X, Z

$$\ln p(X) = \mathcal{L}(q) + KL(q\|p) =$$

max, → min

$$= \int \ln \frac{p(X,Z)}{q(Z)} q(Z)\,dZ - \int \ln \frac{p(Z|X)}{q(Z)} q(Z)\,dZ$$

$p(\theta, Z|X)$

$$\mathcal{L}(q) = \int \ln p(X,Z)\, q(Z)\,dZ - \int \ln q(Z)\, q(Z)\,dZ$$

mean field theory $\quad q(Z) = \prod_i q_i(Z_i) \quad Z_i \cap Z_j = \emptyset$

$$\ln q_j^*(Z_j) = \mathbb{E}_{Z_{-j}}[\ln p(X,Z)] + \text{const}$$

$x_1 \dots x_N$

$$p(D|\mu,\tau) = \left(\frac{\tau}{2\pi}\right)^{N/2} e^{-\frac{\tau}{2}\sum_n (x_n-\mu)^2}$$

$$p(\mu|\tau) = \mathcal{N}(\mu|\mu_0, (\lambda_0\tau)^{-1})$$

$$p(\tau) = \text{Gamma}(\tau|a_0, b_0)$$

$$p(\mu,\tau|D) \propto p(D|\mu,\tau)\cdot p(\mu|\tau)\, p(\tau) \approx q_\mu(\mu) q_\tau(\tau)$$

"$q(\mu,\tau)$"

$KL(q\|p) \to \min$

$$\ln q_\mu^*(\mu) = -\frac{\mathbb{E}\tau}{2}\left(\lambda_0(\mu-\mu_0) + \sum_{n=1}^N (x_n-\mu)^2\right) + \text{const}$$

$q_\mu^*(\mu) = \mathcal{N}(\mu|\mu'_\mu, \tau'_\mu)$

$$\boxed{\mu'_\mu = \frac{\lambda_0\mu_0 + \sum x_n}{\lambda_0 + N}} \qquad \boxed{\tau'_\mu = (\lambda_0 + N)\,\mathbb{E}\tau} = (\lambda_0+N)\frac{a'}{b'}$$

$\tau'_\mu = N \cdot \mathbb{E}\tau = N \cdot a'/b'$

$$\ln q_\tau^*(\tau) = \mathbb{E}_\mu\Big[\ln\tau\left(a_0 - 1 + \frac{1}{2} + \frac{N}{2}\right) - \tau\left(b_0 + \frac{\lambda_0}{2}(\mu-\mu_0)^2 + \frac{1}{2}\sum_n (x_n-\mu)^2\right)\Big]$$

$$\boxed{\mathbb{E}_q\tau = \frac{a'_\tau}{b'_\tau}}$$

$$a'_\tau = a_0 + \frac{N+1}{2}$$

$$b'_\tau = b_0 + \frac{1}{2}\mathbb{E}_\mu\left[\lambda_0(\mu-\mu_0)^2 + \sum_n (x_n-\mu)^2\right]$$

$$b'_\tau = b_0 + \left(\lambda_0\mu_0^2 + \sum_n x_n^2\right)/2 - \frac{\mu^2(\lambda_0+N) - 2\mu\left(\mu_0\lambda_0 + \sum x_n\right) + \lambda_0\mu_0^2 + \sum_n x_n^2}{}$$

$$- (\mathbb{E}\mu)\left(\mu_0\lambda_0 + \sum x_n\right) + \frac{1}{2}\left(\mathbb{E}\mu^2\right)(\lambda_0+N)$$

$$\mathbb{E}\mu^2 = (\mathbb{E}\mu)^2 + \sigma_\mu^2 = (\mathbb{E}\mu)^2 + \frac{1}{\tau'_\mu}$$

$$= b_0 + \frac{\lambda_0\mu_0^2 + \sum_n x_n^2}{2} - 1\,\frac{(\lambda_0\mu_0 + \sum x_n)^2}{\lambda_0+N} + \frac{\lambda_0+N}{2}\left(\left(\frac{\lambda_0\mu_0+\sum x_n}{\lambda_0+N}\right)^2 + \frac{1}{\mathbb{E}\tau}\right)$$

