Without Covariates, the state-space model is given by

$$\underset{\sim}{X}_t = \Phi \underset{\sim}{X}_{t-1} + \underset{\sim}{W}_t$$

$$\underset{\sim}{Y}_t = A_t \underset{\sim}{X}_t + \underset{\sim}{V}_t$$

with $\underset{\sim}{X}_0 \sim N_p(\mu_0, \Sigma_0)$, $\underset{\sim}{W}_t \overset{iid}{\sim} N_p(\underline{0}, Q)$, $\underset{\sim}{V}_t \overset{iid}{\sim} N_q(0, R)$.

The complete data-likelihood is given by

$$p_\Theta(x_{0:T}, y_{1:T}) = p_\Theta(x_0, x_1, \ldots, x_T, y_1, \ldots, y_T)$$

$$= p_\Theta(X_0) \cdot p_\Theta(X_1 | X_0) \cdots p_\Theta(X_T | X_{T-1}) \times p_\Theta(Y_1|X_1) \times \cdots \times p_\Theta(Y_T|X_T)$$

$$= p_\Theta(x_0) \prod_{t=1}^{T} p_\Theta(X_t | X_{t-1}) \prod_{t=1}^{T} p_\Theta(Y_t | X_t)$$

$$\equiv p_{\mu_0, \Sigma_0}(\underline{X}_0) \qquad p_{\Phi, Q}(\underset{\sim}{X}_t | \underset{\sim}{X}_{t-1}) \qquad p_R(\underset{\sim}{Y}_t | \underset{\sim}{X}_t)$$

$$p_{\mu_0, \Sigma_0}(\underline{X}_0) = \frac{1}{|2\pi \Sigma_0|^{1/2}} e^{-\frac{1}{2} \cdot (\underset{\sim}{X}_0 - \underline{\mu}_0)' \Sigma_0^{-1} (\underset{\sim}{X}_0 - \underline{\mu}_0)}$$

$$\Rightarrow \log p_{\mu_0, \Sigma_0}(\underline{X}_0) = -\frac{p}{2} \log(2\pi) - \frac{1}{2} \log|\Sigma_0| - \frac{1}{2} (\underline{X}_0 - \underline{\mu}_0)' \Sigma_0^{-1} (\underline{X}_0 - \underline{\mu}_0)$$

$$p_{\Phi, Q}(\underset{\sim}{X}_t | \underset{\sim}{X}_{t-1}) = \frac{1}{|2\pi Q|^{1/2}} e^{-\frac{1}{2} \cdot (\underline{X}_t - \Phi \underline{X}_{t-1})' Q^{-1} (\underline{X}_t - \Phi \underline{X}_{t-1})}$$

$$\Rightarrow \log p_{\Phi, Q}(\underset{\sim}{X}_t | \underset{\sim}{X}_{t-1}) = -\frac{p}{2} \log(2\pi) - \frac{1}{2} \log|Q| - \frac{1}{2} (\underset{\sim}{X}_t - \Phi \underline{X}_{t-1})' Q^{-1} (\underset{\sim}{X}_t - \Phi \underline{X}_{t-1})$$

$$p_R(\underset{\sim}{Y}_t | \underset{\sim}{X}_t) = \frac{1}{|2\pi R|^{1/2}} \cdot e^{-\frac{1}{2} (\underset{\sim}{Y}_t - A_t \underset{\sim}{X}_t)' R^{-1} (\underset{\sim}{Y}_t - A_t \underset{\sim}{X}_t)}$$

$$\Rightarrow \log p_R(\underset{\sim}{Y}_t | \underset{\sim}{X}_t) = -\frac{q}{2} \log(2\pi) - \frac{1}{2} \log|R| - \frac{1}{2} (\underset{\sim}{Y}_t - A_t \underline{X}_t)' R^{-1} (\underline{Y}_t - A_t \underline{X}_t)$$

