Теоретическое домашнее задание по матричным вычислениям

Солоткий Михаил, 417 группа ВМК МГУ

14 октября 2018 г.

1. Доказать тождество Вудбери:

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

Здесь $A \in \mathbb{R}^{n\times n}$, $C \in \mathbb{R}^{m\times m}$, $U \in \mathbb{R}^{n\times m}$, $V \in \mathbb{R}^{m\times n}$.

Решение:
Т.к. $A + UCV$ – квадратная невырожденная матрица, из равенства $(A + UCV)B = I$ следует, что $B = (A + UCV)^{-1}$
$(A+UCV)(A^{-1}-A^{-1}U(C^{-1}+VA^{-1}U)^{-1}VA^{-1}) = I+UCVA^{-1}-U(I+CVA^{-1}U)(C^{-1}+VA^{-1}U)^{-1}VA^{-1} =$
$= I+UCVA^{-1}-UC(C^{-1}+VA^{-1}U)(C^{-1}+VA^{-1}U)^{-1}VA^{-1} = = I+UCVA^{-1}-UCVA^{-1} = I$

2. Пусть $p(\boldsymbol{x}) = \mathcal{N}(\boldsymbol{x}|\boldsymbol{\mu}, \Sigma)$, $p(\boldsymbol{y}|\boldsymbol{x}) = \mathcal{N}(\boldsymbol{y}|A\boldsymbol{x}, \Gamma)$, $A \in \mathbb{R}^{m\times n}$. Найти распределение $p(\boldsymbol{x}|\boldsymbol{y})$.

Решение:
$p(\boldsymbol{x}|\boldsymbol{y}) = \dfrac{p(\boldsymbol{y}|\boldsymbol{x})p(\boldsymbol{x})}{\int p(\boldsymbol{y}|\boldsymbol{x})p(\boldsymbol{x})d\boldsymbol{x}}$
$p(\boldsymbol{y}|\boldsymbol{x})p(\boldsymbol{x}) = \dfrac{1}{(2\pi)^{\frac{m}{2}}\sqrt{\det(\Gamma)}}\exp\left(\dfrac{1}{2}(\boldsymbol{y}-A\boldsymbol{x})^T\Gamma^{-1}(\boldsymbol{y}-A\boldsymbol{x})\right)\ \dfrac{1}{(2\pi)^{\frac{n}{2}}\sqrt{\det(\Sigma)}}\exp\left(\dfrac{1}{2}(\boldsymbol{x}-\boldsymbol{\mu})^T\Sigma^{-1}(\boldsymbol{x}-\boldsymbol{\mu})\right)$
Распишем показатель экспоненты числителя:

$$\frac{1}{2}\Big(\boldsymbol{y}^T\Gamma^{-1}\boldsymbol{y} + \boldsymbol{\mu}^T\Sigma^{-1}\boldsymbol{\mu} - 2\boldsymbol{y}^T\Gamma^{-1}A\boldsymbol{x} - 2\boldsymbol{\mu}^T\Sigma^{-1}\boldsymbol{x} + \boldsymbol{x}^TA^T\Gamma^{-1}A\boldsymbol{x} + \boldsymbol{x}^T\Sigma^{-1}\boldsymbol{x}\Big) =$$

$$= \frac{1}{2}\Big(\boldsymbol{x}^T(A^T\Gamma^{-1}A+\Sigma^{-1})\boldsymbol{x} - 2(\boldsymbol{y}^T\Gamma^{-1}A + \boldsymbol{\mu}^T\Sigma^{-1})\boldsymbol{x} + const(\boldsymbol{x})\Big)$$

Отсюда можно выделить полный квадрат. В знаменателе стоит $p(\boldsymbol{y})$. Так как в правдоподобии мат. ожидание зависит от апроирного распределения, а ковариационная матрица не зависит, в результате умножения и интегрирования по $\boldsymbol{x}$ функциональный класс останется тот же (отрицательно определённая квадратичная форма под экспонентой). В результате деления числителя на знаменатель в формуле Байеса получится функция, которая по переменной $\boldsymbol{x}$ представляет из себя отрицательную квадратичную форму под экспонентой. Формально даже выделение полного квадрата можно до конца не проводить, главное – понять, какие множители рядом с $\boldsymbol{x}$ и квадратичной частью $\boldsymbol{x}$. Так как матрицы - обобщение чисел, можно воспользоваться формулами для одномерного случая, зная, что в многомерном случае деление заменится на взятие обратной матрицы, а $\boldsymbol{x}^2$ – на $\boldsymbol{x}^T\boldsymbol{x}$.

