Bayesian state space model (BSSM)

R. Maxammadiyev, U. Djumanazarov, Sh. Mahmudov,
S. Abdug‘aniyev

Annotatsiya

Ushbu maqolada qisqa vaqt qatorlari uchun ishonchlilik darajasi yuqori bo‘lgan parametrlarni baholashda samarali hisoblangan Bayesian state space model (BSSM) tahlil qilingan. Jumladan, BSSM ishga tushirish, Variational Bayesian M-step (Maximisation), Variational Bayesian E-step (Expectation), logarifmik marginal ehtimollik quyi chegarasining bahosini hisoblash, giperparametrlarni yangilash, logarifmik marginal ehtimollikning quyi chegarasi bahosi qiymatlarini taqqoslash bosqichlarida amalga oshiriladi. Mazkur maqola O‘zbekiston ko‘chmas mulk bozorida uy-joy narxlarining fundamental qiymatini baholash aniqligini oshirishda modellar qamrovini kengaytirish maqsadida tayyorlangan.

Tayanch so‘zlar: kovariatsiya matritsasi, marginal ehtimollik, nisbiy xatolik, prior taqsimot, posterior taqsimot, vaqt qatorlari.

Bayesian state space model (BSSM)

Bayesian state space model (BSSM) qisqa vaqt qatorlari uchun yuqori ishonchlilik darajasiga ega parametrlarni baholashda state space model (SSM)[1]ga nisbatan samarali hisoblanadi. Xususan, BSSM noaniqliklarni boshqarish va parametrlarni kuzatilgan ma’lumotlarga moslashtirish imkoniyatlarini taqdim etadi. Qisqa vaqt qatorlarida ma’lumotlarning dinamik oʻzgarishlarini yaxshiroq aks ettirish uchun posterior taqsimotlarni (posterior distribution)[2] tez yangilash imkoniyatlarini beradi.

BSSM barcha noaniq miqdorlarni tasodifiy oʻzgaruvchilar sifatida koʻrib chiqadi va ehtimollik qonuniyatlaridan foydalanadi. Ushbu modellar signallarni filtrlash va bashorat qilishda keng qoʻllaniladi. Modeldagi noaniqliklarni aniqlash uchun modeldagi barcha parametrlar boʻyicha posterior taqsimotlar talab qilinadi. Jumladan, ushbu model oʻzaro bogʻliqlik darajasini ifodalovchi kuzatish tenglamasi (observable or measurement equation) hamda dinamik holat oʻzgaruvchilarini ifodalovchi oʻtish tenglamasidan (transition equation) tashkil topgan[3]:

$$y_t = Cx_t + Du_t + v_t$$
$$x_t = Ax_{t-1} + Bu_t + w_t .$$

Bu yerda,

y_t – kuzatilishi mumkin boʻlgan erksiz oʻzgaruvchilar vektori;

x_t – yashirin holat oʻzgaruvchilar vektori;

u_t – kuzatilishi mumkin boʻlgan erkli oʻzgaruvchilar vektori;

A – kxk oʻlchamli holat dinamik matritsasi;

C – pxk oʻlchamli kuzatish matritsasi

B – kxd oʻlchamli holat kirish matritsasi;

D – pxd oʻlchamli kuzatish kirish matritsasi;

$v_t \sim N(0, R)$ – taqsimotdan olingan kuzatish tenglamasi uchun nisbiy xatoliklar vektori;

$w_t \sim N(0, Q)$ – holat tenglamasi uchun nisbiy xatoliklar vektori;

A, B, C va D matritsalar – noma’lum koeffitsientlar matritsalari;

[1] State space model (SSM) metodologiyasi 2023-yil uchun moliyaviy barqarorlik sharhida keltirilgan.

[2] Posterior taqsimot – ma’lumotlar kuzatilgandan keyin parametrning yangilangan taqsimoti.

[3] Beal, M. (2003). Variational Algorithms for Approximate Bayesian Inference. University of Cambridge.

Q va R – nisbiy xatoliklar kovariatsiya matritsalari.

Kuzatish maydonidagi siljishlarni holat kirish va kirish kuzatish matritsalarining parametrlari sifatida o‘rganish mumkin. Yashirin o‘zgaruvchilar kuzatilmaganligi uchun holat tenglamasi uchun nisbiy xatoliklar kovariatsiya matritsasi Q ni holat dinamik matritsasi A ga kiritib yuborish mumkin. Natijada, yashirin holat shunga mos ravishda o‘zgaradi. Ma’lumotlar kuzatilganligi sababli kuzatish tenglamasi uchun nisbiy xatoliklar kovariatsiya matritsasi R ni kuzatish matritsasi C ga kiritib bo‘lmaydi. Jumladan, diagonal deb qabul qilingan nisbiy xatoliklar kovariatsiyasi R, ρ aniqlik vektori orqali quyidagicha aniqlanadi:

$$R^{-1} = diag(\rho)$$

Bu yerda, R – nisbiy xatoliklar kovariatsiyasi, $diag(p)$ –diagonalidan boshqa barcha elementlarini nolga teng bo‘lgan matritsa.

Konjugatsiya (conjugate)[4] uchun aniqlik vektori ρ ning har bir o‘lchami a va b giperparametrlar bilan taqsimlangan Gamma taqsimot sifatida qabul qilinadi. ρ ning prior taqsimoti esa quyidagi ko‘rinishda bo‘ladi:

$$p(\rho|a,b) = \prod_{s=1}^{p} \frac{b^a}{G(a)} \rho_s^{a-1} \exp\{-b\rho_s\}\,.$$

Parametrlar va giperparametrlar quyidagi ko‘rinishdagi ehtimolliklar bilan bog‘lanadi:

$$p\left(a_{(j)}\middle|\alpha\right) = N(a_{(j)}|0, diag(\alpha)^{-1})$$
$$p\left(b_{(j)}\middle|\beta\right) = N(b_{(j)}|0, diag(\beta)^{-1})$$
$$j = 1,2, \dots, k$$
$$p\left(c_{(s)}\middle|\rho_s, \gamma\right) = N(c_{(s)}|0, {\rho_s}^{-1} diag(\gamma)^{-1})$$
$$p\left(d_{(s)}\middle|\rho_s, \delta\right) = N(d_{(s)}|0, {\rho_s}^{-1} diag(\delta)^{-1})$$
$$p(\rho_s|a,b) = Ga(\rho_s|a,b)$$
$$s = 1,2, \dots, p\,.$$

