

Ορίζουμε τον a – posteriori κίνδυνο¹ για έναν κανόνα απόφασης $a = a(x)$ ως την μέση τιμή:

$$PR(a|x) = \mathbb{E}[L(\mathcal{G}, a)|x] = \int_{\Theta} L(\mathcal{G}, a) \Pi(d\mathcal{G}|x),$$

όπου $L(\mathcal{G}, a) \geq 0$ η συνάρτηση απώλειας², που είναι ένα μέτρο της απώλειας που επωμιζόμαστε όταν αντικαθιστούμε την πραγματική τιμή \mathcal{G} με την εκτίμηση a . Ζητάμε τότε την άριστη απόφαση a^* έτσι ώστε

$$PR(a^*|x) = \inf_{a \in \Theta} \mathbb{E}\{L(\mathcal{G}, a)|x\}.$$

Για παράδειγμα θέτοντας $L(\mathcal{G}, a) = (\mathcal{G} - a)^2$ (τετραγωνική συνάρτηση απώλειας) που είναι συμμετρική (ως προς την πραγματική τιμή \mathcal{G}) και κυρτή συνάρτηση του a έχουμε:

Πρόταση

Κάτω από την τετραγωνική συνάρτηση απώλειας $L(\mathcal{G}, a) = (\mathcal{G} - a)^2$, η κατά Bayes εκτίμηση του \mathcal{G} , είναι $a^* = a^*(x) = \mathbb{E}[\mathcal{G}|x]$, με ελάχιστη απώλεια $Var[\mathcal{G}|x]$.

$$\begin{aligned} \frac{\partial}{\partial a} PR(a|x) &= \frac{\partial}{\partial a} \int_{\Theta} (\mathcal{G} - a)^2 \Pi(d\mathcal{G}|x) = -2 \int_{\Theta} (\mathcal{G} - a) \Pi(d\mathcal{G}|x) \\ &= -2 \left\{ \int_{\Theta} \mathcal{G} \Pi(d\mathcal{G}|x) - a \int_{\Theta} \Pi(d\mathcal{G}|x) \right\} = -2 \{ \mathbb{E}[\mathcal{G}|x] - a \} = 0. \end{aligned}$$

$$\frac{\partial^2}{\partial a^2} PR(a|x) = -2 \frac{\partial}{\partial a} \int_{\Theta} (\mathcal{G} - a) \Pi(d\mathcal{G}|x) = 2 \int_{\Theta} \Pi(d\mathcal{G}|x) = 2,$$

και ελάχιστη απώλεια θα είναι:

$$PR(a^*|x) = PR(\mathbb{E}[\mathcal{G}|x]|x) = \int_{\Theta} \{ \mathcal{G} - \mathbb{E}[\mathcal{G}|x] \}^2 \Pi(d\mathcal{G}|x) = Var[\mathcal{G}|x].$$

Εναλλακτικά

¹ Posterior risk.

² Loss function.

$$\begin{aligned}
PR(a|x) &= \int_{\Theta} (\vartheta - a)^2 \Pi(d\vartheta|x) = \int_{\Theta} \vartheta^2 \Pi(d\vartheta|x) - 2a \int_{\Theta} \vartheta \Pi(d\vartheta|x) + a^2 \\
&= \mathbb{E}[\vartheta^2|x] - 2a\mathbb{E}[\vartheta|x] + a^2 = \mathbb{E}[\vartheta|x]^2 - 2a\mathbb{E}[\vartheta|x] + a^2 + \{\mathbb{E}[\vartheta^2|x] - \mathbb{E}[\vartheta|x]^2\} \\
&= \{a - \mathbb{E}[\vartheta|x]\}^2 + Var[\vartheta|x] \Rightarrow a^* = \mathbb{E}[\vartheta|x].
\end{aligned}$$

Πρόταση

Η κατά τόπους γραμμική (piecewise linear) συνάρτηση απώλειας:

$$L(\vartheta, a) = \begin{cases} k_0(a - \vartheta), & a > \vartheta \\ k_1(\vartheta - a), & a \leq \vartheta \end{cases}, k_0 > 0, k_1 > 0,$$

δίνει ποινή $k_0(a - \vartheta)$ όταν $a > \vartheta$ και $k_1(\vartheta - a)$ όταν $a \leq \vartheta$. Η **άριστη επιλογή** a^* είναι το

