به نام زندگی

تخمین Baysian درحالت برداری

$$\underline{\hat{\theta}}_{MMSE} = E\{\underline{\theta} \mid \underline{y} = \underline{y}\} = \text{Conditional Mean } w(\underline{\theta} \mid \underline{y})$$

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

$$\hat{\theta}_{MAP} = \text{Conditional Mode of } w(\underline{\theta} \mid \underline{y})$$

تابع هزینه‌ی متداول دیگری که در حالت برداری از آن استفاده می‌شود، حاصلجمع وزن‌دار خطای مربعی است که به صورت زیر بیان می‌شود.

$$C[\underline{a},\underline{\theta}] = (\underline{a}-\underline{\theta})^T A (\underline{a}-\underline{\theta})$$

که در آن A یک ماتریس متقارن و معین مثبت است.

$$C(\underline{a},\underline{\theta}] = (\underline{a}-\underline{\theta})^T A (\underline{a}-\underline{\theta}) > 0$$

A متقارن و معین مثبت است.

$(\underline{a}-\underline{\theta})^T A (\underline{a}-\underline{\theta})$ را فرم درجه دوم بردار $\underline{a}-\underline{\theta}$ با ماتریس ضرایب A می‌گوییم که همواره یک اسکالر مثبت است

$$(\underline{a}-\underline{\theta})^T A (\underline{a}-\underline{\theta}) = \sum_i \sum_j (a_i-\theta_i)\, a_{ij}\, (a_j-\theta_j)$$

که در آن

$$A = [a_{ij}]$$

در واقع به خطا در المان های مختلف بردار $\underline{\theta}$ ، وزن های مختلفی داده می‌شود.

به طور مثال اگر خطا در یک بعد برای ما حساس تر باشد، می توانیم وزن بیشتری به خطا در آن بعد بدهیم.

در ادامه می خواهیم جواب تخمین بیزین را با این تابع هزینه به دست بیاوریم.

می دانیم که

$$\hat{\underline{\theta}}_B = \arg\min_{\hat{\underline{\theta}}} E\left\{ C[\hat{\underline{\theta}}(\underline{y}), \underline{\theta}] \mid \underline{y} = \underline{y} \right\}$$

از طرف دیگر

$$E\{C[\hat{\underline{\theta}}(\underline{y}),\underline{\theta}] \mid \underline{y}=\underline{y}\} =$$

$$= E\{(\hat{\underline{\theta}}(\underline{y})-\underline{\theta})^T A(\hat{\underline{\theta}}(\underline{y})-\theta) \mid \underline{y}=\underline{y}\}$$

$$= E_{\theta|y}\{\hat{\underline{\theta}}^T(\underline{y}) A \hat{\underline{\theta}}(\underline{y}) + \underline{\theta}^T A \underline{\theta} - 2\hat{\underline{\theta}}^T(y) A \underline{\theta} \mid \underline{y}=\underline{y}\}$$

$$-\underline{\theta}^T A \hat{\underline{\theta}}(y) - \hat{\underline{\theta}}^T(y) A \underline{\theta} = -2\hat{\underline{\theta}}^T(y) A \underline{\theta}$$

$$\xrightarrow[E(\cdot)]{\text{خطی بردن}} E\left\{ C\left[\underline{\hat{\theta}}(\underline{y}), \underline{\theta}\right] \mid \underline{y} = \underline{y} \right\}$$

$$= \underline{\hat{\theta}}^T(\underline{y})\, A\, \underline{\hat{\theta}}(\underline{y}) + E\left\{\underline{\theta}^T A \underline{\theta} \mid \underline{y} = \underline{y}\right\} - 2\, \underline{\hat{\theta}}^T(\underline{y})\, A\, E\left\{\underline{\theta} \mid \underline{y} = \underline{y}\right\}$$

$$= f\left(\underline{\hat{\theta}}(\underline{y})\right)$$

همان طور که می بینیم ، عبارت الا یک عبارت درجه دوم برحسب $\underline{\hat{\theta}}(\underline{y})$ است
با ماتریس ضرایب A که معین مثبت است. بنابراین عبارت الا (تابع هزینه پسین)
دارای یک می نیمم است که با مشتق گیری برداری از عبارت و مساوی صفر قرار دادن آن

