Machine Learning Exercises 1 - Corrected Version
Due: September 18

Tim van Erven

September 13, 2007

Abstract

Two very different exercises. The first is to expose you to some mathematics. The second is more of a conceptual nature, and touches upon the fundamental problem in machine learning of how to represent your data in a machine learning task.

1 Distributivity of Matrix Multiplication

Suppose A and B are arbitrary $k \times m$ matrices and C and D are arbitrary $m \times n$ matrices. Recall that the $k \times n$ matrix product AC is defined component-wise as:

$$(AC)_{ij} = \langle \mathbf{a}_i, \mathbf{c}_j \rangle = \sum_{l=1}^{m} A_{il} C_{lj} \quad \text{for } i = 1, \ldots, k, j = 1, \ldots, n,$$

where $\mathbf{a}_i^\top$ is the ith row vector of matrix A and $\mathbf{c}_j$ is the jth column vector of matrix C.

Exercise 1. *Show that matrix multiplication is distributive.* That means you have to show that

1. $(A + B)C = AC + BC$, and
2. $A(C + D) = AC + AD$.

Hint 1: Both cases can be proved in very similar ways. Your proof of the first case might have the following outline:

1. Show that the matrix that is the result of $(A + B)C$ has the same dimensions as the matrix that is the result of $AC + BC$. In other words, if $(A + B)C$ is an $l \times p$ matrix, then $AC + BC$ should also be an $l \times p$ matrix.
2. Now you need to show that all components of $(A+B)C$ and $AC+BC$ are identical. This can be done by proving that the matrices are identical in an arbitrary component:

   $$((A + B)C)_{ij} = (AC + BC)_{ij} \quad \text{for arbitrary } i \text{ and } j.$$

   This can be done by writing out both sides of the equation using their definitions and then doing some algebraic manipulations.

Hint 2: Try with vectors instead of matrices first if you're having difficulties or work out an example with numbers.

2 Supervised Learning Problems

Exercise 2. *Give an example (not from class or from the book) of each of the following learning problems:*

- *prediction,*
- *classification,*
- *regression.*

How would you represent the labels and feature vectors (if applicable) mathematically? For either the classification or regression case (your choice), give at least two possible representations of your feature vectors.

Your examples do not need to be very realistic, but they do need to be about objects in the real world. For example, the answer "let $\mathbb{R}^6$ be the set of possible feature vectors" is not sufficient. You need to say what the feature vectors represent.

3 Comment on Sources

In general webresources do not meet scientific standards. Nevertheless, I've consulted the ones below during the preparation of these exercises.

References

[1] J. Kolter. Linear algebra review and reference. Available online: `http://www.stanford.edu/class/cs229/section/cs229-linalg.pdf`, October 2006.