Final exam, Statistical Inference II: (2016 Spring): [+32points]

Name: <u>Jia-Han Shih</u>

- Proofs must be understandable to the instructor.
- Avoid typos and undefined notations in your proofs.

Q1 [+8]. Let $X_1,..,X_m \overset{iid}{\sim} N(\xi,\sigma^2)$ and $Y_1,..,Y_n \overset{iid}{\sim} N(\xi,\tau^2)$, where σ^2 and τ^2 are known.

1) **[+2]** Derive the MLE of ξ.

<u>Answer:</u>

Since σ^2 and τ^2 are known, the likelihood function is

$$L(\xi) = \prod_{i=1}^{m} \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left\{-\frac{1}{2\sigma^2}(x_i-\xi)^2\right\} \prod_{i=1}^{n} \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left\{-\frac{1}{2\tau^2}(y_i-\xi)^2\right\}$$

$$= \left(\frac{1}{2\pi\sigma^2}\right)^{\frac{m}{2}} \left(\frac{1}{2\pi\tau^2}\right)^{\frac{n}{2}} \exp\left\{-\frac{1}{2\sigma^2}\sum_{i=1}^{m}(x_i-\xi)^2\right\} \exp\left\{-\frac{1}{2\tau^2}\sum_{i=1}^{n}(y_i-\xi)^2\right\}.$$

Then the log-likelihood function is

$$\ell(\xi) = \text{constant} - \frac{1}{2\sigma^2}\sum_{i=1}^{m}(x_i-\xi)^2 - \frac{1}{2\tau^2}\sum_{i=1}^{n}(y_i-\xi)^2$$

$$\Rightarrow \ell'(\xi) = \frac{1}{\sigma^2}\sum_{i=1}^{m}(x_i-\xi) + \frac{1}{\tau^2}\sum_{i=1}^{n}(y_i-\xi) \overset{\text{set}}{=} 0.$$

Solve the likelihood equation, we obtain the MLE of ξ is

$$\hat{\xi} = \frac{m\bar{X}/\sigma^2 + n\bar{Y}/\tau^2}{m/\sigma^2 + n/\tau^2},$$

where

$$\bar{X} = \frac{1}{m}\sum_{i=1}^{m} x_i, \qquad \bar{Y} = \frac{1}{n}\sum_{i=1}^{n} y_i .$$

2) **[+2]** Show that the above MLE is also UMVUE

Answer:

Since

$$L(\xi)=\left(\frac{1}{2\pi\sigma^2}\right)^{\frac{m}{2}}\left(\frac{1}{2\pi\tau^2}\right)^{\frac{n}{2}}\exp\left\{-\frac{1}{2\sigma^2}\sum_{i=1}^{m}(x_i-\xi)^2\right\}\exp\left\{-\frac{1}{2\tau^2}\sum_{i=1}^{n}(y_i-\xi)^2\right\}$$

$$=\left(\frac{1}{2\pi\sigma^2}\right)^{\frac{m}{2}}\left(\frac{1}{2\pi\tau^2}\right)^{\frac{n}{2}}\exp\left\{-\frac{1}{2\sigma^2}\sum_{i=1}^{m}x_i^{\,2}-\frac{1}{2\tau^2}\sum_{i=1}^{n}y_i^{\,2}\right\}$$

$$\times\exp\left\{\xi\left(\frac{1}{\sigma^2}\sum_{i=1}^{m}x_i+\frac{1}{\tau^2}\sum_{i=1}^{n}y_i\right)-\frac{\xi^2}{2\sigma^2}-\frac{\xi^2}{2\tau^2}\right\}$$

$$=\exp\{\eta T(\mathbf{x},\mathbf{y})-A(\eta)\}h(\mathbf{x},\mathbf{y}),$$

where

$$\eta=\xi,\qquad T(\mathbf{x},\mathbf{y})=\frac{1}{\sigma^2}\sum_{i=1}^{m}x_i+\frac{1}{\tau^2}\sum_{i=1}^{n}y_i=\frac{m}{\sigma^2}\bar{X}+\frac{n}{\tau^2}\bar{Y}\qquad\text{and}$$

$$h(\mathbf{x},\mathbf{y})=\left(\frac{1}{2\pi\sigma^2}\right)^{\frac{m}{2}}\left(\frac{1}{2\pi\tau^2}\right)^{\frac{n}{2}}\exp\left\{-\frac{1}{2\sigma^2}\sum_{i=1}^{m}x_i^{\,2}-\frac{1}{2\tau^2}\sum_{i=1}^{n}y_i^{\,2}\right\}.$$

