Lecture 18: Dimensionality Reduction

Professor Ilias Bilionis

The dimensionality reduction problem

Unsupervised Learning

Your are given n observations:

$$\mathbf{x}_{1:n} = \{\mathbf{x}_1, \ldots, \mathbf{x}_n\}$$

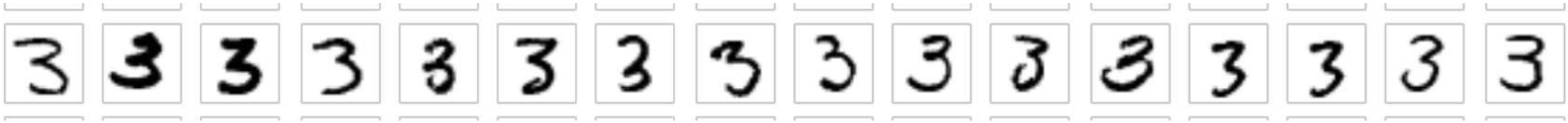
(inputs, features, …)

Assume that the observations are D dimensional ($D \gg 1$).

Problem: Describe the data using d dimensional variables $d \ll D$, $\mathbf{z}_{1:n} = \{\mathbf{z}_1, \ldots, \mathbf{z}_n\}$.

Example: MNIST 3s

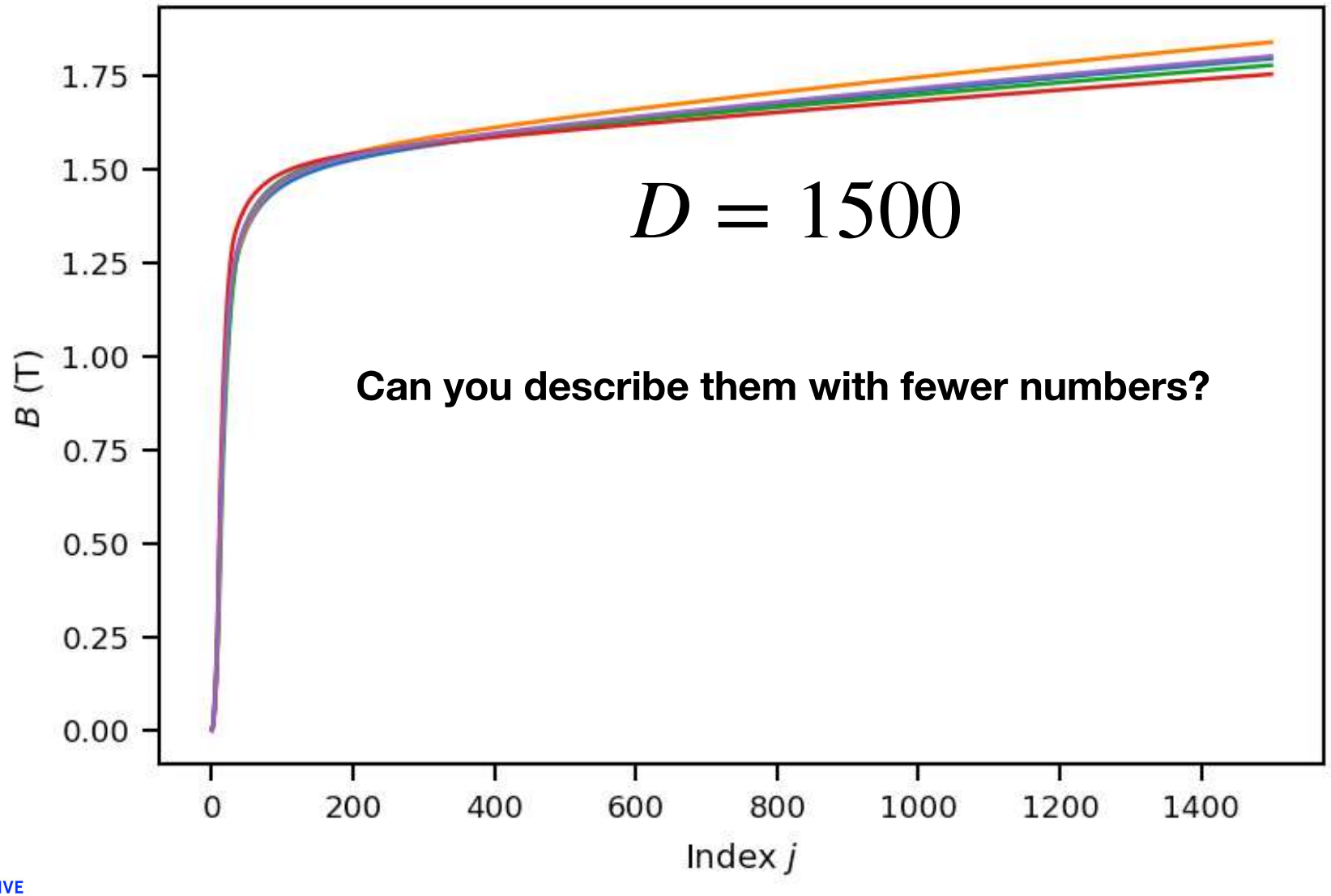
Original data: 28 x 28 pixels -> $D = 784$



