به نام زندگی

مثال: تخمین پارامتر زیع‌های

- تخمین MMSE

$$\hat{\theta}_{MMSE}(y) = E\{\theta | y\} = \frac{2}{\alpha + y}$$

- تمرین MMAE (تابع هزینه قدرمطلق خطا)

$$\hat{\theta}_{MMAE}(y) = \hat{\theta}_{ABS}(y) = \text{Conditional Median of } w(\theta|y)$$

$$w(\theta|y) = \theta(\alpha+y)^2 e^{-\theta(\alpha+y)} \qquad \theta \geq 0 , \quad y \geq 0$$

$$\int_{-\infty}^{\hat{\theta}_{MMAE}} w(\theta|y)\, d\theta = \int_{\hat{\theta}_{MMAE}}^{\infty} w(\theta|y)\, d\theta = \frac{1}{2}$$

با توجه به اینکه $\omega(\theta \mid y)$ در دامنه تعریف پیوسته است، با حل کردن یکی از معادلات قبلی می توانیم جواب مسئله را پیدا کنیم.

$$\int_{\hat{\theta}_{MMAE}}^{\infty} \theta (\alpha+y)^2 e^{-\theta(\alpha+y)} \, d\theta = \frac{1}{2}$$

$$(\alpha+y)^2 \int_{\hat{\theta}_{MMAE}}^{\infty} \underbrace{\theta}_{u} \, \underbrace{e^{-\theta(\alpha+y)} \, d\theta}_{dv} = \frac{1}{2}$$

$$\int_{0}^{\infty} u\,dv = uv\Big|_{0}^{\infty} - \int_{0}^{\infty} v\,du$$

$$(\alpha+y)^2 \left[-\frac{\theta e^{-\theta(\alpha+y)}}{\alpha+y}\Bigg|_{\hat{\theta}_{MMAE}}^{\infty} - \int_{\hat{\theta}_{MMAE}}^{\infty} \frac{-e^{-\theta(\alpha+y)}}{(\alpha+y)}\,d\theta \right] = \frac{1}{2}$$

$$\Rightarrow (\alpha+y)^2 \left[\frac{\hat{\theta}\, e^{-\hat{\theta}(\alpha+y)}}{\alpha+y} + \frac{e^{-\hat{\theta}(\alpha+y)}}{(\alpha+y)^2} \right] = \frac{1}{2}$$

$$\Longrightarrow \quad e^{-\underbrace{\hat{\theta}(\alpha+y)}_{T}} \left[\underbrace{(\alpha+y)\hat{\theta}}_{T} + 1 \right] = \frac{1}{2}$$

$$e^{-T}(T+1) = \frac{1}{2} \quad \overset{\ln}{\Longrightarrow} \quad T \simeq 1.68$$

$$\Longrightarrow \quad \hat{\theta}_{MMAE}(\alpha+y) \simeq 1.68 \quad \Longrightarrow \quad \hat{\theta}_{MMAE}(y) \simeq \frac{1.68}{\alpha+y}$$

- تابع هزینه یکنواخت (تخمین MAP)

$$\hat{\theta}_{MAP}(y) = \text{Conditional Mode of } \omega(\theta|y)$$

$$= \operatorname{Arg}\max_{\theta} \omega(\theta|y)$$

با توجه به اینکه توزیع $\omega(\theta|y)$ فرم نمایی دارد، از رابطه‌ی زیر می‌توانیم به راحتی

$\hat{\theta}_{MAP}$ را به دست بیاوریم.

$$\frac{\partial}{\partial \theta}\left[\log p(y|\theta)\right]_{\hat{\theta}_{MAP}} = -\frac{\partial}{\partial \theta}\left[\log w(\theta)\right]_{\hat{\theta}_{MAP}}$$

$$p(y|\theta) = \theta e^{-\theta y} \qquad y \geq 0$$

$$w(\theta) = \alpha e^{-\alpha\theta} \qquad \theta \geq 0$$

$$\frac{\partial}{\partial \theta}\left(\ln \theta - \theta y\right]_{\hat{\theta}_{MAP}} = -\frac{\partial}{\partial \theta}\left(\ln \alpha - \alpha \theta\right]_{\hat{\theta}_{MAP}}$$

$$\left[\frac{1}{\theta} - y + 0 - \alpha\right]_{\hat{\theta}_{MAP}} = 0 \rightarrow \left.\frac{1}{\theta}\right|_{\hat{\theta}_{MAP}} = (\alpha + y)$$

$$\Rightarrow \hat{\theta}_{MAP} = \frac{1}{\alpha + y}$$

$$\hat{\theta}_{MMSE} = \frac{2}{\alpha + y}$$

$$\hat{\theta}_{MMAE} \cong \frac{1.68}{\alpha + y}$$

تخمین های مختلف در این مسئله در یک ضریب ثابت با هم تفاوت دارند.

