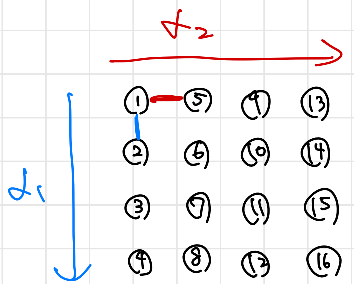
・MRF 精度行列（方向ごとに係数が違う場合）

青：α_1
赤：α_2



$$P(a)=\frac{1}{Z_a}\exp\Big[\underbrace{-\frac{\alpha_1}{2}(a_1-a_2)^2}_{\text{縦方向の結合}}\underbrace{-\frac{\alpha_2}{2}(a_1-a_5)^2}_{\text{横方向の結合}}-\cdots\Big]$$

$$=\frac{1}{Z_a}\Big[-\frac{\alpha_1}{2}a^T P_1 a-\frac{\alpha_2}{2}a^T P_2 a\Big]$$

それぞれの結合の部分だけ要素を持ち、その他は0の行列

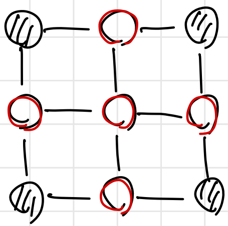
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
1	1,1	-1			-1											
2	-1	2,1	-1			-1										
3		-1	2,1	-1			-1									
4			-1	1,1	-1			-1								
5	-1			-1	1,2	-1			-1							
6		-1			-1	2,2	-1			-1						
7			-1			-1	2,2	-1			-1					
8				-1			-1	1,2	-1			-1				
9					-1			-1	1,2	-1			-1			
10						-1			-1	2,2	-1			-1		
11							-1			-1	2,2	-1			-1	
12								-1			-1	1,2	-1			-1
13									-1			-1	1,1	-1		
14										-1			-1	2,1	-1	
15											-1			-1	2,1	-1
16												-1			-1	1,1

精度行列

$$\Lambda_a=\alpha_1 P_1+\alpha_2 P_2$$

欠損がある場合のガウシアングラフィカルモデル

$(N=N_\ell+N_o)$



○：欠損 ： $a_\ell,\ g_\ell \in \mathbb{R}^{N_\ell}$ (loss)

◎：観測 ： $a_o,\ g_o \in \mathbb{R}^{N_o}$ (observed)

$$a=\begin{bmatrix} a_o \\ a_\ell \end{bmatrix},\quad g=\begin{bmatrix} g_o \\ g_\ell \end{bmatrix}$$

観測モデル

$$p(g\mid a)=\frac{\beta^{\frac{N}{2}}}{(2\pi)^{\frac{N}{2}}}\exp\left[-\frac{\beta}{2}\|g-a\|^2\right]$$

$$=\frac{\beta^{\frac{N}{2}}}{(2\pi)^{\frac{N}{2}}}\exp\left[-\frac{\beta}{2}\|g_o-a_o\|^2-\frac{\beta}{2}\|g_\ell-a_\ell\|^2\right]$$

潜在変数の事前分布

$$p(a\mid\Lambda_a)=\frac{|\Lambda_a|^{\frac{1}{2}}}{(2\pi)^{\frac{N}{2}}}\exp\left[-\frac{1}{2}a^T\Lambda_a a\right]$$

$$\left(\Lambda_a=\begin{bmatrix}\Lambda_{oo} & \Lambda_{o\ell}\\ \Lambda_{\ell o} & \Lambda_{\ell\ell}\end{bmatrix}\right)$$

欠損と観測の対応部分で表せます

同時分布

$$p(a_o,a_\ell,g_o,g_\ell\mid\alpha,\beta)=p(g\mid a,\beta)\,p(a\mid\alpha)$$

$$=\frac{\beta^{\frac{N}{2}}|\Lambda_a|^{\frac{1}{2}}}{(2\pi)^{N}}\exp\left[-\frac{1}{2}a^T\Lambda_a a-\frac{\beta}{2}\|g-a\|^2\right]$$