Therefore, it is an one-dimensional exponential family. Since the parameter space $\Theta=\{\eta:\eta\in(-\infty,\infty)\}$ contains an one-dimensional open rectangle (e.g., $(0,1)\in\Theta$). Hence

$$T(\mathbf{x},\mathbf{y})=\frac{m}{\sigma^2}\bar{X}+\frac{n}{\tau^2}\bar{Y}$$

is the complete sufficient statistics for ξ. Furthermore, we have

$$E\left(\frac{m\bar{X}/\sigma^2+n\bar{Y}/\tau^2}{m/\sigma^2+n/\tau^2}\right)=\xi\,.$$

Thus, we have shown that $\hat{\xi}$ is unbiased and it is a function of complete sufficient statistics. Then we have proven that the above MLE $\hat{\xi}$ is also UMVUE.

3) **[+4]** Derive the asymptotic variance of the MLE under some conditions on m and n.

Answer:

Let $X_1, \cdots, X_m \sim f_1(x)$ and $Y_1, \cdots, Y_n \sim f_2(y)$. Assume that $m+n=N$ and

$$\frac{m}{N} \to \lambda_1, \qquad \frac{n}{N} \to \lambda_2 \ \text{ as } \ m, n \to \infty.$$

Then we define the appropriate information

$$I(\xi) = \sum_{\alpha=1}^{2} \lambda_\alpha I^{(\alpha)}(\xi),$$

where

$$I^{(\alpha)}(\xi) = -E_\xi\left[\frac{\partial^2}{\partial \xi^2} \log f_1(x)\right].$$

Therefore, we have

$$I^{(1)}(\xi) = -E_\xi\left[-\frac{1}{\sigma^2}\right] = \frac{1}{\sigma^2}.$$

Similarly,

$$I^{(2)}(\xi) = \frac{1}{\tau^2}.$$

Thus, we obtain

$$I(\xi) = \frac{\lambda_1}{\sigma^2} + \frac{\lambda_2}{\tau^2} \Rightarrow I^{-1}(\xi) = \frac{\sigma^2\tau^2}{\lambda_2\sigma^2 + \lambda_1\tau^2}.$$

Then the asymptotic distribution of $\hat{\xi}$ is

$$\sqrt{N}(\hat{\xi} - \xi) \xrightarrow{d} N\left(0, \frac{\sigma^2\tau^2}{\lambda_2\sigma^2 + \lambda_1\tau^2}\right).$$

Hence the asymptotic variance is

$$\operatorname{var}(\hat{\xi}) = \frac{\sigma^2\tau^2}{\lambda_2\sigma^2 + \lambda_1\tau^2}.$$

Q2 [+8]. Let $X_1,...,X_n \overset{iid}{\sim} f_{\mu,\sigma^2}(x)$, where

$$f_{\mu,\sigma^2}(x) = \sqrt{\frac{2}{\pi\sigma^2}} \exp\left\{-\frac{(x-\mu)^2}{2\sigma^2}\right\} I(x \geq \mu)$$

is a truncated normal distribution, truncated at unknown value $\mu \in R$.

1) **[+6]** Derive the MLE $\hat{\theta} = (\hat{\mu}, \hat{\sigma}^2)$.

Answer:

The likelihood function is

$$\begin{aligned} L(\boldsymbol{\theta}) &= \prod_{i=1}^{n} \sqrt{\frac{2}{\pi\sigma^2}} \exp\left\{-\frac{1}{2\sigma^2}(x_i-\mu)^2\right\} I(x_i \geq \mu) \\ &= \left(\frac{2}{\pi}\right)^{\frac{n}{2}} \left(\frac{1}{\sigma^2}\right)^{\frac{n}{2}} \exp\left\{-\frac{1}{2\sigma^2}\sum_{i=1}^{n}(x_i-\mu)^2\right\} I(x_{(1)} \geq \mu), \end{aligned}$$

where $x_{(1)} = \min(x_1, x_2, \cdots, x_n)$. The log-likelihood function is

$$\ell(\boldsymbol{\theta}) = \frac{n}{2}\log\left(\frac{2}{\pi}\right) - \frac{n}{2}\log\sigma^2 - \frac{1}{2\sigma^2}\sum_{i=1}^{n}(x_i-\mu)^2 + \log I(x_{(1)} \geq \mu).$$