$$r(\hat{\theta}_{MAP}) = E_Y\left\{ \overbrace{E_{\theta|Y}\left\{ C(\hat{\theta}_{MAP}, \theta] \,\middle|\, Y=y \right\}}^{\text{تابعی از: } Y=y} \right\}$$

$$= E\left\{ E_Y\left\{ C(\hat{\theta}_{MAP}, \theta] \,\middle|\, Y=y \right\} \right]$$

$$= E\left\{ \left[1 - \int_{\hat{\theta}_{MAP}-\Delta}^{\hat{\theta}_{MAP}+\Delta} w(\theta|y)\, d\theta \right] \right\}$$

$$= E\left\{ \left[1 - \int_{\hat{\theta}_{MAP}-\Delta}^{\hat{\theta}_{MAP}+\Delta} \theta(\alpha+y)^2 e^{-\theta(\alpha+y)}\, d\theta \right] \right.$$

$$\Rightarrow r\left(\hat{\theta}_{MAP}\right) = \epsilon\left\{1-(\alpha+y)^2\left[\frac{-\theta\, e^{-\theta(\alpha+y)}}{(\alpha+y)}\Bigg|_{\hat{\theta}-\Delta}^{\hat{\theta}+\Delta} - \frac{e^{-\theta(\alpha+y)}}{(\alpha+y)^2}\Bigg|_{\hat{\theta}-\Delta}^{\hat{\theta}+\Delta}\right]\right\}$$

$$\Rightarrow r\left(\hat{\theta}_{MAP}\right) = \epsilon\Big\{1+(\alpha+y)\left[(\hat{\theta}+\Delta)\,e^{-(\hat{\theta}+\Delta)(\alpha+y)} - (\hat{\theta}-\Delta)\,e^{-(\hat{\theta}-\Delta)(\alpha+y)}\right]$$

$$+\; e^{-(\hat{\theta}+\Delta)(\alpha+y)} - e^{(\hat{\theta}-\Delta)(\alpha+y)}\Big\}$$

$$= \epsilon\left\{e^{-(\hat{\theta}+\Delta)(\alpha+y)}\left[(\alpha+y)(\hat{\theta}+\Delta)+1\right] - e^{-(\hat{\theta}-\Delta)(\alpha+y)}\left((\alpha+y)(\hat{\theta}-\Delta)+1\right)\right\}$$

$$\Rightarrow r(\hat\theta_{MAP}) = E\left\{ e^{-\underbrace{\hat\theta(\alpha+y)}_{1}} \left[e^{-\Delta(\alpha+y)}(2+\Delta(\alpha+y)) - e^{\Delta(\alpha+y)}(2-\Delta(\alpha+y)) \right.\right.$$

$$\left.\begin{aligned} \hat\theta_{MAP} &= \frac{1}{\alpha+y} \\ \hat\theta_{MAP}(\alpha+y) &= 1 \end{aligned}\right\} \Rightarrow r(\hat\theta_{MAP}) = e^{-1}\, E\left\{ 2\cosh \Delta(\alpha+y) + \Delta(\alpha+y)\sinh \Delta(\alpha+y) \right\}$$

$$= e^{-1}\int_0^\infty \cdots\, p(y)\, dy$$

$$p(y) = \frac{\alpha}{(\alpha+y)^2}$$

$$\int_0^\infty x^n e^{-x\alpha}\, dx = \frac{n!}{\alpha^{n+1}}$$

* تخمین Baysian در حالت برداری

در این مسأله می‌خواهیم برداری از پارامترهای نامعلوم را تخمین بزنیم. بنابراین در این حالت $\underline{\theta}$ یک بردار به طول m درنظر گرفته می‌شود.

$$\underline{\theta} \in \Lambda^m \quad , \quad \underline{\theta} = [\theta_1, \theta_2, \dots, \theta_m]^T$$

$$, \ \theta_i \in \Lambda$$

اولین قدم در تعمیم جواب مسئله بنزین به حالت برداری، این است که بتوانیم یک تابع هزینه مناسب در حالت برداری تعریف کنیم (در واقع تابع هزینه را در حالت برداری به صورت تعمیمی از تابع هزینه در حالت تک بعدی بنویسیم) می‌دانیم که