対数同時分布

$$\ln p(a_o,a_\ell,g_o,g_\ell\mid\alpha,\beta)$$

$$=\frac{N}{2}\ln\beta+\frac{1}{2}\ln|\Lambda_a|-N\ln(2\pi)-\frac{1}{2}a^T\Lambda_a a-\frac{\beta}{2}\|g-a\|^2$$

$$=\frac{N}{2}\ln\beta+\frac{1}{2}\ln|\Lambda_a|-N\ln(2\pi)$$

$$-\frac{1}{2}\Big\{a_o^T\Lambda_{oo}a_o+2a_\ell^T\Lambda_{\ell o}a_o+a_\ell^T\Lambda_{\ell\ell}a_\ell+\beta\|g_o-a_o\|^2+\beta\|g_\ell-a_\ell\|^2\Big]$$

$$\left(\beta g_o^Tg_o-2\beta g_o^Ta_o+\beta a_o^Ta_o+\beta g_\ell^Tg_\ell-2\beta g_\ell^Ta_\ell+\beta a_\ell^Ta_\ell\right)$$

変分推論 (EMアルゴリズム) で $a, g_\ell, \Lambda_a, \beta$ を求める.

$$p(a, g_\ell, \Lambda, \beta \mid g_0) \approx \underbrace{q(a)}_{\shortparallel \atop p(a \mid g_0, \Lambda, \beta)} \cdot \underbrace{q(g_\ell)}_{\shortparallel \atop p(g_\ell \mid g_0, \Lambda, \beta)}$$

(Λ, β は固定パラメータとして EMアルゴリズムで求める)

● $q(a)$ の導出

$$\ln q(a) = \langle \ln p(a, g \mid \alpha, \beta) \rangle_{q(g_\ell)} + \text{Const}$$

$$= -\frac{1}{2}\Big\{ a_0^T (\Lambda_{00} + \beta I_{N_0}) a_0 + a_\ell^T (\Lambda_{\ell\ell} + \beta I_{N_\ell}) a_\ell + 2 a_\ell^T \Lambda_{\ell 0} a_0 - 2\beta g_0^T a_0 - 2\beta \langle g_\ell^T \rangle a_\ell \Big\}$$

$$= -\frac{1}{2} \begin{bmatrix} a_0^T & a_\ell^T \end{bmatrix} \begin{bmatrix} \Lambda_{00} + \beta I_{N_0} & \Lambda_{0\ell} \\ \Lambda_{\ell 0} & \Lambda_{\ell\ell} + \beta I_{N_\ell} \end{bmatrix} \begin{bmatrix} a_0 \\ a_\ell \end{bmatrix} + \beta \begin{bmatrix} g_0 & \langle g_\ell \rangle \end{bmatrix} \begin{bmatrix} a_0 \\ a_\ell \end{bmatrix}$$

$$= -\frac{1}{2} a^T \hat{\Lambda}_a a + \beta \langle g \rangle_{q(g_\ell)}^T a$$

$$\left(\begin{aligned} & q(a) = \mathcal{N}(a \mid \hat{\mu}_a, \hat{\Lambda}_a^{-1}) \text{ とおくと} \\ & \ln q(a) = -\frac{1}{2} (a - \hat{\mu}_a)^T \hat{\Lambda}_a (a - \hat{\mu}_a) + \text{Const} \\ & \qquad = -\frac{1}{2} a^T \Lambda_a a + \hat{\mu}_a^T \hat{\Lambda}_a a + \text{Const} \end{aligned} \right)$$

よって

$$\begin{cases} \hat{\Lambda}_a = \begin{bmatrix} \Lambda_{00} + \beta I_{N_0} & \Lambda_{0\ell} \\ \Lambda_{\ell 0} & \Lambda_{\ell\ell} + \beta I_{N_\ell} \end{bmatrix} \\ \hat{\mu}_a = \beta \cdot \hat{\Lambda}_a^{-1} \langle g \rangle_{q(g_\ell)} \end{cases}$$

