Solution for Stats 369, HW1

Chen Cheng

Exercise 1.1.

We copy here the definition of various quantities.

$$\boldsymbol{Y} = \lambda \boldsymbol{v}_0^{\otimes k} + \boldsymbol{W}, \tag{1}$$

$$\boldsymbol{W} = \frac{1}{\sqrt{k!n}} \sum_{\pi \in \Pi} \boldsymbol{G}^{\pi}, \tag{2}$$

$$\mu_{n,\lambda,\beta,h}(\mathrm{d}\boldsymbol{\sigma}) = \frac{1}{Z_n(\beta,\lambda,h)} \exp\left\{ \frac{n\beta}{\sqrt{2(k!)}} \langle \boldsymbol{Y}, \boldsymbol{\sigma}^{\otimes k}\rangle + nh\langle \boldsymbol{v}_0, \boldsymbol{\sigma}\rangle \right\} \nu_0(\mathrm{d}\boldsymbol{\sigma}), \tag{3}$$

$$Z_n(\beta,\lambda,h) = \int_{\mathbb{S}^{n-1}} \exp\left\{ \frac{n\beta}{\sqrt{2(k!)}} \langle \boldsymbol{Y}, \boldsymbol{\sigma}^{\otimes k}\rangle + nh\langle \boldsymbol{v}_0, \boldsymbol{\sigma}\rangle \right\} \nu_0(\mathrm{d}\boldsymbol{\sigma}), \tag{4}$$

$$\Phi_n(\beta,\lambda,h) = \frac{1}{n}\mathbb{E}\log Z_n(\beta,\lambda,h). \tag{5}$$

$$\boldsymbol{X}_0 = \boldsymbol{v}_0^{\otimes k}, \tag{6}$$

$$\hat{\boldsymbol{v}}_{\beta,h}(\boldsymbol{Y}) = \int_{\mathbb{S}^{n-1}} \boldsymbol{\sigma}\mu_{n,\beta,\lambda,h}(\mathrm{d}\boldsymbol{\sigma}), \tag{7}$$

$$\widehat{\boldsymbol{X}}_{\beta,h}(\boldsymbol{Y}) = \int_{\mathbb{S}^{n-1}} \boldsymbol{\sigma}^{\otimes k}\mu_{n,\beta,\lambda,h}(\mathrm{d}\boldsymbol{\sigma}). \tag{8}$$

To verify Equation (1.1.16), we have

$$\begin{aligned}
\frac{\partial \Phi_n}{\partial h}(\beta,\lambda,h) =& \frac{1}{n}\mathbb{E}\{\frac{\partial}{\partial h}\log Z_n(\beta,\lambda,h)\} \\
=& \frac{1}{n}\mathbb{E}\{\frac{1}{Z_n}\frac{\partial}{\partial h} Z_n\} \\
=& \frac{1}{n}\mathbb{E}\{\frac{1}{Z_n}\int_{\mathbb{S}^{n-1}} n\langle \boldsymbol{v}_0, \boldsymbol{\sigma}\rangle \exp\left\{ \frac{n\beta}{\sqrt{2(k!)}} \langle \boldsymbol{Y}, \boldsymbol{\sigma}^{\otimes k}\rangle + h\langle \boldsymbol{v}_0, \boldsymbol{\sigma}\rangle \right\} \nu_0(\mathrm{d}\boldsymbol{\sigma})\} \\
=& \mathbb{E}\{\int_{\mathbb{S}^{n-1}} \langle \boldsymbol{v}_0, \boldsymbol{\sigma}\rangle \mu_{n,\beta,\lambda,h}(\mathrm{d}\boldsymbol{\sigma})\} \\
=& \mathbb{E}\{\langle \boldsymbol{v}_0, \hat{\boldsymbol{v}}_{\beta,h}(\boldsymbol{Y})\rangle\}.
\end{aligned}$$

