Formulas that you are expected to remember

Leandro Minku

March 11, 2024

In general, I do not ask students to memorise formulas for the exam. Rather, students are expected to demonstrate that they are able to understand the formulas. However, some formulas directly represent the core ideas of their underlying approaches. Therefore, if you understand the ideas, you should be able to remember the corresponding formulas. According to that, this document lists the formulas from Weeks 1–4 and 7 that you are expected to remember by heart for the ML exam.

1 Logistic Regression

$$\text{logit}(p_1) = \ln\left(\frac{p_1}{1-p_1}\right) = \mathbf{w}^T\mathbf{x}$$

$$p_y = p(y|\mathbf{x}, \mathbf{w})$$

If $\mathbf{w}^T\mathbf{x} \geq 0$, predict class 1. Otherwise, predict class 0.

$$\text{logit}(p_1) = \ln\left(\frac{p_1}{1-p_1}\right) = \mathbf{w}^T\phi(\mathbf{x})$$

$$p_y = p(y|\phi(\mathbf{x}), \mathbf{w})$$

If $\mathbf{w}^T\phi(\mathbf{x}) \geq 0$, predict class 1. Otherwise, predict class 0.

$$p_0 = 1 - p_1$$

$$\mathcal{L}(\mathbf{w}) = \prod_{i=1}^{N} p_{y^{(i)}}$$

$$\ln(\mathcal{L}(\mathbf{w})) = \sum_{i=1}^{N} \ln\, p_{y^{(i)}}$$

$$E(\mathbf{w}) = -\ln(\mathcal{L}(\mathbf{w}))$$

2 Gradient Descent and IRLS

$$\mathbf{w} = \mathbf{w} - \eta \nabla E(\mathbf{w})$$

$$\mathbf{w} = \mathbf{w} - H_E^{-1}(\mathbf{w}) \nabla E(\mathbf{w})$$

PS: you do not need to remember the equation corresponding to the gradient and Hessian of the cross entropy loss function, just the general equations above.

3 SVM and Soft Margin SVM

$$h(\mathbf{x}) = \mathbf{w}^T \mathbf{x} + b$$

$$h(\mathbf{x}) = \mathbf{w}^T \phi(\mathbf{x}) + b$$

$$h(\mathbf{x}) = \sum_{n \in S} a^{(n)} y^{(n)} k(\mathbf{x}, \mathbf{x}^{(n)}) \ + b$$

When using the linear kernel, $k(\mathbf{x}, \mathbf{x}^{(n)}) = \mathbf{x}^T \mathbf{x}^{(n)}$.

If $h(\mathbf{x}) > 0$, predict class +1. If $h(\mathbf{x}) < 0$, predict class -1.

PS: you do not need to remember the formula for calculating b in the dual representation.