Домашнее задание №2

Иванов Вячеслав, группа 699

6 октября 2018 г.

Оглавление

1 Найти градиент по U и по V функции $J(\mathbf{U}, \mathbf{V}) := \|\mathbf{UV} - \mathbf{Y}\|_F^2 + \frac{\lambda}{2}(\|\mathbf{U}\|_F^2 + \|\mathbf{V}\|_F^2)$

Решение.

Определение. $f : \mathbb{R}^{n\times m} \to \mathbb{R} \implies \frac{df(\mathbf{X})}{d\mathbf{X}} := \left[\frac{\partial f}{dx_{ij}}(\mathbf{X})\right]_{\substack{1\leqslant i\leqslant n\\ 1\leqslant j\leqslant m}}$

Утверждение 1. $\|\mathbf{X}\|_F^2 = \mathrm{tr}(\mathbf{X}^T\mathbf{X})$

Утверждение 2. $\frac{d\ \mathrm{tr}(\mathbf{X}^T\mathbf{AX})}{d\mathbf{X}} = (\mathbf{A} + \mathbf{A}^T)\mathbf{X}$, в частности $\frac{d\|\mathbf{X}\|_F^2}{d\mathbf{X}} = 2\mathbf{X}$, $\frac{d\|\mathbf{XY}\|_F^2}{d\mathbf{Y}} = 2\mathbf{X}^T\mathbf{XY}$.

Утверждение 3 (Свойства следа)**.**

- $\mathrm{tr}(A_1A_2\ldots A_n) = \mathrm{tr}(A_nA_1\ldots A_{n-1})$ (в частности, $\mathrm{tr}(AB) = \mathrm{tr}(BA)$)
- $\mathrm{tr}(A^TA) = \mathrm{tr}(AA^T)$

Утверждение 4.

$$\frac{d\ \mathrm{tr}(\mathbf{AXB})}{d\mathbf{X}} = \mathbf{A}^T\mathbf{B}^T$$

Доказательство. Доказано на семинаре от 11.09.2018. □

Доказательство.

$$\left[\mathbf{X}^T\mathbf{AX}\right]_{lm} = \sum_{i=1}^{n} x_{il}\sum_{j=1}^{n} a_{ij}x_{jm} = \sum_{i=1}^{n}\sum_{j=1}^{n} x_{il}a_{ij}x_{jm}$$

$$\mathrm{tr}(\mathbf{X}^T\mathbf{AX}) = \sum_{l=1}^{n}\left[\mathbf{X}^T\mathbf{AX}\right]_{ll} = \sum_{l=1}^{n}\sum_{i=1}^{n}\sum_{j=1}^{n} x_{il}a_{ij}x_{jl}$$

$$\frac{d\ \mathrm{tr}(\mathbf{X}^T\mathbf{AX})}{dx_{lm}} = \left(\sum_{i=1}^{n}\sum_{j=1}^{n} x_{im}a_{ij}x_{jm}\right)'_{lm} = \sum_{j=1}^{n} a_{lj}x_{jm} + \sum_{i=1}^{n} x_{im}a_{il} = \sum_{i=1}^{n}(a_{li} + a_{il})x_{im} = \left[(A + A^T)X\right]_{lm}$$

□

1.1 $\frac{dJ(\mathbf{U},\mathbf{V})}{d\mathbf{V}}$

Во первых, по утверждению 2: $\frac{d\|\mathbf{V}\|_F^2}{d\mathbf{V}} = 2V$.
Во вторых, по свойствам следа:

$$\begin{aligned}\|\mathbf{UV} - \mathbf{Y}\|_F^2 &= \mathrm{tr}((\mathbf{UV} - \mathbf{Y})^T(\mathbf{UV} - \mathbf{Y})) = \\ &= \|\mathbf{UV}\|_F^2 + \|\mathbf{Y}\|_F^2 - \mathrm{tr}(\mathbf{Y}^T\mathbf{UV}) - \mathrm{tr}(\mathbf{V}^T\mathbf{U}^T\mathbf{Y}) = \\ &= \|\mathbf{UV}\|_F^2 + \|\mathbf{Y}\|_F^2 - 2\cdot\mathrm{tr}(\mathbf{Y}^T\mathbf{UV})\end{aligned}$$

