## 13.33

(13.112)より

$$Q(\theta, \theta^{old}) = -\frac{N-1}{2}\ln|\Gamma| - \mathbb{E}_{p(Z|X,\theta^{*})}\left[\frac{1}{2}\sum_{n=2}^{N}(\mathbf{z}_n - A\mathbf{z}_{n-1})^T\Gamma^{-1}(\mathbf{z}_n - A\mathbf{z}_{n-1})\right] + C,$$ $C$ は $A, \Gamma$ を含まない項

$A$ の最尤推定は

$$0 = \frac{\partial}{\partial A}Q(\theta, \theta^{old}) = \frac{\partial}{\partial A}\left[-\frac{N-1}{2}\ln|\Gamma| - \mathbb{E}_{p(Z|X,\theta^{*})}\left[\frac{1}{2}\sum_{n=2}^{N}(\mathbf{z}_n - A\mathbf{z}_{n-1})^T\Gamma^{-1}(\mathbf{z}_n - A\mathbf{z}_{n-1})\right] + C\right]$$

$$= -\mathbb{E}\left[\frac{1}{2}\sum_{n=2}^{N}\frac{\partial}{\partial A}(\mathbf{z}_n - A\mathbf{z}_{n-1})^T\Gamma^{-1}(\mathbf{z}_n - A\mathbf{z}_{n-1})\right]$$ ← 期待値が存在すると仮定して微分と積分を入れ替えた

$$= -\mathbb{E}\left[\frac{1}{2}\sum_{n=2}^{N}\frac{\partial}{\partial A}\left(\mathbf{z}_n^T\Gamma^{-1}\mathbf{z}_n - \mathbf{z}_n^T\Gamma^{-1}A\mathbf{z}_{n-1} - (A\mathbf{z}_{n-1})^T\Gamma^{-1}\mathbf{z}_n + (A\mathbf{z}_{n-1})^T\Gamma^{-1}A\mathbf{z}_{n-1}\right)\right]$$

$$= -\mathbb{E}\left[\frac{1}{2}\sum_{n=2}^{N}\left(-2\Gamma^{-1}\mathbf{z}_n\mathbf{z}_{n-1}^T + 2\Gamma^{-1}A\mathbf{z}_{n-1}\mathbf{z}_{n-1}^T\right)\right]$$

$$= \Gamma^{-1}\mathbb{E}\left[\sum_{n=2}^{N}\left(\mathbf{z}_n\mathbf{z}_{n-1}^T - A\mathbf{z}_{n-1}\mathbf{z}_{n-1}^T\right)\right]$$

$$= \Gamma^{-1}\left\{\left(\sum_{n=2}^{N}\mathbb{E}[\mathbf{z}_n\mathbf{z}_{n-1}^T]\right) - A\left(\sum_{n=2}^{N}\mathbb{E}[\mathbf{z}_{n-1}\mathbf{z}_{n-1}^T]\right)\right\}$$

よって

$$A = \left(\sum_{n=2}^{N}\mathbb{E}[\mathbf{z}_n\mathbf{z}_{n-1}^T]\right)\left(\sum_{n=2}^{N}\mathbb{E}[\mathbf{z}_{n-1}\mathbf{z}_{n-1}^T]\right)^{-1} \quad \cdots (13.113)$$

を得る。

スカラーなのでTrとっても同じ

$$\frac{\partial}{\partial A}\mathbf{z}_n^T\Gamma^{-1}A\mathbf{z}_{n-1} = \frac{\partial}{\partial A}\mathrm{Tr}(\mathbf{z}_n^T\Gamma^{-1}A\mathbf{z}_{n-1})$$

$$= \frac{\partial}{\partial A}\mathrm{Tr}(A\mathbf{z}_{n-1}\mathbf{z}_n^T\Gamma^{-1})$$ ← (C.9)

$$= (\mathbf{z}_{n-1}\mathbf{z}_n^T\Gamma^{-1})^T$$ ← (C.24) 分母レイアウト

$$= \Gamma^{-1}\mathbf{z}_n\mathbf{z}_{n-1}^T$$ ← $\Gamma^{-1}$ は対称なので

スカラーのTr

$$\frac{\partial}{\partial A}(A\mathbf{z}_{n-1})^T\Gamma^{-1}\mathbf{z}_n = \frac{\partial}{\partial A}\mathrm{Tr}((A\mathbf{z}_{n-1})^T\Gamma^{-1}\mathbf{z}_n)$$

$$= \frac{\partial}{\partial A}\mathrm{Tr}(A^T\Gamma^{-1}\mathbf{z}_n\mathbf{z}_{n-1}^T)$$ ← (C.9)

$$= \Gamma^{-1}\mathbf{z}_n\mathbf{z}_{n-1}^T$$ ← (C.25) 分母レイアウト

スカラーのTr

$$\frac{\partial}{\partial A}(A\mathbf{z}_{n-1})^T\Gamma^{-1}A\mathbf{z}_{n-1} = \frac{\partial}{\partial A}\mathrm{Tr}((A\mathbf{z}_{n-1})^T\Gamma^{-1}A\mathbf{z}_{n-1})$$