به دست می آید.

$$\frac{\partial}{\partial \hat{\underline{\theta}}(\underline{y})} f\left(\hat{\underline{\theta}}(\underline{y})\right) \equiv \nabla_{\hat{\underline{\theta}}(\underline{y})} f\left(\hat{\underline{\theta}}(\underline{y})\right) = 0$$

یادآوری :

$$\frac{\partial}{\partial \underline{z}} \underline{z}^T A \underline{z} = 2 A \underline{z}$$

$$\frac{\partial}{\partial \underline{z}} \underline{z}^T A \underline{v} = A \underline{v}$$

$$\Longrightarrow \frac{\partial}{\partial \hat{\underline{\theta}}(\underline{y})} f\left(\hat{\underline{\theta}}(\underline{y})\right) = 2A\hat{\underline{\theta}}(\underline{y}) + 0 - 2A\,E\{\underline{\theta}\,|\,\underline{y}=y\} \Bigg|_{\hat{\underline{\theta}}_B(y)} = 0$$

جواب تخمین

$$\Longrightarrow$$

بزن

$$A\,\hat{\underline{\theta}}(\underline{y}) \Bigg|_{\hat{\underline{\theta}}_B(\underline{y})} = A\,E\{\underline{\theta}\,|\,\underline{y}=\underline{y}\} \Bigg|_{\hat{\underline{\theta}}_B(\underline{y})}$$

$$\xrightarrow[A^{-1}\ \text{وجود دارد}]{A\ \text{معکوس پذیر}} \qquad \underline{\hat{\theta}}_B(\underline{y}) = E\{\underline{\theta} \mid \underline{Y}=\underline{y}\} \equiv \underline{\hat{\theta}}_{MMSE}(\underline{y})$$

همان طور که می بینیم جواب تخمین بیزین در این حالت با جواب MMSE یکسان است و ماتریس ضرایب A در جواب تخمین تاثیری ندارد.

اما اگر خطای تخمین را حساب کنیم، تاثیر ماتریس A در تابع هزینه را خواهیم دید.

می‌دانیم که فضای تخمین‌زن به صورت زیر است.

$$r\left(\underline{\hat{\theta}}(\underline{y})\right)=E_{\underline{y}}\left\{\overbrace{E_{\theta|y}\left\{C\left[\underline{\hat{\theta}}(\underline{y}),\underline{\theta}\right]\,\middle|\,\underline{y}=y\right\}}^{\text{در حالت کلی - تابعی از } \underline{y}=y \text{ است.}}\right\}$$

$$\underline{\hat{\theta}}_B(y)$$

$$=E\left\{E\left\{\left(\underline{\hat{\theta}}(y)-\underline{\theta}\right)^T A\left(\underline{\hat{\theta}}(y)-\underline{\theta}\right)\,\middle|\,\underline{y}=y\right\}\right\}$$

$$=E\left\{E\left\{\underbrace{\left(\underline{\theta}-E\underline{\theta}|\underline{y}\right)^T A\left(\underline{\theta}-E\underline{\theta}|\underline{y}\right)}_{\mathrm{trace}\left(\left(\underline{\theta}-E\underline{\theta}|\underline{y}\right)^T A\left(\underline{\theta}-E\underline{\theta}|\underline{y}\right)\right)}\,\middle|\,\underline{y}=y\right\}\right\}$$

$$\mathrm{trace}(AB) = \mathrm{trace}(BA)$$

در حالت طلوبی از $\underline{y}=y$ است.

$$\Longrightarrow r(\hat{\underline{\theta}}(\underline{y})) = E_y\left\{ \mathrm{trace}\left(A\, \mathrm{Cov}(\underline{\theta} \mid \underline{y}=y) \right) \right\}$$

$$= \mathrm{trace}\left\{ A\, E_y\left(\mathrm{Cov}(\underline{\theta} \mid y=y) \right) \right\}$$

همان ضرب در ماتریس ضرایب A در رابطه‌ی خطای تخمین تاثیرگذار است.