To verify Equation (1.1.17), we have

$$\begin{aligned}
\frac{\partial \Phi_n}{\partial \lambda}(\beta,\lambda,h) =& \frac{1}{n}\mathbb{E}\{\frac{\partial}{\partial\lambda}\log Z_n(\beta,\lambda,h)\} \\
=& \frac{1}{n}\mathbb{E}\{\frac{1}{Z_n}\frac{\partial}{\partial\lambda}Z_n\} \\
=& \frac{1}{n}\mathbb{E}\{\frac{1}{Z_n}\frac{\partial}{\partial\lambda}\int_{\mathbb{S}^{n-1}}\exp\left\{\frac{n\beta}{\sqrt{2(k!)}}\langle\lambda\boldsymbol{v}_0^{\otimes k}+\boldsymbol{W},\boldsymbol{\sigma}^{\otimes k}\rangle + h\langle\boldsymbol{v}_0,\boldsymbol{\sigma}\rangle\right\}\nu_0(\mathrm{d}\boldsymbol{\sigma})\} \\
=& \frac{1}{n}\mathbb{E}\{\frac{1}{Z_n}\frac{n\beta}{\sqrt{2(k!)}}\langle\boldsymbol{v}_0^{\otimes k},\boldsymbol{\sigma}^{\otimes k}\rangle\int_{\mathbb{S}^{n-1}}\exp\left\{\frac{n\beta}{\sqrt{2(k!)}}\langle\lambda\boldsymbol{v}_0^{\otimes k}+\boldsymbol{W},\boldsymbol{\sigma}^{\otimes k}\rangle + h\langle\boldsymbol{v}_0,\boldsymbol{\sigma}\rangle\right\}\nu_0(\mathrm{d}\boldsymbol{\sigma})\} \\
=& \mathbb{E}\{\int_{\mathbb{S}^{n-1}}\langle\boldsymbol{v}_0,\boldsymbol{\sigma}\rangle^{\otimes k}\mu_{n,\beta,\lambda,h}(\mathrm{d}\boldsymbol{\sigma})\} \\
=& \frac{\beta}{\sqrt{2(k!)}}\mathbb{E}\{\langle\boldsymbol{X}_0,\widehat{\boldsymbol{X}}_{\beta,h}(\boldsymbol{Y})\rangle\}.
\end{aligned}$$

Exercise 1.2.

(a)

Given $\boldsymbol{Y}_1,\ldots,\boldsymbol{Y}_n$, the likelihood function for $(\boldsymbol{v}_0,\boldsymbol{\lambda})$ gives

$$\begin{aligned}
\ell(\boldsymbol{v}_0,\boldsymbol{\lambda};\boldsymbol{Y}_1,\ldots,\boldsymbol{Y}_n) =& \log\left\{\prod_{l=1}^{k}\frac{1}{Z_l}\exp\left\{-\frac{n}{2(l!)}\|\boldsymbol{Y}_l-\lambda_l\boldsymbol{v}_0^{\otimes l}\|_F^2\right\}\right\} && (9)\\
=& \sum_{l=1}^{k}-\frac{n}{2(l!)}\|\boldsymbol{Y}_l-\lambda_l\boldsymbol{v}_0^{\otimes l}\|_F^2+\text{const.} && (10)\\
=& \sum_{l=1}^{k}\frac{n}{2(l!)}(2\lambda_l\langle\boldsymbol{Y}_l,\boldsymbol{v}_0^{\otimes l}\rangle-\lambda_l^2\langle\boldsymbol{v}^{\otimes l},\boldsymbol{v}^{\otimes l}\rangle)+C(Y_l) && (11)\\
=& \sum_{l=1}^{k}\frac{n}{2(l!)}(2\lambda_l\langle\boldsymbol{Y}_l,\boldsymbol{v}_0^{\otimes l}\rangle-\lambda_l^2)+C(Y_l) && (12)
\end{aligned}$$

For $\boldsymbol{\lambda}$ known, MLE solve the following problem

$$\begin{aligned}
\max\quad & \sum_{l=1}^{k}\frac{\lambda_l}{l!}\langle Y_l,\boldsymbol{v}_0^{\otimes l}\rangle && (13)\\
\text{s.t.}\quad & \boldsymbol{v}_0\in\mathbb{S}^{n-1}. && (14)
\end{aligned}$$

For $\boldsymbol{\lambda}$ unknown, maximizing over $\boldsymbol{\lambda}$ is easy. We have

$$\begin{aligned}
\ell(\boldsymbol{v}_0) =& \sup_{\boldsymbol{\lambda}}\ell(\boldsymbol{v}_0,\boldsymbol{\lambda};\boldsymbol{Y}_1,\ldots,\boldsymbol{Y}_n) && (15)\\
=& \sup_{\boldsymbol{\lambda}}\sum_{l=1}^{k}\frac{n}{2(l!)}(2\lambda_l\langle\boldsymbol{Y}_l,\boldsymbol{v}_0^{\otimes l}\rangle-\lambda_l^2)+C(Y_l) && (16)\\
=& \sum_{l=1}^{k}\frac{n}{2(l!)}\langle\boldsymbol{Y}_l,\boldsymbol{v}_0^{\otimes l}\rangle^2+C(Y_l). && (17)
\end{aligned}$$

