№1

$x_i \sim \mathcal{N}(\mu, \sigma^2)$ - н.о.р.

$\mu \sim \mathcal{N}(\mu_0, \sigma_0^2)$

$\frac{1}{\sigma^2} \sim \Gamma(\alpha, \beta)$

а) $p(X, \mu, \sigma^2 | \alpha, \beta, \mu_0, \sigma_0) = \cancel{\prod_{i=1}^{n} \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{(x_i-\mu)}{\ldots}\right.}$

$= \prod_{i=1}^{n} \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{(x_i-\mu)^2}{2\sigma^2}\right) \cdot \frac{1}{\sqrt{2\pi}\,\sigma_0} \exp \cdot$

$\cdot \exp\left(-\frac{(\mu-\mu_0)^2}{2\sigma_0^2}\right) \cdot \frac{\alpha^\beta \cdot \left(\frac{1}{\sigma^2}\right)^{\beta-1}}{\Gamma(\beta)} e^{-\frac{\alpha}{\sigma^2}}$

б) $p(\mu, \sigma^2 | X, \alpha, \beta, \mu_0, \sigma_0) =$

$= \frac{p(X | \mu, \sigma^2, \alpha, \beta, \mu_0, \sigma_0) \cdot p(\mu, \sigma^2 | \alpha, \beta, \mu_0, \sigma_0)}{p(X | \alpha, \beta, \sigma_0, \mu_0)} =$

$= \frac{\prod_{i=1}^{n} p(x_i | \mu, \sigma^2, \ldots)\, p(\mu | \mu_0, \sigma_0) \cdot p\left(\frac{1}{\sigma^2} \middle| \alpha, \beta\right)}{\int_{\mu, \sigma^2} \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{(x_i-\mu)^2}{2\sigma^2}\right) \cdot p(\mu | \mu_0, \sigma_0) \cdot p\left(\frac{1}{\sigma^2} \middle| \alpha, \beta\right) \cdot d\mu\, d\frac{1}{\sigma^2}}$

Интеграл не зависит от μ, σ^2 $\Rightarrow$ он отвечает за нормировочную константу и не влияет на форму распределения.

Вычислим:

$$p(\mu, \sigma^2 | x, \alpha, \beta, \mu_0, \sigma_0) \propto$$

$$\propto \left(\frac{1}{\sqrt{\sigma^2}}\right)^n \exp\left[-\frac{1}{2\sigma^2}\sum_{i=1}^{n}(x_i-\mu)^2\right]\cdot$$

$$\cdot \exp\left[-\frac{1}{2\sigma_0^2}(\mu-\mu_0)^2\right]\cdot\left(\frac{1}{\sigma^2}\right)^{\beta-1}\cdot\exp\left[-\alpha\frac{1}{\sigma^2}\right] =$$

$$= \left|\tau = \frac{1}{\sigma^2}\right| = \tau^{\frac{n}{2}+\beta-1}\cdot\exp\left[-\frac{\tau}{2}\sum_{i=1}^{n}(x_i-\mu)^2 -\right.$$

$$\left.- \frac{1}{2\sigma_0^2}(\mu-\mu_0)^2 - \alpha\tau\right] = \tau^{\frac{n}{2}+\beta-1}\cdot$$

$$\cdot\exp\left(-\frac{\tau}{2}n\cdot((\bar{x}^2 - 2\bar{x}\mu+\mu^2)+s^2) -\right.$$

$$\left.- \frac{1}{2\sigma_0^2}(\mu^2-2\mu\mu_0+\mu_0^2)-\alpha\tau\right] =$$

$$= \tau^{\frac{n}{2}+\beta-1}\cdot\exp\left[-\frac{n}{2}\tau s^2-\alpha\tau\right]\cdot$$

$$\cdot\exp\left(-\left(\frac{n}{2}\tau+\frac{1}{2\sigma_0^2}\right)\mu^2+\left(n\tau\bar{x}+\frac{\mu_0}{\sigma_0^2}\right)\mu -\right.$$

$$-\frac{n\bar{x}^2}{2}\tau - \frac{\mu_0^2}{2\sigma_0^2}\Big\} \sim$$

$$\sim \Gamma\left(\frac{1}{\sigma^2}\,\Big|\,\frac{n}{2}s^2+\alpha,\ \frac{n}{2}+\beta\right)\cdot \mathcal{N}\left(\ldots\right)$$

$$= \mathcal{N}\left(\mu\,\Big|\,\frac{\left(\frac{n\bar{x}}{\sigma^2}+\frac{\mu_0}{\sigma_0^2}\right)}{\frac{n}{\sigma^2}+\frac{1}{\sigma_0^2}},\ \frac{1}{\frac{n}{\sigma^2}+\frac{1}{\sigma_0^2}}\right) = \mathcal{N}\Gamma$$

