Теория 2. Матричные вычисления

Ямаков Дмитрий Спартакович, 417 группа

Задание 1

Покажем по определению, что

$$(A+UCV)(A^{-1}-A^{-1}U(C^{-1}+VA^{-1}U)^{-1}VA^{-1})=I$$

$$(A+UCV)(A^{-1}-A^{-1}U(C^{-1}+VA^{-1}U)^{-1}VA^{-1}) =$$
$$= I - U(C^{-1}+VA^{-1}U)^{-1}VA^{-1} + UCVA^{-1} -$$
$$- UCVA^{-1}U(C^{-1}+VA^{-1}U)^{-1}VA^{-1} = -(U +$$
$$+ UCVA^{-1}U)(C^{-1}+VA^{-1}U)^{-1}VA^{-1} + I + UCVA^{-1} =$$
$$= I + UCVA^{-1} - UC(C^{-1}+VA^{-1}U)(C^{-1} +$$
$$+ VA^{-1}U)^{-1}VA^{-1} = I + UCVA^{-1} - UCVA^{-1} = I$$ – чтд

Задание 4

$$\varphi(X) = \det(X^{-1}+A)$$

выводилось на семинаре

$$g(Y) = \det Y, \quad dg(Y) = \det Y \langle Y^{-T}, dY\rangle$$

$$f(Z) = Z^{-1}+A, \quad df(Z) = -Z^{-1}dZ\,Z^{-1}$$

$$d\varphi(X) = \det(X^{-1}+A)\langle (X^{-1}+A)^{-T}, -X^{-1}dX\,X^{-1}\rangle =$$
$$= -\det(X^{-1}+A)\,\mathrm{tr}[(X^{-1}+A)^{-1}X^{-1}dX\,X^{-1}] =$$
$$= -\det(X^{-1}+A)\,\mathrm{tr}[X^{-1}(X^{-1}+A)^{-1}X^{-1}dX] =$$

$$= -\det(X^{-1}+A)\,\mathrm{tr}\left[(X(X^{-1}+A)X)^{-1}\,dX\right] =$$

$$= -\det(X^{-1}+A)\,\mathrm{tr}\left[(X+XAX)^{-1}\,dX\right] =$$

$$= -\det(X^{-1}+A)\,\langle (X+XAX)^{-T}, dX\rangle$$

$$\Rightarrow \nabla_X \varphi(X) = -\det(X^{-1}+A)\,[X+XAX]^{-T}$$

Задание 5

$$\varphi(X) = \mathrm{tr}(AX^{-T}BXC) = \langle I, AX^{-T}BXC\rangle$$

$$d\varphi(X) = \langle I, d(AX^{-T}B)\cdot XC + AX^{-T}B\,d(XC)\rangle =$$

$$= \langle I, A[dX^{-T}]BXC + AX^{-T}B\,dX\cdot C\rangle =$$

$$= \{ d(X^{-1})^T = (dX^{-1})^T = (-X^{-1}dX\,X^{-1})^T \} =$$

$$= \langle I, A(-X^{-1}dX\,X^{-1})^T BXC + AX^{-T}B\,dX\,C\rangle =$$

$$= \mathrm{tr}(-A(X^{-1}dX\,X^{-1})^T BXC) + \mathrm{tr}(AX^{-T}B\,dX\,C) =$$

$$= -\mathrm{tr}(AX^{-T}(dX)^T X^{-T}BXC) + \mathrm{tr}(CAX^{-T}B\,dX) =$$

$$= -\mathrm{tr}(C^T X^T B^T X^{-1}dX\,X^{-1}A^T) + \mathrm{tr}(CAX^{-T}B\,dX) =$$

$$= -\mathrm{tr}(X^{-1}A^T C^T X^T B^T X^{-1}dX) + \mathrm{tr}(CAX^{-T}B\,dX)$$

$$\Rightarrow \nabla_X \varphi(X) = -X^{-T}BXCAX^{-T} + B^T X^{-1}A^T C^T$$

Задание 3

$p(x) = \mathcal{N}(x \mid \mu, \Sigma)$, $p(y \mid x) \overset{(*)}{=} \mathcal{N}(y \mid Ax, \Gamma)$