Overall, $\log p_\Theta(X_{0:T}, Y_{1:T}) = c - \frac{1}{2} \log|\Sigma_0| - \frac{1}{2} (\underline{X}_0 - \underline{\mu}_0)' \Sigma_0^{-1} (\underline{X}_0 - \underline{\mu}_0)$

$$- \frac{T}{2} \log|Q| - \sum_{t=1}^{T} \frac{1}{2} (\underset{\sim}{X}_t - \Phi \underset{\sim}{X}_{t-1})' Q^{-1} (\underset{\sim}{X}_t - \Phi \underset{\sim}{X}_{t-1}) - \frac{T}{2} \log|R| - \frac{1}{2} \sum_{t=1}^{T} (\underset{\sim}{Y}_t - A_t \underset{\sim}{X}_t)' R^{-1} (\underset{\sim}{Y}_t - A_t \underset{\sim}{X}_t)$$

$$\Rightarrow -2\log p_{\Theta}(x_{0:T}, y_{1:T}) = c^{*} + \log|\Sigma_0| + (\underline{x}_0 - \underline{\mu}_0)'\Sigma_0^{-1}(\underline{x}_0 - \underline{\mu}_0)$$
$$+ T\log|Q| + \sum_{t=1}^{T}(\underset{\sim}{x}_t - \Phi \underline{x}_{t-1})'Q^{-1}(\underset{\sim}{x}_t - \Phi \underline{x}_{t-1})$$
$$+ T\log|R| + \sum_{t=1}^{T}(\underset{\sim}{y}_t - A_t \underset{\sim}{x}_t)'R^{-1}(\underset{\sim}{y}_t - A_t \underset{\sim}{x}_t)$$

Now, $L_{x,y}(\Theta) = p_{\Theta}(x_{0:T}, y_{1:T})$.

$$\Rightarrow -2\log L_{x,y}(\Theta) = -2\log p_{\Theta}(x_{0:T}, y_{1:T})$$

$$\Rightarrow \underset{y_{1:T}}{E}\left(-2\log L_{x,y}(\Theta)\right) = c^{*} + \log|\Sigma_0| + E_{y_{1:T}}\left[(\underline{x}_0 - \underline{\mu}_0)'\Sigma_0^{-1}(\underline{x}_0 - \underline{\mu}_0)\right]$$
$$+ T\log|Q| + \sum_{t=1}^{T} E_{y_{1:T}}(\underset{\sim}{x}_t - \Phi \underline{x}_{t-1})'Q^{-1}(\underline{x}_t - \Phi \underline{x}_{t-1})$$
$$+ T\log|R| + \sum_{t=1}^{T} E_{y_{1:T}}(\underset{\sim}{y}_t - A_t \underset{\sim}{x}_t)'R^{-1}(\underline{y}_t - A_t \underline{x}_t)$$

$$x_0^T = E(x_0 \mid y_{1:T})$$

$$E_{y_{1:T}}\left[(\underline{x}_0 - \underline{\mu}_0)'\Sigma_0^{-1}(\underline{x}_0 - \underline{\mu}_0)\right]$$
$$= E_{y_{1:T}}\left[(\underline{x}_0 - \underline{x}_0^T + \underline{x}_0^T - \underline{\mu}_0)'\Sigma_0^{-1}(x_0 - x_0^T + x_0^T - \underline{\mu}_0)\right]$$
$$= E_{y_{1:T}}\Big[(\underline{x}_0 - \underline{x}_0^T)'\Sigma_0^{-1}(\underline{x}_0 - \underline{x}_0^T) + 2(\underline{x}_0^T - \underline{\mu}_0)'\Sigma_0^{-1}(\underline{x}_0 - \underline{x}_0^T)$$
$$+ (\underline{x}_0^T - \underline{\mu}_0)'\Sigma_0^{-1}(\underline{x}_0^T - \underline{\mu}_0)\Big]$$