Bu yerda,

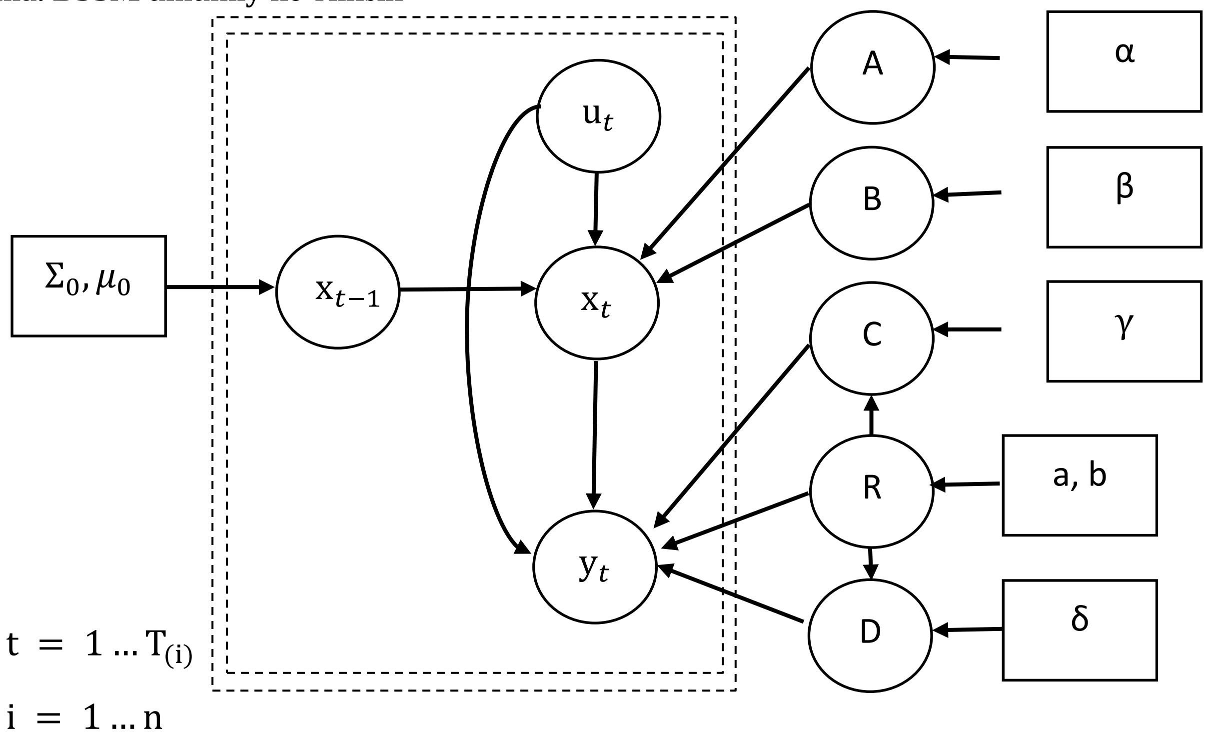
$a_{(j)}$ – A matritsaning j ustuni;

$b_{(j)}$ – B matritsaning j ustuni;

[4] Konjugatsiya – hisob-kitoblarni yengillashtirish uchun oldingi va keyingi taqsimotlarni bir xil ko‘rinishga keltirish.

$c_{(s)}$ – C matritsaning s ustuni;
$d_{(s)}$ – D matritsaning s ustuni.

1-chizma. BSSM umumiy ko‘rinishi



Manba: Beal, M. (2003). Variational Algorithms for Approximate Bayesian Inference. University of Cambridge.

Shuningdek, A matritsaning har bir qatori bo‘yicha o‘rtachasi nolga teng bo‘lgan dastlabki Gauss taqsimotida aniqlik ko‘rsatkichi[5] (precision) $diag(\alpha)$ ga teng. C matritsaning har bir qatori bo‘yicha esa, o‘rtachasi nolga teng bo‘lgan dastlabki Gauss taqsimotida aniqlik ko‘rsatkichi $diag(\rho_s\gamma)$ ga teng. $c_{(s)}$ aniqlik ko‘rsatkichining shovqin chiqish aniqlik ko‘rsatkichiga bog‘liqligi ρ_s konjugatsiya bilan bog‘liq bo‘ladi. Shunga o‘xshash ustunlarni kiritish bilan bog‘liq B va D matritsalar qatoriga mos ravishda yana ikkita giperparametr vektori β va δ kiritiladi. A va B matritsalarning qatorlardagi kovariatsiyalari o‘zgarmaydi. Biroq, C va D matritsalar qatorlaridagi kovariatsiyalar ρ_s ning hisobiga o‘zgarishi mumkin. Agar, A ning j ustuni nolga aylansa $t-1$ vaqtidagi j ustun t vaqtidagi yashirin o‘zgaruvchilarni topishda ishtirok etmaydi. Ammo, j yashirin o‘lchov har bir vaqt bosqichida ma’lumotlarda kovariatsiya strukturasini yaratish uchun foydalaniladi.

[5] Taqsimotlarda aniqlik ko‘rsatkichi dispersiyaning teskarisiga teng, ya’ni ushbu aniqlik ko‘rsatkichi hamda dispersiyalarning o‘zaro ko‘paytmasi birga teng.

