Второе задание

Хохлов Алексей

9 марта 2019 г.

1 Проекция на шар

Если $\|\mathbf{a}\|_2 \leqslant 1$, то, очевидно, проекция $\mathbf{p} = \mathbf{a}$.

Если же $\|\mathbf{a}\|_2 > 1$, очевидно, что искомая точка — точка пересечения прямой, соединяющей центр сферы и $\mathbf{a}$, с поверхностью единичной сферы. Другими словами, искомая точка $\mathbf{p} = \dfrac{\mathbf{a}}{\|\mathbf{a}\|_2}$. Чтобы доказать это, покажем, что $\forall \mathbf{x} \in \mathbf{X}$ верно $\|\mathbf{a} - \mathbf{x}\|_2 \geqslant \|\mathbf{a} - \mathbf{p}\|_2$.

Для правой стороны неравенства

$$\|\mathbf{a} - \mathbf{p}\|_2 = \|\mathbf{a} - \frac{\mathbf{a}}{\|\mathbf{a}\|_2}\|_2 = (1 - \frac{1}{\|\mathbf{a}\|_2})\|\mathbf{a}\|_2 = \|\mathbf{a}\|_2 - 1 \tag{1.1}$$

Для левой стороны неравенства

$$\|\mathbf{a} - \mathbf{x}\|_2 \geqslant \|\mathbf{a}\|_2 - \|\mathbf{x}\|_2 \geqslant \|\mathbf{a}\|_2 - 1 \tag{1.2}$$

Видим, что для любой точки в единичном шаре точка $\mathbf{p} = \dfrac{\mathbf{a}}{\|\mathbf{a}\|_2}$ является решением задачи

$$\|\mathbf{a} - \mathbf{x}\|_2 \to \min \tag{1.3}$$

2 Минимизация скалярного произведения

Искомая точка

$$p_i = \begin{cases} 1,\, i = k \\ 0,\, i \neq k \end{cases} \tag{2.1}$$

$$\tag{2.2}$$

где k — такая, что $c_k = \min(c_1, c_2, ..., c_n)$. Докажем это.

Поскольку $c_k \leqslant c_i$ для любого i, то $c_k x_i \leqslant c_i x_i$. Значит, что

$$\langle \mathbf{c}, \mathbf{p} \rangle = c_k = c_k \sum_{i=1}^{n} x_i = \sum_{i=1}^{n} c_k x_i \leqslant \sum_{i=1}^{n} c_i x_i = \langle \mathbf{c}, \mathbf{x} \rangle \tag{2.3}$$

3 Максимизация функции правдоподобия

В решении использованы формулы

$$\frac{\partial}{\partial \mathbf{A}}(\mathrm{tr}(\mathbf{A}^{-1}\mathbf{B})) = -(\mathbf{A}^{-1}\mathbf{B}\mathbf{A}^{-1})^{\mathrm{T}} \tag{3.1}$$

$$\frac{\partial}{\partial \mathbf{A}}\det(\mathbf{A}) = \det(\mathbf{A})\mathbf{A}^{\text{-T}} \tag{3.2}$$

$$\mathrm{tr}(\mathbf{ABC}) = \mathrm{tr}(\mathbf{BCA}) \tag{3.3}$$

3.1

Разложим функцию на слагаемые

$$f = \frac{m}{2}\log\det\Sigma + \frac{1}{2}\sum_{i=1}^{m}(\mathbf{x}_i - \boldsymbol{\mu})^{\mathrm{T}}\Sigma^{-1}(\mathbf{x}_i - \boldsymbol{\mu}) = f_1 + f_2 \tag{3.4}$$

Найдем производные слагаемых по Σ.
Для первого слагаемого воспользуемся (3.2)

$$\frac{\partial}{\partial\Sigma}f_1 = \frac{m}{2}\frac{1}{\det\Sigma}\det(\Sigma)\Sigma^{-\mathrm{T}} = \frac{m}{2}\Sigma^{-\mathrm{T}} \tag{3.5}$$

Для второго слагаемого воспользуемся (3.3) и (3.1).