в) Да, является, т.к. априорное распределение из гамма-нормального семейств. и апостериорное оказалось из гамма-~~норм~~ нормального

$\mathcal{N}(\mu_0, \sigma^2)\cdot\Gamma(\alpha,\beta) \sim \mathcal{N}\Gamma(\mu, \lambda, \alpha, \beta) \Rightarrow$

$\Rightarrow$ мы перебрали все распределения из гамма-нормального семейства.

Значит оно и сопряжено.

2). $q(\mu, \sigma^2) = q(\mu) q(\sigma^2)$

$\log q(\mu) \sim \mathbb{E}_{q \setminus \mu} \log p(\mu, \sigma^2 | x, a, \beta, \mu_0, \sigma_0^2) \sim$

$\sim \mathbb{E}_{q \setminus \mu} \Big[-\frac{1}{2} \Big(\Big(\frac{n}{\sigma^2} + \frac{1}{\sigma_0^2} \Big) \mu^2 - 2\mu \Big(\frac{n\bar{x}}{\sigma^2} + \frac{\mu_0}{\sigma_0^2} \Big) +$

$+ \Big(\frac{n\bar{x}}{\sigma^2} + \frac{\mu_0}{\sigma_0^2} \Big)^2 \Big(\frac{n}{\sigma^2} + \frac{1}{\sigma_0^2} \Big) \Big) + \log \Big(\frac{n}{\sigma^2} + \frac{1}{\sigma_0^2} \Big) \Big] =$

$= -\frac{1}{2} \Big(\mu^2 \mathbb{E}_{q \setminus \mu} \Big(\frac{n}{\sigma^2} + \frac{1}{\sigma_0^2} \Big) - 2\mu \mathbb{E}_{q \setminus \mu} \Big(\frac{n\bar{x}}{\sigma^2} +$

$+ \frac{\mu_0}{\sigma_0^2} \Big) \Big) = -\frac{1}{2} \mu^2 \Big(\mathbb{E}_{q \setminus \mu} \frac{n}{\sigma^2} + \frac{1}{\sigma_0^2} \Big) +$

$+ \mu \Big(\mathbb{E}_{q \setminus \mu} \frac{n\bar{x}}{\sigma^2} + \frac{\mu_0}{\sigma_0^2} \Big) \Rightarrow$

$\Rightarrow q(\mu) \sim \mathcal{N} \Big(\mu \Big| s^2 \Big(n\bar{x} \mathbb{E} \frac{1}{\sigma^2} + \frac{\mu_0}{\sigma_0^2} \Big), s^2 \Big)$, где

$s^2 = \Big(n \mathbb{E} \frac{1}{\sigma^2} + \frac{1}{\sigma_0^2} \Big)^{-1}$

№2

a) $p(t, \lambda) = \prod_{i=1}^{n} \lambda \exp(-\lambda t_i) =$

$= \lambda^n \exp(-\lambda \sum_{i=1}^{n} t_i)$

$p(\lambda | t, \alpha) = \frac{p(t|\lambda) \cdot p(\lambda|\alpha)}{\int p(t|\lambda) p(\lambda|\alpha) d\lambda} \Rightarrow$

$\Rightarrow p(\lambda | t, \alpha) \sim \lambda^n \exp(-\lambda \sum_{i=1}^{n} t_i) \cdot$

$\cdot p(\lambda|\alpha)$

$p(\lambda|\alpha) \sim \Gamma(\alpha | 1, \frac{1}{\alpha}) \Rightarrow$

$\Rightarrow p(\lambda|\alpha) \sim \alpha \exp(-\alpha\lambda)$

$\Rightarrow p(\lambda | t, \alpha) \sim \lambda^n \exp(-\lambda(\sum_{i=1}^{n} t_i + \alpha))$

$\Rightarrow p(\lambda | t, \alpha) \sim \Gamma(\lambda | n+1, \frac{1}{\sum t_i + n\alpha})$

$\Rightarrow$ Сопряженное.

