Derivations for the Layered Dynamic Texture and Temporally-Switching Layered Dynamic Texture

Antoni B. Chan and Nuno Vasconcelos

Statistical Visual Computing Laboratory

SVCL UCSD

SVCL-TR 2009/01

June 2009

Derivations for the Layered Dynamic Texture and Temporally-Switching Layered Dynamic Texture

Antoni B. Chan and Nuno Vasconcelos

Statistical Visual Computing Lab
Department of Electrical and Computer Engineering
University of California, San Diego

June 2009

Abstract

This is the supplemental material for the paper "Variational Layered Dynamic Textures" in CVPR 2009 [1]. The supplemental contains derivations for the variational approximation of the layered dynamic texture (LDT), and the EM algorithm and variational approximations of the temporally-switching layered dynamic texture (TS-LDT).

Author email: abchan@ucsd.edu

SVCL Technical reports are available on the SVCL's web page at
http://www.svcl.ucsd.edu

University of California, San Diego
Statistical Visual Computing Laboratory
9500 Gilman Drive, Mail code 0407
EBU 1, Room 5512
La Jolla, CA 92093-0407

1 Derivation of the variational approximation for LDT

In this section, we derive a variational approximation for the layered dynamic texture (LDT). Substituting the approximate factorial distribution

$$q(X,Z) = \prod_{j=1}^{K} q(x^{(j)}) \prod_{i=1}^{m} q(z_i) \tag{S.1}$$

into the $\mathcal{L}$ function of (10) yields

$$\mathcal{L}(q(X,Z)) = \int \prod_{j=1}^{K} q(x^{(j)}) \prod_{i=1}^{m} q(z_i) \log \frac{\prod_{j=1}^{K} q(x^{(j)}) \prod_{i=1}^{m} q(z_i)}{p(X,Y,Z)} dX dZ. \tag{S.2}$$

This is minimized by sequentially optimizing each of the factors $q(x^{(j)})$ and $q(z_i)$, while holding the remaining constant [2]. For convenience, we define the variable $W = \{X, Z\}$. Rewriting (S.2) in terms of a single factor $q(w_l)$, while holding all others constant,

$$\begin{aligned} &\mathcal{L}(q(W)) \\ &\quad \propto \int q(w_l) \log q(w_l) dw_l - \int q(w_l) \int \prod_{k \neq l} q(w_k) \log p(W,Y) dW && \text{(S.3)} \\ &\quad = \int q(w_l) \log q(w_l) dw_l - \int q(w_l) \log \tilde{p}(w_l, Y) dw_l && \text{(S.4)} \\ &\quad = D(q(w_l) \,\|\tilde{p}(w_l, Y)) , && \text{(S.5)} \end{aligned}$$

where in (S.3) we have dropped terms that do not depend on $q(w_l)$ (and hence do not affect the optimization), and defined $\tilde{p}(w_l, Y)$ as

$$\log \tilde{p}(w_l, Y) \propto \mathbb{E}_{W_{k \neq l}}[\log p(W, Y)], \tag{S.6}$$

where $\mathbb{E}_{W_{k \neq l}}[\log p(W,Y)] = \int \prod_{k \neq l} q(w_k) \log p(W,Y) dW_{k \neq l}$. Since (S.5) is minimized when $q^*(w_l) = \tilde{p}(w_l, Y)$, the optimal factor $q(w_l)$ is equal to the expectation of the joint log-likelihood with respect to the other factors $W_{k \neq l}$.

We next discuss the joint distribution of the LDT, followed by deriving the forms of the optimal factors $q(x^{(j)})$ and $q(z_i)$. For convenience, we ignore normalization constants during the derivation, and reinstate them after the forms of the factors are known.

1.1 Joint distribution of the LDT

The LDT model assumes that the state processes $X = \{x^{(j)}\}_{j=1}^{K}$ and the layer assignments Z are independent, i.e. the layer dynamics are independent of its location. Under this assumption, the joint distribution factors as

$$\begin{aligned} p(X,Y,Z) &= p(Y|X,Z)p(X)p(Z) && \text{(S.7)} \\ &= \prod_{i=1}^{m} \prod_{j=1}^{K} p(y_i | x^{(j)}, z_i = j)^{z_i^{(j)}} \prod_{j=1}^{K} p(x^{(j)}) p(Z), && \text{(S.8)} \end{aligned}$$

where $Y = \{y_i\}_{i=1}^m$. Each state-sequence is a Gauss-Markov process, with distribution

$$p(x^{(j)}) = p(x_1^{(j)}) \prod_{t=2}^{\tau} p(x_t^{(j)} | x_{t-1}^{(j)}), \tag{S.9}$$

where the individual state densities are

$$p(x_1^{(j)}) = G(x_1^{(j)}, \mu^{(j)}, Q^{(j)}), \qquad p(x_t^{(j)} | x_{t-1}^{(j)}) = G(x_t^{(j)}, A^{(j)} x_{t-1}^{(j)}, Q^{(j)}), \tag{S.10}$$

and $G(x, \mu, \Sigma)$ is a Gaussian of mean μ and covariance Σ. When conditioned on state sequences and layer assignments, pixel values are independent, and pixel trajectories distributed as

$$p(y_i | x^{(j)}, z_i = j) = \prod_{t=1}^{\tau} p(y_{i,t} | x_t^{(j)}, z_i = j), \tag{S.11}$$

where

$$p(y_{i,t} | x_t^{(j)}, z_i = j) = G(y_{i,t}, C_i^{(j)} x_t^{(j)}, r^{(j)}). \tag{S.12}$$

