Exercise Sheet #7

due on Monday, 13 December 2021, 10am sharp

Last week we thought about linear regression which deals with real valued quantities. This week, week's exercise sheet is concerned with logistic regression which deals with binary quantities. In the practical part we also extend our knowledge on regression by applying regularization techniques to multivariate regression problems.

1. **EXAMple: Logistic regression.**
   (a) Show that $\sigma(s) = 1/(1+\exp(-s)) = \exp(s)/(1+\exp(s))$
   (b) Show that the logistic function satisfies $\sigma(-s) + \sigma(s) = 1$
   (c) Show that the first derivative of $\sigma(s)$, $\sigma'(s) = \sigma(s)(1-\sigma(s))$
   (d) Plot $\sigma(s)$ as well as $\log(\sigma(s))$ as a function of s (either using Python or with pen and paper, a rough plot which captures the qualitative features of the functions is sufficient)
   (e) Explain why, for large $s > 0$, $\log(\sigma(s)) \approx 0$ and $\log(\sigma(-s)) \approx -s$.
   (f) The loss function for logistic regression is given by $E(\mathrm{w}) = -\sum_i^N y_i \log \sigma\left(\mathrm{w}^\top \mathrm{x}_i\right) + (1-y_i)\log\left(1-\sigma\left(\mathrm{w}^\top \mathrm{x}_i\right)\right)$. Calculate $\nabla_{\mathrm{w}} E(\mathrm{w})$.

2. **Theory Question: Connection between logistic regression and linear discriminant analysis.**

   Suppose that we have data from two classes, and the data within each class is Gaussian distributed with the same covariance, i.e. $x|t=1 \sim \mathcal{N}(\mu_+, \Sigma)$ and $x|t=-1 \sim \mathcal{N}(\mu_-, \Sigma)$, and that the two classes have the same prior probabilities $\pi_+ = P(t=+1) = \pi_- = P(t=-1) = 0.5$.

   (a) Show that the conditional probability of belonging to the positive class can be written as a logistic function $P(t=1|x) = \sigma(\omega^\top x + \omega_o)$ and identify the corresponding parameters ω and ω_o.
   (b) Now, assume that the two covariances are not the same, but different, i.e. given by Σ_+ and Σ_-. What is the conditional probability in that case? Can you think of an extension of logistic regression that could be applied in that case?

3. **Practical Question:** You can find this week's practical exercise in `Exercise_07.ipynb`