//

◦ $q(g_k)$ の導出

$$\ln q(g_k) = \langle \ln p(a, g \mid \Lambda, \beta) \rangle_{q(a)} + \mathrm{const}$$

$$= -\frac{\beta}{2} \langle \| g_k - a_k \|^2 \rangle_{q(a)}$$

$$= -\frac{\beta}{2} \left(g_k^T g_k - 2 \langle a_k \rangle^T g_k + \underbrace{\langle a_k^T a_k \rangle} \right)$$

$$\left(\begin{aligned} \underbrace{\langle a_k^T a_k \rangle_{q(a_k)}} &= \sum_{k=1}^{N_k} \langle a_{k,k}^2 \rangle_{q(a_k)} \\ &= \sum_{k=1}^{N_k} \left(\mathrm{Var}(a_{k,k}) + \langle a_{k,k} \rangle^2 \right) \\ &= \langle a_k^T \rangle \langle a_k \rangle + \sum_{k=1}^{N_k} \mathrm{Var}_{q(a_k)}(a_{k,k}) \end{aligned} \right)$$

$$= -\frac{\beta}{2} \left(g_k^T g_k - 2 \langle a_k^T \rangle g_k + \langle a_k^T \rangle \langle a_k \rangle \right) + \mathrm{const}$$

$$= -\frac{\beta}{2} \| g_k - \langle a_k \rangle \|^2 + \mathrm{const}$$

$$q(g_k) = \mathcal{N}\left(g_k \mid \langle a_k \rangle_{q(a_k)}, \beta^{-1} I_{N_k} \right)$$

◦ パラメータ Λ_a, β の推定

$$Q(\Lambda_a, \beta) = \langle \ln p(a, g \mid \alpha, \beta) \rangle_{q(a) q(g_k)}$$

$$= \frac{N}{2} \ln \beta + \frac{1}{2} \ln |\Lambda_a| - \frac{1}{2} \langle a^T \Lambda_a a \rangle_{q(a)} - \frac{\beta}{2} \langle \| g - a \|^2 \rangle_{q(g_k) q(a)}$$

Q関数を最大化するパラメータ Λ_a, β を探す.

• $\frac{\partial Q}{\partial \beta} = 0$ を解くと

$$0 = \frac{N}{2} \cdot \frac{1}{\beta} - \frac{1}{2} \langle \| g - a \|^2 \rangle_{q(g_k) q(a)}$$

$$\beta = \left\{ \frac{1}{N} \langle \| g - a \|^2 \rangle_{q(g_k)} \right\}^{-1}$$

• $\frac{\partial Q}{\partial \Lambda_a} = 0$ を解くと

$\boxed{\Lambda_a = \alpha P \text{ の場合}}$

$$Q(\alpha, \beta) = \frac{N}{2} \ln \beta + \frac{1}{2} \ln (\alpha^N P) - \frac{\alpha}{2} \langle a^T P a \rangle_{q(a)} - \frac{\beta}{2} \langle \| g - a \|^2 \rangle_{q(g_k)}$$

$$0 = \frac{\partial Q}{\partial \alpha} = \frac{N}{2} \cdot \frac{1}{\alpha} - \frac{1}{2} \langle a^T P a \rangle_{q(a)}$$

$$\alpha = \left\{ \frac{1}{N} \langle a^T P a \rangle_{q(a)} \right\}^{-1}$$

[$\Lambda_a = \alpha_1 P_1 + \alpha_2 P_2$ の場合]

• $\frac{\partial Q}{\partial \alpha_1} = 0$ を解くと

$$0 = \frac{d}{d\alpha_1}\left\{ \frac{1}{2}\ln|\alpha_1 P_1 + \alpha_2 P_2| - \frac{1}{2}\langle a^T(\alpha_1 P_1 + \alpha_2 P_2) a\rangle \right\}$$

$$= \frac{1}{2}\cdot\frac{\frac{d}{d\alpha_1}|\alpha_1 P_1 + \alpha_2 P_2|}{|\alpha_1 P_1 + \alpha_2 P_2|} - \frac{1}{2}\langle a^T P_1 a\rangle_{q(a)}$$

$$\frac{d}{dx}|A(x)| = |A(x)|\,\mathrm{tr}\left(A(x)^{-1}\frac{\partial A(x)}{\partial x}\right) \text{ より}$$

$$\frac{d}{d\alpha_1}|\alpha_1 P_1 + \alpha_2 P_2| = |\alpha_1 P_1 + \alpha_2 P_2|\,\mathrm{tr}\left\{(\alpha_1 P_1 + \alpha_2 P_2)^{-1} P_1\right\}$$

$$0 = \mathrm{tr}\left\{(\alpha_1 P_1 + \alpha_2 P_2)^{-1} \cdot P_1\right\} - \langle a^T P_1 a\rangle$$