$\frac{k_1}{k_0 + k_1}$ – **ποσοστιαίο σημείο³ της posterior**. Δηλαδή το a^* με την ιδιότητα

$$P\{\vartheta \leq a^* | x\} = \frac{k_1}{k_0 + k_1}.$$

Για να το δούμε αυτό θεωρούμε την διαφορά:

$$L(\vartheta, a^*) - L(\vartheta, a) = \begin{cases} k_0(a^* - \vartheta), & a^* > \vartheta \\ k_1(\vartheta - a^*), & a^* \leq \vartheta \end{cases} - \begin{cases} k_0(a - \vartheta), & a > \vartheta \\ k_1(\vartheta - a), & a \leq \vartheta \end{cases}$$

$$= \begin{cases} k_0(a^* - \vartheta), & \vartheta < a^* < a \\ k_1(\vartheta - a^*), & a^* < \vartheta < a \\ k_1(\vartheta - a^*), & a^* < a < \vartheta \\ k_0(a^* - \vartheta), & \vartheta < a < a^* \\ k_0(a^* - \vartheta), & a < \vartheta < a^* \\ k_1(\vartheta - a^*), & a < a^* < \vartheta \end{cases} - \begin{cases} k_0(a - \vartheta), & \vartheta < a^* < a \\ k_0(a - \vartheta), & a^* < \vartheta < a \\ k_1(\vartheta - a), & a^* < a < \vartheta \\ k_0(a - \vartheta), & \vartheta < a < a^* \\ k_1(\vartheta - a), & a < \vartheta < a^* \\ k_1(\vartheta - a), & a < a^* < \vartheta \end{cases}$$

³ Percentile.

$$= \left\{ \begin{array}{ll} k_0(a^* - a), & \mathcal{G} < a^* < a \\ k_1 \underbrace{(\mathcal{G} - a^*)}_{< a - a^*} + k_0 \underbrace{(\mathcal{G} - a)}_{< 0}, & a^* < \mathcal{G} < a \\ k_1(a - a^*), & a^* < a < \mathcal{G} \\ k_0(a^* - a), & \mathcal{G} < a < a^* \\ k_0 \underbrace{(a^* - \mathcal{G})}_{< a^* - a} + k_1 \underbrace{(a - \mathcal{G})}_{< 0}, & a < \mathcal{G} < a^* \\ k_1(a - \mathcal{G}), & a < a^* < \mathcal{G} \end{array} \right\} \leq \left\{ \begin{array}{ll} k_0(a^* - a), \mathcal{G} < a^* < a \\ k_1(a - a^*), a^* < \mathcal{G} < a \\ k_1(a - a^*), a^* < a < \mathcal{G} \\ k_0(a^* - a), \mathcal{G} < a < a^* \\ k_0(a^* - a), a < \mathcal{G} < a^* \\ k_1(a - a^*), a < a^* < \mathcal{G} \end{array} \right\} = \left\{ \begin{array}{ll} k_0(a^* - a), \mathcal{G} < a^* \\ k_1(a - a^*), \mathcal{G} > a^* \end{array} \right\}$$

Δηλαδή δείξαμε ότι:

$$L(\mathcal{G}, a^*) - L(\mathcal{G}, a) \leq \left\{ \begin{array}{ll} k_0(a^* - a), \mathcal{G} < a^* \\ k_1(a - a^*), \mathcal{G} > a^* \end{array} \right\} = k_0(a^* - a)1(\mathcal{G} < a^*) + k_1(a - a^*)1(\mathcal{G} > a^*).$$

Παίρνοντας την μέση τιμή ως προς το posterior μέτρο $\Pi(d\mathcal{G} | x)$, η προηγούμενη ανισότητα δίνει

$$\begin{aligned} PR(a^* | x) - PR(a | x) &= \int_{\Theta} (L(\mathcal{G}, a^*) - L(\mathcal{G}, a)) \Pi(d\mathcal{G} | x) \\ &\leq \int_{\Theta} \left\{ k_0(a^* - a)1(\mathcal{G} < a^*) + k_1(a - a^*)1(\mathcal{G} > a^*) \right\} \Pi(d\mathcal{G} | x) \\ &= \int_{\mathcal{G}=-\infty}^{a^*} k_0(a^* - a) \Pi(d\mathcal{G} | x) + \int_{\mathcal{G}=a^*}^{\infty} k_1(a - a^*) \Pi(d\mathcal{G} | x) \\ &= k_0(a^* - a) \int_{\mathcal{G}=-\infty}^{a^*} \Pi(d\mathcal{G} | x) + k_1(a - a^*) \int_{\mathcal{G}=a^*}^{\infty} \Pi(d\mathcal{G} | x) \\ &= k_0(a^* - a) P\{\mathcal{G} \leq a^* | x\} + k_1(a - a^*) (1 - P\{\mathcal{G} \leq a^* | x\}) \\ &= (a^* - a) \frac{k_1 k_0}{k_0 + k_1} + k_1(a - a^*) \frac{k_1 k_0}{k_0 + k_1} = 0. \end{aligned}$$

Ισχύει λοιπόν ότι

$$L(\mathcal{G}, a^*) \leq L(\mathcal{G}, a), \forall a \in \Theta.$$

Θέτοντας $k_0 = k_1 = k$ παίρνουμε $P\{\mathcal{G} \leq a^* | x\} = \frac{1}{2}$ δηλαδή το a^* είναι η a – posteriori διάμεσος (posterior median).

