

MCS 549 – Foundations of Data Science  
Fall 2019  
Problem Set 3

Lev Reyzin

**Due:** 11/22/19 at the beginning of class

**Instructions:** Atop your problem set, please write your name and list your collaborators.

## Problems

1. For matrices  $A$  and  $B$  prove that

$$AB = \sum_{k=1}^n A(:, k)B(k, :).$$

2. Give an example of a set  $H$  of hash functions such that  $h(x)$  is equally likely to be any element of  $\{0, \dots, M - 1\}$  but  $H$  is not 2-universal. Prove your answer correct.

3. For the  $k$ -median and the  $k$ -means objectives, prove upper bounds on the ratio between the optimal value when we either require all cluster centers to be data points or allow arbitrary points to be centers.