Bayesian state space model (BSSM)

Р. Махаммадиев, У. Джуманазаров, Ш. Маҳмудов,
С. Абдуғаниев

Аннотация

Ушбу мақолада қисқа вақт қаторлари учун ишончлилик даражаси юқори бўлган параметрларни баҳолашда самарали ҳисобланган Bayesian state space model (BSSM) таҳлил қилинган. Жумладан, BSSM ишга тушириш, Variational Bayesian M-step (Maximisation), Variational Bayesian E-step (Expectation), логарифмик маргинал эҳтимоллик қуйи чегарасининг баҳосини ҳисоблаш, гиперпараметрларни янгилаш, логарифмик маргинал эҳтимолликнинг қуйи чегараси баҳоси қийматларини таққослаш босқичларида амалга оширилади. Мазкур мақола Ўзбекистон кўчмас мулк бозорида уй-жой нархларининг фундаментал қийматини баҳолаш аниқлигини оширишда моделлар қамровини кенгайтириш мақсадида тайёрланган.

Таянч сўзлар: ковариация матрицаси, маргинал эҳтимоллик, нисбий хатолик, приор тақсимот, постериор тақсимот, вақт қаторлари.

Bayesian state space model (BSSM)

Bayesian state space model (BSSM) қисқа вақт қаторлари учун юқори ишончлилик даражасига эга параметрларни баҳолашда state space model (SSM)[1]га нисбатан самарали ҳисобланади. Хусусан, BSSM ноаниқликларни бошқариш ва параметрларни кузатилган маълумотларга мослаштириш имкониятларини тақдим этади. Қисқа вақт қаторларида маълумотларнинг динамик ўзгаришларини яхшироқ акс эттириш учун постериор тақсимотларни (posterior distribution)[2] тез янгилаш имкониятларини беради.

BSSM барча ноаниқ миқдорларни тасодифий ўзгарувчилар сифатида кўриб чиқади ва эҳтимоллик қонуниятларидан фойдаланади. Ушбу моделлар сигналларни фильтрлаш ва башорат қилишда кенг қўлланилади. Моделдаги ноаниқликларни аниқлаш учун моделдаги барча параметрлар бўйича постериор тақсимотлар талаб қилинади. Жумладан, ушбу модел ўзаро боғлиқлик даражасини ифодаловчи кузатиш тенгламаси (observable or measurement equation) ҳамда динамик ҳолат ўзгарувчиларини ифодаловчи ўтиш тенгламасидан (transition equation) ташкил топган[3]:

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

Бу ерда,

y_t – кузатилиши мумкин бўлган эрксиз узгарувчилар вектори;

x_t – яширин ҳолат ўзгарувчилар вектори;

u_t – кузатилиши мумкин булган эркли ўзгарувчилар вектори;

A – kxk ўлчамли ҳолат динамик матрицаси;

C – pxk ўлчамли кузатиш матрицаси

B – kxd ўлчамли ҳолат кириш матрицаси;

D – pxd ўлчамли кузатиш кириш матрицаси;

$v_t \sim N(0, R)$ – тақсимотдан олинган кузатиш тенгламаси учун нисбий хатоликлар вектори;

---

[1] State space model (SSM) методологияси 2023 йил учун молиявий барқарорлик шарҳида келтирилган.

[2] Постериор тақсимот – маълумотлар кузатилгандан кейин параметрнинг янгиланган тақсимоти.

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

$w_t \sim N(0, Q)$ – ҳолат тенгламаси учун нисбий хатоликлар вектори;

A, B, C ва D матрицалар – номаълум коэффицентлар матрицалари;

Q ва R – нисбий хатоликлар ковариация матрицалари.

