5.5. Multiclass logistic regression

$$P(Y = y \mid X, w_{1:c}) = \frac{\exp(w_y^T x)}{\underbrace{\sum_{y'} \exp(w_{y'}^T x)}_{Z(w, x)}}$$

Predict most likely class

$$\hat{y}(x) = \underset{y \in \{1..c\}}{\arg\max} \frac{\exp(w_y^T x)}{Z(w, x)}$$

$$= \underset{y}{\arg\max} \; \exp(w_y^T x)$$

$$= \underset{y}{\arg\max} \; w_y^T x \qquad \Rightarrow \text{linear decision boundaries}$$

However, level sets

$$S_{y, z} = \{ x : P(Y = y \mid X = x, w_{1:c}) = z \}$$

is in general non-linear