 Jason J. Corso, jcorso@buffalo.edu, SUNY Buffalo

Name:

ID#:

Section: 455 or 555

2	8	10

Directions – *The quiz is closed book/notes. You have 10 minutes to complete it; use this paper only.*

Problem 1: Recall (2pts) (Answer in one sentence only.)

What quantity is PCA maximizing during dimension reduction?

Problem 2: Work (8 pts) (Show all derivations/work and explain.)

The problem of principal component analysis ultimately reduces to the eigenproblem:

$$Se = \lambda e$$

First, describe what the variables S, e and λ represent.

Now, given this problem setup for PCA, let's say you are given a very high-dimensional dataset to work with (on the order of 100,000 or even 1,000,000 features)—would it be a good idea to start your analysis by reducing the data's dimensionality with this algorithm? Why or why not?