First, we have

$$\frac{\partial \ell(\boldsymbol{\theta})}{\partial \mu} = \frac{1}{\sigma^2}\sum_{i=1}^{n}(x_i-\mu) \overset{\text{set}}{=} 0 \Rightarrow \hat{\mu} = \frac{1}{n}\sum_{i=1}^{n} x_i = \overline{X}, \qquad \frac{\partial^2 \ell(\boldsymbol{\theta})}{\partial^2 \mu} = -\frac{n}{\sigma^2} < 0.$$

But since $\mu \in (-\infty, x_{(1)}]$, therefore, $\hat{\mu} = x_{(1)}$. Then

$$\frac{\partial^2 \ell(\boldsymbol{\theta})}{\partial \sigma^2} = -\frac{n}{\sigma^2} + \frac{1}{2\sigma^4}\sum_{i=1}^{n}(x_i-\hat{\mu})^2 \overset{\text{set}}{=} 0 \Rightarrow \hat{\sigma}^2 = \frac{1}{n}\sum_{i=1}^{n}(x_i-\hat{\mu})^2.$$

Thus, we obtain the MLE

$$\hat{\theta} = \left(\hat{\mu} = x_{(1)}, \hat{\sigma}^2 = \frac{1}{n}\sum_{i=1}^{n}(x_i - x_{(1)})^2\right).$$

2) **[+2]** Is the asymptotic theory of MLEs apply to this example?

Answer:

No, the support of X is (μ, ∞). It is depend on the parameter μ. Therefore, the common support assumption does not hold in this example.

Q3 [+8]. Let $X_1,..,X_n \overset{iid}{\sim} N(\theta,1)$ and $p = P_\theta(X_1 \le a)$

1) [+3] Derive the UMVUE of $p = P_\theta(X_1 \le a)$ [denoted as δ_{1n}].

Answer:

We have $\bar{X} = \frac{1}{n}\sum_{i=1}^{n} x_i$ is the complete sufficient statistics and

$$X_1 - \bar{X} = \frac{n-1}{n}X_1 + \frac{1}{n}X_2 + \cdots + \frac{1}{n}X_n \sim N\left(0, \left(\frac{n-1}{n}\right)^2 + \frac{n-1}{n^2}\right) = N\left(0, \frac{n-1}{n}\right).$$

Then

$$\begin{aligned}\delta_{1n} &= E[\, I(\, X_1 \le a\,) \,|\, \bar{X}\,] = \Pr(\, X_1 \le a \,|\, \bar{X}\,) = \Pr(\, X_1 - \bar{X} \le a - \bar{X}\,) \\ &= \Pr\left(\sqrt{\frac{n}{n-1}}(\, X_1 - \bar{X}\,) \le \sqrt{\frac{n}{n-1}}(\, a - \bar{X}\,)\right) = \Phi\left(\sqrt{\frac{n}{n-1}}(\, a - \bar{X}\,)\right),\end{aligned}$$

where

$$I(\, X_1 \le a\,) = \begin{cases} 1 & \text{if } X_1 \le a \\ 0 & \text{otherwise} \end{cases}.$$

Hence

$$\Phi\left(\sqrt{\frac{n}{n-1}}(\, a - \bar{X}\,)\right)$$

is the UMVUE of $p = P_\theta(X_1 \le a)$.

2) [+1] Define the nonparametric estimator of $p = P_\theta(X_1 \le a)$ [denoted as δ_{2n}].

Answer:

We define

$$\delta_{2n} = \frac{1}{n}\sum_{i=1}^{n} I(\, X_i \le a\,)$$

is the nonparametric estimator of $p = P_\theta(X_1 \le a)$.

3) [+3] Calculate the ARE $e_{\delta_1\delta_2}$.