* اگر $A=I$ باشد، تخمین بیزین دقیقاً معادل تخمین MMSE است.

$$C(\underline{a},\underline{\theta}) = (\underline{a}-\underline{\theta})^T A (\underline{a}-\underline{\theta}) = (\underline{a}-\underline{\theta})^T (\underline{a}-\underline{\theta}) = \|\underline{a}-\underline{\theta}\|^2$$

مثال: تخمین یک بردار گوسی $\underline{\theta}$ از روی مشاهدات توأماً گوسی $\underline{y}$

$$\Lambda = \mathbb{R}^m, \quad \Gamma = \mathbb{R}^n \quad \longrightarrow \quad \underline{\theta}, \hat{\underline{\theta}} \in \Lambda, \quad \underline{y} \in \Gamma$$

$$\underline{y}_{n\times 1} \quad , \quad \underline{\theta}_{m\times 1}, \; \hat{\underline{\theta}}_{m\times 1}$$

$\underline{\theta}$ یک بردار گوسی است و $\underline{\theta}$ و $\underline{y}$ بردارهای تصادفی توأماً گوسی هستند.

بنابراین $\underline{y}$ نیز یک بردار تصادفی گوسی است. همچنین توابع احتمال شرطی $p(\underline{y}|\underline{\theta})$ و $p(\underline{\theta}|\underline{y})$

نیز دارای فرم گوسی هستند.

(یادآوری از فرآیندهای تصادفی: اگر $\underline{x}$ و $\underline{y}$ به طور توأماً گوسی باشند، $f_{\underline{x}}(\underline{x})$

و $f_{\underline{y}}(\underline{y})$ و $f_{\underline{x}}(\underline{x}|\underline{y})$ و $f_{\underline{y}}(\underline{y}|\underline{x})$ دارای فرم گوسی هستند)

$$\begin{bmatrix} \underline{y} \\ \underline{\theta} \end{bmatrix} \sim N\left(\begin{bmatrix} \underline{\mu}_{\underline{y}} \\ \underline{\mu}_{\underline{\theta}} \end{bmatrix}, \begin{bmatrix} \Sigma_{\underline{y}} & \Sigma_{\underline{y}\underline{\theta}} \\ \Sigma_{\underline{\theta}\underline{y}} & \Sigma_{\underline{\theta}} \end{bmatrix} \right)$$

$$E\,\underline{y} = \underline{m}_y = \underline{\mu}_y$$

$$E\,\underline{\theta} = \underline{m}_\theta = \underline{\mu}_\theta$$

$$\mathrm{Cov}\,\underline{y} = E\,(\underline{y} - \underline{\mu}_y)(\underline{y} - \underline{\mu}_y)^H = \Sigma_y = C_y$$

$$\mathrm{Cov}\,\underline{\theta} = E\,(\underline{\theta} - \underline{\mu}_\theta)(\underline{\theta} - \underline{\mu}_\theta)^H = \Sigma_\theta = C_\theta$$

$$C_{\underline{\theta}\underline{y}} = E\,(\underline{\theta} - \underline{\mu}_\theta)(\underline{y} - \underline{\mu}_y)^H = \Sigma_{\theta y} = \Sigma_{y\theta}^H = \mathrm{Cov}(\underline{\theta}, \underline{y}) = C_{\underline{y}\theta}^H$$

الف) تخمین MMSE

می دانیم که

$$\hat{\underline{\theta}}_{MMSE}(\underline{y}) = E\{\underline{\theta} \mid \underline{Y} = \underline{y}\}$$

برای محاسبه ی $\hat{\underline{\theta}}_{MMSE}$ کافی است $w(\underline{\theta} \mid \underline{y})$ را به دست بیاوریم

(آنچه چگالی احتمال گرسی) و با استفاده از رابطه

$$E\{\underline{\theta} / \underline{Y} = \underline{y}\} = \int_{\underline{\theta}} \underline{\theta}\, w(\underline{\theta} \mid \underline{y})\, d\underline{\theta}$$