Кузатиш майдонидаги силжишларни ҳолат кириш ва кириш кузатиш матрицаларининг параметрлари сифатида ўрганиш мумкин. Яширин ўзгарувчилар кузатилмаганлиги учун ҳолат тенгламаси учун нисбий хатоликлар ковариация матрицаси Q ни ҳолат динамик матрицаси A га киритиб юбориш мумкин. Натижада, яширин ҳолат шунга мос равишда ўзгаради. Маълумотлар кузатилганлиги сабабли кузатиш тенгламаси учун нисбий хатоликлар ковариация матрицаси R ни кузатиш матрицаси C га киритиб бўлмайди. Жумладан, диагонал деб қабул қилинган нисбий хатоликлар ковариацияси R, ρ аниқлик вектори оркали қуйидагича аниқланади:

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

Бу ерда, R – нисбий хатоликлар ковариацияси, $diag(p)$ – диагоналидан бошқа барча элементларини нолга тенг бўлган матрица.

Конжугация (conjugate)[4] учун аниқлик вектори ρ нинг ҳар бир ўлчами a ва b гиперпараметрлар билан тақсимланган Гамма тақсимот сифатида қабул қилинади. ρ нинг приор тақсимоти эса қуйидаги кўринишда бўлади:

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

Параметрлар ва гиперпараметрлар қуйидаги кўринишдаги эҳтимолликлар билан боғланади:

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

[4] Конжугация – ҳисоб-китобларни енгиллаштириш учун олдинги ва кейинги тақсимотларни бир хил кўринишга келтириш.

$$s = 1,2, \dots , p .$$

Бу ерда,

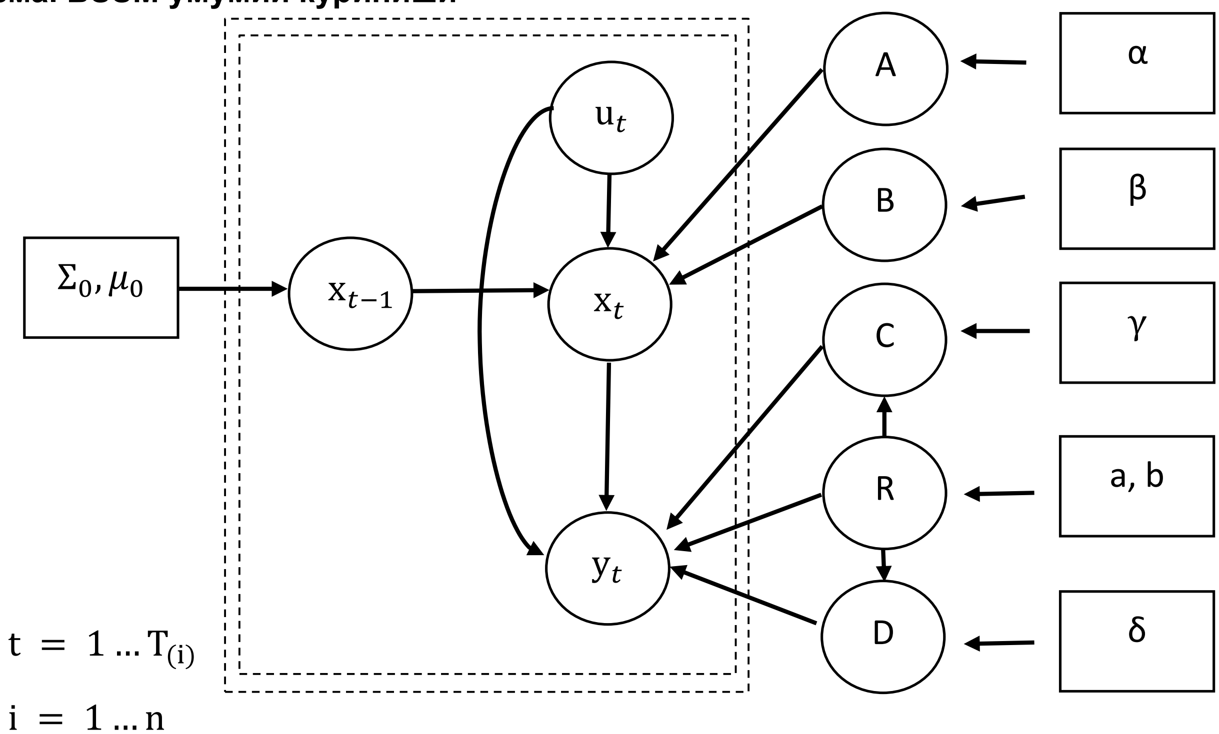
$a_{(j)}$ – A матрицанинг j устуни;

