Chapter 3

Simon Ward-Jones
simonwardjones16@gmail.com

August 24, 2020

Link to chapter

1 Linear Regression

- Linear model
- Mini batch stochastic gradient descent
- Vectorisation of algorithms leeds to massive efficiency gains
- Minimizing the mean squared error is equivalent to maximum likelihood estimation of a linear model under the assumption of additive Gaussian noise

2 Linear Regression Implementation from Scratch

- Iterating the dataset in batches
- Defining the model and the loss
- Initialising params
- Finding the gradient of the loss w.r.t. the params
- Defining the optimisation algorithm that updates the params with the gradients
- Finally with using the above define the training routine looping through epochs
- – getting mini batch
  – find loss
  – optimizer step
  – print progress

3 Concise Implementation of Linear Regression

- Similar as above but use predefined dataset, model, loss and optimiser classes.

4 Softmax Regression

- Softmax transforms logit to probability space
- Log likelihood
- Cross-entropy measures the difference between two probability distributions

5 The Image Classification Dataset

- Fashion-MNIST is an apparel classification dataset consisting of images representing 10 categories
- Rely on well-implemented data iterators that exploit high-performance computing to avoid slowing down your training loop.

References

Zhang, A., Lipton, Z. C., Li, M., & Smola, A. J. (2020). *Dive into deep learning.* (`https://d2l.ai`)