Thus, the MLE solve the following problem

$$\begin{aligned}
\max\quad & \sum_{l=1}^{k}\frac{1}{l!}\langle Y_l,\boldsymbol{v}_0^{\otimes l}\rangle^2 && (18)\\
\text{s.t.}\quad & \boldsymbol{v}_0\in\mathbb{S}^{n-1}. && (19)
\end{aligned}$$

Assuming a uniform prior $\nu_0(\mathrm{d}\boldsymbol{\sigma})$ on the unit sphere, the posterior distribution gives

$$\begin{aligned}\mu_{\text{Bayes}}(\mathrm{d}\boldsymbol{\sigma}) =& \frac{1}{\tilde{Z}_{\text{Bayes}}(\boldsymbol{\lambda})} \int_{\mathbb{S}^{n-1}} \prod_{l=1}^{k} \exp\{-\frac{n}{2(l!)}\|Y_l - \lambda_l \boldsymbol{\sigma}^{\otimes l}\|_F^2\}\nu_0(\mathrm{d}\boldsymbol{\sigma})\\ =& \frac{1}{Z_{\text{Bayes}}(\boldsymbol{\lambda})} \int_{\mathbb{S}^{n-1}} \prod_{l=1}^{k} \exp\{\frac{n\lambda_l}{l!}\langle Y_l, \boldsymbol{\sigma}^{\otimes l}\rangle\}\nu_0(\mathrm{d}\boldsymbol{\sigma})\end{aligned}$$

(b)

We consider the following family of probability measure indexed by $n, \boldsymbol{\lambda}, \beta, h$

$$\mu_{n,\boldsymbol{\lambda},\beta,h}(\mathrm{d}\boldsymbol{\sigma}) = \frac{1}{Z_n(\boldsymbol{\lambda},\beta,h)} \prod_{l=1}^{k} \exp\{\frac{n\lambda_l\beta}{l!}\langle Y_l, \boldsymbol{\sigma}^{\otimes l}\rangle + nh\langle \boldsymbol{v}_0, \boldsymbol{\sigma}\rangle\}\nu_0(\mathrm{d}\boldsymbol{\sigma}). \tag{20}$$

In particular, as $\beta = +\infty$, this measure corresponds to MLE estimator. As $\beta = 1$, this measure corresponds to Bayes estimator.

(c)

First, let's calculate

$$\begin{aligned}\mathbb{E}[Z_n^r] =& \mathbb{E}\left\{\left[\int_{\mathbb{S}^{n-1}} \prod_{l=1}^{k} \exp\{\frac{n\lambda_l\beta}{l!}\langle Y_l, \boldsymbol{\sigma}^{\otimes l}\rangle + nh\langle \boldsymbol{v}_0, \boldsymbol{\sigma}\rangle\}\nu_0(\mathrm{d}\boldsymbol{\sigma})\right]^r\right\}\\ =& \mathbb{E}\left\{\left[\int_{\mathbb{S}^{n-1}} \prod_{l=1}^{k} \exp\{\frac{n\lambda_l^2\beta}{l!}\langle \boldsymbol{v}_0, \boldsymbol{\sigma}\rangle^l + \frac{n\lambda_l\beta}{l!}\langle \boldsymbol{W}, \boldsymbol{\sigma}^{\otimes l}\rangle + nh\langle \boldsymbol{v}_0, \boldsymbol{\sigma}\rangle\}\nu_0(\mathrm{d}\boldsymbol{\sigma})\right]^r\right\}\\ =& \int_{\mathbb{S}^{n-1}} \prod_{l=1}^{k} \exp\{\frac{n\lambda_l^2\beta}{l!}\sum_{a=1}^{r}\langle \boldsymbol{\sigma}^0, \boldsymbol{\sigma}^a\rangle^l + nh\sum_{a=1}^{r}\langle \boldsymbol{\sigma}^0, \boldsymbol{\sigma}^a\rangle\}\mathbb{E}[\exp\{\frac{n\lambda_l\beta}{l!}\sum_{a=1}^{r}\langle \boldsymbol{W}, (\boldsymbol{\sigma}^a)^{\otimes l}\rangle\}]\nu_{0,r+1}(\mathrm{d}\boldsymbol{\sigma}).\end{aligned} \tag{21}$$