$b_{(j)}$ – B матрицанинг j устуни;

$c_{(s)}$ – C матрицанинг s устуни;

$d_{(s)}$ – D матрицанинг s устуни.

1-чизма. BSSM умумий кўриниши



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

Шунингдек, A матрицанинг ҳар бир қатори бўйича ўртачаси нолга тенг бўлган дастлабки Гаусс тақсимотида аниқлик кўрсаткичи[5] (precision) $diag(\alpha)$ га тенг. C матрицанинг ҳар бир қатори бўйича эса, ўртачаси нолга тенг бўлган дастлабки Гаусс тақсимотида аниқлик кўрсаткичи $diag(\rho_s \gamma)$ га тенг. $c_{(s)}$ аниқлик кўрсаткичининг шовқин чиқиш аниқлик кўрсаткичига боғлиқлиги ρ_s конжугация билан боғлиқ бўлади. Шунга ўхшаш устунларни киритиш билан боғлик B ва D матрицалар қаторига мос равишда яна иккита гиперпараметр вектори β ва δ киритилади. A ва B матрицаларнинг

[5] Тақсимотларда аниқлик кўрсаткичи дисперсиянинг тескарисига тенг, яъни ушбу аниқлик кўрсаткичи ҳамда дисперсияларнинг ўзаро кўпайтмаси бирга тенг.

қаторлардаги ковариациялари узгармайди. Бироқ, C ва D матрицалар қаторларидаги ковариациялар ρ_s нинг ҳисобига ўзгариши мумкин. Агар, A нинг j устуни нолга айланса $t-1$ вақтидаги j устун t вақтидаги яширин узгарувчиларни топишда иштирок этмайди. Аммо, j яширин ўлчов хар бир вақт босқичида маълумотларда ковариация структурасини яратиш учун фойдаланилади.

BSSM олти босқичда амалга оширилади. Дастлабги босқич ишга тушириш босқичи ҳисобланиб, бу босқичда ёрдамчи яширин ҳолат ишга туширилади ҳамда параметрларнинг приор тақсимоти (prior distribution)[6] киритилиб, логарифмик маргинал эҳтимоллик[7] (marginal probability) Jensen тенгсизлиги (Jensen's inequality) ёрдамида баҳоланади. Иккинчи босқич Variational Bayesian M-step (Maximisation) босқичи ҳисобланади. Бу босқичда, параметрларнинг постериор тақсимотларининг оптимал шакли топилади ва табиий параметрлар вектори ҳисобланади. Бундан сўнг, Variational Bayesian E-step (Expectation) босқичида яширин ҳолат ўзгарувчилар векторининг тақсимотлари топилади. Тўртинчи босқичда логарифмик маргинал эҳтимолликнинг қуйи чегараси баҳоси яъни F нинг қиймати ҳисобланади. Бешинчи босқич гиперпараметрларни янгилаш босқичи ҳисобланиб, бунда логарифмик маргинал эҳтимолликнинг қуйи чегараси баҳосини ошириш учун гиперпараметрлар янгиланади. Якуний босқичда, логарифмик маргинал эҳтимолликнинг қуйи чегараси баҳоси олдинги қийматига нисбатан солиштирилиб, қуйи чегара баҳоси ортган ҳолатда Variational Bayesian M-step босқичи қайтадан амалга оширилади. Ушбу қуйи чегаранинг баҳоси олдинги қийматига нисбатан ортмай қолгунга қадар ҳисоблаш босқичлари давом эттирилади.