BSSM olti bosqichda amalga oshiriladi. Dastlabki bosqich ishga tushirish bosqichi hisoblanib, bu bosqichda yordamchi yashirin holat ishga tushiriladi hamda parametrlarning prior taqsimoti (prior distribution)[6] kiritilib, logarifmik marginal ehtimollik[7] (marginal probability) Jensen tengsizligi (Jensen’s inequality) yordamida baholanadi. Ikkinchi bosqich Variational Bayesian M-step (Maximisation) bosqichi hisoblanadi. Bu bosqichda, parametrlarning posterior taqsimotlarining optimal shakli topiladi va tabiiy parametrlar vektori hisoblanadi. Bundan so‘ng, Variational Bayesian E-step (Expectation) bosqichida yashirin holat o‘zgaruvchilar vektorining taqsimotlari topiladi. To‘rtinchi bosqichda logarifmik marginal ehtimollikning quyi chegarasi bahosi ya’ni F ning qiymati hisoblanadi. Beshinchi bosqich giperparametrlarni yangilash bosqichi hisoblanib, bunda logarifmik marginal ehtimollikning quyi chegarasi bahosini oshirish uchun giperparametrlar yangilanadi. Yakuniy bosqichda, logarifmik marginal ehtimollikning quyi chegarasi bahosi oldingi qiymatiga nisbatan solishtirilib, quyi chegara bahosi ortgan holatda Variational Bayesian M-step bosqichi qaytadan amalga oshiriladi. Ushbu quyi chegaraning bahosi oldingi qiymatiga nisbatan ortmay qolgunga qadar hisoblash bosqichlari davom ettiriladi.

Ishga tushirish bosqichi

Dastlab $t = 0$ davr uchun yordamchi yashirin holat x_0 ishga tushiriladi. Bunda, yordamchi yashirin holat x_0 Gauss taqsimoti bo‘yicha taqsimlanib, yordamchi yashirin holatning o‘rtacha μ_0 va kovariatsiya Σ_0 qiymatlari, shuningdek prior taqsimoti quyidagi ko‘rinishlarda bo‘ladi:

$$p(x_0 |\, \mu_0, \Sigma_0) = N(x_0 |\, \mu_0, \Sigma_0)$$
$$\mu_0 \sim N(\mu_0 |\, 0, b_{\mu_0} \mathrm{I})$$
$$\Sigma_0 \sim \prod_{j=1}^{k} Ga\,(\Sigma_{0jj}^{-1} |\, a_{\Sigma_0}, b_{\Sigma_0})$$

Bu yerda,

x_0 – yordamchi yashirin holat;

μ_0 – yordamchi yashirin holat o‘rtachasi;

Σ_0 – yordamchi yashirin holat kovariatsiyasi.

Bunda, holat dinamik jarayoni orqali x_1 ning priori quyidagi ko‘rinishda bo‘ladi:

---

[6] Prior taqsimot – parametrlarning ma’lumotlari kuzatilishidan oldingi taqsimoti.

[7] Marginal ehtimollik – boshqa hodisalarga bog‘liq bo‘lmagan bir hodisaning sodir bo‘lish ehtimoli.

$$p(x_1|\,\mu_0,\Sigma_0,\theta) \;=\; \int dx_0\; p(x_0|\,\mu_0,\Sigma_0)\; p(x_1|\,x_0,\theta)$$
$$= N(x_1|\,A\mu_0 + Bu_1,\; A^T\Sigma_0 A \;+\; Q)\;.$$

Soddalashtirish maqsadida parametrlar bitta parametr vektoriga $\theta = (A, B, C, D, R)$ yig'iladi. Shuningdek, A, B, C va D matritsalarning giperparametrlari bilan bog'liqligi quyidagi giperprior taqsimot (hyperprior distribution) orqali beriladi:

$$\alpha \;\sim \prod_{j=1}^{k} Ga\,(\alpha_{(j)}|a_{(\alpha)}, b_{(\alpha)})$$
$$\beta \;\sim \prod_{c=1}^{d} Ga\,(\beta_{(c)}|a_{(\beta)}, b_{(\beta)})$$
$$\gamma \;\sim \prod_{j=1}^{k} Ga\,(\gamma_{(j)}|a_{(\gamma)}, b_{(\gamma)})$$
$$\delta \;\sim \prod_{c=1}^{d} Ga\,(\delta_{(c)}|a_{(\delta)}, b_{(\delta)})$$

Marginal ehtimollik quyidagi ko'rinishda bo'ladi:

$$p(y_{1:T}) \;=\; \int dA\; dB\; dC\; dD\; d\rho\; dx_{0:T}\, p(A, B, C, D, \rho, x_{0:T}, y_{1:T})$$

Modelning parametrlari o'rtasidagi ba'zi bir shartli mustaqillik hisobiga dinamika va chiqish jarayonlari uchun posterior taqsimotlar qanday posterior taqsimotlarga ajralishini ko'rish mumkin. Kirishlarni hisobga olgan holda parametrlar, yashirin o'zgaruvchilar va kuzatilgan ma'lumotlar uchun to'liq birgalikdagi ehtimollik quyidagi ko'rinishda bo'ladi:

$$p(A, B, C, D, \rho, x_{0:T}, y_{1:T}|u_{1:T}) =$$
$$p(A|\alpha)p(B|\beta)p(\rho|a, b)p(C|\rho, \gamma)p(D|\rho, \delta)p(x_0|\,\mu_0, \Sigma_0)\cdot$$
$$\prod_{t=1}^{T} p(x_t|\,x_{t-1}, A, B, u_t)\; p(y_t|\,x_t, C, D, \rho, u_t)$$

Shundan keyin, kirishlarda $u_{1:T}$ bog'liqlik yo'qotiladi va kirishlar yopiq holatda qoldiriladi. Yashirin o'zgaruvchilar va parametrlarga bog'liq $q(\theta, x)$ taqsimot kiritiladi. Jensen tengsizligini qo'llash orqali marginal ehtimollikning quyi chegarasi baholanadi:

$$\ln p(y_{1:T}) = \;\ln\left(\int dA\; dB\; dC\; dD\; d\rho\; dx_{0:T}\; p(A, B, C, D, \rho, x_{0:T}, y_{1:T}\right) \;\geq$$
$$\int dA\; dB\; dC\; dD\; d\rho\; dx_{0:T}\; q(A, B, C, D, \rho, x_{0:T})\; \ln\frac{p(A,B,C,D,\rho,x_{0:T},y_{1:T})}{q(A,B,C,D,\rho,x_{0:T})} \;= F$$