First we calculate the expectation

$$\mathbb{E}[\exp\{\frac{n\lambda_l\beta}{l!}\sum_{a=1}^{r}\langle \boldsymbol{W}, (\boldsymbol{\sigma}^a)^{\otimes l}\rangle\}] \tag{22}$$

$$= \mathbb{E}[\prod_{\pi\in\Pi} \exp\{\frac{n\lambda_l\beta}{l!}\frac{1}{\sqrt{nl!}}\langle \boldsymbol{G}^\pi, \sum_{a=1}^{r}(\boldsymbol{\sigma}^a)^{\otimes l}\rangle\}] \tag{23}$$

$$= \mathbb{E}[\exp\{\frac{\sqrt{n}\lambda_l\beta}{\sqrt{l!}}\langle \boldsymbol{G}, \sum_{a=1}^{r}(\boldsymbol{\sigma}^a)^{\otimes l}\rangle\}] \tag{24}$$

$$= \exp\{\frac{n\lambda_l^2\beta^2}{2l!}\sum_{a,b=1}^{r}\langle \boldsymbol{\sigma}^a, \boldsymbol{\sigma}^b\rangle^l\}. \tag{25}$$

Thus, Equation (21) gives

$$\mathbb{E}[Z_n^r] = \int_{\mathbb{S}^{n-1}} \prod_{l=1}^{k} \exp\{nh\sum_{a=1}^{r}\langle \boldsymbol{\sigma}^0, \boldsymbol{\sigma}^a\rangle + \frac{n\lambda_l^2\beta}{l!}\sum_{a=1}^{r}\langle \boldsymbol{\sigma}^0, \boldsymbol{\sigma}^a\rangle^l + \frac{n\lambda_l^2\beta^2}{2l!}\sum_{a,b=1}^{r}\langle \boldsymbol{\sigma}^a, \boldsymbol{\sigma}^b\rangle^l\}\nu_{0,r+1}(\mathrm{d}\boldsymbol{\sigma}) \tag{26}$$

Consider the $(r+1)\times(r+1)$ matrix $\widehat{\boldsymbol{Q}} = (\widehat{Q}_{ab})_{0\le a,b\le r}$ defined by

$$\widehat{Q}_{ab} = \langle \boldsymbol{\sigma}^a, \boldsymbol{\sigma}^b\rangle, \quad \text{for } a,b \in \{0,1,\ldots,r\}. \tag{27}$$

Denote by $f_{n,r}(\boldsymbol{Q})$ the joint density of the random variables $(\widehat{Q}_{ab})_{0\le ab\le r}$ when $\boldsymbol{\sigma}^0, \boldsymbol{\sigma}^1, \dots, \boldsymbol{\sigma}^r$ are distributed according to $\nu_{0,r+1} = \nu_0 \times \cdots \times \nu_0$. We then have

$$\mathbb{E}[Z_n^r] = \int \prod_{l=1}^{k} \exp\{nh \sum_{a=1}^{r} Q_{0a} + \frac{n\lambda_l^2\beta}{l!} \sum_{a=1}^{r} Q_{0a}^l + \frac{n\lambda_l^2\beta^2}{2l!} \sum_{a,b=1}^{r} Q_{ab}^l\} f_{n,r}(\boldsymbol{Q}) \mathrm{d}\boldsymbol{Q}. \tag{28}$$

In lecture notes we already computed that

$$f_{n,r}(\boldsymbol{Q}) \doteq \det(\boldsymbol{Q})^{n/2}.$$

Thus, we get

$$\begin{aligned}
\mathbb{E}[Z_n^r] &\doteq \sup_{\boldsymbol{Q}} \exp\{\sum_{l=1}^{k}\{nh\sum_{a=1}^{r} Q_{0a} + \frac{n\lambda_l^2\beta}{l!}\sum_{a=1}^{r} Q_{0a}^l + \frac{n\lambda_l^2\beta^2}{2l!}\sum_{a,b=1}^{r} Q_{ab}^l + \frac{n}{2}\mathrm{Tr}\log(\boldsymbol{Q})\}\} && (29)\\
&= \exp\{n \sup_{\boldsymbol{Q}} S(\boldsymbol{Q})\}, && (30)
\end{aligned}$$