در این حالت برداری

$$C[\underline{a}, \underline{\theta}] : \Lambda^m \times \Lambda^m \longrightarrow \mathbb{R}^+$$

یکی از روشهای متداول برای تعریف تابع هزینه، تعمیم تابع هزینه به صورت زیر است.

$$C(\underline{a}, \underline{\theta}) = \sum_{i=1}^{m} C_i(a_i, \theta_i)$$

بدین ترتیب مسأله تخمین بیزین به صورت زیر قابل بیان است.

$$\hat{\underline{\theta}}(\underline{y}) = \underset{\hat{\theta}}{\mathrm{Arg\,Min}}\ \mathrm{E}\left\{ C(\hat{\underline{\theta}}(\underline{y}), \underline{\theta}) \right\}$$

$$\Rightarrow \underline{\hat{\theta}}(y) = \underset{\hat{\theta}}{\operatorname{Arg\,Min}}\; E\left\{ \sum_{i=1}^{m} c_i\left[\hat{\theta}_i(y), \theta_i\right] \right\}$$

$$= \underset{\hat{\theta}}{\operatorname{Arg\,Min}} \left\{ \sum_{i=1}^{m} E\left\{ c_i\left(\hat{\theta}_i(y), \theta_i\right) \right\} \right\}$$

با توجه به اینکه تابع هزینه $c_i(\hat{\theta}_i , \theta_i)$ ، غیرمنفی است در نتیجه امید ریاضی
آن نیز غیرمنفی است و در مسئله بالا ، می‌خواهیم حاصل جمع m مقدار غیرمنفی را

مینیمم کنیم که معادل مینیمم کردن هر یک از عبارت‌های

$E\{C_i[\hat{\theta}_i(y),\theta_i]\}$ (یا به طور معادل مینیمم کردن $E\{C_i[\hat{\theta}_i,\theta_i]|y\}$)

است. بنابراین جواب تخمین بیزین را در حالت برداری می‌توان به صورت زیر نوشت

$$\underline{\hat{\theta}}(\underline{y}) = \begin{bmatrix} \hat{\theta}_{1B}(\underline{y}) \\ \hat{\theta}_{2B}(\underline{y}) \\ \vdots \\ \hat{\theta}_{mB}(\underline{y}) \end{bmatrix} = \begin{bmatrix} \operatorname{Arg\,Min} E\,C_1[\hat{\theta}_1,\theta_1] \\ \operatorname{Arg\,Min} E\,C_2[\hat{\theta}_2,\theta_2] \\ \vdots \\ \operatorname{Arg\,Min} E\,C_m[\hat{\theta}_m,\theta_m] \end{bmatrix}$$

$$\Longrightarrow \underline{\hat{\theta}}(\underline{y}) = \begin{bmatrix} \operatorname{Arg\,Min}\; E\{C_1[\hat{\theta}_1, \theta_1] \mid y\} \\ \operatorname{Arg\,Min}\; E\{C_2[\hat{\theta}_2, \theta_2] \mid y\} \\ \vdots \\ \operatorname{Arg\,Min}\; E\{C_m[\hat{\theta}_m, \theta_m] \mid y\} \end{bmatrix}$$

در ادامه می‌خواهیم تخمین‌های MMSE، MMAE، MAP
را در حالت برداری بررسی کنیم.

الف) تخمین MMSE در حالت برداری

همان طور که بیان کردیم، تابع هزینه در حالت MMSE، برابر مربع خطا است

بنابراین در حالت برداری داریم

$$C[\underline{a}, \underline{\theta}] = \sum_{i=1}^{m} (a_i - \theta_i)^2 = \|\underline{a} - \underline{\theta}\|^2$$

$$\hat{\theta}_{MMSE} = \underset{\hat{\underline{\theta}}}{\mathrm{Arg\,Min}}\ C[\hat{\underline{\theta}}, \underline{\theta}] = \underset{\hat{\underline{\theta}}}{\mathrm{Arg\,Min}}\ \|\hat{\underline{\theta}} - \theta\|^2$$