Answer:

Let

$$c_n = \sqrt{\frac{n}{n-1}} = \left(1-\frac{1}{n}\right)^{-\frac{1}{2}}.$$

By a Taylor expansion, we have

$$c_n = 1 + \frac{1}{2n} + \frac{3}{8n^2}.$$

We also define

$$\overline{Y} = a - \overline{X} \Rightarrow E(Y_i) = a - \theta \equiv \xi \text{ and } \operatorname{var}(Y_i) = 1, \quad \text{for } i = 1, 2, \cdots, n.$$

Then we can apply the general delta method

$$\sqrt{n}(\delta_{1n} - p) = \sqrt{n}\left(\Phi\left(\sqrt{\frac{n}{n-1}}(a - \overline{X})\right) - \Phi(a-\theta)\right)$$

$$= \sqrt{n}(\Phi(c_n\overline{Y}) - \Phi(\xi)) \xrightarrow{d} N(0, \phi(\xi)^2) = N(0, \phi(a-\theta)^2).$$

By CLT, we have

$$\sqrt{n}(\delta_{2n} - p) = N(0, p(1-p)).$$

Therefore, the ARE $e_{\delta_1\delta_2}$ is

$$e_{\delta_1\delta_2} = \frac{p(1-p)}{\phi(a-\theta)^2} = \frac{\Phi(a-\theta)\{1-\Phi(a-\theta)\}}{\phi(a-\theta)^2}.$$

4) [+1] Draw the graph of ARE with respect to θ.

Answer:

Since

$$e_{\delta_2\delta_1} = \frac{\phi(a-\theta)^2}{\Phi(a-\theta)\{1-\Phi(a-\theta)\}}.$$

If $\theta = a$, we have

$$e_{\delta_2\delta_1} = \frac{\left(\frac{1}{\sqrt{2\pi}}\right)^2}{\frac{1}{2}\times\frac{1}{2}} = \frac{2}{\pi} \approx 0.6366 < 1.$$

And $e_{\delta_2\delta_1} \to 0$, as $\theta \to \infty$ or $\theta \to -\infty$. The graph of ARE is shown in Figure 1 with $a = 3$.

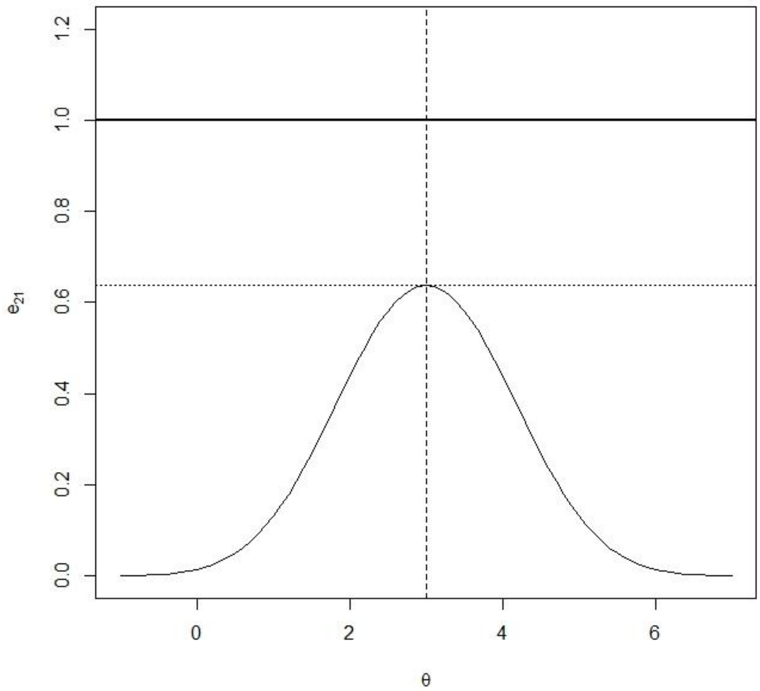

Fig.1 The graph of ARE $e_{\delta_2\delta_1}$ with $a = 3$.

Q4 [+8]. We consider the asymptotic distribution of the MLE under independent but not identically distributed random variables $X_{\alpha 1},..,X_{\alpha n_\alpha} \overset{iid}{\sim} f_\alpha(x|\boldsymbol{\theta}), \alpha = 1,...,k$, $\boldsymbol{\theta} \in \Omega \subset R^s$. Let $\hat{\boldsymbol{\theta}}$ be the solution to the likelihood equation (if exist).