Ишга тушириш босқичи

Дастлаб $t=0$ давр учун ёрдамчи яширин ҳолат x_0 ишга туширилади. Бунда, ёрдамчи яширин ҳолат x_0 Гаусс тақсимоти бўйича тақсимланиб, ёрдамчи яширин ҳолатнинг ўртача μ_0 ва ковариация Σ_0 қийматлари, шунингдек приор тақсимоти қуйидаги кўринишларда бўлади:

$$p(x_0|\,\mu_0,\Sigma_0)=N(x_0|\,\mu_0,\Sigma_0)$$

[6] Приор тақсимот – параметрларнинг маълумотлари кузатилишидан олдинги тақсимоти.

[7] Маргинал эҳтимоллик – бошқа ҳодисаларга боғлиқ бўлмаган бир ҳодисанинг содир бўлиш эҳтимоли.

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

Бу ерда,

x_0 – ёрдамчи яширин ҳолат;

μ_0 – ёрдамчи яширин ҳолат ўртачаси;

Σ_0 – ёрдамчи яширин ҳолат ковариацияси.

Бунда, ҳолат динамик жараёни орқали x_1 нинг приори қуйидаги кўринишда бўлади:

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

Соддалаштириш мақсадида параметрлар битта параметр векторига $\theta = (A, B, C, D, R)$ йиғилади. Шунингдек, A, B, C ва D матрицаларнинг гиперпараметрлари билан боғлиқлиги қуйидаги гиперприор тақсимот (hyperprior distribution) орқали берилади:

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

Маргинал эҳтимоллик қўйидаги кўринишда бўлади:

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

Моделнинг параметрлари ўртасидаги баъзи бир шартли мустақиллик ҳисобига динамика ва чиқиш жараёнлари учун постериор тақсимотлар қандай постериор тақсимотларга ажралишини кўриш мумкин. Киришларни ҳисобга олган ҳолда параметрлар, яширин ўзгарувчилар ва кузатилган маълумотлар учун тўлиқ биргаликдаги эҳтимоллик қуйидаги кўринишда бўлади:

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

$$\prod_{t=1}^{T} p(x_t | \, x_{t-1}, A, B, u_t)\, p(y_t | \, x_t, C, D, \rho, u_t)$$

Шундан кейин, киришларда $u_{1:T}$ боғлиқлик йўқотилади ва киришлар ёпиқ ҳолатда қолдирилади. Яширин ўзгарувчилар ва параметрларга боғлиқ $q(\theta, x)$ тақсимот киритилади. Jensen тенгсизлигини қўллаш орқали маргинал эҳтимолликнинг қуйи чегараси баҳоланади:

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

Вариацион яқинлашишнинг кейинги босқичи параметрлар ва яширин ўзгарувчилар тақсимоти $q(\cdot)$ учун маълум бир тахминий шаклларни қабул қилиш ҳисобланади. Биринчи навбатда, яширин ўзгарувчилар ва параметрлар шартли мустақиллик асосида ажратилади:

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

Логарифмик маргинал эҳтимоллик қуйи чегарасининг баҳосини соддалаштириш учун шартли эҳтимолликнинг қуйидаги хоссаларидан фойдаланилади:

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

Логарифмик маргинал эҳтимоллик қуйи чегарасининг баҳоси F юқоридаги хоссалардан фойдаланиб қуйидаги йиғиндига келтирилади:

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

$$\int dD\, q_D(D|\rho) \int dC\, q_C(C|\rho, D) \int dx_{0:T}\, q_x(x_{0:T}) \ln p(x_{0:T}, y_{1:T}|A, B, C, D, \rho) =$$
$$F(q_x(x_{0:T}), q_B(B), q_A(A|B), q_\rho(\rho), q_D(D|\rho), q_C(C|\rho, D))$$
$$H\big(q_x(x_{0:T})\big) = -\int dx_{0:T} q_x(x_{0:T}) \ln q_x(x_{0:T})$$