$$\begin{aligned}\frac{\partial}{\partial\Sigma}f_2 &= \frac{1}{2}\frac{\partial}{\partial\Sigma}\mathrm{tr}(\sum_{i=1}^{m}(\mathbf{x}_i - \boldsymbol{\mu})^{\mathrm{T}}\Sigma^{-1}(\mathbf{x}_i - \boldsymbol{\mu})) = \frac{1}{2}\frac{\partial}{\partial\Sigma}\mathrm{tr}(\Sigma^{-1}\sum_{i=1}^{m}(\mathbf{x}_i - \boldsymbol{\mu})(\mathbf{x}_i - \boldsymbol{\mu})^{\mathrm{T}}) = \\ &= -\frac{1}{2}\left(\Sigma^{-1}(\sum_{i=1}^{m}(\mathbf{x}_i - \boldsymbol{\mu})(\mathbf{x}_i - \boldsymbol{\mu})^{\mathrm{T}})\Sigma^{-1}\right)^{\mathrm{T}} = -\frac{1}{2}\Sigma^{-T}\left(\sum_{i=1}^{m}(\mathbf{x}_i - \boldsymbol{\mu})(\mathbf{x}_i - \boldsymbol{\mu})^{\mathrm{T}}\right)^{\mathrm{T}}\Sigma^{-T}\end{aligned} \tag{3.6}$$

Приравняем производную функции к нулю

$$\frac{m}{2}\Sigma^{-\mathrm{T}} - \frac{1}{2}\Sigma^{-T}\left(\sum_{i=1}^{m}(\mathbf{x}_i - \boldsymbol{\mu})(\mathbf{x}_i - \boldsymbol{\mu})^{\mathrm{T}}\right)^{\mathrm{T}}\Sigma^{-T} = 0 \tag{3.7}$$

Умножив на Σ и транспонировав, получим

$$\Sigma = \frac{1}{m}\sum_{i=1}^{m}(\mathbf{x}_i - \boldsymbol{\mu})(\mathbf{x}_i - \boldsymbol{\mu})^{\mathrm{T}} \tag{3.8}$$

3.2

Докажем формулу (3.1). Во-первых, из $\mathbf{A}\mathbf{A}^{-1} = \mathbf{E}$ следует

$$\frac{\partial}{\partial a_{pq}}\sum_{j=1}^{n}\tilde{a}_{ij}a_{jk} = \sum_{j=1}^{n}\frac{\partial\tilde{a}_{ij}}{\partial a_{pq}}a_{jk} + \sum_{j=1}^{n}\tilde{a}_{ij}\frac{\partial a_{jk}}{\partial a_{pq}} = 0 \tag{3.9}$$

Домножая справа на обратную матрицу, получим

$$\frac{\partial\tilde{a}_{ij}}{\partial a_{pq}} = -\sum_{k=1}^{n}\sum_{m=1}^{n}\tilde{a}_{im}\frac{\partial a_{mk}}{\partial a_{pq}}\tilde{a}_{kj} = -\tilde{a}_{ip}\tilde{a}_{qj} \tag{3.10}$$

Раскроем теперь производную следа

$$\frac{\partial}{\partial a_{pq}}(\mathrm{tr}(\mathbf{A}^{-1}\mathbf{B})) = \frac{\partial}{\partial a_{pq}}\sum_{i=1}^{n}\sum_{j=1}^{n}\tilde{a}_{ij}b_{ji} = \sum_{i=1}^{n}\sum_{j=1}^{n}\frac{\partial\tilde{a}_{ij}}{\partial a_{pq}}b_{ji} \tag{3.11}$$

Подставляя значение производной из (3.10)

$$\frac{\partial}{\partial a_{pq}}(\mathrm{tr}(\mathbf{A}^{-1}\mathbf{B})) = -\sum_{i=1}^{n}\sum_{j=1}^{n}\tilde{a}_{ip}\tilde{a}_{qj}b_{ji} = -\sum_{i=1}^{n}\sum_{j=1}^{n}\tilde{a}^{\mathrm{T}}_{pi}b^{\mathrm{T}}_{ij}\tilde{a}^{\mathrm{T}}_{jq} \tag{3.12}$$

получив в матричном виде искомую формулу (3.1).

4

4.1

4.2

4.3

4.4

5

5.1

5.2

5.3

6

7