$$= E_{y_{1:T}}\left[(\underline{x}_0 - \underline{x}_0^T)'\Sigma_0^{-1}(\underline{x}_0 - \underline{x}_0^T)\right] + 2(\underline{x}_0^T - \underline{\mu}_0)'\Sigma_0^{-1}E_{y_{1:T}}(x_0 - x_0^T)$$
$$+ (\underline{x}_0^T - \underline{\mu}_0)'\Sigma_0^{-1}(\underline{x}_0^T - \underline{\mu}_0)$$

$$= E_{y_{1:T}}\left[(x_0 - \underline{x}_0^T)'\Sigma_0^{-1}(\underline{x}_0 - \underline{x}_0^T)\right] + (x_0^T - \underline{\mu}_0)'\Sigma_0^{-1}(\underline{x}_0^T - \underline{\mu}_0)$$

because $\left[E_{y_{1:T}}(\underline{x}_0 - x_0^T) = E_{y_{1:T}}(x_0) - x_0^T = x_0^T - x_0^T = 0.\right]$

$$= E_{y_{1:T}} \; tr[\Sigma_0^{-1} (\underline{x}_0 - \underline{x}_0^T)(\underline{x}_0 - \underline{x}_0^T)'] + (x_0^T - \underline{\mu}_0)\Sigma_0^{-1}(x_0^T - \underline{\mu}_0)$$

$$= tr\{\Sigma_0^{-1} \cdot \underbrace{E_{y_{1:T}}[(x_0 - x_0^T)(\underline{x}_0 - \underline{x}_0^T)']}_{=P_0^T}\} + tr\{\Sigma_0^{-1}[(x_0^T - \underline{\mu}_0)(x_0^T - \underline{\mu}_0)']\}$$

$$= tr\{\Sigma_0^{-1} \cdot P_0^T\} + tr\{\Sigma_0^{-1}[(\underline{x}_0 - \underline{\mu}_0)(\underline{x}_0 - \underline{\mu}_0)']\}$$

$$= tr\{\Sigma_0^{-1}[P_0^T + (x_0 - \underline{\mu}_0)(\underline{x}_0 - \underline{\mu}_0)']\}.$$

$$E_{y_{1:T}}[(\underset{\sim}{x}_t - \Phi \underset{\sim}{x}_{t-1})' Q^{-1} (\underset{\sim}{x}_t - \Phi \underline{x}_{t-1})]$$

$$= E_{y_{1:T}}[\underset{\sim}{x}_t' Q^{-1} \underset{\sim}{x}_t - 2\, \underline{x}_{t-1}' \Phi' Q^{-1} \underset{\sim}{x}_t + \underset{\sim}{x}_{t-1}' \Phi' Q \Phi \underline{x}_{t-1}]$$

$$= E_{y_{1:T}}[\underset{\sim}{x}_t' Q^{-1} \underset{\sim}{x}_t] - 2 E_{y_{1:T}}[\underset{\sim}{x}_{t-1}' \Phi' Q^{-1} \underset{\sim}{x}_t] + E_{y_{1:T}}[\underset{\sim}{x}_{t-1}' \Phi' Q \Phi \underline{x}_{t-1}]$$

$\boxed{\text{1st term}}$

$$E_{y_{1:T}}[(\underset{\sim}{x}_t - \underset{\sim}{x}_t^T + \underset{\sim}{x}_t^T)' Q^{-1} (\underset{\sim}{x}_t - \underset{\sim}{x}_t^T + \underset{\sim}{x}_t^T)]$$

$$= E_{y_{1:T}}[(\underset{\sim}{x}_t - \underset{\sim}{x}_t^T)' Q^{-1} (\underset{\sim}{x}_t - \underset{\sim}{x}_t^T)] + \underbrace{E_{y_{1:T}}[\underset{\sim}{x}_t^{T\prime} Q^{-1} (\underset{\sim}{x}_t - \underline{x}_t^T)]}$$