Логарифмик маргинал эҳтимоллик қуйи чегарасининг баҳоси гиперпараметрларга боғлиқ кўринишда бўлади. Ушбу тахминий постериор тақсимотларнинг оптимал шаклларини параметрлар ва яширин ўзгарувчилар кетма-кетликлари бўйича ҳар бир тақсимотга нисбатан F дан олинган функционал хусусий ҳосилаларни нолга тенглаш ёрдамида аниқлаш мумкин.

Variational Bayesian M-step (Maximisation), VBM

Бу босқичда, Bayesian теоремасини қўллаш орқали параметрлар бўйича вариацион постериор тақсимотлар топилади ва булардан кутилган табиий параметрлар вектори ҳисобланади. Постериор тақсимот қуйидаги кўринишда факторларга ажратилади:

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

Бу ерда,

$b_{(j)}$ – B матрицанинг j устуни;

$a_{(j)}$ – A матрицанинг j устуни;

$d_{(s)}$ – D матрицанинг s устуни;

$c_{(s)}$ – C матрицанинг s устуни.

Ҳисоб-китобларда зарур бўладиган кириш ва чиқиш ўзгарувчилар векторларининг баъзи статистик маълумотлари қуйидагича киритилади:

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

Параметрларнинг постериор тақсимотини ва табиий параметрлар векторини ҳисоблашда қўлланиладиган етарли статистик маълумотлар қуйидаги тенгламалар ёрдамида топилади:

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

A ва B учун постериор тақсимотлар қуйидаги тенгламалар билан аниқланади:

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

Бу ерда, $\bar{b}_{(j)}$ – $\bar{B}$ матрицанинг j устунини, $S_{A(j)}$ – S_A матрицанинг j устунини билдиради.

Шунингдек, ρ, C ва D учун постериор тақсимотлар қуйидаги тенгламалар орқали топилади:

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

Бу ерда, $\bar{d}_{(s)}$ – $\bar{D}$ матрицанинг s устунини, $S_{C(s)}$ – S_C матрицанинг s устунини билдиради.

Variational Bayesian M-step (Maximisation) босқичида кутилган табиий параметрлар вектори $\varphi(\theta)$ ҳисобланади. Кейинчалик улар тизимдаги яширин ҳолатлар бўйича $q_x(x_{0:T})$ тақсимотни кўрсатадиган Variational Bayesian E-step (Expectation) босқичида қўлланилади. Тегишли табиий параметрлар вектори эса қуйидаги кўринишда бўлади:

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

Бу ерда, $<\cdot>$ – кейинги вариацияга нисбатан кутишни билдиради.

Variational Bayesian E-step (Expectation), VBE

Яширин ҳолат ўзгарувчилари вектори $x_{0:T}$ нинг постериор тақсимоти вақт босқичлари бўйича биргаликда Гаусс тақсимоти каби бўлиб, яширин ўзгарувчилар векторининг логарифмик постериор тақсимоти қуйидаги кўринишда ифодаланади:

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

Кузатилган маълумотлар берилган t вақтдаги яширин ҳолатга нисбатан $\alpha_t(x_t)$ t вақтгача ($t = \{1,2,\dots,T\}$) постериор бўлади. Шу билан биргаликда, $\alpha_t(x_t)$ яширин ҳолатнинг фильтрланган ўртача μ_t ва ковариацияси Σ_t аниқланади:

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

Шунингдек, $\alpha_t(x_t)$ олдинги $\alpha_{t-1}(x_{t-1})$ билан рекурсияни қуйидаги кўринишда ҳосил қилади:

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

Бу ерда, $\zeta_t(y_t) \equiv p(y_t|\, y_{1:t-1})$ белгилаш фильтрланган чиқиш эҳтимолини беради.