Откуда по утверждению 2 получаем, что:

$$\frac{d\|\mathbf{UV}\|_F^2}{d\mathbf{V}} = 2\mathbf{U}^T\mathbf{UV}$$

Осталось посчитать $\frac{d\mathrm{tr}(\mathbf{Y}^T\mathbf{UV})}{d\mathbf{V}}$. По свойствам следа, $\mathrm{tr}(\mathbf{Y}^T\mathbf{UV}) = \mathrm{tr}(\mathbf{UVY}^T)$. Отсюда по утверждению 3 сразу получаем, что:

$$\frac{d\mathrm{tr}(\mathbf{Y}^T\mathbf{UV})}{d\mathbf{V}} = \mathbf{U}^T\mathbf{Y}$$

Собирая промежуточные результаты в итоговое выражение, получаем, что:

$$\frac{dJ(\mathbf{U},\mathbf{V})}{d\mathbf{V}} = 2\mathbf{U}^T(\mathbf{UV} - \mathbf{Y}) + \lambda\mathbf{V}$$

1.2 $\frac{dJ(\mathbf{U},\mathbf{V})}{d\mathbf{U}}$

И вновь, по утверждению 2: $\frac{d\|\mathbf{U}\|_F^2}{d\mathbf{U}} = 2U$.
По утверждению 3 сразу получаем, что:

$$\frac{d\mathrm{tr}(\mathbf{Y}^T\mathbf{U}\mathbf{V})}{d\mathbf{U}} = \mathbf{Y}\mathbf{V}^{\ T}$$

Из свойств следа (из второго, если быть точным) и утверждения 2 получаем, что:

$$\|\mathbf{U}\mathbf{V}\|_F^2 = \mathrm{tr}(\mathbf{V}^T\mathbf{U}^T\mathbf{U}\mathbf{V}) = \mathrm{tr}(\mathbf{U}\mathbf{V}\mathbf{V}^T\mathbf{U}^T) = \|\mathbf{V}^T\mathbf{U}^T\|_F^2$$
$$\implies \frac{d\|\mathbf{U}\mathbf{V}\|_F^2}{d\mathbf{U}} = \left(\frac{d\|\mathbf{V}^T\mathbf{U}^T\|_F^2}{d\mathbf{U}^T}\right)^T = (2\mathbf{V}\mathbf{V}^T\mathbf{U}^T)^T = 2\mathbf{U}\mathbf{V}\mathbf{V}^T$$

Собирая промежуточные результаты в итоговое выражение, получаем, что:

$$\frac{dJ(\mathbf{U},\mathbf{V})}{d\mathbf{U}} = 2(\mathbf{U}\mathbf{V} - \mathbf{Y})\mathbf{V}^T + \lambda\mathbf{U}$$

□

2 Найти градиент и гессиан функции $f(\mathbf{w}) = \sum_{i=1}^m \log(1 + e^{-y_i\mathbf{w}^T\mathbf{x}_i})$, где $x_i \in \mathbb{R}^n,\ y_i \in \mathbb{R}$

Решение.

$$\frac{\partial f}{\partial w_j}(\mathbf{w}) = \sum_{i=1}^m \frac{-y_i x_{ij} e^{-y_i\mathbf{w}^T\mathbf{x}_i}}{1 + e^{-y_i\mathbf{w}^T\mathbf{x}_i}}$$

Обозначим $\frac{-y_i e^{-y_i\mathbf{w}^T\mathbf{x}_i}}{1+e^{-y_i\mathbf{w}^T\mathbf{x}_i}}$ через $\theta_i(\mathbf{w})$. Тогда $\triangledown f(\mathbf{w}) = \mathbf{X}^T\theta$, где $\mathbf{X} \in \mathbb{R}^{m\times n}$ — матрица, строками которой являются векторы $\mathbf{x}_i$.
Далее,

