\log L(Y_n)) &= G_\theta\left(-\frac{np}{2}\log 2\pi - \frac{1}{2}\sum_{t=1}^{n}\sum_{i=1}^{p}\log F_{t,i} + v_{t,i}^2 F_{t,i}^{-1}\right) \\ &= -\frac{1}{2}\sum_{t=1}^{n}\sum_{i=1}^{p} G_\theta(\log F_{t,i}) + G_\theta(v_{t,i}^2 F_{t,i}^{-1}) \\ &= -\frac{1}{2}\sum_{t=1}^{n}\sum_{i=1}^{p} G_\theta(F_{t,i})\left[F_{t,i}^{-1} - F_{t,i}^{-2}v_{t,i}^2\right] + 2G_\theta(v_{t,i})F_{t,i}^{-1}v_{t,i} \end{aligned}$$

The quantities used in the gradient recursion are computed similarly using the definitions from the univariate filter:

$$
\begin{aligned}
G_\theta(v_{t,i}) &= G_\theta(y_{t,i} - Z_{t,i}a_{t,i} - d_{t,i}) \\
&= -G_\theta(Z_{t,i})a_{t,i} - G_\theta(a_{t,i})Z'_{t,i} - G_\theta(d_{t,i}) \\
G_\theta(F_{t,i}) &= G_\theta(Z_{t,i}P_{t,i}Z'_{t,i} + H_{t,i}) \\
&= G_\theta(Z_{t,i})(P_{t,i}Z'_{t,i})N_m + G_\theta(P_{t,i})(Z'_{t,i} \otimes Z'_{t,i}) + G_\theta(H_{t,i}) \\
G_\theta(K_{t,i}) &= G_\theta(P_{t,i}Z'_{t,i}F_{t,i}^{-1}) \\
&= G_\theta(P_{t,i}Z'_{t,i})(F_{t,i}^{-1} \otimes I_m) + G_\theta(F_{t,i}^{-1})P'_{t,i}Z_{t,i} \\
&= \left[G_\theta(P_{t,i})(Z'_{t,i} \otimes I_m) + G_\theta(Z_{t,i})K_{1m}P_{t,i}\right](F_{t,i}^{-1} \otimes I_m) + G_\theta(F_{t,i})F_{t,i}^{-2}P'_{t,i}Z_{t,i} \\
G_\theta(a_{t+1,1}) &= G_\theta(T_{t+1}a_{t,p+1} + c_{t+1}) \\
&= G_\theta(T_{t+1})(a_{t,p+1} \otimes I_m) + G_\theta(a_{t,p+1})T'_{t+1} + G_\theta(c_{t+1}) \\
G_\theta(a_{t,i+1}) &= G_\theta(a_{t,i} + K_{t,i}v_{t,i}) \\
&= G_\theta(a_{t,i}) + G_\theta(K_{t,i})(v_{t,i} \otimes I_m) + G_\theta(v_{t,i})K'_{t,i} \\
G_\theta(P_{t+1,1}) &= G_\theta(T_{t+1}P_{t,p+1}T'_{t+1} + R_{t+1}Q_{t+1}R'_{t+1}) \\
&= G_\theta(T_{t+1})(P_{t,p+1}T'_{t+1} \otimes I_m)N_m + G_\theta(P_{t,p+1})(T'_{t+1} \otimes T'_{t+1}) \\
&\quad + G_\theta(R_{t+1})(Q_{t+1}R'_{t+1} \otimes I_m)N_m + G_\theta(Q_{t+1})(R'_{t+1} \otimes R'_{t+1}) \\
G_\theta(P_{t,i+1}) &= G_\theta(P_{t,i} - K_{t,i}F_{t,i}K'_{t,i}) \\
&= G_\theta(P_{t,i}) - G_\theta(K_{t,i})(F_{t,i}K'_{t,i} \otimes I_m) - G_\theta(F_{t,i})(K'_{t,i} \otimes K'_{t,i})
\end{aligned}
$$

where $G_\theta(a_{1,1}) = G_\theta(a_1)$ and $G_\theta(P_{1,1}) = G_\theta(P_1)$, which will be given below.