امید ریاضی شرطی را محاسبه کنیم. اما با توجه به اینکه می دانیم $w(\theta|y)$
دارای فرم گوسی است، می توانیم از روابط تابع چگالی احتمال شرطی
گوسی، روابط را بازنویسی کنیم. (θ, y توأماً گوسی هستند بنابراین
$p(y|\theta)$ و $w(\theta|y)$ نیز دارای فرم گوسی هستند)

یادآوری از مرآمند: اگر x و y دو متغیر تصادفی توأماً گوسی باشند، $f_X(x|y)$ نیز دارای خم گوسی است و برای آن داریم:

$$m_{x|y} = \mu_{x|y} = \mu_x + \frac{\sigma_{xy}}{\sigma_y^2}(y-\mu_y) = \mu_x + Cov(x,y)\,Var^{-1}(y)(y-\mu_y)$$

$$\sigma_{x|y}^2 = Var(x|y) = (1-\rho^2)\sigma_x^2 = Var(x) - Cov(x,y)\,Var^{-1}(y)\,Cov(y,x)$$

$$\rho = \frac{\sigma_{xy}}{\sigma_x \sigma_y}$$

در حالت برداری ، $\underline{X}$ و $\underline{Y}$ دو بردار تصادفی توأماً گوسی

$$m_{\underline{x}|\underline{y}} = \mu_{\underline{x}|\underline{y}} = \mu_x + \Sigma_{xy}\,\Sigma_y^{-1}\,(\underline{y}-\underline{\mu}_y)$$

تابعی از $\underline{y}=\underline{y}$

$$\Sigma_{x|y} = Cov(x|y) = \Sigma_x - \Sigma_{xy}\,\Sigma_y^{-1}\,\Sigma_{yx}$$

در حالت توأماً گوسی تابعیت از $\underline{y}=\underline{y}$ ندارد

$$\Rightarrow \quad \hat{\underline{\theta}}_{MMSE}(\underline{y}) = E\{\underline{\theta} \mid \underline{Y} = \underline{y}\}$$

$$\Rightarrow \quad \hat{\underline{\theta}}_{MMSE}(\underline{y}) = \underline{\mu}_{\theta} + \Sigma_{\theta Y} \Sigma_{Y}^{-1} (\underline{y} - \underline{\mu}_{y})$$

تابعی از $\underline{Y}=\underline{y}$ است

خطای تخمین MMSE نیز از رابطه‌ی زیر قابل محاسبه است.

$$r(\hat{\underline{\theta}}_{MMSE}(\underline{y})) = \mathrm{trace}\left(\underbrace{E_{Y}\{Cov(\underline{\theta} \mid \underline{Y} = \underline{y})\}}_{\text{در حالت کلی تابعی از } \underline{Y}=\underline{y} \text{ است}} \right)$$

حالت توأماً گوسی

$$\Longrightarrow \quad r(\hat{\theta}_{MMSE}(\underline{y})) = \underset{\text{trace}}{\Big(} E\Big\{ \underbrace{\Sigma_\theta - \Sigma_{\theta y}\Sigma_y^{-1}\Sigma_{y\theta}}_{\text{تابعیت از } \underline{y} \text{ = } \underline{y} \text{ ندارد}} \Big\}\Big)$$

$$\Longrightarrow r(\hat{\theta}_{mms}(\underline{y})) = \mathrm{trace}\left(\Sigma_\theta - \Sigma_{\theta y}\Sigma_y^{-1}\Sigma_{y\theta}\right)$$

به عنوان یک مثال عملی از حالتی که بردار $\underline{\theta}$ و مشاهدات $\underline{y}$ توأماً گوسی هستند می‌توان به مسأله‌ی زیر اشاره کرد.

$$\underline{y} = H\underline{\theta} + \underline{N} \qquad \Lambda = \mathbb{R}^m \,, \quad \Gamma = \mathbb{R}^n$$

$$\underset{m\times 1}{\underline{\theta}} \in \Lambda \,, \quad \underset{n\times 1}{\underline{y}} \in \Gamma \,, \quad H_{n\times m}$$

$\underline{N} \sim N(\underline{0}, \Sigma_N)$ $\underline{N}$ بردار نویز جمع‌شونده گوسی است.