Variatsion yaqinlashishning keyingi bosqichi parametrlar va yashirin o‘zgaruvchilar taqsimoti $q(\cdot)$ uchun ma’lum bir tahminiy shakllarni qabul qilish hisoblanadi. Birinchi navbatda, yashirin o‘zgaruvchilar va parametrlar shartli mustaqillik asosida ajratiladi:

$$q(A,B,C,D,\rho,x_{0:T}) = q_{\theta}(A,B,C,D,\rho)q_x(x_{0:T})$$
$$q(A,B,C,D,\rho,x_{0:T}) = q_{AB}(A,B)\, q_{CD\rho}(C,D,\rho)\, q_x(x_{0:T})\,.$$

Logarifmik marginal ehtimollik quyi chegarasining bahosini soddalashtirish uchun shartli ehtimollikning quyidagi xossalaridan foydalaniladi:

$$q_{AB}(A,B) = q_B(B)\, q_A(A|B)$$
$$q_{CD\rho}(C,D,\rho) = q_{\rho}(\rho)\, q_D(D|\rho)\, q_C(C|D,\rho)$$

Logarifmik marginal ehtimollik quyi chegarasining bahosi F yuqoridagi xossalardan foydalanib quyidagi yig‘indiga keltiriladi:

$$F = \int dB\, q_B(B) \ln\frac{p(B|\beta)}{q_B(B)} + \int dB\, q_B(B) \int dA\, q_A(A|B) \ln\frac{p(A|\alpha)}{q_A(A|B)} +$$
$$\int d\rho\, q_{\rho}(\rho) \ln\frac{p(\rho|a,b)}{q_{\rho}(\rho)} + \int d\rho\, q_{\rho}(\rho) \int dD q_D(D|\rho) \ln\frac{p(D|\rho,\delta)}{q_D(D|\rho)} +$$
$$\int d\rho\, q_{\rho}(\rho) \int dD\, q_D(D|\rho) \int dC\, q_C(C|\rho,D) \ln\frac{p(C|\rho,\gamma)}{q_C(C|\rho,D)} -$$
$$\int dx_{0:T}\, q_x(x_{0:T}) \ln q_x(x_{0:T}) + \int dB\, q_B(B) \int dA\, q_A(A|B) \int d\rho\, q_{\rho}(\rho) \cdot$$
$$\int dD\, q_D(D|\rho) \int dC\, q_C(C|\rho,D) \int dx_{0:T}\, q_x(x_{0:T}) \ln p(x_{0:T}, y_{1:T}|A,B,C,D,\rho) =$$
$$F(q_x(x_{0:T}), q_B(B), q_A(A|B), q_{\rho}(\rho), q_D(D|\rho), q_C(C|\rho,D))$$
$$H\big(q_x(x_{0:T})\big) = -\int dx_{0:T} q_x(x_{0:T}) \ln q_x(x_{0:T})$$

Logarifmik marginal ehtimollik quyi chegarasining bahosi giperparametrlarga bog‘liq ko‘rinishda bo‘ladi. Ushbu tahminiy posterior taqsimotlarning optimal shakllarini parametrlar va yashirin o‘zgaruvchilar ketma-ketliklari bo‘yicha har bir taqsimotga nisbatan F dan olingan funksional xususiy hosilalarni nolga tenglash yordamida aniqlash mumkin.

Variational Bayesian M-step (Maximisation), VBM

Bu bosqichda, Bayesian teoremasini qo‘llash orqali parametrlar bo‘yicha variatsion posterior taqsimotlar topiladi va bulardan kutilgan tabiiy parametrlar vektori hisoblanadi. Posterior taqsimot quyidagi ko‘rinishda faktorlarga ajratiladi:

$$q_\theta(A,B,C,D,\rho) = \prod_{j=1}^{k} q\big(b_{(j)}\big)\, q\big(a_{(j)}\big|b_{(j)}\big)\ \prod_{s=1}^{p} q(\rho_s)\cdot$$
$$q(d_{(s)}|\rho_s)q(c_{(s)}|\rho_s, d_{(s)})$$

Bu yerda,

$b_{(j)}$ – B matritsaning j ustuni;

$a_{(j)}$ – A matritsaning j ustuni;

$d_{(s)}$ – D matritsaning s ustuni;

$c_{(s)}$ – C matritsaning s ustuni.

Hisob-kitoblarda zarur bo‘ladigan kirish va chiqish o‘zgaruvchilar vektorlarining ba’zi statistik ma’lumotlari quyidagicha kiritiladi:

$$\ddot{U} = \sum_{t=1}^{T} u_t\, u_t^T$$
$$U_y = \sum_{t=1}^{T} u_t\, y_t^T$$
$$\ddot{Y} = \sum_{t=1}^{T} y_t\, y_t^T$$

Parametrlarning posterior taqsimotini va tabiiy parametrlar vektorini hisoblashda qo‘llaniladigan yetarli statistik ma’lumotlar quyidagi tenglamalar yordamida topiladi:

$$W_A = \sum_{t=1}^{T} < x_{t-1}x_{t-1}{}^T > = \sum_{t=1}^{T} \Upsilon_{t-1,t-1} + \omega_{t-1}\omega_{t-1}{}^T$$
$$G_A = \sum_{t=1}^{T} < x_{t-1} > u_t{}^T = \sum_{t=1}^{T} \omega_{t-1}\, u_t{}^T$$
$$\widetilde{M} = \sum_{t=1}^{T} u_t < x_t{}^T > = \sum_{t=1}^{T} u_t\, \omega_t{}^T$$
$$S_A = \sum_{t=1}^{T} < x_{t-1}x_t{}^T > = \sum_{t=1}^{T} \Upsilon_{t-1,t} + \omega_{t-1}\omega_t{}^T$$
$$W_C = \sum_{t=1}^{T} < x_t x_t{}^T > = \sum_{t=1}^{T} \Upsilon_{t,t} + \omega_t\omega_t{}^T$$
$$G_c = \sum_{t=1}^{T} < x_t > u_t{}^T = \sum_{t=1}^{T} \omega_t\, u_t{}^T$$
$$S_c = \sum_{t=1}^{T} < x_t > y_t{}^T = \sum_{t=1}^{T} \omega_t\, y_t{}^T\ .$$