where

$$S(\boldsymbol{Q}) = \sum_{l=1}^{k}\{h\sum_{a=1}^{r} Q_{0a} + \frac{\lambda_l^2\beta}{l!}\sum_{a=1}^{r} Q_{0a}^l + \frac{\lambda_l^2\beta^2}{2l!}\sum_{a,b=1}^{r} Q_{ab}^l\} + \frac{1}{2}\mathrm{Tr}\log(\boldsymbol{Q}).$$

The replica symmetric ansatz gives

$$\begin{aligned}
Q_{aa} &= 1, \quad \text{for } a \in \{0,1,\dots,r\}, && (31)\\
Q_{0a} &= b, \quad \text{for } a \in \{1,\dots,r\}, && (32)\\
Q_{ab} &= q, \quad \text{for } a \neq b, a, b \in \{1,\dots,r\}. && (33)
\end{aligned}$$

Thus, we have

$$\begin{aligned}
S(\boldsymbol{Q}^{\mathrm{RS}}) &= \sum_{l=1}^{k}\{hrb + \frac{\lambda_l^2\beta}{l!}rb^l + \frac{\lambda_l^2\beta^2}{2l!}r + \frac{\lambda_l^2\beta^2}{2l!}r(r-1)q^l\} && (34)\\
&\quad + \frac{1}{2}\log(1 - \frac{rb^2}{1+(r-1)q}) + \frac{1}{2}\log(1 + \frac{rq}{1-q}) + \frac{1}{2}r\log(1-q), && (35)
\end{aligned}$$

and

$$\Phi_{\mathrm{RS}}(b,q;\beta,\boldsymbol{\lambda},h) = \sum_{l=1}^{k}\{hb + \frac{\lambda_l^2\beta}{l!}b^l + \frac{\lambda_l^2\beta^2}{2l!}(1-q^l)\} - \frac{b^2}{2(1-q)} + \frac{q}{2(1-q)} + \frac{1}{2}\log(1-q). \tag{36}$$

The free energy density gives

$$\phi(\beta,\boldsymbol{\lambda},h) = \sup_b \inf_q \Phi_{\mathrm{RS}}(b,q;\beta,\boldsymbol{\lambda},h). \tag{37}$$

Exercise 1.4.

(a)

Due to Markov inequality, we have

$$\mathbb{P}(Z_n \ge \exp(\varepsilon n)\mathbb{E}Z_n) \le \frac{\mathbb{E}Z_n}{\exp(\varepsilon n)\mathbb{E}Z_n} = \exp(-n\varepsilon). \tag{38}$$

Due to Paley-Ziegmund inequality, we have

$$\mathbb{P}(Z_n \ge \exp(-n\varepsilon)\mathbb{E}Z_n) \ge (1 - \exp(-n\varepsilon))^2 \frac{[\mathbb{E}Z_n]^2}{\mathbb{E}[Z_n^2]}. \tag{39}$$

Thus,

$$\frac{1}{n}\log\mathbb{P}(Z_n \geq \exp(-n\varepsilon)\mathbb{E}Z_n) \geq \frac{2}{n}\log(1-\exp(-n\varepsilon)) + \frac{1}{n}\log(\frac{[\mathbb{E}Z_n]^2}{\mathbb{E}[Z_n^2]}) \to 0. \tag{40}$$

The last equality is because equation (1.6.9). This implies that

$$\mathbb{P}(Z_n \geq \exp(-n\varepsilon)\mathbb{E}Z_n) = \exp(-o(n)). \tag{41}$$

The above equation can also be interpreted as, for any $\varepsilon_0 > 0$, we have

$$\lim_{n\to\infty}\exp(n\varepsilon_0)\mathbb{P}(Z_n \geq \exp(-n\varepsilon)\mathbb{E}Z_n) \to \infty. \tag{42}$$

Combining equation (38) and (41), we have

$$\mathbb{P}(\exp(n\varepsilon)\mathbb{E}Z_n \geq Z_n \geq \exp(-n\varepsilon)\mathbb{E}Z_n) = \exp(-o(n)). \tag{43}$$

(b)

