HW 1 Ноутбук 1.

Задача 1.

Задача 1. Свойства ММП оценок для модели нормального распределения (2 балла)

Рассмотрим следующую модель из независимых векторных наблюдений $X_1, \ldots, X_n$:

$$X_i \overset{iid}{\sim} \mathcal{N}(\mu, \Sigma).$$

Постройте и исследуйте оценку для вектора матожиданий μ и ковариационной матрицы Σ методом максимального правдоподобия.

В помощь: Если дифференцировать функцию $f : \mathbb{R}^{m \times n} \to \mathbb{R}$ по ячейкам входной матрицы, то получится матрица из производных, обозначаемая $\nabla_A f(A)$. Условие первого порядка для оптимизации записывается так же, как и в привычных вам случаях, нужно приравнять производные к нулю. Можете использовать таблицу для получения оценок.

1. Правдоподобие:

- $$L(X) = \prod_{i=1}^{n} f(X_i)$$

пусть μ - вектор размера $d \times 1$
Σ - матрица ковариаций размером $d \times d$

Ф-ия плотности многомерного норм. р-ия:

- $$f(X_i \mid \mu, \Sigma) = \frac{1}{(2\pi)^{d/2} \cdot |\Sigma|^{1/2}} \cdot e^{\left(-\frac{1}{2} \cdot (X_i - \mu)^T \cdot \Sigma^{-1} \cdot (X_i - \mu)\right)}$$

- $$L(\mu, \Sigma) = \prod_{i=1}^{n} \left(\frac{1}{(2\pi)^{d/2} \cdot |\Sigma|^{1/2}} \cdot e^{\left(-\frac{1}{2} \cdot (X_i - \mu)^T \cdot \Sigma^{-1} \cdot (X_i - \mu)\right)} \right) =$$

$$= \left(\frac{1}{(2\pi)^{d/2} \cdot |\Sigma|^{1/2}} \right)^n \cdot e^{-1/2 \sum_{i=1}^{n} (X_i - \mu)^T \cdot \Sigma^{-1} \cdot (X_i - \mu)}$$

- Логарифм правдоподобия:

$$l(\mu, \Sigma) = \ln L(\mu, \Sigma) = \sum_{i=1}^{n} \left(-\frac{d}{2} \ln(2\pi) - \frac{1}{2} \ln|\Sigma| - \frac{1}{2} (X_i - \mu)^T \cdot \Sigma^{-1} \cdot (X_i - \mu) \right)$$

$$= -\frac{nd}{2} \ln(2\pi) - \frac{n}{2} \ln|\Sigma| - \frac{1}{2} \sum_{i=1}^{n} (X_i - \mu)^T \cdot \Sigma^{-1} \cdot (X_i - \mu)$$

2. Оценка для μ:

Дифференцируем по μ:

$$\nabla_\mu \ln L(\mu, \Sigma) = \frac{\partial}{\partial \mu} \ln L(\mu, \Sigma) = \frac{\partial}{\partial \mu} \left(-\frac{1}{2} \sum_{i=1}^{n} (X_i - \mu)^T \cdot \Sigma^{-1} \cdot (X_i - \mu) \right)$$

$=$ Σ - симметричная матрица $\Rightarrow$ Σ^{-1} - симметричная матр. $\Rightarrow$

$$\Rightarrow = -(-1) \cdot \frac{1}{2} \cdot 2 \cdot \sum_{i=1}^{n} \Sigma^{-1} \cdot (X_i - \mu) = \Sigma^{-1} \cdot \sum_{i=1}^{n} (X_i - \mu)$$

т.е. использовали формулу из таблицы:

$\mathbf{A}$ is not a function of $\mathbf{x}$ $\mathbf{A}$ is symmetric	$\frac{\partial \mathbf{x}^\top \mathbf{A}\mathbf{x}}{\partial \mathbf{x}} =$	$2\mathbf{x}^\top \mathbf{A}$	$2\mathbf{A}\mathbf{x}$

Максимизируем по μ:

$$\Sigma^{-1}\sum_{i=1}^{n}(X_i-\mu)=0 \Rightarrow n\hat{\mu} = \sum_{i=1}^{n} X_i \Rightarrow \hat{\mu}_{\text{ММП}} = \overline{X}$$

- Оценка для Σ

Дифференцируем по Σ:

$$\nabla_{\Sigma} \ln L(\mu,\Sigma) = -\frac{n}{2}\Sigma^{-1} + \frac{1}{2}\cdot\Sigma^{-1}\cdot\left(\sum_{i=1}^{n}(X_i-\mu)(X_i-\mu)^T\right)\Sigma^{-1}$$

$d\left(\mathbf{X}^{-1}\right) =$	$-\mathbf{X}^{-1}(d\mathbf{X})\,\mathbf{X}^{-1}$