A va B uchun posterior taqsimotlar quyidagi tenglamalar bilan aniqlanadi:

$$q_B(B) = \prod_{j=1}^{k} N(b_{(j)}|\, \Sigma_B \bar{b}_{(j)},\, \Sigma_B)$$

$$q_A(A|B) = \prod_{j=1}^{k} N(a_{(j)} | \Sigma_A \left[S_{A(j)} - G_A b_{(j)}\right], \Sigma_A)$$
$$\Sigma_A^{-1} = diag(\alpha) + W_A$$
$$\Sigma_B^{-1} = diag(\beta) + \ddot{U} - G_A^T \Sigma_A G_A$$
$$\bar{B} = \widehat{M}^T - S_A^T \Sigma_A G_A$$
$$q_A(A) = \prod_{j=1}^{k} N(a_{(j)} | \Sigma_A \left[S_{A(j)} - G_A \Sigma_B \bar{b}_{(j)}\right], \tilde{\Sigma}_A)$$
$$\tilde{\Sigma}_A = \Sigma_A + \Sigma_A G_A \Sigma_B G_A^T \Sigma_A$$

Bu yerda, $\bar{b}_{(j)}$ – $\bar{B}$ matritsaning j ustunini, $S_{A(j)}$ – S_A matritsaning j ustunini bildiradi.

Shuningdek, ρ, C va D uchun posterior taqsimotlar quyidagi tenglamalar orqali topiladi:

$$q_\rho(\rho) = \prod_{s=1}^{p} Ga(\rho_s | a + \frac{T}{2}, b + \frac{1}{2} G_{ss})$$
$$q_D(D|\rho) = \prod_{s=1}^{p} N(d_{(s)} | \Sigma_D \bar{d}_{(s)}, \rho_s^{-1} \Sigma_D)$$
$$q_C(C|D,\rho) = \prod_{s=1}^{p} N(c_{(s)} | \Sigma_C \left[S_{C(s)} - G_C d_{(s)}\right], \rho_s^{-1} \Sigma_C)$$
$$\textstyle\sum_C^{-1} = diag(\gamma) + W_C$$
$$\textstyle\sum_D^{-1} = diag(\delta) + \ddot{U} - G_C^T \Sigma_C G_C$$
$$G = \ddot{Y} - S_C^T \Sigma_C S_C - \bar{D} \Sigma_D \bar{D}^T$$
$$\bar{D} = U_Y{}^T - S_C^T \Sigma_C G_C$$
$$q_C(C|\rho) = \prod_{s=1}^{p} N\left(c_{(s)} \middle| \Sigma_C \left[S_{C(s)} - G_C \Sigma_D \bar{d}_{(s)}\right], \rho_s^{-1} \hat{\Sigma}_C\right)$$
$$\hat{\Sigma}_C = \Sigma_C + \Sigma_C G_C \Sigma_D G_C^T \Sigma_C$$

Bu yerda, $\bar{d}_{(s)}$ – $\bar{D}$ matritsaning s ustunini, $S_{C(s)}$ – S_C matritsaning s ustunini bildiradi.

Variational Bayesian M-step (Maximisation) bosqichida kutilgan tabiiy parametrlar vektori $\varphi(\theta)$ hisoblanadi. Keyinchalik ular tizimdagi yashirin holatlar bo‘yicha $q_x(x_{0:T})$ taqsimotni ko‘rsatadigan Variational Bayesian E-step (Expectation) bosqichida qo‘llaniladi. Tegishli tabiiy parametrlar vektori esa quyidagi ko‘rinishda bo‘ladi:

$$\varphi(\theta) = \varphi(A,B,C,D,R) = [A, A^T A, B, A^T B, C^T R^{-1} C,$$
$$R^{-1} C, C^T R^{-1} D, B^T B, R^{-1}, \ln|R^{-1}|, D^T R^{-1} D, R^{-1} D] .$$
$$< A > = [S_A - G_A \Sigma_B \bar{B}^T]^T \Sigma_A$$
$$< A^T A > = < A >^T < A > + k [\Sigma_A + \Sigma_A G_A \Sigma_B G_A^T \Sigma_A]$$
$$< B > = \bar{B} \Sigma_B$$

$$< A^T B > = \Sigma_A \, [S_A < B > - G_A \, \{< B >^T < B > + k\Sigma_B\}]$$
$$< B^T B > = < B >^T < B > + k\Sigma_B$$
$$< \rho_s > = \bar{\rho_s} = \frac{a_\rho + {}^T/_2}{b_\rho + {}^{G_{ss}}/_2}$$
$$< ln\rho_s > = \overline{ln\rho_s} = \Psi\left(a_\rho + \frac{T}{2}\right) - \ln(b_\rho + \frac{G_{ss}}{2})$$
$$< R^{-1} > = diag(\bar{\rho})$$
$$< C > = [S_C - G_C \Sigma_D \bar{D}^T]^T \, \Sigma_C$$
$$< D > = \bar{D}\Sigma_D$$
$$< C^T R^{-1} C > = < C >^T \; diag(\bar{\rho}) < C > + p \, [\Sigma_C + \Sigma_C G_C \Sigma_D G_C^T \Sigma_C]$$
$$< R^{-1} C > = diag(\bar{\rho}) < C >$$
$$< C^T R^{-1} D > = \Sigma_C \, [S_C diag(\bar{\rho}) < D > - G_C < D >^T \, diag(\bar{\rho}) < D > - pG_C\Sigma_D]$$
$$< R^{-1} D > = diag(\bar{\rho}) < D >$$
$$< D^T R^{-1} D > = < D >^T \, diag(\bar{\rho}) < D > + p\Sigma_D \, .$$

Bu yerda, <·> – keyingi variatsiyaga nisbatan kutishni bildiradi.

