State Space Loglikelihood Gradient Calculation

Using the preferred (Harvey 1989) timing for the state space model,

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

the likelihood of of data $\mathbf{Y}_t$ is given by $L_n \equiv \sum_{t=1}^{n} \ell_t$ where

$$\ell_t = -\frac{p}{2} \log(2\pi) - \frac{1}{2} \log |F_t| + v_t' F_t^{-1} v_t \tag{1}$$

$$\begin{aligned} a_t &= T_t a_{t-1} + c_t + K_t v_t \qquad & P_t &= T_t P_{t-1} L_{t-1} + R_t Q_t R_t' \\ v_t &= y_t - d_t - Z_t a_t \qquad & F_t &= Z_t P_t Z_t' + H_t \\ M_t &= P_t Z_t' F_t^{-1} \qquad & K_t &= T_{t+1} M_t \\ L_t &= T_{t+1} - K_t Z_t \qquad & w_t &= F_t^{-1} v_t \end{aligned}$$

From Nagakura (working paper), we can get the gradient:

$$\begin{aligned} G_\theta(\ell_t) &= G_\theta(a_t) Z_t' w_t \\ &\quad + \frac{1}{2} G_\theta(P_t) \mathrm{vec}(Z_t' w_t w_t' Z_t - Z_t' F_t^{-1} Z_t) \\ &\quad + G_\theta(d_t) w_t \\ &\quad + G_\theta(Z_t) \mathrm{vec}(w_t a_t' + w_t v_t' M_t' - M_t') \\ &\quad + \frac{1}{2} G_\theta(H_t) \mathrm{vec}(w_t w_t' - F_t^{-1}) \end{aligned} \tag{2}$$

where

$$\begin{aligned} G_\theta(a_{t+1}) &= G_\theta(a_t) L_t' \\ &\quad + G_\theta(P_t)(Z_t' w_t \otimes L_t') \\ &\quad + G_\theta(c_{t+1}) \\ &\quad - G_\theta(d_t) K_t' \\ &\quad + G_\theta(Z_t)[P_t L_t' \otimes w_t - (a_t + M_t v_t) \otimes K_t'] \\ &\quad - G_\theta(H_t)(w_t \otimes K_t') \\ &\quad + G_\theta(T_{t+1})[(a_t + M_t v_t) \otimes I_m] \\ G_\theta(P_{t+1}) &= G_\theta(P_t)(L_t' \otimes L_t') \\ &\quad + G_\theta(H_t)(K_t' \otimes K_t') \\ &\quad + G_\theta(Q_{t+1})(R_{t+1}' \otimes R_{t+1}') \\ &\quad + [G_\theta(T_{t+1})(P_t L_t' \otimes I_m) \\ &\qquad - G_\theta(Z_t)(P_t L_t' \otimes K_t') \\ &\qquad + G_\theta(R_{t+1})(Q_{t+1} R_{t+1}' \otimes I_m)] N_m \end{aligned}$$

The initial conditions for the recursion are given by a simplification of the expressions above:

$$\begin{aligned}
G_\theta(a_1) &= G_\theta(a_0)T_1' \\
&\quad + G_\theta(c_1) \\
&\quad + G_\theta(T_1)[a_0 \otimes I_m] \\
G_\theta(P_1) &= G_\theta(P_0)(T_t' \otimes T_t') \\
&\quad + G_\theta(Q_1)(R_1' \otimes R_1') \\
&\quad + [G_\theta(T_1)(P_0T_1' \otimes I_m) \\
&\qquad + G_\theta(R_1)(Q_1R_1' \otimes I_m)]N_m
\end{aligned}$$