Η εύρεση του βέλτιστου a^* όταν η συνάρτηση απώλειας είναι γραμμική μπορεί να γίνει και με παραγωγήσιμη. Έστω $k_0 = k_1 = k$ τότε $L(\mathcal{G}, a) = k|\mathcal{G} - a|, k > 0$.

$$\begin{aligned} \frac{\partial}{\partial a} PR(a | x) &= \frac{\partial}{\partial a} \int_{\Theta} k|\mathcal{G} - a| \Pi(d\mathcal{G} | x) \\ &= k \left\{ \frac{\partial}{\partial a} \int_{\mathcal{G}=a}^{\infty} (\mathcal{G} - a) \Pi(d\mathcal{G} | x) - \frac{\partial}{\partial a} \int_{\mathcal{G}=-\infty}^a (\mathcal{G} - a) \Pi(d\mathcal{G} | x) \right\} \\ &= k \left\{ \lim_{\xi \rightarrow \infty} \frac{\partial}{\partial a} \int_{\mathcal{G}=a}^{\xi} (\mathcal{G} - a) \Pi(d\mathcal{G} | x) - \lim_{\xi \rightarrow -\infty} \frac{\partial}{\partial a} \int_{\mathcal{G}=\xi}^a (\mathcal{G} - a) \Pi(d\mathcal{G} | x) \right\} \end{aligned}$$

Επειδή

$$\frac{\partial}{\partial a} \int_{A(a)}^{B(a)} \varphi(a, \mathcal{G}) d\mathcal{G} = \int_{A(a)}^{B(a)} \frac{\partial \varphi(a, \mathcal{G})}{\partial a} d\mathcal{G} + \varphi(a, B(a)) \frac{dB(a)}{da} - \varphi(a, A(a)) \frac{dA(a)}{da},$$

παίρνουμε

$$\frac{\partial}{\partial a} \int_{\mathcal{G}=a}^{\xi} (\mathcal{G} - a) \Pi(d\mathcal{G} | x) = - \int_{\mathcal{G}=a}^{\xi} \Pi(d\mathcal{G} | x) + 0 - 0,$$

$$\frac{\partial}{\partial a} \int_{\mathcal{G}=\xi}^a (\mathcal{G} - a) \Pi(d\mathcal{G} | x) = - \int_{\mathcal{G}=\xi}^a \Pi(d\mathcal{G} | x) + 0 - 0.$$

Και έτσι

$$\begin{aligned} \frac{\partial}{\partial a} PR(a | x) &= k \left\{ - \int_{\mathcal{G}=a}^{\infty} \Pi(d\mathcal{G} | x) + \int_{\mathcal{G}=-\infty}^a \Pi(d\mathcal{G} | x) \right\} = 0 \\ \Rightarrow \int_{\mathcal{G}=-\infty}^a \Pi(d\mathcal{G} | x) &= \int_{\mathcal{G}=a}^{\infty} \Pi(d\mathcal{G} | x) \Leftrightarrow P\{\mathcal{G} > a | x\} = P\{\mathcal{G} < a | x\}. \end{aligned}$$

Δηλαδή το a^* είναι η a – posteriori διάμεσος (posterior median), εφόσον

$$\begin{aligned} \frac{\partial^2}{\partial a^2} PR(a|x) &= k \frac{\partial}{\partial a} \left\{ - \int_{\mathcal{G}=a}^{\infty} \pi(\mathcal{G}|x) d\mathcal{G} + \int_{\mathcal{G}=-\infty}^a \pi(\mathcal{G}|x) d\mathcal{G} \right\} \\ &= k \left\{ -(0+0-\pi(a|x)) + (0+\pi(a|x)-0) \right\} = 2k\pi(a|x) > 0. \end{aligned}$$