$$= \frac{\partial}{\partial A}\mathrm{Tr}(\mathbf{z}_{n-1}^T A^T\Gamma^{-1}A\mathbf{z}_{n-1})$$

$$= (\Gamma^{-1})^T A\mathbf{z}_{n-1}\mathbf{z}_{n-1}^T + \Gamma^{-1}A\mathbf{z}_{n-1}\mathbf{z}_{n-1}^T$$

$$= 2\Gamma^{-1}A\mathbf{z}_{n-1}\mathbf{z}_{n-1}^T$$

公式 $\dfrac{\partial \mathrm{Tr}(AX^TBXC)}{\partial X} = B^TXA^TC^T + BXCA$ について

Trの循環(C.9)

$$\frac{\partial}{\partial X_{ij}}\mathrm{Tr}(AX^TBXC) = \frac{\partial}{\partial X_{ij}}\mathrm{Tr}(CAX^TBX) = \mathrm{Tr}\left(\frac{\partial}{\partial X_{ij}}CAX^TBX\right)$$ ← スカラー微分とTrは入れ換えても同じ

$$= \mathrm{Tr}\left(CAX^T\frac{\partial BX}{\partial X_{ij}} + \frac{\partial CAX^T}{\partial X_{ij}}BX\right) = \mathrm{Tr}\left(CAX^T\frac{\partial BX}{\partial X_{ij}}\right) + \mathrm{Tr}\left(\frac{\partial CAX^T}{\partial X_{ij}}BX\right)$$

Trの循環(C.9)

$$= \mathrm{Tr}\left(CAX^TB\frac{\partial X}{\partial X_{ij}}\right) + \mathrm{Tr}\left(CA\frac{\partial X^T}{\partial X_{ij}}BX\right) = \mathrm{Tr}\left(CAX^TB\frac{\partial X}{\partial X_{ij}}\right) + \mathrm{Tr}\left(\frac{\partial X^T}{\partial X_{ij}}BXCA\right)$$

転置のTrは元のTrと同じ

$$= \mathrm{Tr}\left(CAX^TB\frac{\partial X}{\partial X_{ij}}\right) + \mathrm{Tr}\left(\left(\frac{\partial X^T}{\partial X_{ij}}BXCA\right)^T\right) = \mathrm{Tr}\left(CAX^TB\frac{\partial X}{\partial X_{ij}}\right) + \mathrm{Tr}\left(A^TC^TX^TB^T\frac{\partial X}{\partial X_{ij}}\right)$$

$$= \mathrm{Tr}\left((CAX^TB + A^TC^TX^TB^T)\frac{\partial X}{\partial X_{ij}}\right) = \mathrm{Tr}\left((CAX^TB + A^TC^TX^TB^T)J^{ij}\right)$$ ← シングルエントリ

$$= (CAX^TB + A^TC^TX^TB^T)_{ji}$$ ← $\mathrm{Tr}(AJ^{12}) = \mathrm{Tr}\left(\begin{pmatrix}A_{11} & A_{12}\\ A_{21} & A_{22}\end{pmatrix}\begin{pmatrix}0 & 1\\ 0 & 0\end{pmatrix}\right) = \mathrm{Tr}\begin{pmatrix}0 & A_{11}\\ 0 & A_{21}\end{pmatrix} = A_{21}$

$$\therefore \frac{\partial \mathrm{Tr}(AXBX^TC)}{\partial X} = (CAX^TB + A^TC^TX^TB^T)^T = B^TXA^TC^T + BXCA$$

分母レイアウト

$\mathbf{\Gamma}$の推定値は

$$\mathbf{0} = \frac{\partial Q}{\partial \mathbf{\Gamma}} = \frac{\partial}{\partial \mathbf{\Gamma}}\left\{ -\frac{N-1}{2}\ln|\mathbf{\Gamma}| - \mathbb{E}_{p(\mathbf{Z}|\mathbf{X},\boldsymbol{\theta}^{*})}\left[\frac{1}{2}\sum_{n=2}^{N}(\mathbf{z}_n - \mathbf{A}\mathbf{z}_{n-1})^{T}\mathbf{\Gamma}^{-1}(\mathbf{z}_n - \mathbf{A}\mathbf{z}_{n-1})\right] + C \right\}$$

$$= -\frac{N-1}{2}\frac{\partial}{\partial \mathbf{\Gamma}}\ln|\mathbf{\Gamma}| - \mathbb{E}\left[\frac{1}{2}\sum_{n=2}^{N}\frac{\partial}{\partial \mathbf{\Gamma}}(\mathbf{z}_n - \mathbf{A}\mathbf{z}_{n-1})^{T}\mathbf{\Gamma}^{-1}(\mathbf{z}_n - \mathbf{A}\mathbf{z}_{n-1})\right]$$