$$\mathbb{E}\tau = \frac{a'_\tau}{b'_\tau}$$

$$= b_0 + \frac{\lambda_0\mu_0^2 + \sum_n x_n^2}{2} - \frac{1}{2}\frac{(\lambda_0\mu_0+\sum x_n)^2}{\lambda_0+N} + \frac{\lambda_0+N}{2(\lambda_0+N)\mathbb{E}\tau}$$

$$b'_\tau = b_0 + \frac{1}{2}\left(\lambda_0\mu_0^2 + \sum_n x_n^2 - \frac{(\lambda_0\mu_0+\sum x_n)^2}{\lambda_0+N} + \frac{1}{\mathbb{E}\tau}\right)$$

$$= b_0 + \frac{1}{2}\left(\ldots + \frac{b'_\tau}{a_0 + \frac{N+1}{2}}\right)$$

$$b'_\tau\left(1 - \frac{1}{2\left(a_0 + \frac{N+1}{2}\right)}\right) = b'_\tau\,\frac{2a_0 + (N+1) - 1}{2\left(a_0+\frac{N+1}{2}\right)} =$$

$$b'_\tau = \frac{a_0 + \frac{N+1}{2}}{a_0 + \frac{N}{2}}\left(b_0 + \frac{1}{2}(\ldots)\right)$$

$$\lambda_0 = \mu_0 = a_0 = b_0 = 0$$

$$\mathbb{E}\mu = \mu' = \frac{1}{N}\sum x_n;\quad \mathbb{E}\tau = \frac{N+1}{2b'_\tau}$$

$$b'_\tau = \frac{1}{2}\mathbb{E}_\mu\left[\sum_n (x_n-\mu)^2\right] = \frac{1}{2}\left(\sum_n x_n^2 - 2\sum x_n\cdot\mathbb{E}\mu + N\cdot\mathbb{E}\mu^2\right)$$

$$= \frac{1}{2}\left(\sum_n x_n^2 - \frac{2}{N}\cdot\left(\sum x_n\right)^2 + N\cdot\frac{(\sum x_n)^2}{N^2} + N\cdot\frac{2b'_\tau}{N\cdot(N+1)}\right)$$

$$2b'_\tau\left(1 - \frac{1}{N+1}\right) = \sum_n x_n^2 - \frac{1}{N}\left(\sum x_n\right)^2$$

$$b'_\tau = \frac{N+1}{2N}\cdot\left(\sum_n x_n^2 - \frac{1}{N}\left(\sum x_n\right)^2\right)$$

$$\sum_n (x_n - \bar{x})^2 = \bar{x}\cdot N = \sum_n x_n^2 - 2\bar{x}\left(\sum_n x_n\right) + N(\bar{x})^2 = \sum_n x_n^2 - \bar{x}^2\cdot N$$

$$\frac{1}{\sigma_\mu'^2} = \tau'_\mu = N\cdot\frac{a'_\tau}{b'_\tau} = N\cdot\frac{N+1}{2}\cdot\frac{2N}{N+1}\cdot(\ldots)^{-1}$$

$$\sigma_\mu'^2 = \frac{1}{\tau'_\mu} = \frac{1}{N^2}\cdot\sum_n (x_n - \bar{x})^2$$

$$\sigma'_\mu = \frac{1}{N}\sqrt{\sum_n (x_n-\bar{x})^2}$$

$\mu'_\tau = \bar{x}$

$$\mathbb{E}[\mu^2] = \bar{x}^2 + \frac{1}{N \cdot \mathbb{E}\tau}$$

$$\mathbb{E}\tau = \frac{a'}{b'} ; \quad \frac{1}{\mathbb{E}\tau} = \frac{b'}{a'} = \frac{1}{N+1}\mathbb{E}_\mu\left[\sum_n (x_n-\mu)^2\right] =$$

$$= \frac{1}{N+1}\left(\sum_n x_n^2 - \frac{2}{N}\left(\sum x_n\right)^2 + N\,\mathbb{E}\mu^2\right)$$

