

HW06: Linear models

Hand in via moodle at: <https://moodle.umass.edu/course/view.php?id=20836>. Remember that only PDF submissions are accepted. We encourage using L^AT_EX to produce your writeups. See `hw00.tex` for an example of how to do so. You can make a `.pdf` out of the `.tex` by running “`pdflatex hw00.tex`”. You’ll need `mydefs.sty` and `notes.sty` which can be downloaded from the course page.

1. What values of λ in the regularized loss objective (Slide #4) will lead to overfitting? What values will lead to underfitting?
2. The solution to the least-squares regression problem involves inverting the matrix $(\mathbf{X}^T \mathbf{X} + \lambda \mathbf{I}_D)$ which might not be invertible. Is this actually a problem?
3. Explain why the squared loss is not suitable for binary classification problems.
4. One disadvantage of the squared loss is that it has a tendency to be dominated by outliers – the overall loss $\sum_n (y_n - \hat{y}_n)^2$, is influenced too much by points that have high $|y_n - \hat{y}_n|$. Suggest a modification to the squared loss that remedies this.