$$+ E_{y_{1:T}}[\underset{\sim}{x}_t^T Q^{-1} \underset{\sim}{x}_t^T]$$

$$= x_t^{T\prime} Q^{-1} [E_{y_{1:T}}(x_t) - \underline{x}_t^T]$$

$$= x_t^{T\prime} Q^{-1} (\underline{x}_t^T - \underline{x}_t^T) = 0$$

$$= tr\{Q^{-1} \cdot E_{y_{1:T}}[(\underset{\sim}{x}_t - \underset{\sim}{x}_t^T)(\underset{\sim}{x}_t - \underset{\sim}{x}_t^T)']\}$$

$$+ [\underset{\sim}{x}_t^T]' Q^{-1} \underset{\sim}{x}_t^T$$

$$= tr\{Q^{-1} \cdot P_t^T\} + \cancel{[\underset{\sim}{x}_t^T]' Q^{-1} \underset{\sim}{x}_t^T} \; tr\{Q^{-1}([\underset{\sim}{x}_t^T][\underset{\sim}{x}_t^T]')\}$$

$$= tr\{Q^{-1}(P_t^T + \underset{\sim}{x}_t^T [\underset{\sim}{x}_t^T]')\}$$

$$\sum_{t=1}^{T} E_{y_{1:T}}(\underset{\sim}{x}_t' Q^{-1} \underset{\sim}{x}_t) = tr\{Q^{-1} \cdot \overbrace{\sum_{t=1}^{T} (P_t^T + \underset{\sim}{x}_t^T (\underset{\sim}{x}_t^T)')}^{=S_{11}}\}$$

$$= tr\{Q^{-1} S_{11}\}$$

Complete the rest of the steps.

Overall,

$E_{y_{1:T}}(Y_t) = Y_t$, naturally.

$$E_{y_{1:T}}\left[(Y_t - A_t X_t)' R^{-1} (Y_t - A_t X_t)\right]$$

$$= E_{y_{1:T}}\left[(Y_t - A_t X_t^T + A_t X_t^T - A_t X_t)' R^{-1} (Y_t - A_t X_t^T + A_t X_t^T - A_t X_t)\right]$$

$$= E_{y_{1:T}}\left[(Y_t - A_t X_t^T)' R^{-1} (Y_t - A_t X_t^T)\right]$$

$$+ E_{y_{1:T}}\left[(A_t X_t^T - A_t X_t)' R^{-1} (Y_t - A_t X_t^T)\right]$$

$$+ E_{y_{1:T}}\left[(A_t X_t^T - A_t X_t)' R^{-1} (A_t X_t^T - A_t X_t)\right]$$

$$= (Y_t - A_t X_t^T)' R^{-1} (Y_t - A_t X_t^T) + \overbrace{(A_t X_t^T - A_t E(X_t | y_{1:T}))'}^{= A_t X_t^T - A_t X_t^T = 0} R^{-1} (Y_t - A_t X_t^T)$$

$$+ E_{y_{1:T}}\left[(X_t - X_t^T)' A_t' R^{-1} A_t (X_t - X_t^T)\right]$$

$$= (Y_t - A_t X_t^T)' R^{-1} (Y_t - A_t X_t^T) + 0 + E_{y_{1:T}} \, tr\left\{R^{-1} \left[A_t (X_t - X_t^T)(X_t - X_t^T)' A_t'\right]\right\}$$

$$= tr\left\{R^{-1}\left[(Y_t - A_t X_t^T)(Y_t - A_t X_t^T)'\right]\right\} + tr\left\{R^{-1} A_t \underbrace{E_{y_{1:T}}\left[(X_t - X_t^T)(X_t - X_t^T)'\right]}_{= P_t^T} A_t'\right\}$$

$$= tr\left\{R^{-1} \left[(Y_t - A_t X_t^T)(Y_t - A_t X_t^T)' + A_t P_t^T A_t'\right]\right\}$$

Overall, in $E[-2\log L_{x,y}(\Theta) \mid y_{1:T}]$, the terms involving the following parameters are as follows.