1. [+2] State the necessary assumption about the sample size $n_\alpha, \alpha = 1,...,k$.

Answer:

Assumption (E):

Let $\sum_{\alpha=1}^{k} n_\alpha = N$, we define $\lim_{N\to\infty} \frac{n_\alpha}{N} = \lambda_\alpha > 0$, for $\alpha = 1, 2, ..., k$.

2. [+1] Define the log-likelihood function

Answer:

The log-likelihood function is

$$\ell(\boldsymbol{\theta}) = \sum_{\alpha=1}^{k}\sum_{i=1}^{n_\alpha} \log f_\alpha(x_{\alpha i} | \boldsymbol{\theta}).$$

3. [+1] Define the appropriate Fisher information matrix

Answer:

The appropriate Fisher information matrix is

$$I(\boldsymbol{\theta}) = \sum_{\alpha=1}^{l} \lambda_\alpha I^{(\alpha)}(\boldsymbol{\theta}),$$

where

$$I^{(\alpha)}(\boldsymbol{\theta}) = E_{\boldsymbol{\theta}}\left[\frac{\partial}{\partial \boldsymbol{\theta}} \log f_\alpha(x|\boldsymbol{\theta}) \frac{\partial}{\partial \boldsymbol{\theta}'} \log f_\alpha(x|\boldsymbol{\theta})\right].$$

4. [+4] Provide the outline of the proof of the asymptotic normality of $\hat{\boldsymbol{\theta}}$.
(explain how the assumption about the sample size $n_\alpha, \alpha = 1, \ldots, k$ is used)

Answer:

We define

$$\ell'_j(\boldsymbol{\theta}) = \frac{\partial \ell(\boldsymbol{\theta})}{\partial \theta_j}, \quad \ell''_{jr}(\boldsymbol{\theta}) = \frac{\partial^2 \ell(\boldsymbol{\theta})}{\partial \theta_j \partial \theta_k}, \quad \ell'''_{jrl}(\boldsymbol{\theta}) = \frac{\partial^3 \ell(\boldsymbol{\theta})}{\partial \theta_j \partial \theta_r \partial \theta_l}.$$

Let $\hat{\boldsymbol{\theta}} = (\hat{\theta}_1, \hat{\theta}_2, \cdots, \hat{\theta}_s)$ be the solution of the likelihood equations

$$\frac{\partial \ell(\boldsymbol{\theta})}{\partial \theta_j} = 0, \text{ for } j = 1, 2, \cdots, n.$$

By a Taylor expansion around the true parameter $\boldsymbol{\theta}^0 = (\theta_1^0, \theta_2^0, \cdots, \theta_s^0)$, we have

$$\ell'_j(\hat{\boldsymbol{\theta}}) = \ell'_j(\boldsymbol{\theta}^0) + \sum_{r=1}^{s} (\hat{\theta}_r - \theta_r^0) \ell''_{jr}(\boldsymbol{\theta}^0) + \frac{1}{2} \sum_{l=1}^{s} \sum_{r=1}^{s} (\hat{\theta}_r - \theta_r^0)(\hat{\theta}_l - \theta_l^0) \ell'''_{jrl}(\boldsymbol{\theta}^*),$$

where $\boldsymbol{\theta}^*$ is on the line between $\hat{\boldsymbol{\theta}}$ and $\boldsymbol{\theta}^0$. Then we have

$$0 = \ell'_j(\boldsymbol{\theta}^0) + \sum_{r=1}^{s} (\hat{\theta}_r - \theta_r^0) \left\{ \ell''_{jr}(\boldsymbol{\theta}^0) + \frac{1}{2} \sum_{l=1}^{s} (\hat{\theta}_l - \theta_l^0) \ell'''_{jrl}(\boldsymbol{\theta}^*) \right\}$$

$$\Rightarrow \frac{1}{\sqrt{N}} \ell'_j(\boldsymbol{\theta}^0) = \sum_{r=1}^{s} \sqrt{N} (\hat{\theta}_r - \theta_r^0) \left\{ -\frac{1}{N} \ell''_{jr}(\boldsymbol{\theta}^0) - \frac{1}{2N} \sum_{l=1}^{s} (\hat{\theta}_l - \theta_l^0) \ell'''_{jrl}(\boldsymbol{\theta}^*) \right\} \cdots (*)$$

