Submitted to the Annals of Applied Statistics

APPENDICES OF FAST DYNAMIC NONPARAMETRIC DISTRIBUTION TRACKING IN ELECTRON MICROSCOPIC DATA

BY YANJUN QIAN*, JIANHUA Z. HUANG†, CHIWOO PARK‡ AND YU DING†

*Virginia Commonwealth University**
Texas A & M University†
Florida State University‡

APPENDIX A: GAUSSIAN APPROXIMATION OF POISSON DISTRIBUTION

The Poisson distributed observation equation can be written as

$$Y_{it} \sim \text{Poisson}\{(\exp[\mathbf{B}\boldsymbol{\alpha}_t])_i\},\ i = 1, \dots, m, \tag{A.1}$$

and we would like to find a Gaussian distribution

$$\mathbf{Y}_t \sim \text{normal}(\mathbf{B}\boldsymbol{\alpha}_t + \boldsymbol{\mu}_t, \mathbf{H}_t), \tag{A.2}$$

to approximate it. Durbin and Koopman (1997) proposed that if the probability distribution functions (pdfs) in Equation (A.1) and (A.2) have the same first and second derivatives w.r.t the state $\boldsymbol{\alpha}_t$, then Equation (A.2) can serve as a good approximation of Equation (A.1) in updating the state space model. We can use this idea to calculate $\boldsymbol{\mu}_t$ and $\mathbf{H}_t$ in Equation (A.2). To simplify the derivation, we use $\mathbf{B}\boldsymbol{\alpha}_t$ instead of $\boldsymbol{\alpha}_t$ as the variable to calculate those derivatives.

The logarithm of the pdfs in Equation (A.1) and Equation (A.2), as a function of $\mathbf{B}\boldsymbol{\alpha}_t$, can be expressed, respectively, as

$$\log p_{\text{poi}}([\mathbf{B}\boldsymbol{\alpha}_t]_i) = Y_{it}[\mathbf{B}\boldsymbol{\alpha}_t]_i - \exp[\mathbf{B}\boldsymbol{\alpha}_t]_i, i = 1, \dots, m, \tag{A.3}$$

and

$$\log p_{\text{nor}}(\mathbf{B}\boldsymbol{\alpha}_t) = -\frac{1}{2}(\mathbf{Y}_t - \mathbf{B}\boldsymbol{\alpha}_t - \boldsymbol{\mu}_t)^T \mathbf{H}_t^{-1}(\mathbf{Y}_t - \mathbf{B}\boldsymbol{\alpha}_t - \boldsymbol{\mu}_t) + \text{const}, \tag{A.4}$$

where 'const' is a term unrelated to $\boldsymbol{\alpha}_t$.

In Equation (A.3), the pdf of each coordinate of $\mathbf{B}\boldsymbol{\alpha}_t$ is independent to each other, Equation (A.4) should have the same property, meaning that $\mathbf{H}_t$ should be a diagonal matrix. We can then rewrite Equation (A.4) as:

$$\log p_{\text{nor}}([\mathbf{B}\boldsymbol{\alpha}_t]_i) = -\frac{1}{2[\mathbf{H}_t]_{ii}}(Y_{it} - [\mathbf{B}\boldsymbol{\alpha}_t]_i - [\boldsymbol{\mu}_t]_i)^2 + \text{const}, \tag{A.5}$$

Then calculating the first and second derivatives of Equation (A.3) and (A.5) w.r.t $[\mathbf{B}\boldsymbol{\alpha}_t]_i$ and equating them at the estimated $\hat{\boldsymbol{\alpha}}_t$, we get the following two equations:

$$Y_{it} - \exp[\mathbf{B}\hat{\boldsymbol{\alpha}}_t]_i = \frac{1}{[\mathbf{H}_t]_{ii}}(Y_{it} - [\mathbf{B}\hat{\boldsymbol{\alpha}}_t]_i - [\boldsymbol{\mu}_t]_i), \tag{A.6}$$

and

$$\exp[\mathbf{B}\hat{\boldsymbol{\alpha}}_t]_i = \frac{1}{[\mathbf{H}_t]_{ii}}. \tag{A.7}$$