Exact Initial Gradient

The process for computing the gradient for use with the exact initial filter is similar. Throughout, it is assumed that the boolean outcome of the tests $F_{*,t,i} = 0$ or $F_{\infty,t,i} = 0$ are unaffected by changes to θ. Beginning with the likelihood function,

$$
\begin{aligned}
G_\theta(\log L_d(Y_n)) &= G_\theta\left(-\frac{1}{2}\sum_{t=1}^{n}\sum_{i=1}^{p}\iota_{t,i}\log 2\pi - \frac{1}{2}\sum_{t=1}^{d}\sum_{i=1}^{p} w_{t,i} - \frac{1}{2}\sum_{t=d}^{n}\sum_{i=1}^{p}\iota_{t,i}\left(\log F_{t,i} + v_{t,i}^2/F_{t,i}\right)\right) \\
&= -\frac{1}{2}\sum_{t=1}^{d}\sum_{i=1}^{p} G_\theta(w_{t,i}) - \frac{1}{2}\sum_{t=d}^{n}\sum_{i=1}^{p} G_\theta(F_{t,i})\left[F_{t,i}^{-1} - F_{t,i}^{-2}v_{t,i}^2\right] + 2G_\theta(v_{t,i})F_{t,i}^{-1}v_{t,i}
\end{aligned}
$$

where

$$
\begin{aligned}
G_\theta(w_{t,i}) &= \begin{cases} G_\theta(\iota_{t,i}(\log(F_{*,t,i}) + v_{t,i}^{(0)2}/F_{*,t,i})) & F_{\infty,t,i} = 0 \\ G_\theta(\log F_{\infty,t,i}) & F_{\infty,t,i} \neq 0 \end{cases} \\
&= \begin{cases} \iota_{t,i}G_\theta(F_{*,t,i})\left[F_{*,t,i}^{-1} - F_{*,t,i}^{-2}v_{t,i}^{(0)2}\right] + \iota_{t,i}2G_\theta(v_{t,i}^{(0)})F_{*,t,i}^{-1}v_{t,i}^{(0)} & F_{\infty,t,i} = 0 \\ G_\theta(F_{\infty,t,i})F_{\infty,t,i}^{-1} & F_{\infty,t,i} \neq 0 \end{cases}
\end{aligned}
$$

Again, the required quantities in the recursion come from the definition of the filter:

$$
\begin{aligned}
G_\theta(F_{*,t,i}) &= G_\theta(Z_{t,i}P_{*,t,i}Z'_{t,i} + H_{t,i}) \\
&= G_\theta(Z_{t,i})(P_{*,t,i}Z'_{t,i})N_m + G_\theta(P_{*,t,i})(Z'_{t,i} \otimes Z'_{t,i}) + G_\theta(H_{t,i}) \\
G_\theta(F_{\infty,t,i}) &= G_\theta(Z_{t,i}P_{\infty,t,i}Z'_{t,i} + H_{t,i}) \\
&= G_\theta(Z_{t,i})(P_{\infty,t,i}Z'_{t,i})N_m + G_\theta(P_{\infty,t,i})(Z'_{t,i} \otimes Z'_{t,i}) + G_\theta(H_{t,i}) \\
G_\theta(K_{*,t,i}) &= G_\theta(P_{*,t,i}Z'_{t,i}) \\
&= G_\theta(P_{*,t,i})(Z'_{t,i} \otimes I_m) + G_\theta(Z_{t,i})K_{m1}P_{*,t,i} \\
G_\theta(K_{\infty,t,i}) &= G_\theta(P_{\infty,t,i}Z'_{t,i}) \\
&= G_\theta(P_{\infty,t,i})(Z'_{t,i} \otimes I_m) + G_\theta(Z_{t,i})K_{m1}P_{\infty,t,i}
\end{aligned}
$$

Between periods,

$$
\begin{aligned}
G_\theta(a_{t+1,1}) &= G_\theta(T_{t+1}a_{t,p+1} + c_{t+1}) \\
&= G_\theta(T_{t+1})(a_{t,p+1} \otimes I_m) + G_\theta(a_{t,p+1})T'_{t+1} + G_\theta(c_{t+1}) \\
G_\theta(P_{*,t+1,1}) &= G_\theta(T_{t+1}P_{*,t,p+1}T'_{t+1} + R_{t+1}Q_{t+1}R'_{t+1}) \\
&= G_\theta(T_{t+1})(P_{*,t,p+1}T'_{t+1} \otimes I_m)N_m + G_\theta(P_{*,t,p+1})(T'_{t+1} \otimes T'_{t+1}) \\
&\quad + G_\theta(R_{t+1})(Q_{t+1}R'_{t+1} \otimes I_m)N_m + G_\theta(Q_{t+1})(R'_{t+1} \otimes R'_{t+1}) \\
G_\theta(P_{\infty,t+1,1}) &= G_\theta(T_{t+1}P_{\infty,t,p+1}T'_{t+1}) \\
&= G_\theta(T_{t+1})(P_{\infty,t,p+1}T'_{t+1} \otimes I_m)N_m + G_\theta(P_{\infty,t,p+1})(T'_{t+1} \otimes T'_{t+1})
\end{aligned}
$$