https://en.wikipedia.org/wiki/MNIST_database#/media/File:MnistExamples.png

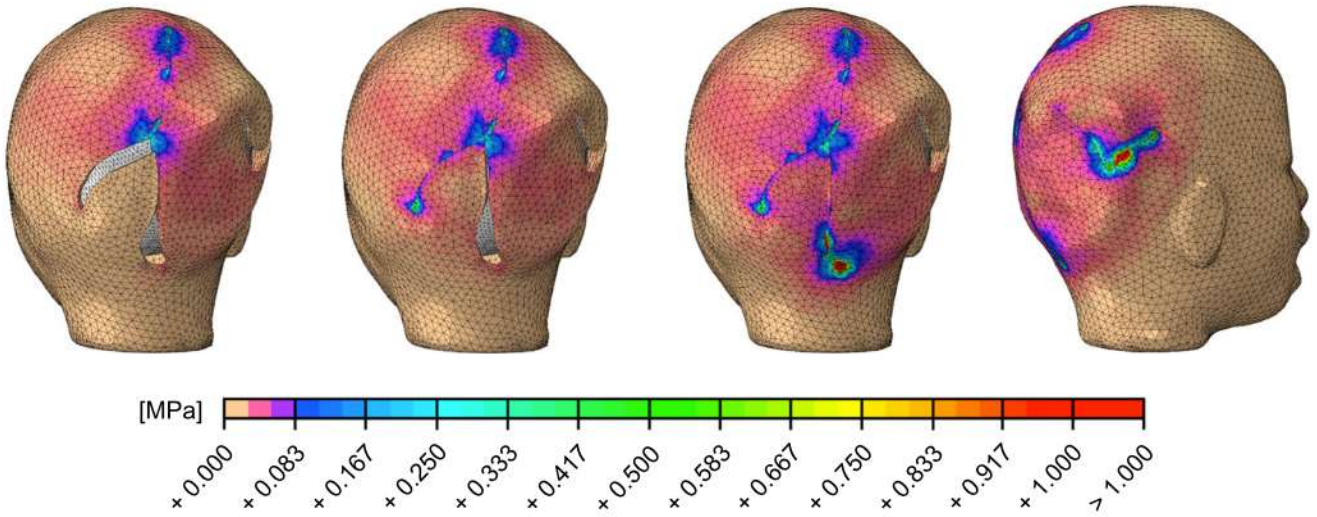
Can you describe them with fewer numbers?

Example: B-H curves of steel



PREDICTIVE
SCIENCE LABORATORY

Example: Von Mises stress on human skull



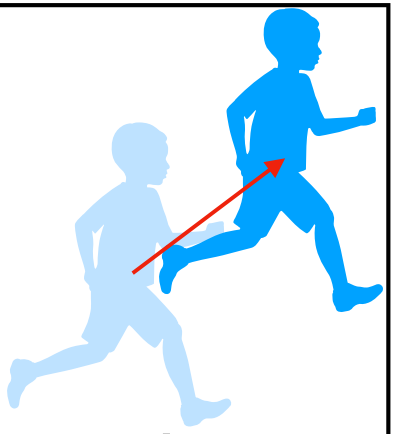
Lee, Turin, Gosain, Bilionis, Tepole, 2018

$$D = 6807$$

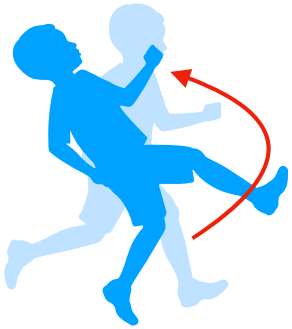
Can you describe them with fewer numbers?

Why is dimensionality reduction possible?

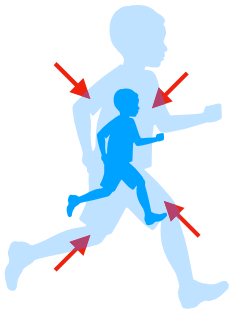
Thousands of numbers



Translations: 2 numbers



Rotations: 1 number



Size: 1 number

You can describe all the possible positions of this boy using only 4 numbers!

Dimensionality reduction is possible because of symmetries and physical constraints!

PREDICTIVE SCIENCE LABORATORY