$$\Sigma_{t-1}^{*} = (\Sigma_{t-1}^{-1} + A^TA)^{-1}$$
$$x_{t-1}^{*} = \Sigma_{t-1}^{*}[\Sigma_{t-1}^{-1}\mu_{t-1} + A^Tx_t]$$
$$\Sigma_t = [\mathrm{I} + C^TR^{-1}C - A\Sigma_{t-1}^{*}A^T]^{-1}$$
$$\mu_t = \Sigma_t\,[C^TR^{-1}y_t + A\Sigma_{t-1}^{*}\Sigma_{t-1}^{-1}\mu_{t-1}] .$$

Ҳар бир босқичда $\alpha_t(x_t)$ нинг маҳражи сифатида олинган нормаллаштирувчи доимий $\zeta_t(y_t)$ маълумотларнинг эҳтимоллигини ҳисоблашга ёрдам беради. Бунда, доимий маълумотлар эҳтимоллигини қуйидаги шаклда ёзиш мумкин:

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

Ушбу тақсимот ёрдамида кетма-кетликда олдинги кузатувларни ҳисобга олган холда, ҳар бир кузатув y_t эҳтимоллиги кузатилади ва маълумотлар келиши билан ҳар бир вақт босқичида прогнозли ўртача ва дисперсия олинади. Бироқ, бу тақсимот яширин ҳолат кетма-кетлиги текислангандан сунг ўзгаради. Барча маълумотлар берилган яширин ҳолат устидан постериор бўлиши учун яширин ҳолат текисланган баҳоси γ аниқланади. Шу билан биргаликда, $\gamma_t(x_t) = p(x_t|y_{1:T})$ яширин ҳолатнинг текисланган баҳоси қуйидагича ифодаланади:

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

Бунда, t, y_t вақтидаги маълумотлар фильтрлангандан сўнг яширин ўзгарувчилар ўрнига $\alpha_t(x_t)$ дан фойдаланилади.

Турли вақт нуқталаридаги ҳолатларни баҳолаш учун параллел рекурсия $\beta_t(x_t) = p(y_{t+1:T}|x_t)$ ҳисобланади:

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

$\beta_T(x_T) = 1$ якуний шартни қаноатлантириш учун ${\Psi_T}^{-1} = 0$ кўринишида ифодаланади. Рекурсия $t = T$ дан $t = 1$ гача давом этади. Бу рекурсиянинг охирги босқичида, ёрдамчи яширин ҳолат ўзгарувчиси x_0 га боғлиқ бўлган барча маълумотларнинг эҳтимоллиги аниқланади:

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

Юқоридаги формулалардан фойдаланиб, $t = 0$ дан $t = T$ гача $\{\eta_t, \Psi_t\}_{t=0}^{T}$ лар ҳисобланилади.

Variational Bayesian ёндашувида маргиналларни кетма-кет ўтиш билан осонгина олиш мумкин эмас ва улар ўрнига α ва β маълумотлар қуйидаги тарзда бирлаштириш орқали ҳисобланади. Хусусан, $t = \{0, 1, 2, \dots, T-1\}$ учун ω_t, $\Upsilon_{t,t}$ ларни қуйидагича ҳисобланади:

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

$t = \{0, 1, 2, \dots, T-1\}$ учун барча кузатувларни ҳисобга олган ҳолда яширин ўзгарувчилар векторининг ўзаро ковариацияси $\Upsilon_{t,t+1}$ эса, қуйидагича аниқланади:

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

Логарифмик маргинал эҳтимоллик қуйи чегарасининг баҳосини ҳисоблаш

Логарифмик маргинал эҳтимолликни тўғридан-тўғри ҳисоблаш мураккаблиги сабабли, унинг қуйи чегарасини баҳолаб логарифмик маргинал эҳтимоллик қуйи чегарасининг баҳоси сифатида F олинади.

