Теория 2. Матричные вычисления

Дятлова Дарья

1. Доказать тождество Вудбери:

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

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

Доказательство:

(a) по определению, если $A^{-1} = B$, то $A \cdot A^{-1} = A \cdot B = I$;

(b) $(A+UCV)\cdot(A^{-1} - A^{-1}U(C^{-1}+VA^{-1}U)^{-1}VA^{-1}) =$
$I + UCVA^{-1} - (U + UCVA^{-1}U)\cdot(C^{-1}+VA^{-1}U)^{-1}VA^{-1} =$
$I + UCVA^{-1} - (UC)(C^{-1}+VA^{-1}U)(C^{-1}+VA^{-1}U)^{-1}VA^{-1} =$
$I + UCVA^{-1} - (UC)VA^{-1} = I.$

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: $p(\boldsymbol{x}|\boldsymbol{y}) = \frac{p(y|x)p(x)}{\int p(y|x)p(x)d(x)}$

$\boldsymbol{p(y|x)p(x)} = \int \frac{\exp((y-Ax)^T(y-Ax)\cdot(2\Gamma)^{-1})}{(2\pi)^{\frac{m}{2}}|\Gamma|^{\frac{1}{2}}} \cdot \frac{\exp((x-\mu)^T(x-\mu)\cdot(2\Sigma)^{-1})}{(2\pi)^{\frac{n}{2}}|\Sigma|^{\frac{1}{2}}} =$
$\frac{\exp((y-Ax)^T(y-Ax)\cdot(2\Gamma)^{-1}+(x-\mu)^T(x-\mu)\cdot(2\Sigma)^{-1})}{(2\pi)^{\frac{m}{2}}|\Gamma|^{\frac{1}{2}}\cdot 2\pi^{\frac{n}{2}}|\Sigma|^{\frac{1}{2}}} =$

раскроем скобки в показателе экспоненты:

$\frac{exp(y^Ty\cdot(2\Gamma)^{-1} - y^TAx\cdot(2\Gamma)^{-1} - (Ax)^Ty\cdot(2\Gamma)^{-1} - (Ax)^TAx\cdot(2\Gamma)^{-1} + x^Tx\cdot(2\Sigma)^{-1} - x^T\mu\cdot(2\Sigma)^{-1} - \mu^Tx\cdot(2\Sigma)^{-1} + \mu^T\mu\cdot(2\Sigma)^{-1})}{(2\pi)^{\frac{m}{2}}|\Gamma|^{\frac{1}{2}}\cdot 2\pi^{\frac{n}{2}}|\Sigma|^{\frac{1}{2}}} =$

сгруппируем подобные слагаемые в показателе экспоненты:

$\frac{exp(2^{-1}\cdot(\boldsymbol{x^T}(\Sigma^{-1} - A^TA\Gamma^{-1})\boldsymbol{x} - \boldsymbol{2x}(y^TA\Gamma^{-1}+\mu\Sigma^{-1}) + y^Ty\Gamma^{-1} + \mu^T\mu\Sigma^{-1}))}{(2\pi)^{\frac{m}{2}}|\Gamma|^{\frac{1}{2}}\cdot 2\pi^{\frac{n}{2}}|\Sigma|^{\frac{1}{2}}};$

проведем замену переменных: пусть $H = \Sigma^{-1} - A^TA\Gamma^{-1}, K = y^TA\Gamma^{-1} + \mu\Sigma^{-1}, C = y^Ty\Gamma^{-1} + \mu^T\mu\Sigma^{-1}$, тогда интеграл в знаменателе можно переписать следующим образом:

$\int p(y|x)p(x)d(x) = 2^{-1}\int \exp(x^THx - 2xK + C)dx =$
можем воспользоваться формулой для вычисления интеграла плотности гауссианы:
$= e^C \cdot (\frac{(2\pi)^n}{|H|})^{\frac{1}{2}} \cdot e^{\frac{1}{2}K^TH^{-1}K} =$
$= e^{(y^Ty\Gamma^{-1}+\mu^T\mu\Sigma^{-1})} \cdot (\frac{(2\pi)^n}{|\Sigma^{-1}-A^TA\Gamma^{-1}|})^{\frac{1}{2}} \cdot e^{\frac{1}{2}(y^TA\Gamma^{-1}+\mu\Sigma^{-1})^T(\Sigma^{-1}-A^TA\Gamma^{-1})^{-1}(y^TA\Gamma^{-1}+\mu\Sigma^{-1})}$, обозначим получившуюся нормировочную константу $\boldsymbol{z}$.