When $F_{\infty,t,i} = 0$,

$$
\begin{aligned}
G_\theta(a_{t,i+1}) &= G_\theta(a_{t,i} + K_{*,t,i}F^{-1}_{*,t,i}v_{t,i}) \\
&= G_\theta(a_{t,i}) + G_\theta(K_{*,t,i})(F^{-1}_{*,t,i}v_{t,i} \otimes I_m) + G_\theta(F^{-1}_{*,t,i}v_{t,i})K_{*,t,i} \\
&= G_\theta(a_{t,i}) + G_\theta(K_{*,t,i})(F^{-1}_{*,t,i}v_{t,i} \otimes I_m) - G_\theta(F_{*,t,i})F^{-2}_{*,t,i}K_{*,t,i} + G_\theta(v_{t,i})F^{-1}_{*,t,i}K_{*,t,i} \\
G_\theta(P_{*,t,i}) &= G_\theta(P_{*,t,i} - K_{*,t,i}K'_{*,t,i}F^{-1}_{*,t,i}) \\
&= G_\theta(P_{*,t,i}) - G_\theta(K_{*,t,i})(F_{*,t,i}K'_{*,t,i} \otimes I_m) - G_\theta(F_{*,t,i})(K'_{*,t,i} \otimes K'_{*,t,i}) \\
G_\theta(P_{\infty,t,i+1}) &= G_\theta(P_{\infty,t,i})
\end{aligned}
$$

and when $F_{\infty,t,i} \neq 0$,

$$
\begin{aligned}
G_\theta(a_{t,i+1}) &= G_\theta(a_{t,i} + K_{\infty,t,i}F_{\infty,t,i}^{-1}v_{t,i}) \\
&= G_\theta(a_{t,i}) + G_\theta(K_{\infty,t,i})(F_{\infty,t,i}^{-1}v_{t,i} \otimes I_m) - G_\theta(F_{\infty,t,i})F_{\infty,t,i}^{-2}K_{\infty,t,i} + G_\theta(v_{t,i})F_{\infty,t,i}^{-1}K_{\infty,t,i} \\
G_\theta(P_{*,t,i+1}) &= G_\theta\left(P_{*,t,i} - K_{\infty,t,i}K'_{\infty,t,i}F_{*,t,i}F_{\infty,t,i}^{-2} - \left(K_{*,t,i}K'_{\infty,t,i} + K_{\infty,t,i}K'_{*,t,i}\right)F_{\infty,t,i}^{-1}\right) \\
&= G_\theta(P_{*,t,i}) - G_\theta\left(K_{\infty,t,i}K'_{\infty,t,i}F_{*,t,i}F_{\infty,t,i}^{-2}\right) - G_\theta\left((K_{*,t,i}K'_{\infty,t,i} + K_{\infty,t,i}K'_{*,t,i})F_{\infty,t,i}^{-1}\right) \\
&= G_\theta(P_{*,t,i}) - G_\theta(K_{\infty,t,i})(K'_{\infty,t,i} \otimes I_m)(F_{*,t,i}F_{\infty,t,i}^{-2} \otimes I_m) \\
&\quad - \left[G_\theta(F_{*,t,i})F_{\infty,t,i}^{-2} - 2G_\theta(F_{\infty,t,i})F_{\infty,t,i}^{-3}F_{*,t,i})\right]K_{\infty,t,i}K'_{\infty,t,i} \\
&\quad - [G_\theta(K_{*,t,i})\left[(K'_{\infty,t,i} \otimes I_m) + K_{1m}K'_{\infty,t,i}\right] \\
&\quad\quad + G_\theta(K_{\infty,t,i})\left[K_{1m}K'_{*,t,i} + (K'_{*,t,i} \otimes I_m)\right]](F_{\infty,t,i}^{-1} \otimes I_m) \\
&\quad - G_\theta(F_{\infty,t,i})F_{\infty,t,i}^{-2}(K_{*,t,i}K'_{\infty,t,i} + K_{\infty,t,i}K'_{*,t,i}) \\
G_\theta(P_{\infty,t,i+1}) &= G_\theta(P_{\infty,t,i} - K_{\infty,t,i}K'_{\infty,t,i}F_{\infty,t,i}^{-1}) \\
&= G_\theta(P_{\infty,t,i}) - G_\theta(K_{\infty,t,i}K'_{\infty,t,i})(F_{\infty,t,i}^{-1} \otimes I_m) - G_\theta(F_{\infty,t,i}^{-1})(K_{\infty,t,i}K'_{\infty,t,i}) \\
&= G_\theta(P_{\infty,t,i}) - G_\theta(K_{\infty,t,i})(K'_{\infty,t,i} \otimes I_m)N_m(F_{\infty,t,i}^{-1} \otimes I_m) + G_\theta(F_{\infty,t,i})F_{\infty,t,i}^{-2}(K_{\infty,t,i}K'_{\infty,t,i})
\end{aligned}
$$