$$\frac{\partial^2 f}{\partial w_i w_j}(\mathbf{w}) = \sum_{k=1}^m \left(\frac{-y_k x_{kj} e^{-y_k\mathbf{w}^T\mathbf{x}_k}}{1 + e^{-y_k\mathbf{w}^T\mathbf{x}_k}}\right)'_i = \sum_{k=1}^m \left(\frac{y_k^2 e^{-y_k\mathbf{w}^T\mathbf{x}_k}}{1 + e^{-y_k\mathbf{w}^T\mathbf{x}_k}} - \left(\frac{-y_k e^{-y_k\mathbf{w}^T\mathbf{x}_k}}{1 + e^{-y_k\mathbf{w}^T\mathbf{x}_k}}\right)^2\right) x_{ki}x_{kj} =$$
$$= \sum_{k=1}^m \frac{y_k^2 e^{-y_k\mathbf{w}^T\mathbf{x}_k}}{1 + e^{-y_k\mathbf{w}^T\mathbf{x}_k}}\left(1 - \frac{e^{-y_k\mathbf{w}^T\mathbf{x}_k}}{1 + e^{-y_k\mathbf{w}^T\mathbf{x}_k}}\right) x_{ki}x_{kj} =$$
$$= \sum_{k=1}^m \frac{y_k^2 e^{-y_k\mathbf{w}^T\mathbf{x}_k}}{(1 + e^{-y_k\mathbf{w}^T\mathbf{x}_k})^2} x_{ki}x_{kj}$$

Если положить $\psi_k(\mathbf{w}) := \frac{y_k^2 e^{-y_k\mathbf{w}^T\mathbf{x}_k}}{(1+e^{-y_k\mathbf{w}^T\mathbf{x}_k})^2}$, то в матричном виде выражение можно переписать так:

$$\mathbf{H}(f)(\mathbf{w}) = \mathbf{X}^T\mathrm{diag}(\psi_1(\mathbf{w}), \ldots, \psi_m(\mathbf{w}))\mathbf{X}$$

. □

3 Найти матрицу Якоби для функции $f : \mathbb{R}^n \to \mathbb{R}^n,\ f_j(\mathbf{w}) := \frac{e^{w_j}}{\sum_{k=1}^n e^{w_k}}$

Утверждение 5.

$$\mathbf{J}(\mathbf{f})(\mathbf{w}) = \mathrm{diag}(f_1, \ldots f_n)(\mathbf{w}) - \mathbf{f}(\mathbf{w})(\mathbf{f}(\mathbf{w}))^T$$

Доказательство.

$$\frac{\partial \mathbf{f}_l}{\partial \mathbf{w}_l}(\mathbf{w}) = \frac{e^{w_l}}{\sum_{k=1}^n e^{w_k}} - \frac{e^{2w_l}}{\left(\sum_{k=1}^n e^{w_k}\right)^2} = \frac{e^{w_l}(\sum_{k\neq l} e^{w_k})}{(\sum_{k=1}^n e^{w_k})^2}$$
$$\frac{\partial \mathbf{f}_l}{\partial \mathbf{w}_m}(\mathbf{w}) = -\frac{e^{w_l+w_m}}{\left(\sum_{k=1}^n e^{w_k}\right)^2},\ l \neq m$$

В матричной записи:

$$\mathbf{J}(\mathbf{f})(\mathbf{w}) = \mathbf{f}(\mathbf{w})(\mathbf{f}(\mathbf{w}))^T - \operatorname{diag}(f_1, \dots f_n)(\mathbf{w})$$

□

4 Доказать тождество Шермана-Моррисона-Вудбери

Утверждение 6.

$$(\mathbf{A} + \mathbf{UCV})^{-1} = \mathbf{A}^{-1} - \mathbf{A}^{-1}\mathbf{U}(\mathbf{C}^{-1} + \mathbf{V}\mathbf{A}^{-1}\mathbf{U})^{-1}\mathbf{V}\mathbf{A}^{-1}$$
$$\mathbf{A} \in \mathbb{R}^{n\times n},\ \mathbf{U} \in \mathbb{R}^{n\times m},\ \mathbf{C} \in \mathbb{R}^{m\times m},\ \mathbf{V} \in \mathbb{R}^{m\times n}$$

Доказательство.