Finally, the layer assignments are jointly distributed as

$$p(Z) = \frac{1}{\mathcal{Z}_Z} \prod_{i=1}^{m} V_i(z_i) \prod_{(i,i') \in \mathcal{E}} V_{i,i'}(z_i, z_{i'}), \tag{S.13}$$

where $\mathcal{E}$ is the set of edges of the MRF, $\mathcal{Z}_Z$ a normalization constant (partition function), and V_i and $V_{i,i'}$ potential functions of the form

$$\begin{aligned} V_i(z_i) &= \prod_{j=1}^{K} (\alpha_i^{(j)})^{z_i^{(j)}} = \begin{cases} \alpha_i^{(1)}, z_i = 1 \\ \vdots \\ \alpha_i^{(K)}, z_i = K \end{cases}, \\ V_{i,i'}(z_i, z_{i'}) &= \gamma_2 \prod_{j=1}^{K} \left(\frac{\gamma_1}{\gamma_2}\right)^{z_i^{(j)} z_{i'}^{(j)}} = \begin{cases} \gamma_1, z_i = z_{i'} \\ \gamma_2, z_i \neq z_{i'} \end{cases}. \end{aligned} \tag{S.14}$$

V_i is the prior probability of each layer, while $V_{i,i'}$ attributes higher probability to configurations with neighboring pixels in the same layer.

1.2 Optimization of $q(x^{(j)})$

Rewriting (S.6) with $w_l = x^{(j)}$,

$$\begin{aligned} \log q^*(x^{(j)}) &\propto \log \tilde{p}(x^{(j)}, Y) = \mathbb{E}_{Z, X_{k \neq j}} [\log p(X, Y, Z)] & \text{(S.15)} \\ &\propto \mathbb{E}_{Z, X_{k \neq j}} \left[\sum_{i=1}^{m} z_i^{(j)} \log p(y_i | x^{(j)}, z_i = j) + \log p(x^{(j)}) \right] & \text{(S.16)} \\ &= \sum_{i=1}^{m} \mathbb{E}_{z_i} [z_i^{(j)}] \log p(y_i | x^{(j)}, z_i = j) + \log p(x^{(j)}), & \text{(S.17)} \end{aligned}$$

where in (S.16) we have dropped the terms of the joint log-likelihood (S.8) that are not a function of $x^{(j)}$. Finally, defining $h_i^{(j)} = \mathbb{E}_{z_i}[z_i^{(j)}] = \int q(z_i) z_i^{(j)} dz_i$, and the normalization constant

$$\mathcal{Z}_q^{(j)} = \int p(x^{(j)}) \prod_{i=1}^{m} p(y_i | x^{(j)}, z_i = j)^{h_i^{(j)}} dx^{(j)}, \tag{S.18}$$

the optimal $q(x^{(j)})$ is given by (12).

1.3 Optimization of $q(z_i)$

Rewriting (S.6) with $w_l = z_i$ and dropping terms that do not depend on z_i,

$$\begin{aligned}
\log q^*(z_i) &\propto \log \tilde{p}(z_i, Y) = \mathbb{E}_{X, Z_{k \neq i}}[\log p(X, Y, Z)] && \text{(S.19)} \\
&\propto \mathbb{E}_{X, Z_{k \neq i}} \left[\sum_{j=1}^{K} z_i^{(j)} \log p(y_i | x^{(j)}, z_i = j) + \log p(Z) \right] \\
&= \sum_{j=1}^{K} z_i^{(j)} \mathbb{E}_{x^{(j)}}[\log p(y_i | x^{(j)}, z_i = j)] + \mathbb{E}_{Z_{k \neq i}}[\log p(Z)]. && \text{(S.20)}
\end{aligned}$$

For the last term, we have

$$\begin{aligned}
&\mathbb{E}_{Z_{k \neq i}}[\log p(Z)] \\
&\propto \mathbb{E}_{Z_{k \neq i}}[\log(V_i(z_i) \prod_{(i,i') \in \mathcal{E}} V_{i,i'}(z_i, z_{i'}))] && \text{(S.21)} \\
&= \log V_i(z_i) + \sum_{(i,i') \in \mathcal{E}} \mathbb{E}_{z_{i'}}[\log V_{i,i'}(z_i, z_{i'})] && \text{(S.22)} \\
&= \sum_{j=1}^{K} z_i^{(j)} \log \alpha_i^{(j)} + \sum_{(i,i') \in \mathcal{E}} \mathbb{E}_{z_{i'}}[\sum_{j=1}^{K} z_i^{(j)} z_{i'}^{(j)} \log \frac{\gamma_1}{\gamma_2} + \log \gamma_2] && \text{(S.23)} \\
&\propto \sum_{j=1}^{K} z_i^{(j)} \log \alpha_i^{(j)} + \sum_{j=1}^{K} z_i^{(j)} \sum_{(i,i') \in \mathcal{E}} \mathbb{E}_{z_{i'}}[z_{i'}^{(j)}] \log \frac{\gamma_1}{\gamma_2} && \text{(S.24)} \\
&= \sum_{j=1}^{K} z_i^{(j)} \left(\log \alpha_i^{(j)} + \sum_{(i,i') \in \mathcal{E}} h_{i'}^{(j)} \log \frac{\gamma_1}{\gamma_2} \right). && \text{(S.25)}
\end{aligned}$$

Hence,

$$\begin{aligned}
&\log q^*(z_i) \\
&\propto \sum_{j=1}^{K} z_i^{(j)} \left(\mathbb{E}_{x^{(j)}}[\log p(y_i | x^{(j)}, z_i = j)] + \sum_{(i,i') \in \mathcal{E}} h_{i'}^{(j)} \log \frac{\gamma_1}{\gamma_2} + \log \alpha_i^{(j)} \right) \\
&= \sum_{j=1}^{K} z_i^{(j)} \log(g_i^{(j)} \alpha_i^{(j)}), && \text{(S.26)}
\end{aligned}$$

where $g_i^{(j)}$ is defined in (15). This is a multinomial distribution of normalization constant $\sum_{j=1}^{K}(\alpha_i^{(j)} g_i^{(j)})$, leading to (13) with $h_i^{(j)}$ as given in (14).

