INFO251 – Applied Machine Learning

Lab 5

Emily Aiken

Announcements

- PS4 posted, due Monday March 14. Start early!
- Quiz 1 solutions in lecture tomorrow
- Midsemester lab feedback: https://forms.gle/uCFCfpmDH791bhSG8

Today's Topics

1. Steps for training an ML algorithm from scratch
2. Common loss functions
3. Practice
   - Bivariate OLS, squared error loss
   - Multivariate OLS, squared error loss
   - Multivariate OLS, squared error loss with Ridge regularization
4. Cross validation for optimal regularization parameter

Training a Parametric ML Algorithm

1. Define a model
2. Define a loss function
3. Optionally add regularization to the loss function
4. Calculate partial derivatives
5. Run gradient descent

1. Define a Model

- Choosing which model to use for prediction is entirely up to you.
- For today: **Linear regression models**, of the form $y = ax + b$
  - Multivariate models: $y = ax_1 + bx_2 + cx_3 + d$
  - Nonlinearities: $y = ax^k + b$
  - Interaction terms: $y = ax_1x_2 + b$

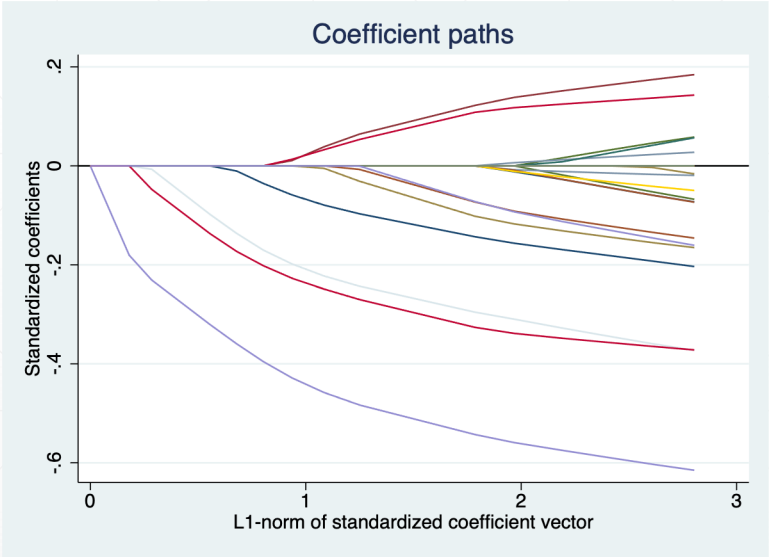
2. Define a Loss Function

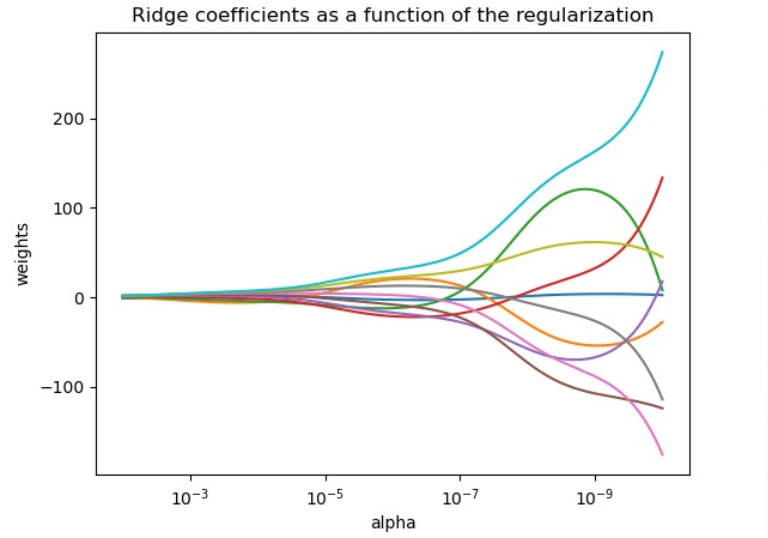
- The choice of loss function is also entirely up to you!
- Common loss functions for **regression**:
  - Squared error loss: $J(y, \hat{y}) = (y - \hat{y})^2$
  - Absolute error loss: $J(y, \hat{y}) = |y - \hat{y}|$
- Common loss functions for **binary classification**:
  - Logistic loss: $J(y, \hat{p}) = -(y \log(\hat{p}) + (1 - y)\log(1 - \hat{p}))$
  - Hinge loss: $J(y, \hat{p}) = \max(0, 1 - \hat{p}y)$
- Common loss functions for **multivariate classification**:
  - Cross-entropy loss: $J(y, \hat{p}) = \sum_{c=1}^{M} y_c \log(\hat{p}_c))$

3. Optionally Add Regularization to the Loss

- Regularization? Also up to you.
- **LASSO**: $J(\theta) \mathrel{+}= \|\theta\|_1 = \sum_{j=1}^{k} |\theta_k|$

Ridge: $J(\theta) \mathrel{+}= \|\theta\|_2 = \sum_{j=1}^{k} \theta_k^2$





4. Partial Derivatives and 5. Gradient Descent

- Implement as discussed in last week's lab.

Model: Univariate least squares

Cost: Squared error

1. Define the model
2. Define the loss function
3. ~~Optionally add regularization to the loss function~~
4. Calculate partial derivatives
5. Write psuedocode

Model: Multivariate least squares

Cost: Squared error

1. Define the model
2. Define the loss function
3. ~~Optionally add regularization to the loss function~~
4. Calculate partial derivatives
5. Write psuedocode

Model: Multivariate least squares

Cost: Squared error + Ridge regularization

1. Define the model
2. Define the loss function
3. Optionally add regularization to the loss function
4. Calculate partial derivatives
5. Write psuedocode