期待値が存在すると仮定して積分と微分を入れ換えた

$$= -\frac{N-1}{2}(\mathbf{\Gamma}^{-1})^{T} - \mathbb{E}\left[\frac{1}{2}\sum_{n=2}^{N}\left\{-(\mathbf{\Gamma}^{-1})^{T}(\mathbf{z}_n - \mathbf{A}\mathbf{z}_{n-1})(\mathbf{z}_n - \mathbf{A}\mathbf{z}_{n-1})^{T}(\mathbf{\Gamma}^{-1})^{T}\right\}\right]$$

$\mathbf{\Gamma}^{-1}$は対称なので

$$= -\frac{N-1}{2}\mathbf{\Gamma}^{-1} + \frac{1}{2}\sum_{n=2}^{N}\mathbb{E}\left[\mathbf{\Gamma}^{-1}(\mathbf{z}_n - \mathbf{A}\mathbf{z}_{n-1})(\mathbf{z}_n - \mathbf{A}\mathbf{z}_{n-1})^{T}\mathbf{\Gamma}^{-1}\right]$$

$$= -\frac{N-1}{2}\mathbf{\Gamma}^{-1} + \frac{1}{2}\mathbf{\Gamma}^{-1}\sum_{n=2}^{N}\mathbb{E}\left[(\mathbf{z}_n - \mathbf{A}\mathbf{z}_{n-1})(\mathbf{z}_n - \mathbf{A}\mathbf{z}_{n-1})^{T}\right]\mathbf{\Gamma}^{-1}$$

$$\frac{\partial \mathbf{a}^{T}\mathbf{X}^{-1}\mathbf{b}}{\partial X_{ij}} = \mathbf{a}^{T}\frac{\partial \mathbf{X}^{-1}}{\partial X_{ij}}\mathbf{b} = \mathbf{a}^{T}\left(-\mathbf{X}^{-1}\frac{\partial \mathbf{X}}{\partial X_{ij}}\mathbf{X}^{-1}\right)\mathbf{b}$$

(C.21)

$$= -\mathbf{a}^{T}\mathbf{X}^{-1}\mathbf{J}^{ij}\mathbf{X}^{-1}\mathbf{b} = -\left\{(\mathbf{a}^{T}\mathbf{X}^{-1})^{T}(\mathbf{X}^{-1}\mathbf{b})^{T}\right\}_{ij}$$

$$= -\left\{(\mathbf{X}^{-1})^{T}\mathbf{a}\mathbf{b}^{T}(\mathbf{X}^{-1})^{T}\right\}_{ij}$$

分母レイアウト

$$\therefore \frac{\partial \mathbf{a}^{T}\mathbf{X}^{-1}\mathbf{b}}{\partial \mathbf{X}} = -(\mathbf{X}^{-1})^{T}\mathbf{a}\mathbf{b}^{T}(\mathbf{X}^{-1})^{T}$$

$$\mathbf{a}^{T}\mathbf{J}^{12}\mathbf{b} = (a_1, a_2)\begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}\begin{pmatrix} b_1 \\ b_2 \end{pmatrix} = a_1 b_2 = (\mathbf{a}\mathbf{b}^{T})_{12}$$

$$\mathbf{a}\mathbf{b}^{T} = \begin{pmatrix} a_1 \\ a_2 \end{pmatrix}(b_1\ b_2) = \begin{pmatrix} a_1 b_1 & a_1 b_2 \\ a_2 b_1 & a_2 b_2 \end{pmatrix}$$

左右から$\mathbf{\Gamma}$をかけると

$$(N-1)\mathbf{\Gamma} = \sum_{n=2}^{N}\mathbb{E}\left[(\mathbf{z}_n - \mathbf{A}\mathbf{z}_{n-1})(\mathbf{z}_n - \mathbf{A}\mathbf{z}_{n-1})^{T}\right]$$

$$= \sum_{n=2}^{N}\left\{\mathbb{E}[\mathbf{z}_n\mathbf{z}_n^{T}] - \mathbb{E}[\mathbf{z}_n\mathbf{z}_{n-1}^{T}]\mathbf{A}^{T} - \mathbf{A}\mathbb{E}[\mathbf{z}_{n-1}\mathbf{z}_n^{T}] + \mathbf{A}\mathbb{E}[\mathbf{z}_{n-1}\mathbf{z}_{n-1}^{T}]\mathbf{A}^{T}\right\}$$

よって

$$\mathbf{\Gamma} = \frac{1}{N-1}\sum_{n=2}^{N}\left\{\mathbb{E}[\mathbf{z}_n\mathbf{z}_n^{T}] - \mathbb{E}[\mathbf{z}_n\mathbf{z}_{n-1}^{T}]\mathbf{A}^{T} - \mathbf{A}\mathbb{E}[\mathbf{z}_{n-1}\mathbf{z}_n^{T}] + \mathbf{A}\mathbb{E}[\mathbf{z}_{n-1}\mathbf{z}_{n-1}^{T}]\mathbf{A}^{T}\right\} \quad \cdots (13.114)$$

を得る。