1.4 Normalization constant for $q(x^{(j)})$

Taking the log of (S.18),

$$\log \mathcal{Z}_q^{(j)} = \log \int p(x^{(j)}) \prod_{i=1}^{m} p(y_i | x^{(j)}, z_i = j)^{h_i^{(j)}} dx^{(j)} \tag{S.27}$$

$$= \log \int p(x^{(j)}) \prod_{i=1}^{m} \prod_{t=1}^{\tau} p(y_{i,t} | x_t^{(j)}, z_i = j)^{h_i^{(j)}} dx^{(j)}. \tag{S.28}$$

Note that the term $p(y_{i,t}|x_t^{(j)}, z_i = j)^{h_i^{(j)}}$ does not affect the integral when $h_i^{(j)} = 0$. Defining $\mathcal{I}_j$ as the set of indices with non-zero $h_i^{(j)}$, i.e. $\mathcal{I}_j = \{i | h_i^{(j)} > 0\}$, (S.28) becomes

$$\log \mathcal{Z}_q^{(j)} = \log \int p(x^{(j)}) \prod_{i \in \mathcal{I}_j} \prod_{t=1}^{\tau} p(y_{i,t} | x_t^{(j)}, z_i = j)^{h_i^{(j)}} dx^{(j)}, \tag{S.29}$$

where

$$p(y_{i,t} | x_t^{(j)}, z_i = j)^{h_i^{(j)}} = G(y_{i,t}, C_i^{(j)} x_t^{(j)}, r^{(j)})^{h_i^{(j)}} \tag{S.30}$$

$$= (2\pi r^{(j)})^{-\frac{1}{2} h_i^{(j)}} \left(\frac{2\pi r^{(j)}}{h_i^{(j)}} \right)^{\frac{1}{2}} G\left(y_{i,t}, C_i^{(j)} x_t^{(j)}, \frac{r^{(j)}}{h_i^{(j)}} \right). \tag{S.31}$$

For convenience, we will define an LDS over the subset $\mathcal{I}_j$ parameterized by $\tilde{\Theta}_j = \{A^{(j)}, Q^{(j)}, \tilde{C}^{(j)}, \tilde{R}_j, \mu^{(j)}\}$, where $\tilde{C}^{(j)} = [C_i^{(j)}]_{i \in \mathcal{I}_j}$, and $\tilde{R}_j$ is diagonal with entries $\tilde{r}_i^{(j)} = \frac{r^{(j)}}{h_i^{(j)}}$ for $i \in \mathcal{I}_j$. Noting that this LDS has conditional observation likelihood $\tilde{p}(y_{i,t}|x_t^{(j)}, z_i = j) = G(y_{i,t}, C_i^{(j)} x_t^{(j)}, \tilde{r}_i^{(j)})$, we can rewrite

$$p(y_{i,t} | x_t^{(j)}, z_i = j)^{h_i^{(j)}} = \left(2\pi r^{(j)}\right)^{\frac{1}{2}(1 - h_i^{(j)})} (h_i^{(j)})^{-\frac{1}{2}} \tilde{p}(y_{i,t} | x_t^{(j)}, z_i = j) \tag{S.32}$$

and, from (S.29),

$$\log \mathcal{Z}_q^{(j)} \tag{S.33}$$
$$= \log \int p(x^{(j)}) \prod_{i \in \mathcal{I}_j} \prod_{t=1}^{\tau} \left[\left(2\pi r^{(j)}\right)^{\frac{1}{2}(1 - h_i^{(j)})} (h_i^{(j)})^{-\frac{1}{2}} \tilde{p}(y_{i,t} | x_t^{(j)}, z_i = j) \right] dx^{(j)}.$$

Since, under the restricted LDS, the likelihood of $Y_j = [y_i]_{i \in \mathcal{I}_j}$ is

$$\tilde{p}_j(Y_j) = \int p(x^{(j)}) \prod_{i \in \mathcal{I}_j} \prod_{t=1}^{\tau} \tilde{p}(y_{i,t} | x_t^{(j)}, z_i = j) dx^{(j)}, \tag{S.34}$$

it follows that

$$\log \mathcal{Z}_q^{(j)} = \log \left[\tilde{p}_j(Y_j) \prod_{i \in \mathcal{I}_j} \prod_{t=1}^{\tau} \left(2\pi r^{(j)}\right)^{\frac{1}{2}(1-h_i^{(j)})} (h_i^{(j)})^{-\frac{1}{2}} \right] \quad \text{(S.35)}$$

$$= \frac{\tau}{2} \sum_{i \in \mathcal{I}_j} (1 - h_i^{(j)}) \log(2\pi r^{(j)}) - \frac{\tau}{2} \sum_{i \in \mathcal{I}_j} \log h_i^{(j)} + \log \tilde{p}_j(Y_j). \quad \text{(S.36)}$$

2 Derivation of the EM algorithm for the TS-LDT

In this section, we derive the EM algorithm for the temporally-switching layered dynamic texture (TS-LDT). We begin by deriving the complete data log-likelihood, followed by the E and M steps.

