# Теоретическое задание 2

Королев Кирилл

28 октября 2023 г.

## Задание 1

Доказать тождество Вудбери:

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

Проверим по определению.

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

Правая обратная матрица является также и обратной левой и совпадает просто с обратной матрицей, которая единственна.

## Задание 2

Пусть $p(x) = \mathcal{N}(x \mid \mu, \Sigma),\ p(y \mid x) = \mathcal{N}(y \mid Ax, \Gamma),\ A \in \mathbb{R}^{mxn}$.

Для нормального правдоподобия нормальный prior будет сопряжен с нормальным апостериорным, осталось понять, как пересчитать его параметры.

$$p(x \mid y) = \frac{p(y \mid x)p(x)}{C} \propto \exp{(-\frac{1}{2}[(y - Ax)^t\Gamma^{-1}(y - Ax) + (x - \mu)^t\Sigma^{-1}(x - \mu)])}$$

По функциональному виду это перевернутая парабола от x под экспонентной, значит квадратичная форма будет обратной к матрице ковариации, а вектор при линейном члене будет ожиданием. Рассмотрим выражение под экспонентой без -1/2 и опустим члены не зависящие от x.

$$(y - Ax)^t\Gamma^{-1}(y - Ax) + (x - \mu)^t\Sigma^{-1}(x - \mu) = x^t(A^t\Gamma^{-1}A + \Sigma^{-1})x - 2x^t(A^t\Gamma^{-1}y + \Sigma^{-1}\mu) + const$$

Видим сразу, что матрица ковариации получится $R = (A^t\Gamma^{-1}A + \Sigma^{-1})^{-1} = \Sigma - \Sigma A^t(\Gamma + A\Sigma A^t)^{-1}A\Sigma$. Найдем ожидание, рассмотрев линейный член.

$$-2x^t(A^t\Gamma^{-1}y + \Sigma^{-1}\mu) = -2x^tR^{-1}R(A^t\Gamma^{-1}y + \Sigma^{-1}\mu) \Rightarrow$$

$$m = (A^t\Gamma^{-1}A + \Sigma^{-1})^{-1}(A^t\Gamma^{-1}y + \Sigma^{-1}\mu)$$

В итоге апостериорное распределение имеет вид

$$p(x \mid y) = \mathcal{N}(x \mid (A^t\Gamma^{-1}A + \Sigma^{-1})^{-1}(A^t\Gamma^{-1}y + \Sigma^{-1}\mu), (A^t\Gamma^{-1}A + \Sigma^{-1})^{-1})$$

## Задание 3

Пусть $p(x) = \mathcal{N}(x \mid \mu, \Sigma),\ p(y \mid x) = \mathcal{N}(y \mid Ax, \Gamma),\ A \in \mathbb{R}^{mxn}$.

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

Под интегралом стоит произведение нормальных. Если вынесем зависящее от y из интеграла, умножим и разделим на соответствующие экспоненты и константы, то под интегралом будет стоять плотность нормального по $x$, которая проинтегрируется в 1, и то что умножали и делили поможет собрать плотность нормального по $y$. Так что делаем вывод, что $y$ имеет нормальное распределение, но пойдем по пути попроще. Сделаем reparametrization trick

$$x = \mu + \Sigma^{1/2}\epsilon \quad \epsilon \sim \mathcal{N}(0, I)$$

$$y = Ax + \Gamma^{1/2}\hat{\epsilon} \quad \hat{\epsilon} \sim \mathcal{N}(0, I)$$

$$\Rightarrow y = A(\mu + \Sigma^{1/2}\epsilon) + \Gamma^{1/2}\hat{\epsilon}$$

Посчитаем ожидание и дисперсию и таким образом найдем параметры. $\epsilon$ и $\hat{\epsilon}$ независимы.

$$\mathbb{E}y = \mathbb{E}[A(\mu + \Sigma^{1/2}\epsilon) + \Gamma^{1/2}\hat{\epsilon}] = A\mu$$

$$\mathbb{D}y = \mathbb{D}[A(\mu + \Sigma^{1/2}\epsilon) + \Gamma^{1/2}\hat{\epsilon}] = \mathbb{D}[A\Sigma^{1/2}\epsilon] + \mathbb{D}[\Gamma^{1/2}\hat{\epsilon}] = A\Sigma^{1/2}\mathbb{D}[\epsilon](\Sigma^{1/2})^t A^t + \Gamma^{1/2}\mathbb{D}[\hat{\epsilon}](\Gamma^{1/2})^t =$$
$$= A\Sigma A^t + \Gamma$$

Таким образом, $p(y) = \mathcal{N}(y \mid A\mu, A\Sigma A^t + \Gamma)$.

## Задание 4

Найти $\dfrac{\partial}{\partial X} det(X^{-1} + A)$ (все матрицы не являются симметричными)

Известный дифференциал определителя.

$$d(det(X))[H] = det(X)\langle X^{-t}, H\rangle$$

Известный дифференциал обратной матрицы.

$$d(X^{-1})[H] = -X^{-1}HX^{-1}$$

Ищем дифференциал сложной функции.

$$d(det(X^{-1} + A))[H] = det(X^{-1} + A)\langle (X^{-1} + A)^{-t}, -X^{-1}HX^{-1}\rangle = -det(X^{-1} + A)\langle X^{-t}(X^{-1} + A)^{-t}X^{-t}, H\rangle$$

Следовательно, производная имеет вид

$$\frac{\partial}{\partial X} det(X^{-1} + A) = -det(X^{-1} + A)X^{-t}(X^{-1} + A)^{-t}X^{-t}$$

## Задание 5

Найти $\frac{\partial}{\partial X} tr(AX^{-t}BXC)$ (все матрицы не являются симметричными, матрицы $A, C$ не являются квадратными.

Воспользуемся линейностью следа и дифференциала, а так же свойством циклической перестановки под следом.

$$d(tr(AX^{-t}BXC)) = tr(d(AX^{-t}BXC)) = tr(Ad(X^{-t}BX)C) = tr(A[-X^{-t}(dX^t)X^{-t}BX + X^{-t}BdX]C) =$$

$$= tr((dX)^t X^{-t}BXCAX^{-t}) + tr(CAX^{-t}BdX) = \langle X^{-t}BXCAX^{-t} + B^tX^{-1}A^tC^t, dX\rangle$$

$$\frac{\partial}{\partial X} tr(AX^{-t}BXC) = X^{-t}BXCAX^{-t} + B^tX^{-1}A^tC^t$$