The two equations further yield:

$$[\mathbf{H}_t]_{ii} = \frac{1}{\exp[\mathbf{B}\hat{\boldsymbol{\alpha}}_t]_i} = \exp[-\mathbf{B}\hat{\boldsymbol{\alpha}}_t]_i, \tag{A.8}$$

and

$$[\boldsymbol{\mu}_t]_i = Y_{it} - [\mathbf{B}\hat{\boldsymbol{\alpha}}_t]_i - \exp[-\mathbf{B}\hat{\boldsymbol{\alpha}}_t]_i(Y_{it} - \exp[\mathbf{B}\hat{\boldsymbol{\alpha}}_t]_i). \tag{A.9}$$

Rewriting Equation (A.8) and (A.9) in a matrix form, we finally obtain $\boldsymbol{\mu}_t$ and $\mathbf{H}$ as:

$$\begin{aligned} &\boldsymbol{\mu}_t = \mathbf{Y} - \mathbf{B}\hat{\boldsymbol{\alpha}}_t - \exp(-\mathbf{B}\hat{\boldsymbol{\alpha}}_t)[\mathbf{Y} - \exp(\mathbf{B}\hat{\boldsymbol{\alpha}}_t)], \\ &\mathbf{H} = \mathrm{diag}[\exp(-\mathbf{B}\hat{\boldsymbol{\alpha}}_t)]. \end{aligned} \tag{A.10}$$

APPENDIX B: DETAILED STEPS OF KALMAN FILTER

Given a linear Gaussian state space model

$$\begin{aligned} &\mathbf{Y}_t \sim \mathrm{normal}(\mathbf{B}\boldsymbol{\alpha}_t + \boldsymbol{\mu}_t, \mathbf{H}_t), \\ &\boldsymbol{\alpha}_{t+1} = \boldsymbol{\alpha}_t + \mathbf{w}_t,\ \ \mathbf{w}_t \sim \mathrm{normal}(\mathbf{0}, \mathbf{Q}), \end{aligned} \tag{B.1}$$

the Kalman filter can estimate the state $\boldsymbol{\alpha}_t$ in a recursive way from $t = 1$ to time T. First we need to predict $\boldsymbol{\alpha}_t$ and its covariance according to the estimation of the previous step as

$$\begin{aligned} &\hat{\boldsymbol{\alpha}}_t^- = \hat{\boldsymbol{\alpha}}_{t-1}, \\ &\mathbf{P}_t^- = \mathbf{P}_{t-1} + \mathbf{Q}, \end{aligned} \tag{B.2}$$

where $\hat{\boldsymbol{\alpha}}_t^-$ is called the prior estimator and $\mathbf{P}_t^-$ is the prior covariance matrix. The two equations above can be derived from the distribution of $p(\boldsymbol{\alpha}_t|\mathbf{Y}_1, \cdots, \mathbf{Y}_{t-1})$ (Durbin and Koopman, 2012).

When a new $\mathbf{Y}_t$ is coming, we calculate the innovation $\boldsymbol{\nu}_t$ and its covariance matrix according to the previous prediction $\hat{\boldsymbol{\alpha}}_t^-$ and the new input $\mathbf{Y}_t$ as

$$\begin{aligned} &\boldsymbol{\nu}_t = \mathbf{Y}_t - \mathbf{B}\hat{\boldsymbol{\alpha}}_t^- - \boldsymbol{\mu}_t; \\ &\mathbf{F}_t = \mathbf{B}\mathbf{P}_t^-\mathbf{B}^T + \mathbf{H}_t. \end{aligned} \tag{B.3}$$

Then the Kalman gain will be calculated as:

$$\mathbf{K}_t = \mathbf{P}_t^-\mathbf{B}^T\mathbf{F}_t^{-1}. \tag{B.4}$$

At last we update the estimator of state $\boldsymbol{\alpha}_t$ and its covariance matrix as:

$$\begin{aligned} &\hat{\boldsymbol{\alpha}}_t = \hat{\boldsymbol{\alpha}}_t^- + \mathbf{K}_t\boldsymbol{\nu}_t, \\ &\mathbf{P}_t = \mathbf{P}_t^-(\mathbf{I} - \mathbf{K}_t\mathbf{B})^T, \end{aligned} \tag{B.5}$$

where $\hat{\boldsymbol{\alpha}}_t$ is called the posterior estimator and $\mathbf{P}_t^-$ is the posterior covariance matrix. Those equations can be derived from the distribution of $p(\boldsymbol{\alpha}_t|\mathbf{Y}_1, \cdots, \mathbf{Y}_t)$ (Durbin and Koopman, 2012).