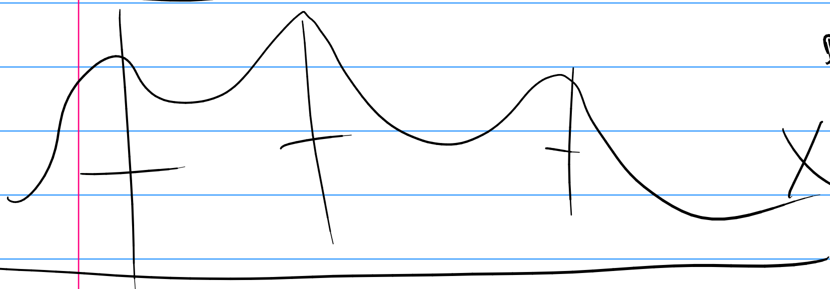
$$= \frac{1}{N+1}\left(\sum_n x_n^2 - \frac{2}{N}\left(\sum x_n\right)^2 + \cancel{N\cdot(\bar{x})^2} + \frac{1}{\mathbb{E}\tau}\right)$$

$$\frac{1}{\mathbb{E}\tau}\left(1 - \frac{1}{N+1}\right) = \frac{1}{N+1}\left(\sum_n (x_n-\bar{x})^2\right)$$

$$\frac{1}{\mathbb{E}\tau} = \frac{1}{N}\sum_n (x_n-\bar{x})^2$$

$$\frac{1}{N}\tau'_\mu = \mathbb{E}\tau = \left(\frac{1}{N}\sum_n (x_n-\bar{x})^2\right)^{-1}$$

$$e^{-\frac{1}{2}(\ )^T\Lambda(\ )} \qquad \Lambda = \Sigma^{-1}$$



$$p(\bar{x}) = \sum_i \pi_i \, N(\bar{x}|\bar{\mu}_i, \Lambda_i)$$

$$X, \quad Z = (z_{nk}) \quad (0 \ldots 1 \ldots 0)$$

$$p(X, Z|\bar{\pi}, \bar{\mu}, \Lambda) = p(Z|\bar{\pi}) \cdot p(X|Z, \bar{\mu}, \Lambda) =$$

$$= \left(\prod_{n,k} \pi_k^{z_{nk}}\right)\left(\prod_{n,k} N(\bar{x}_n|\bar{\mu}_k, \Lambda_k)^{z_{nk}}\right)$$

$$p(\bar{\pi}|\alpha) = Dir(\bar{\pi}|\bar{\alpha}) \propto \pi_1^{\alpha_1-1}\pi_2^{\alpha_2-1}\ldots\pi_K^{\alpha_K-1}$$

$$\prod_k p(\bar{\mu}_k, \Lambda_k \ldots) = \prod_k p(\Lambda_k)\, p(\mu_k|\Lambda_k) = \prod_k W(\Lambda_k|\nu_0) \cdot N(\bar{\mu}_k|\bar{\mu}_0, (\beta_0\Lambda_k)^{-1})$$

$$p(X|\bar{\alpha}, \mu_0, \nu_0, \Lambda_0, \beta_0) = \int p(X, Z, \bar{\pi}, \bar{\mu}, \Lambda)\, dZ\, d\pi\, d\mu\, d\Lambda$$

$$p(\pi, \mu, \Lambda, Z|X) \propto p(X, Z, \pi, \mu, \Lambda) = p(\pi)\cdot p(\mu, \Lambda) \cdot p(Z|\pi) \cdot p(X|Z, \mu, \Lambda)$$

$$\approx q(Z, \pi, \mu, \Lambda) = q_Z(Z)\, q_\theta(\pi, \mu, \Lambda)$$

$$\ln q_z^*(z) = \mathbb{E}_\theta\left[\ln p(X, z, \pi, \mu, \Lambda)\right] \overset{+\text{const}}{=}$$

$$= \mathbb{E}_\pi\left[\ln p(z|\pi)\right] + \mathbb{E}_{\mu,\Lambda}\left[\ln p(X|z, \mu, \Lambda)\right] + \text{const}$$

$$\ln q_\theta^*(\pi, \mu, \Lambda) = \mathbb{E}_z\left[\ln p(X, z, \pi, \mu, \Lambda)\right] + \text{const}$$

$$= \ln p(\pi) + \ln p(\mu, \Lambda) + \mathbb{E}_z\left[\ln p(z|\pi)\right] + \mathbb{E}_z\left[\ln p(X|z, \mu, \Lambda)\right] + \text{const}$$