統計學習初論（105-2）

Self-Assessment Homework 3B

國立台灣大學資管系

Question 1 (Exercise 4.17)

4.17 (⋆) **www** Show that the derivatives of the softmax activation function (4.104), where the a_k are defined by (4.105), are given by (4.106).

Question 2 (Exercise 4.18)

4.18 (⋆) Using the result (4.91) for the derivatives of the softmax activation function, show that the gradients of the cross-entropy error (4.108) are given by (4.109).

Question 3 (Exercise 4.19)

4.19 (⋆) **www** Write down expressions for the gradient of the log likelihood, as well as the corresponding Hessian matrix, for the probit regression model defined in Section 4.3.5. These are the quantities that would be required to train such a model using IRLS.