↑

解析的に解けない？ (α_2も同様)

EMアルゴリズム全体の流れ ($\Lambda_a = \alpha P$ の場合)

• a の推定

$$q(a) = \mathcal{N}(a \mid \hat{\mu}_a, \hat{\Lambda}_a^{-1})$$

αが含まれる

$$\begin{cases} \hat{\Lambda}_a = \begin{bmatrix} \Lambda_{oo} + \beta I_{N_o} & \Lambda_{o\ell} \\ \Lambda_{\ell o} & \Lambda_{\ell\ell} + \beta I_{N_\ell} \end{bmatrix} \\ \hat{\mu}_a = \beta \cdot \hat{\Lambda}_a^{-1}\langle g\rangle_{q(g_\ell)} \end{cases}$$

• g_ℓ の推定

$$q(g_\ell) = \mathcal{N}\Big(g_\ell \,\Big|\, \underbrace{\langle a_\ell\rangle_{q(a_\ell)}}_{\hat{\mu}_{g_\ell}}, \underbrace{\beta^{-1} I_{N_\ell}}_{\hat{\Lambda}_{g_\ell}^{-1}}\Big)$$

• β の推定

$$\beta = \left\{\frac{1}{N}\langle\|g - a\|^2\rangle_{q(g_\ell)q(a)}\right\}^{-1} = \frac{N}{\|g_o\|^2 + \|\hat{\mu}_{g_\ell}\|^2 + \|\hat{\mu}_a\|^2 + \mathrm{tr}(\hat{\Lambda}_{g_\ell}^{-1}) + \mathrm{tr}(\hat{\Lambda}_a^{-1}) - 2\langle g^T\rangle_{q(g_\ell)}\hat{\mu}_a}$$

(計算は次ページ)

• α の推定

$$\alpha = \left\{\frac{1}{N}\langle a^T P a\rangle_{q(a)}\right\}^{-1} = \frac{N}{\mathrm{tr}(P\hat{\Lambda}_a^{-1}) + \hat{\mu}_a^T P \hat{\mu}_a}$$

● $\langle a^T P a \rangle_{q(a)}$ の計算

「2次形式の期待値」を参照。

$$\boxed{\begin{array}{l} E_x\left[x^T A x\right] = \mathrm{Tr}(A C_x) + \mu_x^T A \mu_x \\ \left(p(x) = N(x \mid \mu_x, C_x),\quad \begin{array}{ll} x \in \mathbb{R}^{N\times 1} & \mu_x \in \mathbb{R}^{N\times N} \\ A \in \mathbb{R}^{N\times N} & C_x \in \mathbb{R}^{N\times N} \end{array} \right) \end{array}}$$ より

$$\langle a^T P a \rangle_{q(a)} = \mathrm{Tr}(P \hat{\Lambda}_a^{-1}) + \hat{\mu}_a^T P \hat{\mu}_a \quad //$$

● $\langle \| g - a \|^2 \rangle_{q(g_1)q(a)}$ の計算

変分近似のおかげでうまく分離できる

上の式を利用

$$\langle \| g - a \|^2 \rangle = \langle g^T g \rangle - 2 \langle g^T \rangle \langle a \rangle + \langle a^T a \rangle$$

$$= g_0^T g_0 + \langle g_1^T g_1 \rangle_{q(g_1)} - 2 \langle g^T \rangle \langle a \rangle + \mathrm{tr}(\hat{\Lambda}_a^{-1}) + \hat{\mu}_a^T \hat{\mu}_a$$

欠損部分のみ期待値とる

$$= \|g_0\|^2 + \mathrm{tr}(\hat{\Lambda}_{g_1}^{-1}) + \|\hat{\mu}_{g_1}\|^2 - 2 \langle g^T \rangle_{q(g_1)} \hat{\mu}_a + \mathrm{Tr}(\hat{\Lambda}_a^{-1}) + \|\hat{\mu}_a\|^2 \quad //$$