$\underline{\mu}_0$: only appears in $tr\left[\Sigma_0^{-1}\left\{P_0^T + (\underline{x}_0^T - \underline{\mu}_0)(\underline{x}_0^T - \underline{\mu}_0)'\right\}\right]$

$$= tr\{\Sigma_0^{-1} P_0^T\} + (\underline{x}_0^T - \underline{\mu}_0)' \Sigma_0^{-1} (\underline{x}_0^T - \underline{\mu}_0)$$

So, we need to minimize. $(\underline{x}_0^T - \underline{\mu}_0)' \Sigma_0^{-1} (\underline{x}_0^T - \underline{\mu}_0)$.

$$f(\underline{\mu}_0) = (x_0^T - \underline{\mu}_0)' \Sigma_0^{-1} (x_0^T - \underline{\mu}_0).$$

$$f'(\underline{\mu}_0) = -\Sigma_0^{-1} (x_0^T - \underline{\mu}_0).$$

$$f''(\underline{\mu}_0) = \Sigma_0^{-1} \text{ is p.d.}$$

$f'(\underline{\mu}_0) = 0 \Rightarrow \underline{\mu}_0 = x_0^T$. Thus, the updated value $\hat{\mu}_0 = x_0^T$

$$\underline{\mu}_0^{(i)} = x_0^T.$$

minimizing $tr\left[\Sigma_0^{-1} \cdot \left\{P_0^T + (\underline{x}_0^T - \underline{\mu}_0)(\underline{x}_0^T - \underline{\mu}_0)'\right\}\right]$ w.r.t. Σ_0

we obtain $\hat{\Sigma}_0 = P_0^T + (x_0^T - \hat{\mu}_0)(x_0^T - \hat{\mu}_0)'$

$$= P_0^T + \underbrace{(x_0^T - x_0^T)(x_0^T - x_0^T)'}_{=0} = P_0^T.$$

Φ appears within $tr\left\{\left[Q^{-1}\left\{S_{11} - S_{10}\Phi' - \Phi S_{10} + \Phi S_{00}\Phi'\right\}\right]\right.$

So, derivative w.r.t. Φ gives $-2S_{10} + 2\Phi S_{00} = 0$.

$$\Rightarrow S_{10} = \Phi S_{00} \Rightarrow S_{10} S_{00}^{-1} = \Phi S_{00} S_{00}^{-1}$$

$$\Rightarrow \Phi = S_{10} S_{00}^{-1}.$$

~~minimizing w.r.t. Q, we have~~ $\hat{Q} = \cancel{S_{11} - S_{10}\Phi' - \Phi S_{10} + \Phi S_{00}\Phi'}$

$= \cancel{S_{11} - S_{10} S_{00}^{-1} S_{10}}$

Q appears in the expressions,

$$T \log|Q| + \cancel{\sum} \, tr\{Q^{-1}.\{S_{11} - S_{10}\Phi' - \Phi S_{10} + \Phi S_{00}\Phi'\}\}$$

this function is minimized when.

$$Q = T^{-1}\{S_{11} - S_{10}\hat{\Phi}' - \hat{\Phi} S_{10} + \hat{\Phi} S_{00}\hat{\Phi}'\}.$$

$$= T^{-1}\{S_{11} - S_{10} S_{00}^{-1} S_{10} - S_{10} S_{00}^{-1} S_{10} + S_{10} S_{00}^{-1} S_{00} S_{00}^{-1} S_{10}\}$$

$$= T^{-1}\{S_{11} - 2S_{10} S_{00}^{-1} S_{10} + S_{10} S_{00}^{-1} S_{10}\}$$

$$= T^{-1}\{S_{11} - S_{10} S_{00}^{-1} S_{10}\}.$$

Similarly, $\hat{R} = T^{-1}\{\sum_{t=1}^{T}\{(Y_t - A_t X_t^T)(Y_t - A_t X_t^T)' + A_t P_t^T A_t'\}\}$