$$\begin{aligned}
&(\mathbf{A} + \mathbf{UCV})\left(\mathbf{A}^{-1} - \mathbf{A}^{-1}\mathbf{U}(\mathbf{C}^{-1} + \mathbf{V}\mathbf{A}^{-1}\mathbf{U})^{-1}\mathbf{V}\mathbf{A}^{-1}\right) = \\
&= (\mathbf{E} + \mathbf{UCVA}^{-1})\left(\mathbf{E} - \mathbf{U}(\mathbf{C}^{-1} + \mathbf{V}\mathbf{A}^{-1}\mathbf{U})^{-1}\mathbf{V}\mathbf{A}^{-1}\right) = \\
&= \mathbf{E} + \mathbf{UCVA}^{-1} - (\mathbf{E} + \mathbf{UCVA}^{-1})\mathbf{U}\left(\mathbf{C}^{-1} + \mathbf{V}\mathbf{A}^{-1}\mathbf{U}\right)^{-1}\mathbf{V}\mathbf{A}^{-1} = \\
&= \mathbf{E} + \mathbf{UCVA}^{-1} - \mathbf{UC}(\mathbf{C}^{-1} + \mathbf{VA}^{-1}\mathbf{U})\left(\mathbf{C}^{-1} + \mathbf{V}\mathbf{A}^{-1}\mathbf{U}\right)^{-1}\mathbf{V}\mathbf{A}^{-1} = \\
&= \mathbf{E}
\end{aligned}$$

$$\begin{aligned}
&\left(\mathbf{A}^{-1} - \mathbf{A}^{-1}\mathbf{U}(\mathbf{C}^{-1} + \mathbf{V}\mathbf{A}^{-1}\mathbf{U})^{-1}\mathbf{V}\mathbf{A}^{-1}\right)(\mathbf{A} + \mathbf{UCV}) = \\
&= \left(\mathbf{E} - \mathbf{A}^{-1}\mathbf{U}(\mathbf{C}^{-1} + \mathbf{V}\mathbf{A}^{-1}\mathbf{U})^{-1}\mathbf{V}\right)(\mathbf{E} + \mathbf{A}^{-1}\mathbf{UCV}) = \\
&= \mathbf{E} + \mathbf{A}^{-1}\mathbf{UCV} - \mathbf{A}^{-1}\mathbf{U}(\mathbf{C}^{-1} + \mathbf{V}\mathbf{A}^{-1}\mathbf{U})^{-1}\mathbf{V}(\mathbf{E} + \mathbf{A}^{-1}\mathbf{UCV}) = \\
&= \mathbf{E} + \mathbf{A}^{-1}\mathbf{UCV} - \mathbf{A}^{-1}\mathbf{U}(\mathbf{C}^{-1} + \mathbf{V}\mathbf{A}^{-1}\mathbf{U})^{-1}(\mathbf{C}^{-1} + \mathbf{VA}^{-1}\mathbf{U})\mathbf{CV} = \\
&= \mathbf{E}
\end{aligned}$$

□

5 Найти градиент отношения Релея для заданной матрицы $\mathbf{A} \in \mathbf{S}^n$. Исследовать связь между отношением Релея и задачей поиска собственного вектора, соответствующего максимальному собственному значению.

Определение. *Отношением Релея* для заданной матрицы $A \in S^n$ называют выражение вида:

$$R(\mathbf{x}|\mathbf{A}) := \frac{\mathbf{x}^T\mathbf{A}\mathbf{x}}{\mathbf{x}^T\mathbf{x}}$$

Утверждение 7. Собственный вектор, соответствующий максимальному собственному значению матрицы A, максимизирует отношение Релея для этой матрицы.

Доказательство.
Перепишем отношение Релея в бескоординатном виде через скалярное произведение:

$$R(\mathbf{x}|\mathbf{A}) = \frac{(\mathbf{x}, \mathbf{Ax})}{(\mathbf{x}, \mathbf{x})} = \frac{(\mathbf{x}, \mathbf{Ax})}{\|\mathbf{x}\|_2^2} = \left(\frac{\mathbf{x}}{\|\mathbf{x}\|_2}, \mathbf{A}\frac{\mathbf{x}}{\|\mathbf{x}\|_2}\right)$$

Из такой записи видно, что отношение Релея есть $(\mathbf{y}, \mathbf{Ay})$ для $\mathbf{y}$, лежащих на единичной сфере. На семинаре обсуждалось, что для операторов, порождающих ортонормированный базис из собственных векторов (к таковым относятся эрмитовые, подмножеством которых является S^n), данный функционал на единичной сфере достигает максимума на собственном векторе, принадлежащем максимальному собственному значению. □

Утверждение 8. $\nabla R(\mathbf{x}|\mathbf{A}) = (\mathbf{A} - R(\mathbf{x}|\mathbf{A})\mathbf{E})\frac{2\mathbf{x}}{\mathbf{x}^T\mathbf{x}}$