2.1 Complete data log-likelihood of the TS-LDT

We introduce an indicator variable $z_{i,t}^{(j)}$ of value 1 if and only if $z_{i,t} = j$, and 0 otherwise. Under the assumption that the state processes X and layer assignments Z are independent, the joint distribution factors as

$$p(X, Y, Z) = p(Y|X, Z)p(X)p(Z) \quad \text{(S.37)}$$

$$= \prod_{i=1}^{m} \prod_{j=1}^{K} \prod_{t=1}^{\tau} p(y_{i,t}|x_t^{(j)}, z_{i,t} = j)^{z_{i,t}^{(j)}} \prod_{j=1}^{K} p(x^{(j)})p(Z), \quad \text{(S.38)}$$

where the conditional observation likelihood is

$$p(y_{i,t}|x_t^{(j)}, z_{i,t} = j) = G(y_{i,t}, C_i^{(j)} x_t^{(j)} + \gamma_i^{(j)}, r^{(j)}), \quad \text{(S.39)}$$

and the distribution for $p(x^{(j)})$ is the same as the LDT, given in (S.9, S.10). Finally, for the layer assignments Z, we assume that each frame $Z_t = \{z_{i,t}\}_{i=1}^{m}$ has the same MRF structure, with temporal edges only connecting nodes corresponding to the same pixel (e.g. $z_{i,t}$ and $z_{i,t+1}$). The layer assignments are then jointly distributed as

$$p(Z) = \frac{1}{\mathcal{Z}_Z} \left[\prod_{t=1}^{\tau} \prod_{i=1}^{m} V_{i,t}(z_{i,t}) \right] \left[\prod_{t=1}^{\tau} \prod_{(i,i') \in \mathcal{E}_t} V_{i,i'}(z_{i,t}, z_{i',t}) \right] \quad \text{(S.40)}$$

$$\cdot \left[\prod_{i=1}^{m} \prod_{(t,t') \in \mathcal{E}_i} V_{t,t'}(z_{i,t}, z_{i,t'}) \right],$$

where $\mathcal{E}_t$ is the set of MRF in frame t, $\mathcal{E}_i$ is the set of MRF edges between frames for pixel i, and $\mathcal{Z}_Z$ a normalization constant (partition function). The potential functions $V_{i,t}$, $V_{i,i'}$, $V_{t,t'}$ are of the form:

$$V_{i,t}(z_{i,t}) = \prod_{j=1}^{K} (\alpha_{i,t}^{(j)})^{z_{i,t}^{(j)}} = \begin{cases} \alpha_{i,t}^{(1)}, z_{i,t} = 1 \\ \vdots \\ \alpha_{i,t}^{(K)}, z_{i,t} = K \end{cases}, \quad \text{(S.41)}$$

$$V_{i,i'}(z_{i,t}, z_{i',t}) = \gamma_2 \prod_{j=1}^{K} \left(\frac{\gamma_1}{\gamma_2}\right)^{z_{i,t}^{(j)} z_{i',t}^{(j)}} = \begin{cases} \gamma_1, z_{i,t} = z_{i',t} \\ \gamma_2, z_{i,t} \neq z_{i',t} \end{cases},$$
$$V_{t,t'}(z_{i,t}, z_{i,t'}) = \beta_2 \prod_{j=1}^{K} \left(\frac{\beta_1}{\beta_2}\right)^{z_{i,t}^{(j)} z_{i,t'}^{(j)}} = \begin{cases} \beta_1, z_{i,t} = z_{i,t'} \\ \beta_2, z_{i,t} \neq z_{i,t'} \end{cases}, \quad \text{(S.42)}$$

where $V_{i,t}$ is the prior probability of each layer in each frame t, and $V_{i,i'}$ and $V_{t,t'}$ attributes higher probability to configurations with neighboring pixels (both spatially and temporally) in the same layer.

Taking the logarithm of (S.38), the complete data log-likelihood is

$$\log p(X, Y, Z) = \sum_{i=1}^{m} \sum_{j=1}^{K} \sum_{t=1}^{\tau} z_{i,t}^{(j)} \log p(y_{i,t} | x_t^{(j)}, z_{i,t} = j) \quad \text{(S.43)}$$
$$+ \sum_{j=1}^{K} \left(\log p(x_1^{(j)}) + \sum_{t=2}^{\tau} \log p(x_t^{(j)} | x_{t-1}^{(j)}) \right) + \log p(Z).$$

Using (S.10) and (S.39) and dropping terms that do not depend on the parameters Θ (and thus play no role in the M-step),

$$\log p(X, Y, Z) = \quad \text{(S.44)}$$
$$-\frac{1}{2} \sum_{j=1}^{K} \sum_{i=1}^{m} \sum_{t=1}^{\tau} z_{i,t}^{(j)} \left(\left\| y_{i,t} - C_i^{(j)} x_t^{(j)} - \gamma_i^{(j)} \right\|_{r^{(j)}}^2 + \log r^{(j)} \right)$$
$$-\frac{1}{2} \sum_{j=1}^{K} \left(\left\| x_1^{(j)} - \mu^{(j)} \right\|_{Q^{(j)}}^2 + \sum_{t=2}^{\tau} \left\| x_t^{(j)} - A^{(j)} x_{t-1}^{(j)} \right\|_{Q^{(j)}}^2 + \tau \log \left| Q^{(j)} \right| \right).$$

Note that $p(Z)$ can be ignored since the parameters of the MRF are constants. Finally, the complete data log-likelihood is

$$\log p(X, Y, Z) = \quad \text{(S.45)}$$
$$-\frac{1}{2} \sum_{j=1}^{K} \sum_{i=1}^{m} \sum_{t=1}^{\tau} z_{i,t}^{(j)} \frac{1}{r^{(j)}} \left((y_{i,t} - \gamma_i^{(j)})^2 - 2(y_{i,t} - \gamma_i^{(j)}) C_i^{(j)} x_t^{(j)} \right.$$
$$\left. + C_i^{(j)} P_{t,t}^{(j)} {C_i^{(j)}}^T \right)$$
$$-\frac{1}{2} \sum_{j=1}^{K} \operatorname{tr} \left({Q^{(j)}}^{-1} \left(P_{1,1}^{(j)} - x_1^{(j)} {\mu^{(j)}}^T - \mu^{(j)} {x_1^{(j)}}^T + \mu^{(j)} {\mu^{(j)}}^T \right) \right)$$
$$-\frac{1}{2} \sum_{j=1}^{K} \sum_{t=2}^{\tau} \operatorname{tr} \left({Q^{(j)}}^{-1} \left(P_{t,t}^{(j)} - P_{t,t-1}^{(j)} {A^{(j)}}^T - A^{(j)} {P_{t,t-1}^{(j)}}^T \right. \right.$$
$$\left. \left. + A^{(j)} P_{t-1,t-1}^{(j)} {A^{(j)}}^T \right) \right)$$

