[4] - MAP prediction when costs are equal

$$\text{optimal } j^* = \underset{j}{\text{argmax}}\ P(j \mid x)$$

$$= \underset{j}{\text{argmax}}\ \frac{\Pi_j \cdot f_j(x)}{\Pi_1 f_1(x) + \Pi_2 f_2(x) + \cdots + \Pi_K f_K(x)}$$

$$= \underset{j}{\text{argmax}}\ \frac{\Pi_j \cdot f_j(x)}{Z(x)}$$

important

$$= \underset{j}{\text{argmax}}\ \Pi_j \cdot f_j(x)$$

where $Z(x) = \Pi_1 f_1(x) + \Pi_2 f_2(x) + \cdots + \Pi_K f_K(x)$, which does not depend on the label j.

[5] Let $c(i \mid j)$ = cost of saying i when the truth is j.

- If our policy is to predict i whenever we get a particular data vector x, then the expected cost on x-like vectors is:

$$\mathbb{E}(\text{cost} \mid X = x) = \sum_{j=1}^{K} c(i \mid j) P(j \mid x)$$

$$= \sum_{j=1}^{K} c(i \mid j) \frac{\Pi_j \cdot f_j(x)}{Z(x)}$$

- Hence optimal $j^* = \underset{j=1\ldots K}{\text{argmin}} \sum_{i=1}^{K} c(j \mid i) \Pi_i f_i(x)$

* note how we drop $Z(x)$ again *