since

$$
\begin{aligned}
G_\theta\left(K_{\infty,t,i}K'_{\infty,t,i}F_{*,t,i}F_{\infty,t,i}^{-2}\right) &= G_\theta(K_{\infty,t,i}K'_{\infty,t,i})(F_{*,t,i}F_{\infty,t,i}^{-2} \otimes I_m) \\
&\quad + G_\theta(F_{*,t,i}F_{\infty,t,i}^{-2})K_{\infty,t,i}K'_{\infty,t,i} \\
&= G_\theta(K_{\infty,t,i})(K'_{\infty,t,i} \otimes I_m)(F_{*,t,i}F_{\infty,t,i}^{-2} \otimes I_m) \\
&\quad + \left[G_\theta(F_{*,t,i})F_{\infty,t,i}^{-2} - 2G_\theta(F_{\infty,t,i})F_{\infty,t,i}^{-3}F_{*,t,i})\right]K_{\infty,t,i}K'_{\infty,t,i} \\
G_\theta\left((K_{*,t,i}K'_{\infty,t,i} + K_{\infty,t,i}K'_{*,t,i})F_{\infty,t,i}^{-1}\right) &= G_\theta\left((K_{*,t,i}K'_{\infty,t,i} + K_{\infty,t,i}K'_{*,t,i})\right)(F_{\infty,t,i}^{-1} \otimes I_m) \\
&\quad + G_\theta(F_{\infty,t,i}^{-1})(K_{*,t,i}K'_{\infty,t,i} + K_{\infty,t,i}K'_{*,t,i}) \\
&= [G_\theta(K_{*,t,i})(K'_{\infty,t,i} \otimes I_m) + G_\theta(K_{\infty,t,i})K_{1m}K'_{*,t,i} \\
&\quad + G_\theta(K_{\infty,t,i})(K'_{*,t,i} \otimes I_m) + G_\theta(K_{*,t,i})K_{1m}K'_{\infty,t,i}](F_{\infty,t,i}^{-1} \otimes I_m) \\
&\quad + G_\theta(F_{\infty,t,i})F_{\infty,t,i}^{-2}(K_{*,t,i}K'_{\infty,t,i} + K_{\infty,t,i}K'_{*,t,i}) \\
&= [G_\theta(K_{*,t,i})\left[(K'_{\infty,t,i} \otimes I_m) + K_{1m}K'_{\infty,t,i}\right] \\
&\quad + G_\theta(K_{\infty,t,i})\left[K_{1m}K'_{*,t,i} + (K'_{*,t,i} \otimes I_m)\right]](F_{\infty,t,i}^{-1} \otimes I_m) \\
&\quad + G_\theta(F_{\infty,t,i})F_{\infty,t,i}^{-2}(K_{*,t,i}K'_{\infty,t,i} + K_{\infty,t,i}K'_{*,t,i})
\end{aligned}
$$