$$- \frac{1}{2} \sum_{j=1}^{K} \sum_{i=1}^{m} \sum_{t=1}^{\tau} z_{i,t}^{(j)} \log r^{(j)} - \frac{\tau}{2} \sum_{j=1}^{K} \log \left| Q^{(j)} \right|,$$

where we define $P_{t,t}^{(j)} = x_t^{(j)} x_t^{(j)T}$ and $P_{t,t-1}^{(j)} = x_t^{(j)} x_{t-1}^{(j)}{}^T$.

2.2 E-step

From (S.45), it follows that the E-step of (4) requires conditional expectations of two forms:

$$\mathbb{E}_{X,Z|Y}[f(x^{(j)})] = \mathbb{E}_{X|Y}[f(x^{(j)})], \tag{S.46}$$

$$\mathbb{E}_{X,Z|Y}[z_{i,t}^{(j)} f(x^{(j)})] = \mathbb{E}_{Z|Y}[z_{i,t}^{(j)}] \mathbb{E}_{X|Y,z_{i,t}=j}[f(x^{(j)})] \tag{S.47}$$

for some function f of $x^{(j)}$, and where $\mathbb{E}_{X|Y,z_{i,t}=j}$ is the conditional expectation of X given the observation Y and that the i-th pixel at time t belongs to layer j. Defining the conditional expectations in (18) and aggregated statistics in (19), substituting (S.45) into (4), leads to the $\mathcal{Q}$ function

$$\begin{aligned}
\mathcal{Q}(\Theta; \hat{\Theta}) = & \\
& -\frac{1}{2} \sum_{j=1}^{K} \frac{1}{r^{(j)}} \sum_{i=1}^{m} \left(\sum_{t=1}^{\tau} \hat{z}_{i,t}^{(j)} (y_{i,t} - \gamma_i^{(j)})^2 - 2 C_i^{(j)} \Gamma_i^{(j)} + C_i^{(j)} \Phi_i^{(j)} C_i^{(j)T} \right) \\
& - \frac{1}{2} \sum_{j=1}^{K} \mathrm{tr} \Big(Q^{(j)^{-1}} \Big(\hat{P}_{1,1}^{(j)} - \hat{x}_1^{(j)} \mu^{(j)T} - \mu^{(j)} (\hat{x}_1^{(j)})^T + \mu^{(j)} \mu^{(j)T} + \phi_2^{(j)} \\
& - \psi^{(j)} A^{(j)T} - A^{(j)} \psi^{(j)T} + A^{(j)} \phi_1^{(j)} A^{(j)T} \Big) \Big) - \frac{1}{2} \sum_{j=1}^{K} \hat{N}_j \log r^{(j)} \\
& - \frac{\tau}{2} \sum_{j=1}^{K} \log \left| Q^{(j)} \right|.
\end{aligned} \tag{S.48}$$

2.3 M-step

The maximization of the $\mathcal{Q}$ function with respect to the TS-LDT parameters leads to two optimization problems. The first is a maximization with respect to a square matrix X of the form

$$X^* = \underset{X}{\operatorname{argmax}} - \frac{1}{2} \mathrm{tr} \left(X^{-1} A \right) - \frac{b}{2} \log |X| \quad \Rightarrow \quad X^* = \frac{1}{b} A. \tag{S.49}$$

The second is a maximization with respect to a matrix X with the form

$$X^* = \underset{X}{\operatorname{argmax}} - \frac{1}{2} \mathrm{tr} \left[D(-BX^T - XB^T + XCX^T) \right] \quad \Rightarrow \quad X^* = BC^{-1}, \tag{S.50}$$

where D and C are symmetric and invertible matrices.

The optimal parameters are found by collecting the relevant terms in (S.49) and maximizing. This leads to a number of problems of the form of (S.50), namely

$$\begin{aligned} A^{(j)*} &= \operatorname*{argmax}_{A^{(j)}} -\frac{1}{2}\mathrm{tr}\left(Q^{(j)^{-1}}(-\psi^{(j)}A^{(j)^T} - A^{(j)}\psi^{(j)^T}\right. \\ &\quad \left. + A^{(j)}\phi_1^{(j)}A^{(j)^T})\right), \end{aligned} \tag{S.51}$$

$$\begin{aligned} \mu^{(j)*} &= \operatorname*{argmax}_{\mu^{(j)}} -\frac{1}{2}\mathrm{tr}\left(Q^{(j)^{-1}}\left(-\hat{x}_1^{(j)}\mu^{(j)^T} - \mu^{(j)}(\hat{x}_1^{(j)})^T\right.\right. \\ &\quad \left.\left. + \mu^{(j)}\mu^{(j)^T}\right)\right), \end{aligned} \tag{S.52}$$

$$C_i^{(j)*} = \operatorname*{argmax}_{C_i^{(j)}} -\frac{1}{2}\frac{1}{r^{(j)}}\hat{z}_i^{(j)}\left(-2C_i^{(j)}\Gamma_i^{(j)} + C_i^{(j)}\Phi_i^{(j)}C_i^{(j)^T}\right). \tag{S.53}$$