в числителе выделим полный квадрат под экспонентой, получим:
$\mu' = (\Sigma^{-1} - A^T\Gamma^{-1}A)^{-1}\cdot(y^TA\Gamma^{-1} + \mu\Sigma^{-1}), \Sigma' = (\Sigma^{-1} - A^T\Gamma^{-1}A)^{-1} \Rightarrow$
$\Rightarrow$p$(\boldsymbol{x}|\boldsymbol{y}) = N(x|\mu',\Sigma') = N(x|(\Sigma^{-1} - A^T\Gamma^{-1}A)^{-1}\cdot(y^TA\Gamma^{-1} + \mu\Sigma^{-1}), (\Sigma^{-1} - A^T\Gamma^{-1}A)^{-1}).$

3. Пусть $p(\boldsymbol{x}) = \mathcal{N}(\boldsymbol{x}|\boldsymbol{\mu},\Sigma) =$, $p(\boldsymbol{y}|\boldsymbol{x}) = \mathcal{N}(\boldsymbol{y}|A\boldsymbol{x},\Gamma)$. Доказать, что $p(\boldsymbol{y}) = \mathcal{N}(\boldsymbol{y}|A\boldsymbol{\mu},\Gamma + A\Sigma A^T)$.

$p(\boldsymbol{y}) = p(y|x)\cdot p(x)$ – два нормальных распределения, которые между собой сопрягаются. Найдем дисперсию и математическое ожидание, чтобы восстановить плотность для y.

$y = Ax + \epsilon$, где $\epsilon\ N(0,\Gamma)$.
$Ey = A\mu$
$Dy = [(y-A\mu)(y-A\mu)^T] = E[(A(x-\mu)+\epsilon)(A(x-\mu)+\epsilon)^T] = A\cdot E((x-\mu)(x-\mu)^T)^T\cdot A^T + E(\epsilon\epsilon^T) = A\Sigma A^T + \Gamma.$

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

   $d(det(X^{-1}+A)) = det(X^{-1}+A) < (X^{-1}+A)^{-T}, d(X^{-1}+A) >=$
   $= det(X^{-1}+A) < (X^{-1}+A)^{-T}, -X^{-1}(dX)X^{-1} >=$
   $= det(X^{-1}+A) < -X^{-T}(X^{-1}+A)^{-T}(dX)X^{-T}) >.$

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

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

   Пусть размерности матриц: $A : [m,n]; X, B : [n,n]; C : [n,m]$.
   $\frac{\partial}{\partial X}\operatorname{tr}(AX^{-T}BXC) = \frac{\partial}{\partial X}\operatorname{tr}(A^TC^T, X^{-T}BX) =< A^TC^T, d(X^{-T}BX) >$

   $d(X^{-T}BX) = d(X^{-T}B)X + X^{-T}B(dX) = d(X^{-T})BX + X^{-T}B(dX) = -X^{-T}(dX^T)X^{-T}BX + X^{-T}B(dX) =$
   $= -X^{-T}((dx^T)x^{-T}BX - (dx^TB^T)^T) = -X^{-T}((X^TB^TX^{-1}(dx))^T - B(dx))$

   $\frac{\partial}{\partial X}\operatorname{tr}(AX^{-T}BXC) =< A^TC^T, -X^{-T}(X^TB^TX^{-1}(dx))^T > + < A^TC^T, +X^{-T}B(dx) >=$
   $=< A^TC^TX^TB^TX^{-1}dx, -X^{-T} > + < B^TX^{-1}A^TC^T, (dx) >=$
   $< -X^{-T}XBCAX^{-T}, dx > + < B^TX^{-1}A^TC^T, (dx) >=< B^TX^{-1}A^TC^T - X^{-T}XBCAX^{-T}, dx >=$
   $= B^TX^{-1}A^TC^T - X^{-T}XBCAX^{-T}.$