Московский Государственный Университет

Матричные вычисления

Выполнил: Курцев Д.В.

Группа: 417

Факультет Вычислительной математики и кибернетики

Кафедра Математических методов прогнозирования

Октябрь 2022

Задача 1

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

Воспользуемся Теоремой. Если A - квадратная матрица и $AB = I$, то $B = A^{-1}$.

Несложно увидеть, что $(A+UCV) \subset \mathbb{R}^{n\times n}$ Т.е. достаточно доказать, что $(A+UCV)(A^{-1} - A^{-1}U(C^{-1} + VA^{-1}U)^{-1}VA^{-1}) = I$.

Обозначим $M = U(C^{-1} + VA^{-1}U)^{-1}$

$(A+UCV)(A^{-1} - A^{-1}MVA^{-1}) = I - MVA^{-1} + UCVA^{-1} - UCVA^{-1}MVA^{-1}$

Таким образом, достаточно показать, что $UCVA^{-1} = MVA^{-1} + UCVA^{-1}MVA^{-1}$

Но $MVA^{-1} + UCVA^{-1}MVA^{-1} = (I + UCVA^{-1})MVA^{-1}$

То есть, достаточно показать, что $UC = (I + UCVA^{-1})M$

$(I + UCVA^{-1})M = (I + UCVA^{-1})(U(C^{-1} + VA^{-1}U)^{-1}) = (U + UCVA^{-1}U)(C^{-1}+VA^{-1}U)^{-1} = U(I+CVA^{-1}U)(C^{-1}+VA^{-1}U)^{-1} = UC(C^{-1} + VA^{-1}U)(C^{-1} + VA^{-1}U)^{-1} = UC$.

Ч.т.д.

Задача 2

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

$$p(x|y) = \frac{p(y|x)p(x)}{\int p(y|x)p(x)dx} = \frac{1}{Z}p(y|x)p(x)$$

Так как $p(x)$ и $p(y|x)$ оказываются сопряжёнными,то $p(x|y)$ тоже будет из нормального распределения $\Rightarrow p(x|y) \sim \mathcal{N}(x|m, C)$. Найдём его мат ожиданние и дисперсию:

$$p(y|x)p(x) = \frac{1}{(2\pi)^{\frac{m}{2}}} \frac{1}{\sqrt{\det \Gamma}} \frac{1}{(2\pi)^{\frac{n}{2}}} \frac{1}{\sqrt{\det \Sigma}} \exp(-\frac{1}{2}(x-\mu)^T \Sigma^{-1}(x-\mu)-$$

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

$$+ y^T\Gamma^{-1}y - 2x^T A^T \Gamma^{-1} y + x^T A^T \Gamma^{-1} Ax)) =$$

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

$$+ y^T\Gamma^{-1}y))$$

Коэффициент при $x^T x$ соответствует обратной ковариационной матрице $\Rightarrow \quad C = (\Sigma^{-1} + A^T\Gamma^{-1}A)^{-1}$

Так как мат ожидание нормальной случайной величины совпадает с её модой, то достаточно найти аргмаксимум плотности. Для этого приравняем производную выражения, стоящего под экспонентой, к нулю. ($\arg\max_x e^{f(x)} = \arg\max_x f(x)$).

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

$$= -(\Sigma^{-1} + A^T\Gamma^{-1}A)x + \Sigma^{-1}\mu + A^T\Gamma^{-1}y = 0$$

$$\Rightarrow \quad m = (\Sigma^{-1} + A^T\Gamma^{-1}A)^{-1}(\Sigma^{-1}\mu + A^T\Gamma^{-1}y) = C(\Sigma^{-1}\mu + A^T\Gamma^{-1}y)$$

Таким образом получим:

$$p(x|y) \sim \mathcal{N}(x|C(\Sigma^{-1}\mu + A^T\Gamma^{-1}y), C), C = (\Sigma^{-1} + A^T\Gamma^{-1}A)^{-1}$$

Задача 3

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

$p(y) \sim \mathcal{N}(y|A\mu, \Gamma + A\Sigma A^T) - ?$

$$p(y) = \int\limits_{\mathbb{R}} p(x, y)dx = \int\limits_{\mathbb{R}} p(y|x)p(x)dx$$

Так как $p(x)$ и $p(y|x)$ оба из нормального распределения, то в интеграле будут стоять экспоненты, которые квадратично зависят от x и y. После перемножения их аргументы сложатся, следовательно останется квадратичная зависимость от y. Таким образом, можно утвержать, что $p(y)$ также из нормального распределения.

Тогда из вида $p(y|x)$ при фиксированном x, y можно представить следующим образом: $y = Ax + \varepsilon$, $\varepsilon \sim \mathcal{N}(0, \Gamma)$. Тогда для $p(y)$ x является случайной величиной. Получим в силу независимости x и ε:

$$\mathbb{E}\, y = \mathbb{E}\, Ax + \mathbb{E}\, \varepsilon = A\mathbb{E}\, x + 0 = A\mu$$

$$\mathbb{D}\, y = \mathbb{D}\, Ax + \mathbb{D}\, \varepsilon = A\mathbb{D}x\, A^T + \Gamma = A\Sigma A^T + \Gamma$$

Задача 4

$f(X) = \det(X^{-1} + A)$. Будем пользоваться стандартными формулами дифференциалов:

$dtr(X) = tr(dX);\ d\det X = \det X tr(X^{-1}dX);\ dX^{-1} = -X^{-1}dXX^{-1}$

$$df(x) = d\det(X^{-1}+A) = \det(X^{-1}+A)\,tr[(X^{-1}+A)^{-1}d(X^{-1}+A)] =$$

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

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

$$= -\det(X^{-1}+A)\,tr[X^{-1}(X^{-1}+A)^{-1}X^{-1}dX] \;\;\Rightarrow$$

$$\Rightarrow\;\; \nabla f(x) = -\det(X^{-1}+A)\;X^{-T}(X^{-1}+A)^{-T}X^{-T}$$

Задача 5

$f(X) = tr(AX^{-T}BXC)$

$$df(x) = d\,tr(AX^{-T}BXC) = tr\,d(AX^{-T}BXC) = tr[AdX^{-T}\,BXC+$$

$$+AX^{-T}BdX\,C] = trAX^{-T}BdX\,C - trAX^{-T}(dX)^TX^{-T}BXC =$$

$$= tr\,CAX^{-T}BdX - tr\,X^{-T}BXCAX^{-T}(dX)^T =$$

$$\langle (CAX^{-T}B)^T,\ dX\rangle - \langle X^{-T}BXCAX^{-T},\ dX\rangle \;\;\Rightarrow$$

$$\Rightarrow\;\; \nabla f(x) = B^TX^{-1}A^TC^T - X^{-T}BXCAX^{-T}$$