Using (S.50) leads to the solutions

$$A^{(j)^*} = \psi^{(j)}\phi_1^{(j)^{-1}}, \quad \mu^{(j)*} = \hat{x}_1^{(j)}, \quad C_i^{(j)^*} = \Gamma_i^{(j)^T}\Phi_i^{(j)^{-1}}. \tag{S.54}$$

The remaining problems are of the form of (S.49)

$$\begin{aligned} Q^{(j)*} &= \operatorname*{argmax}_{Q^{(j)}} -\frac{1}{2}\mathrm{tr}Q^{(j)^{-1}}\left(\hat{P}_{1,1}^{(j)} - \hat{x}_1^{(j)}\mu^{(j)^T} - \mu^{(j)}(\hat{x}_1^{(j)})^T\right. \\ &\quad \left. +\mu^{(j)}\mu^{(j)^T} + \phi_2^{(j)} - \psi^{(j)}A^{(j)^T} - A^{(j)}\psi^{(j)^T} + A^{(j)}\phi_1^{(j)}A^{(j)^T}\right) \\ &\quad - \frac{\tau}{2}\log\left|Q^{(j)}\right|, \end{aligned} \tag{S.55}$$

$$\begin{aligned} r^{(j)*} &= \operatorname*{argmax}_{r^{(j)}} -\frac{1}{2}\frac{1}{r^{(j)}}\sum_{i=1}^{m}\left(\sum_{t=1}^{\tau}\hat{z}_{i,t}^{(j)}(y_{i,t} - \gamma_i^{(j)})^2 - 2C_i^{(j)}\Gamma_i^{(j)}\right. \\ &\quad \left. + C_i^{(j)}\Phi_i^{(j)}C_i^{(j)^T}\right) - \frac{1}{2}\hat{N}_j\log r^{(j)} \end{aligned} \tag{S.56}$$

In the first case, it follows from (S.49) that

$$\begin{aligned} Q^{(j)*} &= \frac{1}{\tau}\left(\hat{P}_{1,1}^{(j)} - \hat{x}_1^{(j)}\mu^{(j)^T} - \mu^{(j)}(\hat{x}_1^{(j)})^T + \mu^{(j)}\mu^{(j)^T} + \phi_2^{(j)}\right. \\ &\quad \left. -\psi^{(j)}A^{(j)^T} - A^{(j)}\psi^{(j)^T} + A^{(j)}\phi_1^{(j)}A^{(j)^T}\right) \end{aligned} \tag{S.57}$$

$$= \frac{1}{\tau}\left(\hat{P}_{1,1}^{(j)} - \mu^{(j)*}\mu^{(j)*^T} + \phi_2^{(j)} - A^{(j)*}\psi^{(j)^T}\right). \tag{S.58}$$

In the second case,

$$r^{(j)*} = \frac{1}{\hat{N}_j}\sum_{i=1}^{m}\left[\sum_{t=1}^{\tau}\hat{z}_{i,t}^{(j)}(y_{i,t} - \gamma_i^{(j)})^2 - 2C_i^{(j)}\Gamma_i^{(j)} + C_i^{(j)}\Phi_i^{(j)}C_i^{(j)^T}\right] \tag{S.59}$$

$$= \frac{1}{\hat{N}_j}\sum_{i=1}^{m}\left[\sum_{t=1}^{\tau}\hat{z}_{i,t}^{(j)}(y_{i,t} - \gamma_i^{(j)})^2 - C_i^{(j)*}\Gamma_i^{(j)}\right]. \tag{S.60}$$

Finally, noting that $\frac{\partial}{\partial \gamma_i^{(j)}}\Gamma_i^{(j)} = -\sum_{t=1}^{\tau} \hat{z}_{i,t}^{(j)} \hat{x}_{t|i}^{(j)} = -\xi_i^{(j)}$, the estimate of the mean parameters are

$$\frac{\partial \mathcal{Q}}{\partial \gamma_i^{(j)}} = \frac{1}{r^{(j)}} \left(\sum_{t=1}^{\tau} -2\hat{z}_{i,t}^{(j)}(y_{i,t} - \gamma_i^{(j)}) + 2C_i^{(j)}\xi_i^{(j)} \right) = 0, \tag{S.61}$$

$$\Rightarrow \gamma_i^{(j)} = \frac{1}{\sum_{t=1}^{\tau} \hat{z}_{i,t}^{(j)}} \left(\sum_{t=1}^{\tau} \hat{z}_{i,t}^{(j)} y_{i,t} - C_i^{(j)}\xi_i^{(j)} \right). \tag{S.62}$$

3 Variational approximation for the TS-LDT

In this section, we derive the variational approximation for the TS-LDT, which follows closely to that of the LDT. Substituting (20) into (10), leads to

$$\mathcal{L}(q(X,Z)) = \int \prod_j q(x^{(j)}) \prod_{i,t} q(z_{i,t}) \log \frac{\prod_j q(x^{(j)}) \prod_{i,t} q(z_{i,t})}{p(X,Y,Z)} dX dZ. \tag{S.63}$$

The $\mathcal{L}$ function (S.63) is minimized by sequentially optimizing each of the factors $q(x^{(j)})$ and $q(z_{i,t})$, while holding the others constant [2].

