

# Cross-Entropy-Loss

dl勉強会@Osk

10/5

経営学部1年 村上広樹

# 今回の発表で取り扱うこと

- 損失関数とは
- 損失関数の種類
- Cross-entropy lossの概要
- Cross-entropy lossの優れている点
- Cross-entropy lossとSoftmax関数との関係

# 損失関数とは

- 学習に使われる関数
- 予測した結果と正解データとの差を求める
- 損失関数の値を減らすように学習する

# 損失関数の種類

L1 (絶対誤差)

$$L1LossFunction = \sum_{i=1}^n |y_{true} - y_{predicted}|$$

L2 (二乗誤差)

$$L2LossFunction = \sum_{i=1}^n (y_{true} - y_{predicted})^2$$

---

Cross-entropy loss  
(交差エントロピー誤差)

次のスライドで説明

# Cross-entropy loss

予測した結果に対して、対数をとる

予測した  
結果

正解ラベ  
ル

正解ラベ  
ル

予測した  
結果

$$L_{\text{cross-entropy}}(\hat{\mathbf{y}}, \mathbf{y}) = - \sum_i y_i \log(\hat{y}_i)$$

分類問題で、正解のクラス  
を $y_k$ とすると

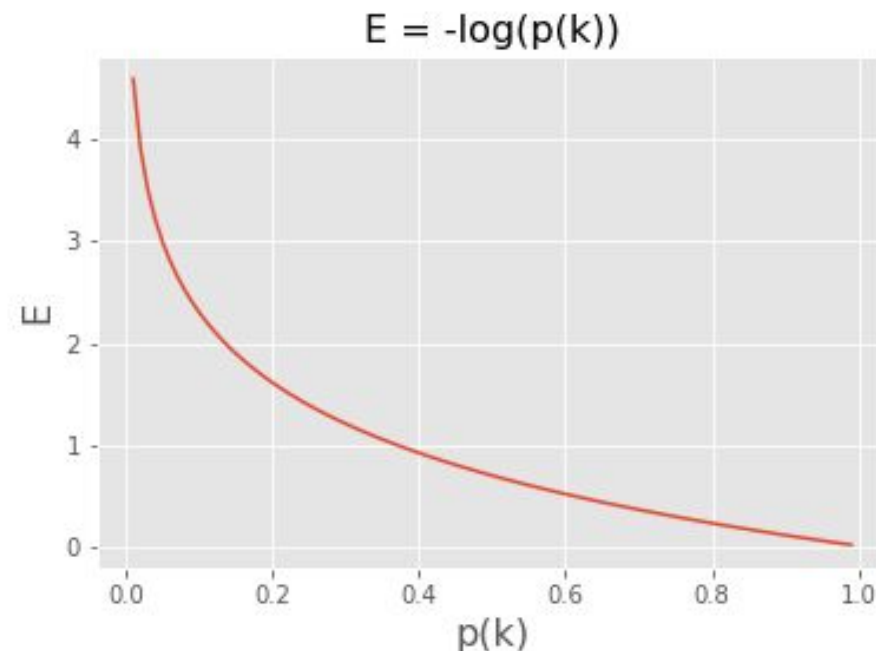
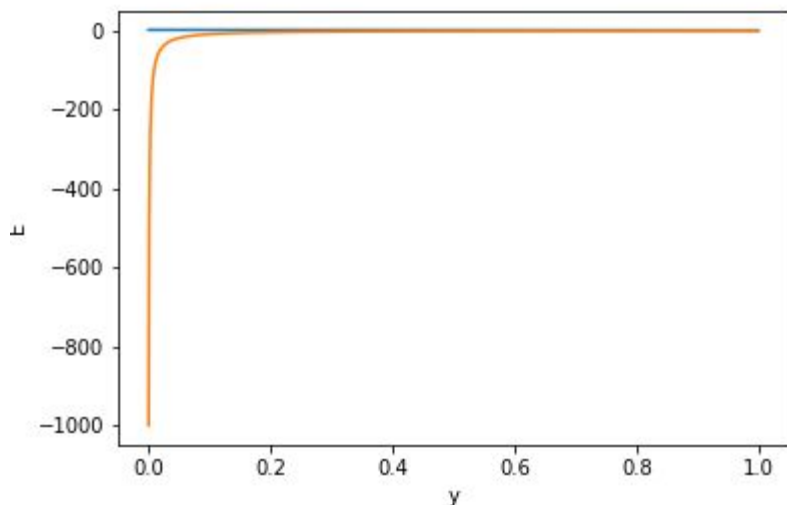
$$= - \left( 1 * \log(y_{k_c}) + 0 * \sum_{k \neq k_c} \log(y_k) \right) = - \log(y_{k_c})$$

# Cross-entropy lossの優れている点

## 学習が早い

なぜ学習が早い？

→大きく外れて予測したときに、Cross-entropy lossを微分した値が大きくなり、一回当たりの損失関数の減少幅が大きくなるから



# Cross-entropy lossとSoftmax関数との関係

相性◎ 微分して掛けたときに式がきれいになる☆彡

Cross-entropy lossにsoftmaxを代入したものを微分してみる

$$\text{式}) - \log(\exp(a_k) / \sum \exp(a_i))$$

$$= -\log(\exp(a_k)) + \log(\sum \exp(a_i))$$

$$= \log(\sum \exp(a_i)) - \log(\exp(a_k))$$

$$= \log(\sum \exp(a_i)) - a_k \quad \leftarrow \text{これをsoftmaxの各要素}(a_k)\text{で微分!}$$

$$\rightarrow 1 / \sum \exp(a_i) * \exp(a_k) - a_k$$

$$= y_k - a_k$$

# まとめ

- Cross-entropy lossの特徴

- ①学習が早い

- ②Softmax関数と相性がいい

このような理由から、分類問題でよく使われています。