CS 480/680 Homework 3

Released Jan 28, 2026

Due Feb 4, 2026 - Before 11:59PM EST

©Marina Meilă
mmp@stat.washington.edu

Problem 1 – Bias and Variance Again

The questions in this problem refer to **Problem 2 from Homework 1**.

a. Consider all the quantities you are asked to calculate or plot in Homework 1, Problem 2 (e.g., $\hat{l}_b, L_b, \hat{l}, V, \ldots$). List 2 of these which are *statistical approximations*.

Note: For example, computing a mean from samples is a *statistical approximation*, whereas computing an integral by discretization is a *numerical approximation*. We assume that all floating point computations and function calls are exact.

Explain (1 line or less) in each case what is (are) the approximation(s) made.

b. For one of your answers above, explain how you could increase the approximation accuracy.

c. [**680 only**]Assume that in Homework 1, Problem 2, $\tilde{n} \to \infty$. Will the error bars on $L \to 0$? Explain (1 line).

d. Assume that in Homework 1, Problem 2, $n \to \infty$. Will the error bars on $\hat{l} \to 0$? Explain (1 line).

[Problem 2 – Logistic Regression Basic Properties – NOT GRADED]

Notations follow the `lgood2-linear.pdf` *course notes.*

We consider the following definitions for the Sigmoid function $\sigma(u)$, the Logit function $\psi(\theta)$, and the Logistic Regression model:

$$\text{Sigmoid:} \quad \sigma(u) = \frac{1}{1+e^{-u}}$$
$$\text{Logit:} \quad \psi(\theta) = \ln(1+e^{\theta})$$
$$\text{Model:} \quad \ln \frac{P[Y=1|X=x]}{P[Y=0|X=x]} = f(x) \equiv \beta^T \mathbf{x}$$

Here $Y \in \{0,1\}$.

a. Symmetry: Prove that $\sigma(-u) = 1-\sigma(u)$. Using this, show that $\sigma'(-u) = \sigma'(u)$.

b. Calculus Properties:

- **(i)** Show that $\text{argmax}_{u\in\mathbb{R}}\, \sigma'(u) = 0$.
- **(ii)** Show that $\max_u \sigma'(u) = 1/4$.
- **(iii)** Prove that if $u \to \pm\infty$, the derivative $\sigma'(u) \to 0$.

Interpretation: Based on result (iii), explain why correctly classified data points away from the boundary have little influence on the log-likelihood. (i.e., Why do the parameters β change little when points far away from the decision boundary move?)

c. Alternative Form: Verify that $\sigma(u) = \frac{e^{u/2}}{e^{u/2}+e^{-u/2}} = \frac{e^u}{1+e^u}$.

d. Normalization: Let the probability mass function be $P_\theta(y) = e^{\theta y - \psi_0(\theta)}$, where $\psi_0(\theta) = \ln Z(\theta)$ and $Z(\theta)$ is the normalization constant. Show that $\psi_0 = \psi$ (the logit function defined above).

e. [680 only] Expectation: Calculate $\mathbb{E}_\beta[Y]$ and identify it with one of the three functions defined at the start of the problem. $\mathbb{E}_\beta[Y]$ is the expectation of Y under the model $P[Y|X=x]$ defined by the logistic regression. Note that when $Y \in \{0,1\}$, the expectation is equal to the probability that $Y=1$.

Problem 3 – Implementing Logistic Regression

In this question, you will implement logistic regression from scratch. Download the following files into your working directory. Note that these files should in the same directory.

- `hw3q3.py`: The script where you will implement the main optimization algorithm.

- `hw3q3vis.py`: The script where you will implement the visualization functions.
- `hw3_X_logistic.dat`: Data matrix $X \in \mathbb{R}^{n\times 2}$, where n is the number of samples.
- `hw3_y_logistic.dat`: Label vector $\mathbf{y} \in \{-1, 1\}^n$.
- `hw3q3.sh`: A shell script to automate the execution. You should be able to execute this script with no errors after finishing the problem. **Do not modify this file.**

a. The logistic regression model you will train is

$$f(x_1, x_2) \; = \; \begin{bmatrix}\beta_1 & \beta_2\end{bmatrix}\begin{bmatrix}x_1 \\ x_2\end{bmatrix} + \beta_0 \tag{1}$$

You will fit a parameter vector $\boldsymbol{\beta} = \begin{bmatrix}\beta_0 \\ \beta_1 \\ \beta_2\end{bmatrix} \in \mathbb{R}^3$. What is the vector $\mathbf{x} \in \mathbb{R}^3$ so that $f(x_1, x_2) = \boldsymbol{\beta}^T\mathbf{x}$?

b. Write down the mathematical expression for the Log-likelihood $\ell(\boldsymbol{\beta})$ and its gradient $\nabla_{\boldsymbol{\beta}}\ell$.

c. Implement and run the logistic regression model on the dataset given in files `hw3q3.py`. Use **Gradient Ascent** for optimization. Terminate the training when the relative change in the log-likelihood is less than 10^{-3}, i.e.,

$$\frac{|\ell_{\text{new}} - \ell_{\text{old}}|}{|\ell_{\text{old}}|} \leq 10^{-3}$$

Report the learning rate η you selected. In how many iterations did the training terminate?

d. Report the final fitted values of $\boldsymbol{\beta}$.

Create a visualization of the training process in `hw3q3vis.py` by implementing the functions listed below. **Note:** Ensure your previous script is named exactly `hw3q3.py` so it can be imported correctly.

e. Complete the function `plot_ll`, and use it to create a plot of the log-likelihood ℓ and one of the magnitude of the change in log-likelihood $|\Delta\ell|$ with respect to the number of iterations. Include these plots here.

f. Complete the function `plot_betas`, and use it to create a plot of the parameter values $\beta_0, \beta_1, \beta_2$ with respect to the number of iterations (on the same graph). Include this plot here.

g. Complete the function `plot_gradients`, and use it to create a plot of the gradient components $\frac{\partial \ell}{\partial \beta_0}, \frac{\partial \ell}{\partial \beta_1}, \frac{\partial \ell}{\partial \beta_2}$ with respect to the number of iterations. Include this plot here.

h. Complete the function `plot_decision_boundary`, and use it to create a scatter plot of the data samples (color-coded by class -1 and 1) and overlay the final linear decision boundary. Include this plot here.

Instructions for Code Submission:
Ensure you can run `./hw3q3.sh` without errors. This script should be able to create a zip file named `hw3q3-submission.zip` containing only `hw3q3.py` and `hw3q3vis.py`. **Submit** `hw3q3-submission.zip` **to LEARN.**