3.1 Optimization of $q(x^{(j)})$

Rewriting (S.6) with $w_l = x^{(j)}$,

$$\log q^*(x^{(j)}) \propto \log \tilde{p}(x^{(j)}, Y) = \mathbb{E}_{Z, X_{k \neq j}}[\log p(X,Y,Z)] \tag{S.64}$$

$$\propto \quad \mathbb{E}_{Z, X_{k \neq j}} \left[\sum_{t=1}^{\tau} \sum_{i=1}^{m} z_{i,t}^{(j)} \log p(y_{i,t} | x_t^{(j)}, z_{i,t} = j) + \log p(x^{(j)}) \right] \tag{S.65}$$

$$= \quad \sum_{t=1}^{\tau} \sum_{i=1}^{m} \mathbb{E}_{z_{i,t}}[z_{i,t}^{(j)}] \log p(y_{i,t} | x_t^{(j)}, z_{i,t} = j) + \log p(x^{(j)}), \tag{S.66}$$

where in (S.65) we have dropped the terms of the complete data log-likelihood (S.43) that are not a function of $x^{(j)}$. Finally, defining $h_{i,t}^{(j)} = \mathbb{E}_{z_{i,t}}[z_{i,t}^{(j)}] = \int q(z_{i,t}) z_{i,t}^{(j)} dz_{i,t}$, and the normalization constant

$$\mathcal{Z}_q^{(j)} = \int p(x^{(j)}) \prod_{t=1}^{\tau} \prod_{i=1}^{m} p(y_{i,t} | x_t^{(j)}, z_{i,t} = j)^{h_{i,t}^{(j)}} dx^{(j)}, \tag{S.67}$$

the optimal $q(x^{(j)})$ is given by (21).

3.2 Optimization of $q(z_{i,t})$

Rewriting (S.6) with $w_l = z_{i,t}$ and dropping terms that do not depend on $z_{i,t}$,

$$\log q^*(z_{i,t}) \propto \log \tilde{p}(z_{i,t}, Y) = \mathbb{E}_{X, Z_{k \neq i, s \neq t}}[\log p(X,Y,Z)] \tag{S.68}$$

$$\propto \mathbb{E}_{X,Z_{k\neq i,s\neq t}}\left[\sum_{j=1}^{K} z_{i,t}^{(j)} \log p(y_{i,t}|x_t^{(j)}, z_{i,t}=j) + \log p(Z)\right] \tag{S.69}$$

$$= \sum_{j=1}^{K} z_{i,t}^{(j)} \mathbb{E}_{x_t^{(j)}}[\log p(y_{i,t}|x_t^{(j)}, z_{i,t}=j)] + \mathbb{E}_{Z_{k\neq i,s\neq t}}[\log p(Z)]. \tag{S.70}$$

For the last term,

$$\mathbb{E}_{Z_{k\neq i,s\neq t}}[\log p(Z)] \tag{S.71}$$

$$\propto \mathbb{E}_{Z_{k\neq i,s\neq t}}[\log(V_{i,t}(z_{i,t}) \prod_{(i,i')\in\mathcal{E}_t} V_{i,i'}(z_{i,t}, z_{i',t}) \prod_{(t,t')\in\mathcal{E}_i} V_{t,t'}(z_{i,t}, z_{i,t',}))]$$

$$= \log V_{i,t}(z_{i,t}) + \sum_{(i,i')\in\mathcal{E}_t} \mathbb{E}_{z_{i',t}}[\log V_{i,i'}(z_{i,t}, z_{i',t})] \tag{S.72}$$

$$+ \sum_{(t,t')\in\mathcal{E}_i} \mathbb{E}_{z_{i,t'}}[\log V_{t,t'}(z_{i,t}, z_{i,t'})]$$

$$= \sum_{j=1}^{K} z_{i,t}^{(j)} \log \alpha_{i,t}^{(j)} + \sum_{(i,i')\in\mathcal{E}_t} \mathbb{E}_{z_{i',t}}[\sum_{j=1}^{K} z_{i,t}^{(j)} z_{i',t}^{(j)} \log\frac{\gamma_1}{\gamma_2} + \log\gamma_2] \tag{S.73}$$

$$+ \sum_{(t,t')\in\mathcal{E}_i} \mathbb{E}_{z_{i,t'}}[\sum_{j=1}^{K} z_{i,t}^{(j)} z_{i,t'}^{(j)} \log\frac{\beta_1}{\beta_2} + \log\beta_2]$$

$$\propto \sum_{j=1}^{K} z_{i,t}^{(j)} \log \alpha_{i,t}^{(j)} + \sum_{j=1}^{K} z_{i,t}^{(j)} \sum_{(i,i')\in\mathcal{E}_t} \mathbb{E}_{z_{i',t}}[z_{i',t}^{(j)}] \log\frac{\gamma_1}{\gamma_2} \tag{S.74}$$

$$+ \sum_{j=1}^{K} z_{i,t}^{(j)} \sum_{(t,t')\in\mathcal{E}_i} \mathbb{E}_{z_{i,t'}}[z_{i,t'}^{(j)}] \log\frac{\beta_1}{\beta_2}$$

$$= \sum_{j=1}^{K} z_{i,t}^{(j)} \left[\log \alpha_{i,t}^{(j)} + \sum_{(i,i')\in\mathcal{E}_t} h_{i',t}^{(j)} \log\frac{\gamma_1}{\gamma_2} + \sum_{(t,t')\in\mathcal{E}_i} h_{i,t'}^{(j)} \log\frac{\beta_1}{\beta_2}\right]. \tag{S.75}$$

Hence,

$$\log q^*(z_{i,t}) \propto \sum_{j=1}^{K} z_{i,t}^{(j)} \Bigg(\mathbb{E}_{x_t^{(j)}}[\log p(y_{i,t}|x_t^{(j)}, z_{i,t}=j)] + \sum_{(i,i')\in\mathcal{E}_t} h_{i',t}^{(j)} \log\frac{\gamma_1}{\gamma_2}$$

$$+ \sum_{(t,t')\in\mathcal{E}_i} h_{i,t'}^{(j)} \log\frac{\beta_1}{\beta_2} + \log \alpha_{i,t}^{(j)}\Bigg) \tag{S.76}$$

$$= \sum_{j=1}^{K} z_{i,t}^{(j)} \log(g_{i,t}^{(j)} \alpha_{i,t}^{(j)}), \tag{S.77}$$

where $g_{i,t}^{(j)}$ is defined in (24). This is a multinomial distribution of normalization constant $\sum_{j=1}^{K}(\alpha_{i,t}^{(j)} g_{i,t}^{(j)})$, leading to (22) with $h_{i,t}^{(j)}$ as given in (23).