Использовала 2 ф-лы из таблицы

$\mathbf{a}$ and $\mathbf{b}$ are not functions of $\mathbf{X}$	$\frac{\partial \mathbf{a}^\top \mathbf{X}\mathbf{b}}{\partial \mathbf{X}} =$	$\mathbf{b}\mathbf{a}^\top$	$\mathbf{a}\mathbf{b}^\top$

Максимизируем по Σ.

$$-\frac{n}{2}\Sigma^{-1} + \frac{1}{2}\cdot\Sigma^{-1}\cdot\left(\sum_{i=1}^{n}(X_i-\mu)(X_i-\mu)^T\right)\Sigma^{-1} \Big/ \cdot\,\Sigma$$

$$-\frac{n}{2}\cdot I_d + \frac{1}{2}\cdot\Sigma^{-1}\left(\sum_{i=1}^{n}(X_i-\mu)(X_i-\mu)^T\right)\cdot I_d = 0$$

$$n = \Sigma^{-1}\left(\sum_{i=1}^{n}(X_i-\mu)(X_i-\mu)^T\right)$$

$$n\Sigma = \sum_{i=1}^{n}(X_i-\mu)(X_i-\mu)^T$$

$$\hat{\Sigma}_{\text{ммп}} = \frac{1}{n}\sum_{i=1}^{n}(X_i-\mu)(X_i-\mu)^T$$

3. Сравнение оценок с одномерном случаем:

Знаем, что в одномерном случае: $\hat{\mu}_{\text{ммп}} = \overline{X}$

для $X_i \sim N(\mu, \sigma^2)$ $\quad \hat{\sigma}^2_{\text{ммп}} = \frac{1}{n}\sum_{i=1}^{n}(X_i-\overline{X})^2$

Видим, что в обоих случаях оценки для мат. ожид. μ – это выборочное среднее. В одномерном случае дисперсия оценивается выборочной дисперсией, а в многомерном случае оценка является суммой матричных произведений, включающей в себя ковариации между различными компонентами векторов x_i.

Задача 2.

Задача 2. Свойства ММП оценок для модели равномерного распределения (2 балла)

Рассмотрим модель выборки $X_1, .., X_n$ из наблюдений

$X_i \overset{iid}{\sim} U[0, a]$

Правдоподобие

плотность:

$$f(X_i) = \begin{cases} \frac{1}{a}, & 0 \le X_i \le a \\ 0, & \text{иначе} \end{cases}$$

$$L(a) = \prod_{i=1}^{n} f(X_i) = \frac{1}{a^n} \, Ind\{X_{(n)} \le a\}$$

$$\ln L(a) = \ln\left(\frac{1}{a^n}\right) = -n\ln a, \quad \text{при } X_{(n)} \le a$$

Максимизируем:

С увеличением a функция лог-правдоподобия убывает, поэтому, чтобы ф-я принимала макс зн-е, значение a должно быть минимально возможным, т.е.

$$\hat{a}_{ммп} = X_{(n)}$$

Задача 3.

Задача 2. Свойства ММП оценок для модели равномерного распределения (2 балла)

Рассмотрим модель выборки $X_1, .., X_n$ из наблюдений

$X_i \overset{iid}{\sim} U[0, a]$

Правдоподобие

1. Для оценки остаются п-р r - константа и п-р σ^2 – дисперсия ошибки ε_i

$\Rightarrow$ Вектор параметров ОМП : $\theta = (r, \sigma^2)$

* Рассмотрим модель авторегрессии: $X_{i+1} = \alpha X_i + \varepsilon_{i+1}$, где $\varepsilon_i \sim N(0, \sigma^2)$

Тогда правдоподобие : $X_{i+1} = \alpha X_i + \varepsilon_{i+1}$; $X_{i+1} | X_i \sim N(\alpha X_i, \sigma^2)$

$$L(\theta) = p_\theta(\hat{X}_1, \dots, \hat{X}_n) = p_\theta(\hat{X}_1) \cdot p_\theta(\hat{X}_2 | \hat{X}_1) \dots p_\theta(\hat{X}_n | \hat{X}_{n-1}, \dots, \hat{X}_1) = \prod_{i=1} p_\theta(\hat{X}_{i+1} | \hat{X}_i) =$$

$$= \prod_{i=1}^{n} p_{N(\alpha \hat{X}_i, \sigma^2)}(\hat{X}_{i+1})$$

$$L(\theta) = p_\theta(\hat{\varepsilon}_1, \dots, \hat{\varepsilon}_n) = \prod_{i=1}^{n} p_\theta(\hat{\varepsilon}_i) = \prod_{i=1}^{n} p_{N(0, \sigma^2)}(\hat{X}_{i+1} - \alpha \hat{X}_i)$$