با توجه به نکات گفته شده ، جواب تخمین MMSE در حالت برداری به صورت زیر خواهد بود

$$\hat{\underline{\theta}}_{MMSE}(\underline{y}) = \begin{bmatrix} \hat{\theta}_{1_{MMSE}}(\underline{y}) \\ \hat{\theta}_{2_{MMSE}}(\underline{y}) \\ \vdots \\ \hat{\theta}_{m_{MMSE}}(\underline{y}) \end{bmatrix} = \begin{bmatrix} \text{Conditional Mean of } w(\theta_1|\underline{y}) \\ \text{Conditional Mean of } w(\theta_2|\underline{y}) \\ \vdots \\ \text{Conditional Mean of } w(\theta_n|\underline{y}) \end{bmatrix}$$

$$\Rightarrow \hat{\theta}_{MMSE}(\underline{y}) = \begin{bmatrix} E\,\theta_1|\underline{y} \\ E\,\theta_2|\underline{y} \\ \vdots \\ E\,\theta_m|\underline{y} \end{bmatrix} = E\,\underline{\theta}|\underline{y} =$$

Conditional Mean of $w(\underline{\theta}|\underline{y})$

ب) تخمین MMAE

در حالت MMAE ، تابع هزینه برابر قدرمطلق خطا است

$$C(\underline{a},\underline{\theta}) = \sum_{i=1}^{m} |a_i - \theta_i|$$

با توجه به نکات گفته شده ، جواب تخمین MMAE در حالت برداری به صورت زیر قابل بیان است.

$$\hat{\underline{\theta}}_{MMAE}(\underline{y}) = \begin{bmatrix} \hat{\theta}_{1_{MMAE}}(\underline{y}) \\ \hat{\theta}_{2_{MMAE}}(\underline{y}) \\ \vdots \\ \hat{\theta}_{m_{MMAE}}(\underline{y}) \end{bmatrix} = \begin{bmatrix} \text{Conditional Median of } w(\theta_1|\underline{y}) \\ \text{Conditional Median of } w(\theta_2|\underline{y}) \\ \vdots \\ \text{Conditional Median of } w(\theta_m|\underline{y}) \end{bmatrix}$$

$$= \text{Conditional Median of } w(\underline{\theta}|\underline{y})$$

ج) تخمین MAP در حالت برداری

همان طور که بالا داریم ، در حالت MAP ، تابع هزینه ، تابع هزینه یکنواخت است.

$$C(\underline{a},\underline{\theta}) = \sum_{i=1}^{m} C(a_i,\theta_i)$$

که در آن

$$C(a_i,\theta_i) = \begin{cases} 1 & |a_i-\theta_i| > \Delta \\ 0 & |a_i-\theta_i| \leq \Delta \end{cases}$$

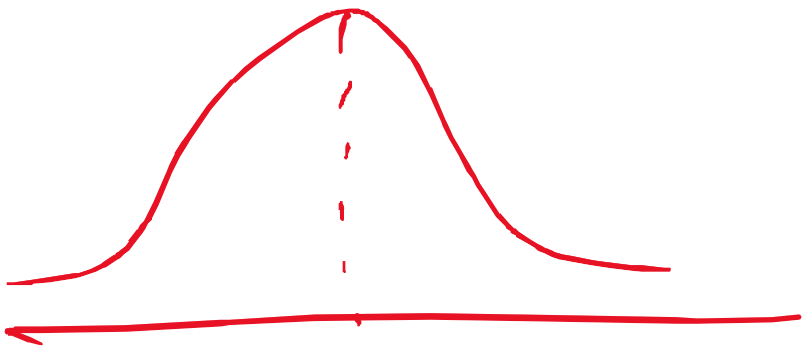
$$\Longrightarrow \hat{\theta}_{MAP}(\underline{y}) = \begin{bmatrix} \hat{\theta}_{1_{MAP}}(\underline{y}) \\ \hat{\theta}_{2_{MAP}}(\underline{y}) \\ \vdots \\ \hat{\theta}_{m_{MAP}}(\underline{y}) \end{bmatrix} = \begin{bmatrix} \text{Conditional Mode of } w(\theta_1|\underline{y}) \\ \text{Conditional Mode of } w(\theta_2|\underline{y}) \\ \vdots \\ \text{Conditional Mode of } w(\theta_m|\underline{y}) \end{bmatrix}$$

