Lecture 2: logistic regression and cross-entropy

Helen Yannakoudakis and Oana Cocarascu

Department of Informatics
King's College London

Logistic regression update rule

- Use the logistic function as a threshold: $h_{\mathbf{w}}(\mathbf{x}) = \frac{1}{1+e^{-\mathbf{w}\cdot\mathbf{x}}}$
- Update rule: $w_i \leftarrow w_i - \alpha \frac{\partial}{\partial w_i} Loss(\mathbf{w})$
- Equations for a single training example:

$$\frac{\partial}{\partial w_i}(y - h_{\mathbf{w}}(\mathbf{x}))^2$$

$$2\,(y - h_{\mathbf{w}}(\mathbf{x})) \times \frac{\partial}{\partial w_i}(y - h_{\mathbf{w}}(\mathbf{x}))$$

$$-2\,(y - h_{\mathbf{w}}(\mathbf{x})) \times g'(\mathbf{w}\cdot\mathbf{x}) \times x_i$$

where g' is the derivative of the logistic function.

- Update rule becomes:

$$w_i \leftarrow w_i + \alpha(y - h_w(x)).h_w(x)(1 - h_w(x)).x_i$$

Cross-entropy

Logistic regression commonly uses the **cross-entropy loss** function.
(note: the update rule in Lecture 2, slide 79 uses the L2 loss function – see previous slide here).

$p(y|x) = \hat{y}^{y}\,(1-\hat{y})^{1-y}$ where $y \in \{0,1\}$ and $\hat{y} = h(x)$

Rewrite as:

$$log\, p(y|x) = y\, log\, \hat{y} + (1-y)\, log\,(1-\hat{y})$$

Flip sign to turn into loss:

$$-log\, p(y|x) = -[y\, log\, \hat{y} + (1-y)\, log\,(1-\hat{y})]$$