$$\frac{1}{2}ax^2 - bx = \frac{1}{2}a\Big(x^2 - 2\frac{b}{a}x\Big) = \frac{1}{2}a\Big(x-\frac{b}{a}\Big)^2 - \frac{1}{2}a\frac{b^2}{a^2}$$
$$p(\boldsymbol{x}|\boldsymbol{y}) = \mathcal{N}\Big(\boldsymbol{x}|(A^T\Gamma^{-1}A+\Sigma^{-1})^{-1}(\boldsymbol{y}^T\Gamma^{-1}A+\boldsymbol{\mu}^T\Sigma^{-1}), (A^T\Gamma^{-1}A+\Sigma^{-1})^{-1}\Big)$$

3. Вычислить $\mathbb{E}_{\mathcal{N}(\boldsymbol{x}|\boldsymbol{\mu},\Sigma)}(\boldsymbol{x}-\boldsymbol{a})^TB(\boldsymbol{x}-\boldsymbol{a})$;

Решение:
$\mathbb{E}_{\mathcal{N}(\boldsymbol{x}|\boldsymbol{\mu},\Sigma)}(\boldsymbol{x}-\boldsymbol{a})^TB(\boldsymbol{x}-\boldsymbol{a}) = \mathbb{E}_{\mathcal{N}(\boldsymbol{x}|\boldsymbol{\mu},\Sigma)}\Big[\boldsymbol{x}^TB\boldsymbol{x} - 2\boldsymbol{x}^TBa + a^TBa\Big]$

$\mathbb{E}_{\mathcal{N}(\boldsymbol{x}|\boldsymbol{\mu},\Sigma)}\boldsymbol{x}^T B\boldsymbol{x} = \mathbb{E}_{\mathcal{N}(\boldsymbol{x}|\boldsymbol{\mu},\Sigma)}\sum\limits_{i=1}^n\sum\limits_{j=1}^n x_i B_{ij} x_j = \Big\{cov(x_i, x_j) = \mathbb{E}(x_i x_j) - \mathbb{E}x_i \mathbb{E}x_j\Big\} = \sum\limits_{i=1}^n\sum\limits_{j=1}^n (\Sigma_{ij} + \mu_i\mu_j)B_{ij} =$
$= \mathrm{tr}\Big((\Sigma + \boldsymbol{\mu}\boldsymbol{\mu}^T)^T B\Big) = \mathrm{tr}\Big((\Sigma + \boldsymbol{\mu}\boldsymbol{\mu}^T)B\Big)$
$\mathbb{E}_{\mathcal{N}(\boldsymbol{x}|\boldsymbol{\mu},\Sigma)} 2\boldsymbol{x}^T Ba = 2\boldsymbol{\mu}^T Ba$
$\mathbb{E}_{\mathcal{N}(\boldsymbol{x}|\boldsymbol{\mu},\Sigma)}(\boldsymbol{x} - \boldsymbol{a})^T B(\boldsymbol{x} - \boldsymbol{a}) = \mathrm{tr}\Big((\Sigma + \boldsymbol{\mu}\boldsymbol{\mu}^T)B\Big) + 2\boldsymbol{\mu}^T Ba$

4. Вычислить $\frac{\partial}{\partial X}\det(X^{-1} + A)$ (все матрицы не являются симметричными);

   Решение:
   $X^{-1} + A = X^{-1}(I + XA)$
   $\det(X^{-1} + A) = \det(X^{-1})\det(I + XA)$
   $\frac{\partial}{\partial X_{ij}}\det(X) = \sum\limits_{k=1}^n \frac{\partial x_{kj}}{\partial X_{ij}}\hat{A}_{jk}$, где $\hat{A}$ – матрица алгебраических дополнений матрицы X.
   $X^{-1}_{ij} = \frac{1}{\det(X)}\hat{A}_{ji}$
   $\frac{\partial}{\partial X_{ij}}\det(X) = det(X)X^{-1}_{ji}$
   $\frac{\partial}{\partial X_{ij}}\det(I + XA) = \sum\limits_{k=1}^n\sum\limits_{l=1}^n \frac{\partial\det(I+XA)}{\partial(I+XA)_{kl}}\frac{\partial(I+XA)_{kl}}{\partial X_{ij}}$
   $\frac{\partial\det(I+XA)}{\partial(I+XA)} = \det(I + XA)(I + XA)^{-T}$
   $\frac{\partial(I+XA)_{kl}}{\partial X_{ij}} = 0$, если $i \neq k$
   $(I + XA)_{il} = (I_{il} + \sum\limits_{p=1}^n X_{ip}A_{pl})$
   $\frac{\partial(I+XA)_{il}}{\partial X_{ij}} = A_{jl}$
   $\frac{\partial}{\partial X_{ij}}\det(I + XA) = A(I + XA)^{-1}$
   $\frac{\partial}{\partial X_{ij}}\Big[\det(X)\det(I + XA)\Big] = \dfrac{1}{\det(X)^2}\det(X)\det(I + XA)X^{-T} + \det(X)\det(I + XA)A(I + XA)^{-1}$