$$= \begin{bmatrix} \operatorname{Arg\,max} w(\theta_1|\underline{y}) \\ \operatorname{Arg\,max} w(\theta_2|\underline{y}) \\ \vdots \\ \operatorname{Arg\,max} w(\theta_m|\underline{y}) \end{bmatrix} \not\equiv \underbrace{\operatorname{Arg\,Max} w(\underline{\theta}|\underline{y})}_{\text{MAP جواب تخمین}}$$

$$f_{X_i}(x_i) \longrightarrow \max$$

$$\neq$$

$$\max f_{\underline{X}}(\underline{x})$$



$$F_X(x) = P_r\{X \le x\}, \quad f_X(x) = \frac{\partial}{\partial x} F_X(x)$$

$$F_{\underline{X}}(\underline{x}) = P_r\{X_1 \le x_1, \dots, X_n \le x_n\}, \quad f_{\underline{X}}(\underline{x}) = \frac{\partial}{\partial \underline{x}} F_{\underline{X}}(\underline{x})$$

همان طور که می‌بینیم، جواب به دست آمده با جواب تخمین MAP در حالت برداری همخوانی ندارد.

جواب تخمین MAP در حالت برداری برابر $\text{Arg Max } \omega(\underline{\theta} \mid \underline{y})$ است. در صورتی که اگر ماکزیمم هر یک از توابع چگالی احتمال نشانی $\omega(\theta_i \mid \underline{y})$ را به دست بیاوریم و بردار حاصل را به عنوان جواب تخمین بهترین در نظر بگیریم با نقطه‌ی ماکزیمم $\omega(\underline{\theta} \mid \underline{y})$ همخوانی ندارد. (دو بردار معادل نیستند)

بنابراین برای اینکه با روش بیزین به جواب تخمین MAP برسیم، لازم است که تابع هزینه را در حالت برداری به صورت زیر تعمیم دهیم (از ماکزیمم استفاده کنیم)

$$C(\underline{a}, \underline{\theta}) = \begin{cases} 1 & \max |a_i - \theta_i| > \Delta \\ 0 & \max |a_i - \theta_i| \leq \Delta \end{cases}$$

به عبارت دیگر تابع هزینه زمانی صفر است که تمامی $|a_i - \theta_i|_{i=1}^{m}$ ها از Δ کوچکتر باشند.

بنابراین تخمین بیزین به صورت زیر بیان خواهد شد

مساله

$$\hat{\underline{\theta}}_B(\underline{y}) = \operatorname{Arg\,Min} \; E\left\{ C(\hat{\underline{\theta}}(\underline{y}), \underline{\theta}) \,\middle|\, \underline{Y} = \underline{y} \right\}$$

$$= \operatorname{Arg\,Min} \left\{ \int \cdots \int_{\max_i |\hat{\theta}_i - \theta| > \Delta} w(\underline{\theta} \,|\, \underline{y}) \, d\underline{\theta} \right\}$$

$$= \operatorname{Arg\,Min} \left\{ 1 - \int \cdots \int_{\max_i |\hat{\theta}_i - \theta| \leqslant \Delta} w(\underline{\theta} \,|\, \underline{y}) \, d\underline{\theta} \right\}$$

$$\hat{\underline{\theta}}_B(\underline{y}) = \operatorname{Arg\,Min}\left\{ 1 - \underset{\forall i,\ |\hat{\theta}_i - \theta| \leq \Delta}{\int \cdots \int} w(\underline{\theta} \mid \underline{y})\, d\underline{\theta} \right\}$$

$$|\hat{\theta}_1 - \theta_1| \leq \Delta,\ |\hat{\theta}_2 - \theta_2| \leq \Delta,\ \dots,\ |\hat{\theta}_m - \theta_m| \leq \Delta$$

$$= \operatorname{Arg\,Min}\left\{ 1 - P_r\left\{ |\hat{\theta}_1 - \theta_1| \leq \Delta, \dots, |\hat{\theta}_m - \theta_m| \leq \Delta \right\} \right\}$$

$$= \operatorname{Arg\,Min}\left\{ 1 - \Delta^m\, w(\underline{\theta} \mid \underline{y}) \right\}$$

تعمیم مقدار میانگین

برای مینیمم شدن عبارت قبل، لازم است که نقطه‌ی ماکزیمم $w(\underline{\theta} \mid \underline{y})$ را پیدا کنیم.

$$\Longrightarrow \hat{\theta}_{MAP-B}(\underline{y}) = \underset{\underline{\theta}}{\text{Arg Max}} \; w(\underline{\theta} \mid \underline{y})$$