Вернёмся к нашей модели:

Распределение T_{i+1} при данном T_i будет нормальным:

$$T_{i+1} | T_i \sim N\left((1 - r\Delta t) T_i + r T_{env} \Delta t, \sigma^2\right)$$

Тогда плотность :

$$p(T_{i+1} | T_i ; r, \sigma^2) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(T_{i+1} - (1 - r\Delta t) T_i - r T_{env} \Delta t)^2}{2\sigma^2}}$$

Функция правдоподобия :

$$L(r, \sigma^2) = \prod_{i=0}^{n-1} \frac{1}{\sqrt{2\pi\sigma^2}} \cdot e^{-\frac{(T_{i+1} - (1 - r\Delta t) T_i - r T_{env} \Delta t)^2}{2\sigma^2}}$$

начальное зн-е T_0

Лог-правдоподобие:

$$\ln L(r, \sigma^2) = -\frac{1}{2}\sum_{i=0}^{n-1}\left(\ln(2\pi\sigma^2) - \frac{(T_{i+1} - (1-r\Delta t)T_i - rT_{env}\Delta t)^2}{2\sigma^2}\right) =$$

$$= -\frac{n}{2}\ln(2\pi\sigma^2) - \frac{1}{2\sigma^2}\sum_{i=0}^{n-1}\left(T_{i+1} - (1-r\Delta t)T_i - rT_{env}\Delta t\right)^2$$

2. Оценка для r:

$$\frac{\partial \ell}{\partial r} = -\frac{1}{2\sigma^2}\sum_{i=0}^{n-1}\frac{\partial}{\partial r}\left(T_{i+1} - (1-r\Delta t)T_i - rT_{env}\Delta t\right)^2 =$$

$$= -\frac{1}{2\sigma^2}\cdot 2\cdot\sum_{i=0}^{n-1}\left(T_{i+1} - (1-r\Delta t)T_i - rT_{env}\Delta t\right)\cdot\left(-\Delta t T_i - T_{env}\cdot\Delta t\right)$$

Максимизируем:

$$\sum_{i=0}^{n-1}\left(T_{i+1} - (1-r\Delta t)T_i - rT_{env}\Delta t\right)\cdot\left(-\Delta t T_i - T_{env}\cdot\Delta t\right) = 0$$

Упростим. $-\Delta t\cdot T_i\cdot T_{i+1} - T_{i+1}\cdot T_{env}\cdot\Delta t + \Delta t\cdot T_i^2(1-r\Delta t) +$

$+ T_{env}\cdot\Delta t\cdot T_i(1-r\Delta t) + r\Delta t^2 T_i\cdot T_{env} + r\Delta t^2\cdot T_{env}^2$

$$\sum_{i=0}^{n-1}\Big(-\Delta t\cdot T_i\cdot T_{i+1} - T_{i+1}\cdot T_{env}\cdot\Delta t + \Delta t\cdot T_i^2(1-r\Delta t) +$$
$$+ T_{env}\cdot\Delta t\cdot T_i(1-r\Delta t) + r\Delta t^2 T_i\cdot T_{env} + r\Delta t^2\cdot T_{env}^2\Big) = 0$$

Выразим r:

$$r\Delta t\left(\sum_{i=0}^{n-1}T_i(T_i - T_{env}) - T_{env}\sum_{i=0}^{n-1}(T_i - T_{env})\right) = -\sum_{i=0}^{n-1}(T_{i+1} - T_i)(T_i - T_{env})$$

$$\hat{r}_{\text{ММП}} = -\frac{\sum_{i=0}^{n-1}(T_{i+1} - T_i)(T_i - T_{env})}{\Delta t\cdot\sum_{i=0}^{n-1}\underbrace{T_i(T_i - T_{env}) - T_{env}(T_i - T_{env})}_{=(T_i - T_{env})^2}}$$

Оценка для σ^2:

$$\frac{\partial l}{\partial \sigma^2} = -\frac{n}{2\sigma^2} + \frac{1}{2\sigma^4}\sum_{i=0}^{n-1}\left(T_{i+1} - (1-r\Delta t)T_i - rT_{env}\Delta t\right)^2$$

Максимизируем: $\quad (\cdot\ 2\sigma^4)$

$$\sum_{i=0}^{n-1}\left(T_{i+1} - (1-r\Delta t)T_i - rT_{env}\Delta t\right)^2 = n\cdot\sigma^2$$

$$\hat{\sigma}^2_{ммп} = \frac{1}{n}\sum_{i=0}^{n-1}\left(T_{i+1} - (1-r\Delta t)T_i - rT_{env}\Delta t\right)^2$$