Доказательство.

$$\nabla\left(\frac{\mathbf{x}^T\mathbf{Ax}}{\mathbf{x}^T\mathbf{x}}\right) = \frac{\nabla(\mathbf{x}^T\mathbf{Ax})\mathbf{x}^T\mathbf{x} - (\mathbf{x}^T\mathbf{Ax})\nabla(\mathbf{x}^T\mathbf{x})}{(\mathbf{x}^T\mathbf{x})^2} = \frac{2\mathbf{Ax}(\mathbf{x}^T\mathbf{x}) - \mathbf{x}^T\mathbf{Ax}(2\mathbf{x})}{(\mathbf{x}^T\mathbf{x})^2} =$$
$$= 2\frac{(\mathbf{x}^T\mathbf{x})\mathbf{A} - \mathbf{x}^T\mathbf{Ax}}{(\mathbf{x}^T\mathbf{x})^2}\mathbf{x} = \frac{2}{\mathbf{x}^T\mathbf{x}}\left(\mathbf{A} - \frac{\mathbf{x}^TA\mathbf{x}}{\mathbf{x}^T\mathbf{x}}\mathbf{E}\right)\mathbf{x} = (\mathbf{A} - R(\mathbf{x}|\mathbf{A})\mathbf{E})\frac{2\mathbf{x}}{\mathbf{x}^T\mathbf{x}}$$

□

6 Найти градиенты функций от собственных значений матрицы:

6.1 $f(\mathbf{X}) := \sum_{i=1}^n \lambda_i(\mathbf{X})$

Утверждение 9. След матрицы инвариантен относительно смены базиса.

Доказательство. Пусть C — матрица перехода. По свойству следа:

$$\mathrm{tr}(\mathbf{CXC}^{-1}) = \mathrm{tr}(\mathbf{C}^{-1}\mathbf{CX}) = \mathrm{tr}(\mathbf{X})$$

□

Утверждение 10. $f(\mathbf{X}) = \mathrm{tr}(\mathbf{X})$.

Доказательство.
Равенство верно по определению, если $\mathbf{X}$ приведена к Жордановой форме.
Оно является тождеством в силу предыдущего утверждения. □

Утверждение 11. $\nabla f(\mathbf{X}) = \mathbf{E}$

Доказательство.
Тривиально следует из определения матричной производной скалярной функции. □

6.2 $f(\mathbf{X}) := \prod_{i=1}^n \lambda_i(\mathbf{X})$

Утверждение 12. Определитель матрицы инвариантен относительно смены базиса.

Доказательство. Тривиально следует из мультипликативности определителя и выражения для определителя обратной матрицы. □

Утверждение 13. $f(\mathbf{X}) = \det(\mathbf{X})$.

Доказательство.
Равенство верно по определению, если $\mathbf{X}$ приведена к Жордановой форме.
Оно является тождеством в силу предыдущего утверждения. □

Утверждение 14.

$$\nabla \det(\mathbf{X}) = \mathrm{C}(\mathbf{X})$$

где $\mathrm{C}(\mathbf{X}) = [(-1)^{i+j} M_{ij}]_{\substack{1\leqslant i\leqslant n\\ 1\leqslant j\leqslant m}}$, а M_{ij} — дополнительный минор для позиции x_{ij}.

Доказательство. Доказано на семинаре. □

7 Найти градиент логарифма функции правдоподобия для многомерного нормального распределения по вектору средних значений и по матрице ковариаций

Определение. Функция правдоподобия для многомерного нормального распределения:

$$p(\mathbf{X}|\mu, \Sigma) = \prod_{i=1}^{N} \frac{1}{(2\pi)^{n/2}\sqrt{\det \Sigma}} \exp\left(-\frac{1}{2}(\mathbf{x}_i - \mu)^T \Sigma^{-1}(\mathbf{x}_i - \mu)\right),\ \mathbf{x}_i \in \mathbb{R}^n$$

Утверждение 15. $\frac{\partial}{\partial \mathbf{A}} \mathrm{tr}(\mathbf{A}^{-1}\mathbf{B}) = -(\mathbf{A}^{-1}\mathbf{B}\mathbf{A}^{-1})^T$