б) Обоснованность?

$$p(t|\alpha) = \int p(t|\lambda)\, p(\lambda|\alpha)\, d\lambda =$$

$$= \int_0^\infty \lambda^n \exp\left(-\lambda \sum_{i=1}^n t_i\right) \alpha e^{-\alpha\lambda}\, d\lambda =$$

$$\frac{\partial p}{\partial \alpha} = \int_0^\infty \lambda^n \exp\left(-\lambda \sum_{i=1}^n t_i\right)\left(e^{-\alpha\lambda} - \alpha^2 e^{-\alpha\lambda}\right) =$$

$$= (1-\alpha^2) \int_0^\infty \lambda^n \exp\left(-\lambda\left(\sum_{i=1}^n t_i + \alpha\right)\right) d\lambda =$$

$$= \frac{(1-\alpha^2)}{\left(\sum_{i=1}^n t_i + \alpha\right)^{n+1}} \int_0^\infty u^n \exp(-u)\, du =$$

$$= \frac{\Gamma(n+1)(1-\alpha^2)}{\left(\sum_{i=1}^n t_i+\alpha\right)^{n+1}} = 0 \Rightarrow \alpha = 1.$$

$\cancel{\mathcal{N}}$

$$p(y, w \mid A, X) = \mathcal{N}(w \mid 0, A^{-1}) \prod_j p(y_j \mid x_j, w),$$

где

$$\mathbb{P}(y_j = 1) = \frac{1}{1 + \exp(-w^T x_j)} = \sigma(w^T x_j)$$

$A = \mathrm{diag}(\alpha_j)$

$$A = \arg\max_A (y_0 \mid X_1, A)$$

а) $p(y, w \mid A, X) = p(y \mid \cancel{X} w, A) \cdot p(\cancel{w} \mid A) =$

$$= \left[\prod_{j=1}^{m} p(y_j \mid x_j, w)\right] \cdot \mathcal{N}(w \mid 0, A^{-1}) =$$

$$= \left[\prod_{j=1}^{m} (1 - \mathbb{P}(y_j = 1))^{1-y_j} (\mathbb{P}(y_j = 1))^{y_j}\right] \cdot$$

$$\cdot \mathcal{N}(w \mid 0, A^{-1}) =$$

$$= \frac{\sqrt{\det A}}{(\sqrt{2\pi})^{[redacted]}} \exp\left(-\frac{w^T A w}{2}\right) \cdot \cancel{\prod_{j=1}^{m} (A}$$

$$\cdot \prod_{j=1}^{m} \left(\frac{1}{1 + \exp(-w^T x_j)}\right)^{y_j} \cdot \left(\frac{\exp(-w^T x_j)}{1 + \exp(-w^T x_j)}\right)^{1-y_j}$$

б) j признак не значим $\Leftrightarrow$ ~~модель~~ ~~интерпретация~~ $p(w_j, \cdot) = p(\cdot)$

$\Rightarrow$ Раскладываем единичный вектор глоб. с w_j

$$w^T(\operatorname{diag}\alpha)\, w = (\alpha_1 w_1, \dots, \alpha_n w_n)^T w = \sum_{j=1}^{n} \alpha_j w_j^2$$

$\Rightarrow$ При $\alpha_j = 0$ модель инвариантна отн. w_j $\Rightarrow$ $\boxed{\alpha_j = 0}$

в) $$p(y, w \mid X, A) \geq L(w, A, \xi) = \frac{\sqrt{\det A}}{(\sqrt{2\pi})^n} \cdot \exp\left(-\frac{1}{2} w^T A w\right) \prod_{j=1}^{n} \Big[\sigma(\xi_j) \cdot \cancel{\exp(\dots)}$$

$$\cdot \exp\left\{-\frac{2\sigma(\xi_j) - 1}{4\xi_j}\left(w^T x_j x_j^T w - \xi_j^2\right) + \frac{y_j w^T x_j - \xi_j}{2}\right\}\Big] = \frac{\sqrt{\det A}}{(\sqrt{2\pi})^n} \prod_{j=1}^{n} \sigma(\xi_j) \cdot$$

$$\cdot \exp\left(\frac{2\sigma(\xi_j) - 1}{4\xi_j}\left(\xi_j^2 - \frac{\xi_j}{2}\right)\right) \exp\left\{-\frac{1}{2} w^T \tilde{A} w + w^T R\right\}$$

$$\tilde{A} = A + \sum_{j=1}^{n} \frac{2\sigma(\xi_j) - 1}{2\xi_j} x_j x_j^T; \quad R = \frac{1}{2}\sum_{j=1}^{n} y_j x_j$$