APPENDIX C: POSTERIOR DISTRIBUTION OF σ_α^2 AND σ_ϵ^2

Here we want to show the derivation of the posterior distribution of σ_α^2 and σ_ϵ^2 in our Bayesian model:

$$\begin{aligned} &Y_{it} \sim \mathrm{Poisson}\{(\exp[\mathbf{B}\mathbf{C}\boldsymbol{\gamma}_t])_i\}, \\ &\boldsymbol{\gamma}_{t+1} - \boldsymbol{\gamma}_t = \mathbf{w}_t \sim \mathrm{normal}(\mathbf{0}, \mathbf{Q}),\ \ \mathbf{Q} = \mathrm{diag}(\sigma_\alpha^2, \sigma_\alpha^2, \sigma_\epsilon^2, \cdots, \sigma_\epsilon^2), \\ &\sigma_\alpha^2 \sim \text{inverse-gamma}(a_1, b_1),\ \ \sigma_\epsilon^2 \sim \text{inverse-gamma}(a_2, b_2). \end{aligned} \tag{C.1}$$

Since $\mathbf{Q}$ is a covariance matrix, we can rewrite the second layer of the model (C.1) as:

$$\gamma_{jt} - \gamma_{j(t-1)} \sim \text{normal}(0, \sigma_\alpha^2),\ j = 1, 2,\ t = 2, \cdots, T, \tag{C.2}$$

and

$$\gamma_{jt} - \gamma_{j(t-1)} \sim \text{normal}(0, \sigma_\epsilon^2),\ j = 3, \cdots, n,\ t = 2, \cdots, T. \tag{C.3}$$

For $j = 1, 2$, $\gamma_{jt} - \gamma_{j(t-1)}$ are regarded as $2(T-1)$ i.i.d variables following normal$(0, \sigma_\alpha^2)$. Since we choose the conjugate prior $\sigma_\alpha^2 \sim$ inverse-gamma(a_1, b_1), its posterior distribution has the same formation as inverse-gamma$(a_1^{\text{post}}, b_1^{\text{post}})$. As derived in Bolstad and Curran (2016), a_1^{post} and b_1^{post} are calculated as:

$$a_1^{\text{post}} = a_1 + \frac{1}{2} 2(T-1) = a_1 + (T-1), \tag{C.4}$$

and

$$b_1^{\text{post}} = b_1 + \frac{1}{2} \sum_{j=1}^{2} \sum_{t=2}^{T} [\gamma_{jt} - \gamma_{j(t-1)}]^2. \tag{C.5}$$

For $j = 3, \cdots, n$, $\gamma_{jt} - \gamma_{j(t-1)}$ are regarded as $(n-2)(T-1)$ i.i.d variables following normal$(0, \sigma_\epsilon^2)$. Following the same derivation, the posterior distribution of σ_ϵ^2 is written as inverse-gamma$(a_2^{\text{post}}, b_2^{\text{post}})$ with

$$a_2^{\text{post}} = a_2 + \frac{1}{2}(n-2)(T-1), \tag{C.6}$$

and

$$b_2^{\text{post}} = b_2 + \frac{1}{2} \sum_{j=3}^{n} \sum_{t=2}^{T} [\gamma_{jt} - \gamma_{j(t-1)}]^2. \tag{C.7}$$

REFERENCES

Bolstad, W. M. and J. M. Curran (2016). *Introduction to Bayesian Statistics*. John Wiley & Sons.

Durbin, J. and S. J. Koopman (1997). Monte Carlo maximum likelihood estimation for non-Gaussian state space models. *Biometrika 84*(3), 669–684.

Durbin, J. and S. J. Koopman (2012). *Time Series Analysis by State Space Methods*. Oxford University Press.

Yanjun Qian
Department of Statistical Sciences
& Operations Research
Virginia Commonwealth University
Richmond, VA 23220
E-mail: yqian3@vcu.edu

Jianhua Z. Huang
Department of Statistics
Texas A & M University
College Station, TX 77843
E-mail: jianhua@stat.tamu.edu

Chiwoo Park
Department of Industrial &
Manufacturing Engineering
Florida State University
Tallahassee, FL 32301
E-mail: chiwoo.park@eng.famu.fsu.edu

Yu Ding
Department of Industrial &
Systems Engineering
Texas A & M University
College Station, TX 77843
E-mail: yuding@tamu.edu