Consider equation $(*)$ separately. First, by W.L.L.N., we have

$$\begin{aligned} -\frac{1}{N} \ell''_{jr}(\boldsymbol{\theta}^0) &= -\frac{1}{N} \frac{\partial^2}{\partial \theta_j \partial \theta_k} \sum_{\alpha=1}^{k} \sum_{i=1}^{n_\alpha} \log f_\alpha(x_{\alpha i} \mid \boldsymbol{\theta}) \\ &= \sum_{\alpha=1}^{k} \frac{n_\alpha}{N} \left(-\frac{1}{n_\alpha} \sum_{i=1}^{n_\alpha} \frac{\partial^2}{\partial \theta_j \partial \theta_k} \log f_\alpha(x_{\alpha i} \mid \boldsymbol{\theta}) \right) \\ &\xrightarrow{p} \sum_{\alpha=1}^{k} \lambda_\alpha I_{jr}^{(\alpha)}(\boldsymbol{\theta}^0) = I_{jr}(\boldsymbol{\theta}^0). \end{aligned}$$

Second, since $\hat{\theta}_l - \theta_l^0 \xrightarrow{p} 0$ and $\ell'''_{jrl}(\boldsymbol{\theta}^*) \xrightarrow{p} \text{constant}$, we have

$$\frac{1}{2N} \sum_{l=1}^{s} (\hat{\theta}_l - \theta_l^0) \ell'''_{jrl}(\boldsymbol{\theta}^*) \xrightarrow{p} 0.$$

Finally,

$$\begin{aligned} \frac{1}{\sqrt{N}} \ell'_j(\boldsymbol{\theta}^0) &= \frac{1}{\sqrt{N}} \frac{\partial}{\partial \theta_j} \sum_{\alpha=1}^{k} \sum_{i=1}^{n_\alpha} \log f_\alpha(x_{\alpha i} \mid \boldsymbol{\theta}) \\ &= \sum_{\alpha=1}^{k} \sqrt{\frac{n_\alpha}{N}} \frac{1}{\sqrt{n_\alpha}} \sum_{i=1}^{n_\alpha} \frac{\partial}{\partial \theta_j} \log f_\alpha(x_{\alpha i} \mid \boldsymbol{\theta}). \end{aligned}$$

In the vector form, by the Multivariate CLT, we obtain

$$\frac{1}{\sqrt{N}}\begin{bmatrix} \ell'_1(\boldsymbol{\theta}^0) \\ \ell'_2(\boldsymbol{\theta}^0) \\ \vdots \\ \ell'_s(\boldsymbol{\theta}^0) \end{bmatrix} = \sum_{\alpha=1}^{k}\sqrt{\frac{n_\alpha}{N}}\begin{bmatrix} \frac{1}{\sqrt{n_\alpha}}\sum_{i=1}^{n_\alpha}\frac{\partial}{\partial\theta_1}\log f_\alpha(x_{\alpha i}\,|\,\boldsymbol{\theta}) \\ \frac{1}{\sqrt{n_\alpha}}\sum_{i=1}^{n_\alpha}\frac{\partial}{\partial\theta_2}\log f_\alpha(x_{\alpha i}\,|\,\boldsymbol{\theta}) \\ \vdots \\ \frac{1}{\sqrt{n_\alpha}}\sum_{i=1}^{n_\alpha}\frac{\partial}{\partial\theta_s}\log f_\alpha(x_{\alpha i}\,|\,\boldsymbol{\theta}) \end{bmatrix}$$

$$\xrightarrow{d}\sum_{\alpha=1}^{k}\sqrt{\lambda_\alpha}\,N(0, I^{(\alpha)}(\boldsymbol{\theta}^0)) = N(0, I(\boldsymbol{\theta}^0)).$$

Therefore, the equation $(*)$ becomes

$$I(\boldsymbol{\theta}^0)\sqrt{N}(\hat{\boldsymbol{\theta}}-\boldsymbol{\theta}^0)\xrightarrow{d}N(0, I(\boldsymbol{\theta}^0))$$

$$\Rightarrow \sqrt{N}(\hat{\boldsymbol{\theta}}-\boldsymbol{\theta}^0)\xrightarrow{d}N(0, I^{-1}(\boldsymbol{\theta}^0)).$$