$$c)\ p(y|X,A) = \int p(y,w|X,A)\,dw \geq$$

$$\geq \int \mathcal{L}(w,A,\xi)\,dw = \tilde{\mathcal{L}}(A,\xi)$$

$$\tilde{\mathcal{L}}(A,\xi) \to \max_{A,\xi}$$

E-шаг:

$$q(y) = p(y|X,w) = \prod_{n=1}^{m} p(y_n|x_n,\theta)$$

M-шаг:

$$\mathbb{E}_{q(y)} \log p(X,y|w) \to \max_{\theta}, \quad \theta = (A,\xi)$$

E-шаг:

$$p(y|X,w) = \frac{\prod_{n=1}^{m} p(x_n|y_n,w)\,p(y_n)}{\int \prod_{n=1}^{m} p(x_n|y_n,\theta)\,p(y_n)\,dy_n} =$$

$$= \prod_{n=1}^{m} p(y_n,A_n,w) \Rightarrow$$

$$\Rightarrow q(y_n) = \mathcal{N}(y_n|\mu_n, \Sigma_n)$$

М-шаг:

$$\mathbb{E}_{q(y)} \log p(x, y \mid A, \sigma) \to \max_{A, \sigma}$$

$$\theta(A, \xi) = (V, \sigma)$$

$$\log p(x, y \mid V, \sigma) = \log \prod_{n=1}^{m} p(x_n, y_n \mid V, \sigma) =$$

$$= \log \prod_{n=1}^{m} \left\{ \frac{1}{R^2} \exp\left(-\frac{1}{2\sigma^2}(x_n - V y_n)^T (x_n - V y_n)\right) \right\} p(y_n) =$$

$$= \sum_{n=1}^{m} \log \left\{ \frac{1}{R^2} \exp\left(-\frac{1}{2\sigma^2}\left(x_n^T x_n - 2 y_n^T V^T x_n + \right.\right.\right.$$

$$\left.\left.\left. + y_n^T V^T V y_n\right)\right) \right\} p(y_n) = \log p(y) + \text{const} +$$

$$+ \sum_{n=1}^{m} \left(-\frac{1}{2\sigma^2}\left(-2 y_n^T V^T x_n + y_n^T V^T V y_n\right)\right)$$

$$\frac{\partial}{\partial V} \log p(y, x \mid V, \sigma) = \sum \left(-\frac{1}{2\sigma^2}\left(-2 x_n y_n^T + \right.\right.$$

$$\left.\left. + 2 V y_n y_n^T\right)\right) = 0$$

$$\mathbb{E}_{q(y)} \left\{ \sum_{n=1}^{m} \left(-\frac{1}{2\sigma^2}\left(-2 x_n y_n^T + 2 V y_n y_n^T\right)\right) \right\} = 0$$

$$\sum_{n=1}^{m} \mathbb{E}_{q(y)} (-x_n y_n)^T + \mathbb{E}_{q(y)} (V y_n y_n^T) = 0$$

$$\sum_{n=1}^{N} \mathbb{E}_{q(y_n)} x_n y_n^T = V \left[\sum_{n=1}^{N} \mathbb{E}_{q(y_n)} y_n y_n^T\right] =$$

$$V = \left(\sum_{n=1}^{N} \mathbb{E}_{q(y_n)} x_n y_n^T\right)\left(\sum_{n=1}^{N} \mathbb{E}_{q} y_n y_n^T\right)^{-1}$$

$$V = \left(\sum_{n=1}^{N} x_n \mathbb{E}_{q(y_n)} y_n^T\right)\left(\sum \mathbb{E}_{q(y_n)} y_n y_n^T\right)^{-1}$$

Из в-шага:

$$q(y_n) = \mathcal{N}(y_n \mid \mu_n, \Sigma_n)$$

$$\mathbb{E}_{q(y_n)} y_n = \mu_n$$

$$\mathbb{E}_{q(y_n)} y_n y_n^T = \Sigma_n + \mu_n \mu_n^T$$

$$\frac{\partial}{\partial \sigma}: \mathbb{E}_{q(y_n)} y_n^T V^T V y_n = \mathbb{E}_{q(y_n)} \operatorname{tr}(y_n^T V^T V y_n) =$$

$$= \operatorname{tr}(\mathbb{E}_{q(y_n)} y_n y_n^T V^T V)$$