

CISC 271 Class 37

Course Summary – Linear Data Analysis

In this course, we have touched on many technical matters related to linear algebra. We can summarize the themes of the course, briefly, by viewing linear data analysis as:

- Linear regression
- Principal Components Analysis (PCA), especially for dimensionality reduction
- Clustering, or unsupervised learning

Our main tool for exploring PCA was the singular value decomposition (SVD). This powerful result allowed us to understand that any matrix can be written as a series; dimensionality reduction is the truncation of this series. PCA is the SVD of zero-mean data, so the SVD can be used simply and directly for many purposes.

Linear data analysis in general, and linear algebra in particular, are abundantly found in current machine learning. Linear methods, combined with nonlinear optimization and statistical reasoning, are essential tools in understanding how to analyze data and how to learn from data.