و $\underline{\Theta}$ بردار پارامترهای مورد نظر برای تخمین است و داریم

$$\underline{\Theta} \sim N(\underline{\mu}_{\Theta}, \Sigma_{\Theta}) \quad , \quad \underline{\Theta} \perp\!\!\!\perp \underline{N}$$

به علت وجود رابطه‌ی خطی $\underline{Y} = H\underline{\Theta} + \underline{N}$ و

استقلال $\underline{\Theta}$ و $\underline{N}$ ، بردار تصادفی $\underline{Y}$ نیز یک بردار تصادفی گوسی است

و بردارهای تصادفی $\underline{\Theta}$ و $\underline{Y}$ نیز توأماً گوسی هستند.

$$H_{n\times m}=\begin{bmatrix} h_{11} & h_{12} & \cdots & h_{1m}\\ h_{21} & h_{22} & \cdots & h_{2m}\\ \vdots & & & \\ h_{n1} & h_{n2} & \cdots & h_{nm}\end{bmatrix}_{n\times m}=\left[h_{ij}\right]_{n\times m}$$

ماتریس H می‌تواند ماتریس یک کانال چندتیک یا یک کانال چندورودی-چندخروجی باشد. یا در حالت کلی می‌تواند یک فیلتر باشد.

برای اینکه جواب تخمین MMSE را به دست بیاوریم، کافی است که $w(\underline{\theta}|\underline{y})$ را به دست بیاوریم و $E\{\underline{\theta}|\underline{Y}=\underline{y}\}$ را محاسبه کنیم. می‌دانیم که $\underline{\theta}$ و $\underline{Y}$ بردار تصادفی توأماً گوسی هستند، بنابراین روابط مورد نیاز برای محاسبه تخمین MMSE را در اختیار داریم.

$$\hat{\underline{\theta}}_{MMSE}(\underline{y}) = E\{\underline{\theta}|\underline{Y}=\underline{y}\} = m_{\underline{\theta}|\underline{Y}} = \mu_{\underline{\theta}|\underline{Y}}$$

$$= \mu_{\theta} + \Sigma_{\theta Y}\Sigma_{Y}^{-1}(\underline{y} - \underline{\mu}_{Y})$$

بنابراین کافی است $\underline{\mu}_y$ ، $\Sigma_{\theta y}$ ، Σ_y را بدست بیاوریم:

$$\underline{Y} = H\underline{\theta} + \underline{N} \longrightarrow \mathbb{E}\,\underline{Y} = \mathbb{E}\left(H\underline{\theta} + \underline{N}\right) = H\,\mathbb{E}\,\underline{\theta} + \underbrace{\mathbb{E}\,\underline{N}}_{\underline{0}}$$

$$\Rightarrow \underline{\mu}_y = H\underline{\mu}_\theta$$

$$\Sigma_{\theta y} = \mathbb{E}\left(\underline{\theta} - \underline{\mu}_\theta\right)\left(\underline{y} - \underline{\mu}_y\right)^H = \mathbb{E}\left(\underline{\theta} - \underline{\mu}_\theta\right)\left(H\underline{\theta} + \underline{N} - H\underline{\mu}_\theta\right)^H$$

$$= \mathbb{E}\left(\underline{\theta} - \underline{\mu}_\theta\right)\left(H\left(\underline{\theta} - \underline{\mu}_\theta\right) + \underline{N}\right)^H =$$

$$\Rightarrow \Sigma_{\theta y} = E\left((\underline{\theta} - \mu_\theta)(\underline{\theta} - \mu_\theta)^H H^H + (\underline{\theta} - \mu_\theta)\underline{N} \right)$$

$$= \Sigma_\theta H^H + \underbrace{E(\theta - \mu_\theta)}_{\underline{0}} \underbrace{E N}_{\underline{0}}$$

$$\underline{\theta} \perp\!\!\!\perp \underline{N}$$

$$\Sigma_{\theta y} = \Sigma_\theta H^H$$

$$\Sigma_y = E\,(\underline{y} - \underline{\mu}_y)(\underline{y} - \underline{\mu}_y)^H$$