Д-ть, что $p(y) = \mathcal{N}(y \mid A\mu, \Gamma + A\Sigma A^T)$

▲ $p(y) = \int p(x,y)\,dx = \int p(y \mid x)\,p(x)\,dx$

Под интегралом две параболы: одна зависит от "x", другая от "y". Квадратичную часть, завислагаемую от "y" можно вынести из-под интеграла, а оставшаяся часть будет также норм. распределением, интеграл которого будет равен 1

$\Rightarrow \quad p(y) = \mathcal{N}(y \mid \mathbb{E}y, \mathbb{D}y)$

Из $(*)$ $\Rightarrow$ $y = Ax + \xi, \ \xi \sim \mathcal{N}(0, \Gamma)$

Тогда $\mathbb{E}y = \mathbb{E}(Ax) + \mathbb{E}\xi = A\mu + 0 = A\mu$

$$\mathbb{D}y = \mathbb{D}(Ax) + \mathbb{D}\xi + \underbrace{\operatorname{cov}(Ax, \xi)}_{=0,\ \text{т.к. независимы}} = A\mathbb{D}x A^T + \Gamma =$$

$$= A\Sigma A^T + \Gamma$$

$\Rightarrow p(y) \sim \mathcal{N}(y \mid A\mu, \Gamma + A\Sigma A^T)$ — ч.т.д. ■

$\boxed{\text{Задание 2}}$

$p(x) = \mathcal{N}(x \mid \mu, \Sigma)$, $p(y|x) = \mathcal{N}(y \mid Ax, \Gamma)$

$p(x|y)$ — ?

$$p(x|y) = \frac{p(y|x)p(x)}{\int p(y|x)p(x)} = \{\text{сопр. распр}\} = \mathcal{N}(x \mid a, P)$$

Мат. ожидание норм. распределения = моде норм. распределения $= x_{MP}$

$$x_{MP} = \arg\max_x p(x|y) = \arg\max_x p(y|x)p(x) =$$

$$= \arg\max_x \left[\log p(y|x) + \log p(x)\right] = \arg\max_x \left[-\frac{1}{2}(y - Ax)^T \cdot \right.$$

$$\left. \cdot \Gamma^{-1}(y - Ax) - \frac{1}{2}(x-\mu)^T \Sigma^{-1}(x-\mu)\right] = \arg\max_x \left[-\frac{1}{2}y^T\Gamma^{-1}y + \right.$$

$$+ (Ax)^T\Gamma^{-1}y - \frac{1}{2}(Ax)^T\Gamma^{-1}Ax - \frac{1}{2}x^T\Sigma^{-1}x + \mu^T\Sigma^{-1}x -$$

$$\left. - \frac{1}{2}\mu^T\Sigma^{-1}\mu\right] \ (*) \ \Big| \ \frac{\partial}{\partial x}$$

$$A^T\Gamma^{-1}y - A^T\Gamma^{-1}Ax_{MP} - \Sigma^{-1}x_{MP} + \Sigma^{-1}\mu = 0$$

$$\left(A^T\Gamma^{-1}A + \Sigma^{-1}\right)x_{MP} = A^T\Gamma^{-1}y + \Sigma^{-1}\mu$$

$$x_{MP} = \left(A^T\Gamma^{-1}A + \Sigma^{-1}\right)^{-1}\left(A^T\Gamma^{-1}y + \Sigma^{-1}\mu\right)$$

В выражении (*) под argmax выделим квадратичную часть:

$$-\frac{1}{2}(Ax)^T\Gamma^{-1}Ax - \frac{1}{2}x^T\Sigma^{-1}x = -\frac{1}{2}x^T\left(A^T\Gamma^{-1}A + \Sigma^{-1}\right)x$$

$$\Rightarrow P = \mathbb{D}[x|y] = \left(A^T\Gamma^{-1}A + \Sigma^{-1}\right)^{-1}$$

<u>Ответ:</u> $p(x|y) = \mathcal{N}\left(x \,\middle|\, \left(A^T\Gamma^{-1}A + \Sigma^{-1}\right)^{-1}\left(A^T\Gamma^{-1}y + \Sigma^{-1}\mu\right), \left(A^T\Gamma^{-1}A + \Sigma^{-1}\right)^{-1}\right)$