Variational Bayesian E-step (Expectation), VBE

Yashirin holat o‘zgaruvchilari vektori $x_{0:T}$ ning posterior taqsimoti vaqt bosqichlari bo‘yicha birgalikda Gauss taqsimoti kabi bo‘lib, yashirin o‘zgaruvchilar vektorining logarifmik posterior taqsimoti quyidagi ko‘rinishda ifodalanadi:

$$\ln q_x(x_{0:T}) = -\ln Z + < \ln p(A, B, C, D, \rho, x_{0:T}, y_{1:T}) >$$
$$= -\ln Z' + < \ln p(x_{0:T}, y_{1:T} | \, A, B, C, D, \rho) >$$
$$Z' = \int dx_{0:T} \; exp < \ln p(x_{0:T}, y_{1:T} | \, A, B, C, D, \rho) > .$$
$$\ln Z' = \sum_{t=1}^{T} \ln \zeta_t{}'(y_t)$$
$$\ln\zeta_t{}'(y_t) = -\frac{1}{2}[< \ln|2\pi R| > - \ln|\Sigma_{t-1}^{-1}\Sigma_{t-1}^{*}\Sigma_t| + \mu_{t-1}^T \Sigma_{t-1}^{-1} \mu_{t-1} - \mu_t^T \Sigma_t^{-1} \mu_t$$
$$+ y_t^T < R^{-1} > y_t - 2y_t^T < R^{-1} D > u_t + u_t^T < D^T R^{-1} D > u_t$$
$$- (\Sigma_{t-1}^{-1}\mu_{t-1} - < A^T B > u_t)^T \Sigma_{t-1}^{*} (\Sigma_{t-1}^{-1}\mu_{t-1} - < A^T B > u_t) \, .$$

Kuzatilgan ma’lumotlar berilgan t vaqtdagi yashirin holatga nisbatan $\alpha_t(x_t)$ t vaqtgacha ($t = \{1, 2, \ldots, T\}$) posterior bo‘ladi. Shu bilan birgalikda, $\alpha_t(x_t)$ yashirin holatning filtrlangan o‘rtacha μ_t va kovariatsiyasi Σ_t aniqlanadi:

$$\alpha_t(x_t) \equiv p(x_t|\, y_{1:t}) .$$

Shuningdek, $\alpha_t(x_t)$ oldingi $\alpha_{t-1}(x_{t-1})$ bilan rekursiyani quyidagi ko‘rinishda hosil qiladi:

$$\alpha_t(x_t) = \int dx_{t-1} \frac{p(x_{t-1}|\, y_{1:t-1})\, p(x_t|\, x_{t-1}) p(y_t|\, x_t)}{p(y_t|\, y_{1:t-1})}$$
$$= \frac{1}{\zeta_t(y_t)} \int dx_{t-1} \alpha_{t-1}(x_{t-1})\, p\,(x_t|\, x_{t-1})\, p(y_t|\, x_t)$$
$$= \frac{1}{\zeta_t(y_t)} \int dx_{t-1} N(x_{t-1}|\, \mu_{t-1}, \Sigma_{t-1})\, N(x_t|\, Ax_{t-1}, \mathrm{I})\, N(y_t|\, Cx_t, R)$$
$$= N(x_t|\, \mu_t, \Sigma_t)$$

Bu yerda, $\zeta_t(y_t) \equiv p(y_t|\, y_{1:t-1})$ belgilash filtrlangan chiqish ehtimolini beradi.

$$\Sigma^*_{t-1} = (\Sigma^{-1}_{t-1} + A^T A)^{-1}$$
$$x^*_{t-1} = \Sigma^*_{t-1}[\Sigma^{-1}_{t-1}\mu_{t-1} + A^T x_t]$$
$$\Sigma_t = [\mathrm{I} + C^T R^{-1} C - A\Sigma^*_{t-1} A^T]^{-1}$$
$$\mu_t = \Sigma_t\, [C^T R^{-1} y_t + A\Sigma^*_{t-1}\Sigma^{-1}_{t-1}\mu_{t-1}] .$$

Har bir bosqichda $\alpha_t(x_t)$ ning mahraji sifatida olingan normallashtiruvchi doimiy $\zeta_t(y_t)$ ma’lumotlarning ehtimolligini hisoblashga yordam beradi. Bunda, doimiy ma’lumotlar ehtimolligini quyidagi shaklda yozish mumkin:

$$p(y_{1:T}) = p(y_1) p(y_2|\, y_1) \ldots p(y_t|\, y_{1:t-1}) \ldots p(y_T|\, y_{1:T-1}) =$$
$$p(y_1) \prod_{t=2}^{T} p(y_t|\, y_{1:t-1}) = \prod_{t=1}^{T} \zeta_t(y_t)$$
$$\zeta_t(y_t) = N(y_t|\, \varpi_t, \xi_t)$$
$$\xi_t = (R^{-1} - R^{-1} C \Sigma_t C^T R^{-1})^{-1}$$
$$\varpi_t = \xi_t R^{-1} C \Sigma_t A \Sigma^*_{t-1} \Sigma^{-1}_{t-1} \mu_{t-1} .$$