$$= E\,(\underbrace{H\underline{\theta} + \underline{N} - H\underline{\mu}_\theta}_{H(\underline{\theta} - \underline{\mu}_\theta)})(H\underline{\theta} + \underline{N} - H\underline{\mu}_\theta)^H$$

$$\Sigma_y = H\,E\,(\underline{\theta} - \underline{\mu}_\theta)(\underline{\theta} - \underline{\mu}_\theta)^H H^H + \overbrace{E\underline{N}}^{\underline{0}}\,\overbrace{E(\underline{\theta} - \underline{\mu}_\theta)^H}^{\underline{0}} H^H$$

$$+ H\underbrace{E\,(\underline{\theta} - \underline{\mu}_\theta)}_{\underline{0}}\underbrace{E\underline{N}^H}_{\underline{0}} + E\,\underline{N}\,\underline{N}^H = H\,\Sigma_\theta\,H^H + \Sigma_N$$

$$\Rightarrow \hat{\theta}_{MMSE}(\underline{y}) = \underline{\mu}_{\theta} + \Sigma_{\theta y}\Sigma_{y}^{-1}(\underline{y} - \underline{\mu}_{y})$$

$$\Sigma_{\theta y} = \Sigma_{\theta} H^{H}, \quad \Sigma_{y} = H\Sigma_{\theta}H^{H} + \Sigma_{N}$$

$$\mu_{y} = H\underline{\mu}_{\theta}$$

خطای تخمین MMSE

$$r(\hat{\theta}_{MMSE}(y)) = \operatorname{trace}\left(\Sigma_{\theta} - \Sigma_{\theta y}\Sigma_{y}^{-1}\Sigma_{y\theta}\right)$$

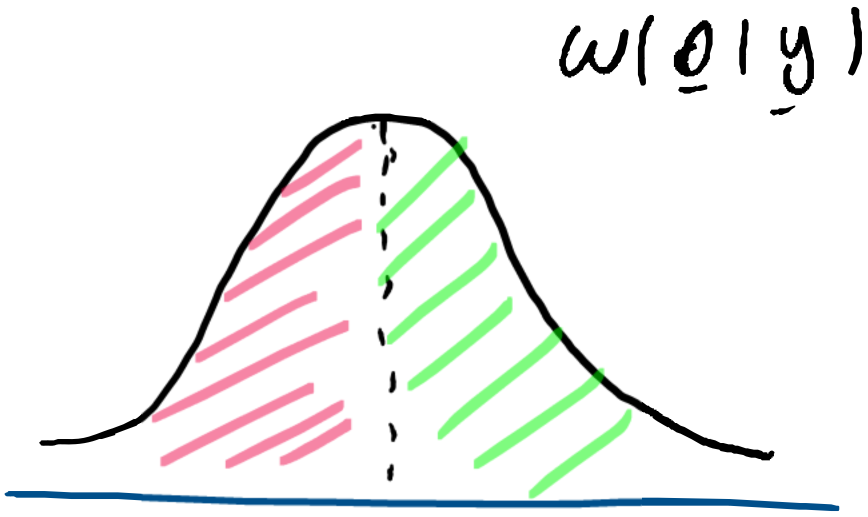
$$\Sigma_{Y\theta} = \Sigma_{\theta Y}^{H} = H\,\Sigma_{\theta}$$

ب) تخمین MMAE

با توجه به اینکه $W(\underline{\theta}|\underline{y})$ درحالیکه $\underline{y}$ و $\underline{\theta}$ توأماً گوسی هستند این تابع چگالی احتمال گوسی است و تقارن دارد، میانگین، میانه و Mode آن با هم منطبق هستند.

$$\hat{\theta}_{MMAE}(\underline{y}) = \hat{\theta}_{MMSE}(\underline{y}) = E(\underline{\theta}|\underline{y})$$ بنابراین داریم

$$\hat{\theta}_{MAP}(\underline{y}) = \hat{\theta}_{MMSE}(\underline{y}) = E\,\underline{\theta}\,|\,\underline{y}$$



Median of $w(\underline{\theta}|\underline{y}) = \mu_{\underline{\theta}|\underline{y}}$ = Mode of $w(\underline{\theta}|\underline{y})$