According to equation (1.6.7), for any $\varepsilon > 0$, we have

$$\mathbb{P}(\frac{1}{n}|\log Z_n - \mathbb{E}\log Z_n| \geq \varepsilon) \leq 2\exp(-\frac{\varepsilon}{\beta^2}n) \to 0. \tag{44}$$

But we know

$$\mathbb{P}(\frac{1}{n}|\log Z_n - \log\mathbb{E}Z_n| \leq \varepsilon) \geq \exp(-o(n)). \tag{45}$$

Thus, we must have

$$\lim_{n\to\infty}\frac{1}{n}\mathbb{E}\log Z_n = \lim_{n\to\infty}\frac{1}{n}\log\mathbb{E}Z_n = \phi. \tag{46}$$

Thus, equation (44) implies that $\frac{1}{n}\log Z_n$ converges to ϕ in probability.

Remark 1. *It is essential to have the concentration of $\frac{1}{n}\log Z_n$ itself. If we only have (1.6.8) and (1.6.9), this claim is not true. An example gives $Z_n \sim 1/2\cdot\delta(1) + 1/2\cdot\delta(\exp(n\phi))$.*

(c)

Note that we have

$$\mathbb{E}[Z_n^r] \doteq \exp(n\sup_{\boldsymbol{Q}} S_r(\boldsymbol{Q})). \tag{47}$$

We need to compute $\sup_{\boldsymbol{Q}} S_1(\boldsymbol{Q})$ and $\sup_{\boldsymbol{Q}} S_2(\boldsymbol{Q})$. As $r \leq 2$, the replica symmetric ansatz becomes the only possible solution. So we borrow the replica symmetric ansats calculation result here

$$\begin{aligned} S_r(\boldsymbol{Q}) =& hrb + \frac{\beta\lambda}{\sqrt{2(k!)}}rb^k + \frac{\beta^2 r}{4} + \frac{\beta^2 r(r-1)}{4}q^k \\ &+ \frac{1}{2}\log(1-\frac{rb^2}{1+(r-1)q}) + \frac{1}{2}\log(1+\frac{rq}{1-q}) + \frac{1}{2}r\log(1-q). \end{aligned} \tag{48}$$

For $r = 1$, taking $h = 0$, we have

$$S_1(\boldsymbol{Q}) = \frac{\beta\lambda}{\sqrt{2(k!)}}b^k + \frac{\beta^2}{4} + \frac{1}{2}\log(1-b^2). \tag{49}$$

For $r = 2$, taking $h = 0$, we have

$$S_2(\boldsymbol{Q}) = 2\frac{\beta\lambda}{\sqrt{2(k!)}}b^k + \frac{\beta^2}{2} + \frac{\beta^2}{2}q^k + \frac{1}{2}\log(1-\frac{2b^2}{1+q}) + \frac{1}{2}\log(1-q^2). \tag{50}$$

Left hand side of equation (1.6.8) gives

$$\sup_{b,q} S_2(b,q) - 2\sup_b S_1(b). \tag{51}$$

(d)

Expanding $S_1(b)$ around $b = 0$ gives

$$S_1(b) = \frac{\beta^2}{4} - \frac{1}{2}b^2 + \frac{\beta\lambda}{\sqrt{2(k!)}}b^k + o(b^2). \tag{52}$$

As β and λ is small, $b = 0$ is a local maximum of S_1, and is also the global maximum of S_1.

For $S_2(b, q)$, as β and λ small, $(b, q) = (0, 0)$ is the global maximum of S_2. Thus, we have

$$\sup_{b,q} S_2(b, q) - 2\sup_b S_1(b) = S_2(0, 0) - 2S_1(0) = 0. \tag{53}$$

(e)

As $k \geq 2$, we have as long as $\beta\lambda \leq \sqrt{k!/2}$ it holds that

$$\sup_b S_1(b) = S_1(0) = \frac{\beta^2}{4}. \tag{54}$$

However, for β large enough, we can also have

$$\begin{aligned}
\sup_{b,q} S_2(b, q) &\geq \sup_q \left(\frac{\beta^2}{2} + \frac{\beta^2}{2}q^k + \frac{1}{2}\log(1 - q^2) \right) \\
&= \frac{\beta^2}{2} + \sup_q \left(\frac{\beta^2}{2}q^k + \frac{1}{2}\log(1 - q^2) \right) \\
&> \frac{\beta^2}{2}.
\end{aligned} \tag{55}$$