Ushbu taqsimot yordamida ketma-ketlikda oldingi kuzatuvlarni hisobga olgan holda, har bir kuzatuv y_t ehtimolligi kuzatiladi va ma’lumotlar kelishi bilan har bir vaqt bosqichida prognozli o‘rtacha va dispersiya olinadi. Biroq, bu taqsimot yashirin holat ketma-ketligi tekislangandan so‘ng o‘zgaradi. Barcha ma’lumotlar berilgan yashirin holat ustidan posterior bo‘lishi uchun yashirin holat tekislangan bahosi γ aniqlanadi. Shu bilan birgalikda, $\gamma_t(x_t) = p(x_t|y_{1:T})$ yashirin holatning tekislangan bahosi quyidagicha ifodalanadi:

$$\gamma_t(x_t) = p(x_t|\, y_{1:T}) = \int dx_{t+1}\, p(x_t, x_{t+1}|\, y_{1:T})$$
$$= \int dx_{t+1}\, p(x_t|\, x_{t+1}, y_{1:T}) p(x_{t+1}|\, y_{1:T})$$
$$= \int dx_{t+1}\, p(x_t|\, x_{t+1}, y_{1:t}) p(x_{t+1}|\, y_{1:T})$$
$$= \int dx_{t+1} \frac{p(x_t|\, y_{1:t}) p(x_{t+1}|\, x_t)}{\int dx'_t\, p(x'_t|\, y_{1:t}) p(x_{t+1}|\, x'_t)} p(x_{t+1}|\, y_{1:T})$$
$$= \int dx_{t+1} \left[\frac{\alpha_t(x_t)\, p(x_{t+1}|x_t)}{\int dx'_t\, \alpha_t(x'_t)\ p(x_{t+1}|x'_t)}\right] \gamma_{t+1}(x_{t+1})$$
$$\gamma_t(x_t) \ = \ N(x_t|\, \omega_t, \Upsilon_{tt})$$
$$K_t \ = (\Upsilon_{t+1,t+1}^{-1} \ + \ A\Sigma_t^{*} A^T)^{-1}$$
$$\Upsilon_{tt} \ = \ [\Sigma_t^{*-1} - A^T K_t A]^{-1}$$
$$\omega_t \ = \ \Upsilon_{tt}\, [\Sigma_t^{-1}\mu_t + \ A^T K_t (\Upsilon_{t+1,t+1}^{-1}\omega_{t+1} - A\Sigma_t^{*}\Sigma_t^{-1}\mu_t)]\ .$$

Bunda, t, y_t vaqtidagi ma’lumotlar filtrlangandan so‘ng yashirin o‘zgaruvchilar o‘rniga $\alpha_t(x_t)$ dan foydalaniladi.

Turli vaqt nuqtalaridagi holatlarni baholash uchun parallel rekursiya $\beta_t(x_t) \ = p(y_{t+1:T}|x_t)$ hisoblanadi:

$$\beta_{t-1}(x_{t-1}) = \ \int dx_t\, p(x_t|\, x_{t-1})\, p(y_t|\, x_t) p(y_{t+1:T}|\, x_t) \ =$$
$$\int dx_t\, p(x_t|\, x_{t-1})\ \ p(y_t|\, x_t)\, \beta_t(x_t) \propto N(x_{t-1}|\, \eta_{t-1}, \Psi_{t-1})\ .$$

$\beta_T(x_T) = 1$ yakuniy shartni qanoatlantirish uchun $\Psi_T{}^{-1} \ = \ 0$ ko‘rinishida ifodalanadi. Rekursiya $t = T$ dan $t = 1$ gacha davom etadi. Bu rekursiyaning oxirgi bosqichida, yordamchi yashirin holat o‘zgaruvchisi x_0 ga bog‘liq bo‘lgan barcha ma’lumotlarning ehtimolligi aniqlanadi:

$$\Psi_t{}^{*} \ = \ (\mathrm{I} + C^T R^{-1} C + \Psi_t{}^{-1})^{-1}$$
$$\Psi_{t-1} \ \ = [A^T A - A^T\, \Psi_t{}^{*} A]^{-1}$$
$$\eta_{t-1} \ = \Psi_{t-1}\, A^T \Psi_t{}^{*} [C^T R^{-1} y_t \ + \ \Psi_t{}^{-1} \eta_t]\ .$$

Yuqoridagi formulalardan foydalanib, $t = 0$ dan $t = T$ gacha $\{\eta_t, \Psi_t\}_{t=0}^{T}$ lar hisoblaniladi.

Variational Bayesian yondashuvida marginallarni ketma-ket o‘tish bilan osongina olish mumkin emas va ular o‘rniga α va β ma’lumotlar quyidagi tarzda birlashtirish orqali hisoblanadi. Xususan, $t = \{0, 1, 2, \dots, T-1\}$ uchun ω_t, $\Upsilon_{t,t}$ larni quyidagicha hisoblanadi:

$$
\begin{gathered}
p(x_t |\, y_{1:T}) \propto p(x_t |\, y_{1:t}) p(y_{t+1:T} |\, x_t) = \\
\alpha_t(x_t)\, \beta_t(x_t) = N(x_t |\, \omega_t, \Upsilon_{tt}) \\
\Upsilon_{t,t} = [\Sigma_t^{-1} + {\Psi_t}^{-1}]^{-1} \\
\omega_t = \Upsilon_{t,t}\, [\Sigma_t^{-1} \mu_t + {\Psi_t}^{-1} \eta_t]\,.
\end{gathered}
$$

$t = \{0, 1, 2, \dots, T-1\}$ uchun barcha kuzatuvlarni hisobga olgan holda yashirin o‘zgaruvchilar vektorining o‘zaro kovariatsiyasi $\Upsilon_{t,t+1}$ esa, quyidagicha aniqlanadi:

$$
\Upsilon_{t,t+1} = \Sigma_t^* A^T \Upsilon_{t+1,t+1}.
$$

Logarifmik marginal ehtimollik quyi chegarasining bahosini hisoblash

Logarifmik marginal ehtimollikni to‘g‘ridan-to‘g‘ri hisoblash murakkabligi sababli, uning quyi chegarasini baholab logarifmik marginal ehtimollik quyi chegarasining bahosi sifatida F olinadi. Logarifmik marginal ehtimollik quyi chegarasining bahosi ham murakkab funksiya bo‘lganligi uchun bir nechta bosqichdan o‘tish orqali, logarifmik marginal ehtimollik quyi chegarasining bahosi soddaroq ko‘rinishdagi yig‘indiga keltiriladi.