Логарифмик маргинал эҳтимоллик қуйи чегарасининг баҳоси ҳам мураккаб функция бўлганлиги учун бир нечта босқичдан ўтиш орқали, логарифмик маргинал эҳтимоллик қуйи чегарасининг баҳоси соддароқ кўринишдаги йиғиндига келтирилади.

KL (Kullback-Leibler) бир хил ўзгарувчилар бўйича икки нормал ёки икки гамма тақсимот орасидаги фарқ ҳисобланиб, J ва K ўзгарувчилар жуфтлиги учун қуйидаги тенгликлардан фойдаланилади:

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

Бу ерда,
$KL(J)$ – KL фарқи,
$KL(J|K)$ – шартли *KL*,
$< KL(J|K) >_{q(K)}$ – кутилган шартли *KL*.

Бу тенгликлардан фойдаланган ҳолда, F ни қуйидаги кўринишда қайта ифодалаш мумкин:

$$F = -KL(B) - < KL(A|\, B) >_{q(B)} - KL(\rho) - < KL(D|\, \rho) >_{q(\rho)} - < KL(C|\, \rho, D) >_{q(\rho,D)} + H(q_x(x_{0:T})) + < \ln p(x_{1:T}, y_{1:T}|\, A, B, C, D, \rho) >_{q(A,B,C,D,\rho)q(x_{1:T})}.$$

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

$H(q_x(x_{0:T}))$ дан фойдаланган ҳолда F нинг тенгламаси қуйидаги содда кўринишга келтирилади. Бу эса, F ни ҳисоблашда қулайлик яратади:

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

Гиперпараметрларни янгилаш

Гиперпараметрлар α, β, γ, δ, a, b, шунингдек Σ_0 ва μ_0 кўринишидаги олдинги гиперпараметрлар маргинал эҳтимоллик буйича пастки чегара баҳосини максимал даражага кўтариш учун янгиланади. Бунда, гиперпараметрлар қуйидагича янгиланади:

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

Бу ерда, $[\cdot]_{jj}$ – $[\cdot]$ матрицанинг (j, j) элементини билдиради.

Олдинги приор тақсимот остидаги яширин ҳолат кетма-кетлигининг эҳтимолини максимал даражада ошириш учун, ёрдамчи яширин ҳолатнинг гиперпараметрлари x_0 нинг текисланган баҳосининг тақсимланишига мувофиқ ўрнатилади:

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

Чиқиш шовқини бўйича приор тақсимотни бошқарувчи гиперпараметрлар a ва b эса $R = diag(\rho)$ тенгламанинг белгиланган нуқталарига ўрнатилади:

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

Логарифмик маргинал эҳтимолликнинг қуйи чегараси баҳоси қийматларини таққослаш

Логарифмик маргинал эҳтимолликнинг қуйи чегараси баҳоси аниқлангандан сўнг, якуний босқичда ушбу баҳонинг бир давр олдинги қиймати билан ўзаро солиштирилади. Агар, қуйи чегара баҳоси бир давр олдинги қийматига нисбатан ортган бўлса, Variational Bayesian M-step босқичидан бошлаб қайтадан ҳисоб-китоблар амалга оширилади. Ушбу такрорланиш қуйи чегара баҳоси ортмай қолгунга қадар давом этади. Бунда, логарифмик маргинал эҳтимолликнинг қуйи чегараси баҳоси олдинги қиймати билан солиштирилган ҳолатда улар орасида фарқ бўлмаса, логарифмик маргинал эҳтимолликнинг қуйи чегараси баҳоси оптимал қийматга эришган ҳисобланади. Бу даврга келиб, гиперпараметрлар орқали параметрлар ҳам ўзининг оптимал қийматига эришади. Аниқланган параметрлар ва мавҳум ўзгарувчилардан келиб чиқиб, кузатилиши мумкин бўлган эрксиз ўзгарувчи Y баҳоланади.