Утверждение 16. $\frac{\partial \log p(\mathbf{X}|\mu,\Sigma)}{\partial \mu} = \sum_{i=1}^{N} \Sigma^{-1}(\mathbf{x}_i - \mu)$

Доказательство.

$$\begin{aligned}
\frac{\partial \log p(\mathbf{X}|\mu, \Sigma)}{\partial \mu} &= \frac{\partial}{\partial \mu} \sum_{i=1}^{N} \log\left[\frac{1}{(2\pi)^{n/2}\sqrt{\det \Sigma}} \exp\left(-\frac{1}{2}(\mathbf{x}_i - \mu)^T \Sigma^{-1}(\mathbf{x}_i - \mu)\right)\right] = \\
&= \frac{\partial}{\partial \mu}\left(-\frac{1}{2}\sum_{i=1}^{n}(\mathbf{x}_i - \mu)^T \Sigma^{-1}(\mathbf{x}_i - \mu)\right) = \\
&= -\frac{1}{2}\frac{\partial}{\partial \mu}\sum_{i=1}^{n}\left(\mathbf{x}_i^T\Sigma^{-1}\mathbf{x}_i - \mu^T\Sigma^{-1}\mathbf{x}_i - \mathbf{x}_i^T\Sigma^{-1}\mu + \mu^T\Sigma^{-1}\mu\right) = \\
&= \sum_{i=1}^{N}\Sigma^{-1}(\mathbf{x}_i - \mu)
\end{aligned}$$

□

Утверждение 17. $\frac{\partial \log p(\mathbf{X}|\mu,\Sigma)}{\partial \Sigma} = -\frac{1}{2}(N\Sigma^{-1} - \Sigma^{-1}(\mathbf{x}_i - \mu)(\mathbf{x}_i - \mu)^T\Sigma^{-1})$

Доказательство.

$$\begin{aligned}
\frac{\partial \log p(\mathbf{X}|\mu, \Sigma)}{\partial \Sigma} &= \frac{\partial}{\partial \Sigma} \sum_{i=1}^{N} \log\left[\frac{1}{(2\pi)^{n/2}\sqrt{\det \Sigma}} \exp\left(-\frac{1}{2}(\mathbf{x}_i - \mu)^T \Sigma^{-1}(\mathbf{x}_i - \mu)\right)\right] = \\
&= \frac{\partial}{\partial \Sigma}\left(-\frac{N}{2}\log\det\Sigma - \frac{1}{2}\sum_{i=1}^{n}(\mathbf{x}_i - \mu)^T \Sigma^{-1}(\mathbf{x}_i - \mu)\right)
\end{aligned}$$

Перепишем сумму таким образом, чтобы можно было воспользоваться утверждением 15:

$$(\mathbf{x}_i - \mu)^T\Sigma^{-1}(\mathbf{x}_i - \mu) = \mathrm{tr}\left(\Sigma^{-1}(\mathbf{x}_i - \mu)(\mathbf{x}_i - \mu)^T\right)$$

Тогда:

$$\begin{aligned}
\frac{\partial}{\partial \Sigma}(\mathbf{x}_i - \mu)^T\Sigma^{-1}(\mathbf{x}_i - \mu) &= -(\Sigma^{-1}(\mathbf{x}_i - \mu)(\mathbf{x}_i - \mu)^T\Sigma^{-1})^T = \\
&= -\Sigma^{-1}(\mathbf{x}_i - \mu)(\mathbf{x}_i - \mu)^T\Sigma^{-1}
\end{aligned}$$

Т.к. матрица ковариаций симметрична.
Также, по правилу дифференцирования сложной функции получаем :

$$\frac{d}{d\Sigma}\log\det\Sigma = \frac{1}{\det\Sigma}\frac{d\det\Sigma}{d\Sigma} = \frac{C(\Sigma)}{\det\Sigma} = \frac{(\det\Sigma)\Sigma^{-T}}{\det\Sigma} = \Sigma^{-T} = \Sigma^{-1}$$

Итого получаем, что:

$$\frac{\partial\log p(\mathbf{X}|\mu,\Sigma)}{\partial\Sigma} = -\frac{1}{2}(N\Sigma^{-1} - \Sigma^{-1}(\mathbf{x}_i-\mu)(\mathbf{x}_i-\mu)^T\Sigma^{-1})$$

$\square$