KL (Kullback-Leibler) bir xil o‘zgaruvchilar bo‘yicha ikki normal yoki ikki gamma taqsimot orasidagi farq hisoblanib, J va K o‘zgaruvchilar juftligi uchun quyidagi tengliklardan foydalaniladi:

$$
\begin{gathered}
KL(J) = \int dJ\, q(J)\, \ln \frac{q(J)}{p(J)} \\
KL(J|K) = \int dJ\, q(J|K)\, \ln \frac{q(J|K)}{p(J|K)} \\
< KL(J|K) >_{q(K)} = \int dK\, q(K)\, KL(J|K)
\end{gathered}
$$

Bu yerda,

$KL(J)$ – KL farqi,

$KL(J|K)$ – shartli *KL*,

$< KL(J|K) >_{q(K)}$ – kutilgan shartli KL.

Bu tengliklardan foydalangan holda, F ni quyidagi ko‘rinishda qayta ifodalash mumkin:

$$F = -KL(B) - < KL(A|\, B) >_{q(B)} - KL(\rho) - < KL(D|\, \rho) >_{q(\rho)} -$$
$$< KL(C|\, \rho, D) >_{q(\rho,D)} + H(q_x(x_{0:T})) +$$
$$< \ln p(x_{1:T}, y_{1:T}|\, A, B, C, D, \rho) >_{q(A,B,C,D,\rho)q(x_{1:T})} .$$
$$H\big(q_x(x_{0:T})\big) = -\int dx_{0:T}\, q_x(x_{0:T}) \ln q_x(x_{0:T}) =$$
$$-\int dx_{0:T}\, q_x(x_{0:T})[-\ln Z' + < \ln p(x_{0:T}, y_{1:T}|\, A, B, C, D, \rho, \mu_0, \Sigma_0) >_{q_\theta(A,B,C,D,\rho)}]$$
$$= \ln Z' - < \ln p(x_{0:T}, y_{1:T}|\, A, B, C, D, \rho, \mu_0, \Sigma_0) >_{q_\theta(A,B,C,D,\rho) q_x(x_{0:T})}$$

$H(q_x(x_{0:T}))$ dan foydalangan holda F ning tenglamasi quyidagi sodda ko‘rinishga keltiriladi. Bu esa, F ni hisoblashda qulaylik yaratadi:

$$F = -KL(B) - < KL(A|\, B) >_{q(B)} - KL(\rho) - < KL(D|\, \rho) >_{q(\rho)} -$$
$$< KL(C|\, \rho, D) >_{q(\rho,D)} + ln\, Z'.$$

Giperparametrlarni yangilash

Giperparametrlar α, β, γ, δ, a, b, shuningdek Σ_0 va μ_0 ko‘rinishidagi oldingi giperparametrlar marginal ehtimollik bo‘yicha pastki chegara bahosini maksimal darajaga ko‘tarish uchun yangilanadi. Bunda, giperparametrlar quyidagicha yangilanadi:

$$\alpha_j^{-1} \leftarrow \frac{1}{k}\big[k\Sigma_A + \Sigma_A[S_A {S_A}^T - 2G_A < B >^T {S_A}^T +$$
$$G_A\{k\Sigma_B + < B >^T < B >\}{G_A}^T]\, \Sigma_A]_{jj}$$
$$\beta_j^{-1} \leftarrow \frac{1}{k}[k\Sigma_B + < B >^T < B >]_{jj}$$
$$\gamma_j^{-1} \leftarrow \frac{1}{p}[p\Sigma_C + \Sigma_C[S_C diag(\bar{\rho}) {S_C}^T - 2S_C diag(\bar{\rho}) < D > {G_C}^T + pG_C\Sigma_D {G_C}' +$$
$$G_C < D >^T diag(\bar{\rho}) < D > {G_C}^T]\, \Sigma_C]_{jj}$$
$$\delta_j^{-1} \leftarrow \frac{1}{p}[p\Sigma_D + < D >^T diag(\bar{\rho}) < D >]_{jj}$$

Bu yerda, $[\cdot]_{jj}$ – $[\cdot]$ matritsaning (j, j) elementini bildiradi.

Oldingi prior taqsimot ostidagi yashirin holat ketma-ketligining ehtimolini maksimal darajada oshirish uchun, yordamchi yashirin holatning giperparametrlari x_0 ning tekislangan bahosining taqsimlanishiga muvofiq o‘rnatiladi:

$$\Sigma_0 \leftarrow \Upsilon_{0,0}$$
$$\mu_{\mathbf{0}} \leftarrow \omega_0$$

Chiqish shovqini bo‘yicha prior taqsimotni boshqaruvchi giperparametrlar a va b esa $R = diag(\rho)$ tenglamaning belgilangan nuqtalariga o‘rnatiladi:

$$\Psi(a) \; = \; \ln b \; + \frac{1}{p} \sum_{s=1}^{p} \overline{ln \; \rho_s}$$
$$\frac{1}{b} = \frac{1}{pa} \sum_{s=1}^{p} \overline{\rho_s}$$

Logarifmik marginal ehtimollikning quyi chegarasi bahosi qiymatlarini taqqoslash

Logarifmik marginal ehtimollikning quyi chegarasi bahosi aniqlangandan so‘ng, yakuniy bosqichda ushbu bahoning bir davr oldingi qiymati bilan o‘zaro solishtiriladi. Agar, quyi chegara bahosi bir davr oldingi qiymatiga nisbatan ortgan bo‘lsa, Variational Bayesian M-step bosqichidan boshlab qaytadan hisob-kitoblar amalga oshiriladi. Ushbu takrorlanish quyi chegara bahosi ortmay qolgunga qadar davom etadi. Bunda, logarifmik marginal ehtimollikning quyi chegarasi bahosi oldingi qiymati bilan solishtirilgan holatda ular orasida farq bo‘lmasa, logarifmik marginal ehtimollikning quyi chegarasi bahosi optimal qiymatga erishgan hisoblanadi. Bu davrga kelib, giperparametrlar orqali parametrlar ham o‘zining optimal qiymatiga erishadi. Aniqlangan parametrlar va mavhum o‘zgaruvchilardan kelib chiqib, kuzatilishi mumkin bo‘lgan erksiz o‘zgaruvchi Y baholanadi.