3.3 Normalization constant for $q(x^{(j)})$

Taking the log of (S.67)

$$\log \mathcal{Z}_q^{(j)} = \log \int p(x^{(j)}) \prod_{i=1}^{m} \prod_{t=1}^{\tau} p(y_{i,t}|x_t^{(j)}, z_{i,t} = j)^{h_{i,t}^{(j)}} dx^{(j)}. \quad \text{(S.78)}$$

Note that the term $p(y_{i,t}|x_t^{(j)}, z_{i,t} = j)^{h_{i,t}^{(j)}}$ does not affect the integral when $h_{i,t}^{(j)} = 0$. Defining $\mathcal{I}_j$ as the set of indices (i,t) with non-zero $h_{i,t}^{(j)}$, i.e. $\mathcal{I}_j = \{(i,t)|h_{i,t}^{(j)} > 0\}$, (S.78) becomes

$$\log \mathcal{Z}_q^{(j)} = \log \int p(x^{(j)}) \prod_{(i,t)\in\mathcal{I}_j} p(y_{i,t}|x_t^{(j)}, z_{i,t} = j)^{h_{i,t}^{(j)}} dx^{(j)}, \quad \text{(S.79)}$$

where

$$p(y_{i,t}|x_t^{(j)}, z_{i,t} = j)^{h_{i,t}^{(j)}} = G(y_{i,t}, C_i^{(j)} x_t^{(j)}, r^{(j)})^{h_{i,t}^{(j)}} \quad \text{(S.80)}$$

$$= (2\pi r^{(j)})^{-\frac{1}{2} h_{i,t}^{(j)}} \left(\frac{2\pi r^{(j)}}{h_{i,t}^{(j)}} \right)^{\frac{1}{2}} G\left(y_{i,t}, C_i^{(j)} x_t^{(j)}, \frac{r^{(j)}}{h_{i,t}^{(j)}} \right). \quad \text{(S.81)}$$

For convenience, we define an LDS over the subset of observations indexed by $\mathcal{I}_j$. Note that the dimension of the observation y_t changes over time, depending on how many $h_{i,t}^{(j)}$ are active in each frame, and hence the LDS is parameterized by $\tilde{\Theta}_j = \{A^{(j)}, Q^{(j)}, \tilde{C}_t^{(j)}, \tilde{R}_t^{(j)}, \mu^{(j)}\}$, where $\tilde{C}_t^{(j)} = [C_i^{(j)}]_{(i,t)\in\mathcal{I}_j}$ is a time-varying observation matrix, and $\tilde{R}_t^{(j)}$ is time-varying diagonal covariance matrix with diagonal entries $[\frac{r^{(j)}}{h_{i,t}^{(j)}}]_{(i,t)\in\mathcal{I}_j}$. This LDS has conditional observation likelihood $\tilde{p}(y_{i,t}|x_t^{(j)}, z_{i,t} = j) = G(y_{i,t}, C_i^{(j)} x_t^{(j)}, \tilde{r}_{i,t}^{(j)})$, we can rewrite

$$p(y_{i,t}|x_t^{(j)}, z_{i,t} = j)^{h_{i,t}^{(j)}} = \left(2\pi r^{(j)}\right)^{\frac{1}{2}(1-h_{i,t}^{(j)})} (h_{i,t}^{(j)})^{-\frac{1}{2}} \tilde{p}(y_{i,t}|x_t^{(j)}, z_{i,t} = j), \quad \text{(S.82)}$$

and, from (S.79),

$$\log \mathcal{Z}_q^{(j)} = \quad \text{(S.83)}$$

$$\log \int p(x^{(j)}) \prod_{(i,t)\in\mathcal{I}_j} \left[\left(2\pi r^{(j)}\right)^{\frac{1}{2}(1-h_{i,t}^{(j)})} (h_{i,t}^{(j)})^{-\frac{1}{2}} \tilde{p}(y_{i,t}|x_t^{(j)}, z_{i,t} = j) \right] dx^{(j)}.$$

Since, under the restricted LDS, the likelihood of the observation $Y_j = [y_{i,t}]_{(i,t)\in\mathcal{I}_j}$ is

$$\tilde{p}_j(Y_j) = \int p(x^{(j)}) \prod_{(i,t)\in\mathcal{I}_j} \tilde{p}(y_{i,t}|x_t^{(j)}, z_{i,t} = j) dx^{(j)}, \quad \text{(S.84)}$$

it follows that

$$\log \mathcal{Z}_q^{(j)} = \log \left[\tilde{p}_j(Y_j) \prod_{(i,t)\in\mathcal{I}_j} \left(2\pi r^{(j)}\right)^{\frac{1}{2}(1-h_{i,t}^{(j)})} (h_{i,t}^{(j)})^{-\frac{1}{2}} \right] \quad \text{(S.85)}$$

$$= \frac{1}{2} \sum_{(i,t)\in\mathcal{I}_j} (1-h_{i,t}^{(j)}) \log(2\pi r^{(j)}) - \frac{1}{2} \sum_{(i,t)\in\mathcal{I}_j} \log h_{i,t}^{(j)} + \log \tilde{p}_j(Y_j). \quad \text{(S.86)}$$

References

[1] A. B. Chan and N. Vasconcelos, “Variational layered dynamic textures,” in *IEEE Conf. Computer Vision and Pattern Recognition*, 2009.

[2] C. M. Bishop, *Pattern Recognition and Machine Learning*. Springer, 2006.

SVCL-TR 2009/01

June 2009

Derivations for the Layered Dynamic Texture and Temporally-Switching Layered Dynamic Texture

Antoni B. Chan