Ενώ έχουμε ότι

$$\begin{aligned} \frac{\partial^2}{\partial a^2} PR(a|x) &= k \frac{\partial}{\partial a} \left\{ - \int_{\mathcal{G}=a}^{\infty} \pi(\mathcal{G}|x) d\mathcal{G} + \int_{\mathcal{G}=-\infty}^a \pi(\mathcal{G}|x) d\mathcal{G} \right\} \\ &= k \left\{ -(0+0-\pi(a|x)) + (0+\pi(a|x)-0) \right\} = 2k\pi(a|x) > 0 \end{aligned}$$

Άσκηση

Δείξτε την γενική περίπτωση χρησιμοποιώντας παραγωγήση.

$$\frac{\partial}{\partial a} PR(a|x) = \frac{\partial}{\partial a} \int_{\Theta} \{k_0(a-\mathcal{G})1(a>\mathcal{G}) + k_1(\mathcal{G}-a)1(a\leq\mathcal{G})\} \Pi(d\mathcal{G}|x)$$

$$= k_0 \int_{\mathcal{G}=-\infty}^a \Pi(d\mathcal{G}|x) - k_1 \int_{\mathcal{G}=a}^{\infty} \Pi(d\mathcal{G}|x) = 0$$

$$\Rightarrow k_0 P\{\mathcal{G} < a | x\} = k_1 P\{\mathcal{G} \geq a | x\} \Rightarrow \begin{cases} k_0 P\{\mathcal{G} < a | x\} = k_1 (1 - P\{\mathcal{G} < a | x\}) \\ k_0 (1 - P\{\mathcal{G} \geq a | x\}) = k_1 P\{\mathcal{G} \geq a | x\} \end{cases}$$

$$\Rightarrow \begin{cases} P\{\mathcal{G} < a | x\} = \frac{k_1}{k_0 + k_1} \\ P\{\mathcal{G} \geq a | x\} = \frac{k_0}{k_0 + k_1} \end{cases}$$

Δηλαδή το a^* είναι το a – posteriori $\frac{k_1}{k_0 + k_1}$ ποσοστιαίο σημείο.

$$\frac{\partial^2}{\partial a^2} PR(a|x) = \frac{\partial}{\partial a} \left\{ k_0 \int_{\mathcal{G}=-\infty}^a \pi(\mathcal{G}|x) d\mathcal{G} - k_1 \int_{\mathcal{G}=a}^{\infty} \pi(\mathcal{G}|x) d\mathcal{G} \right\} = (k_0 + k_1) \pi(a|x) > 0.$$

Η 0-1 συνάρτηση απώλειας.

Κάτω από τη συνάρτηση απώλειας

$$L(\vartheta, a) = \lim_{\varepsilon \rightarrow 0^+} \mathcal{I}(|\vartheta - a| \geq \varepsilon) = 1 - \lim_{\varepsilon \rightarrow 0^+} \mathcal{I}(|\vartheta - a| < \varepsilon) = 1 - \delta(\vartheta - a),$$

η κατά Bayes σημειακή εκτίμηση a^* του ϑ είναι το posterior mode (maximum a-posteriori MAP estimate)

$$a^* = a^*(x) = \max_{\vartheta \in \Theta} \pi(\vartheta | x).$$

Πράγματι

$$PR(a^* | x) = \inf_{a \in \Theta} \int_{\Theta} L(\vartheta, a) \pi(\vartheta | x) d\vartheta = \inf_{a \in \Theta} \int_{\Theta} \{1 - \delta(\vartheta - a)\} \pi(\vartheta | x) d\vartheta = \inf_{a \in \Theta} \{1 - \pi(a | x)\}$$

από όπου $a^* = \max_{a \in \Theta} \pi(a | x)$.

Έστω τώρα ότι $L_\varepsilon(\vartheta, a) = \mathcal{I}(|\vartheta - a| \geq \varepsilon) = 1 - \mathcal{I}(|\vartheta - a| < \varepsilon)$ για μικρό ε , τότε

$$PR_\varepsilon(a^* | x) = \inf_{a \in \Theta} \int_{\Theta} L_\varepsilon(\vartheta, a) \pi(\vartheta | x) d\vartheta = \inf_{a \in \Theta} \int_{\Theta} \{1 - \mathcal{I}(|\vartheta - a| < \varepsilon)\} \pi(\vartheta | x) d\vartheta$$

$$= 1 - \sup_{a \in \Theta} \int_{\Theta} \mathcal{I}(|\vartheta - a| < \varepsilon) \pi(\vartheta | x) d\vartheta = 1 - \sup_{a \in \Theta} \int_{a-\varepsilon}^{a+\varepsilon} \pi(\vartheta | x) d\vartheta,$$

όπου $\int_{a-\varepsilon}^{a+\varepsilon} \pi(\vartheta | x) d\vartheta = P\{|\vartheta - a| < \varepsilon | x\}$.