Initial Conditions

The initial conditions for the recursion are given via the expressions for a_1 and P_1. We assume that the rank of A and R_0 do not affected by θ:

$$
\begin{aligned}
G_\theta(a_1) &= G_\theta(a) + G_\theta(R_0\eta_0) + G_\theta(A\delta) = G_\theta(a) \\
G_\theta(P_1) &= G_\theta(\tilde{R}\tilde{Q}\tilde{R}) + G_\theta(\kappa AA') = G_\theta(\tilde{R}\tilde{Q}\tilde{R})
\end{aligned}
$$

From the definitions of a and $\tilde{R}\tilde{Q}\tilde{R}$ above,

$$\begin{aligned}
G_\theta(a) &= G_\theta([I_m - T_1]^{-1}c_1) \\
&= G_\theta([I_m - T_1]^{-1})(c_1 \otimes I_m) + G_\theta(c_1)(I_m - T_1)^{-1})' \\
&= -G_\theta(I_m - T_1)[(I_m - T_1)^{-1} \otimes (I_m - T_1)'^{-1}](c_1 \otimes I_m) + G_\theta(c_1)(I_m - T_1)'^{-1} \\
&= G_\theta(T_1)[(I_m - T_1)^{-1} \otimes (I_m - T_1)'^{-1}](c_1 \otimes I_m) + G_\theta(c_1)(I_m - T_1)'^{-1} \\
G_\theta(\tilde{R}\tilde{Q}\tilde{R}') &= G_\theta(S \text{ vec}(R_1 Q_1 R_1')) \\
&= G_\theta(S)[\text{vec}(R_1 Q_1 R_1') \otimes I_{m^2}] + G_\theta(\text{vec}(R_1 Q_1 R_1'))(I_1 \otimes S') \\
&= -G_\theta(I_{m^2} - T_1 \otimes T_1)(S \otimes S')[\text{vec}(R_1 Q_1 R_1') \otimes I_{m^2}] \\
&\quad + [G_\theta(R_1)(Q_1 R_1' \otimes I_m)N_m + G_\theta(Q_1)(R_1' \otimes R_1')]S' \\
&= G_\theta(T_1 \otimes T_1)(S \otimes S')[\text{vec}(R_1 Q_1 R_1') \otimes I_{m^2}] \\
&\quad + [G_\theta(R_1)(Q_1 R_1' \otimes I_m)N_m + G_\theta(Q_1)(R_1' \otimes R_1')]S'
\end{aligned}$$

where $S = (I_{m^2} - T_1 \otimes T_1)^{-1}$. Note that $G_\theta(T_1 \otimes T_1)$ must be computed separately. For the diffuse filter, note that $G_\theta(P_{*,1}) = G_\theta(P_1)$ and $G_\theta(P_{\infty,1}) = 0$.

A Gradient Derivation

Denote $\partial\,[\text{vec}(A)']\,/\partial\theta$ by $G_\theta(A)$, where for an $m \times n$ matrix A and a θ of length n_θ, $G_\theta(A)$ will be a $n_\theta \times mn$ matrix of partial derivatives. Several identities are needed in computing the gradient (assume matrices are of dimensions such that the expression on the left-hand side exists):

(a) $G_\theta(A + B) = G_\theta(A) + G_\theta(B)$

(b) $G_\theta(AB) = G_\theta(A)(B \otimes I_{p_A}) + G_\theta(B)(I_{q_B} \otimes A)$

(c) $G_\theta(A') = G_\theta(A)K_{nm}$

(d) $G_\theta(ACA') = G_\theta(A)(CA' \otimes I_m)N_m + G_\theta(C)(A' \otimes A')$

(e) $G_\theta(D^{-1}) = -G_\theta(D)(D^{-1} \otimes D^{-1})$

(f) $G_\theta(\log|D|) = G_\theta(D)D'^{-1}$

where C is symmetric, D is nonsingular, K_{mn} is a commutation matrix of size mn, and $N_m = I_{m^2} + K_{mm}$. Note that $G_\theta(AA')$ can easily be computed from (d) by treating C as the identity matrix. Also note that $A \otimes I_1 = A$ and $\log|B| = \log B$ if B is a scalar:

References

[1] James Durbin and Siem Jan Koopman. *Time Series Analysis by State Space Methods: Second Edition (Oxford Statistical Science Series)*. Oxford University Press, 2012.

[2] Daisuke Nagakura. Exact gradient vector of loglikelihood for linear gaussian state space model. 2013.