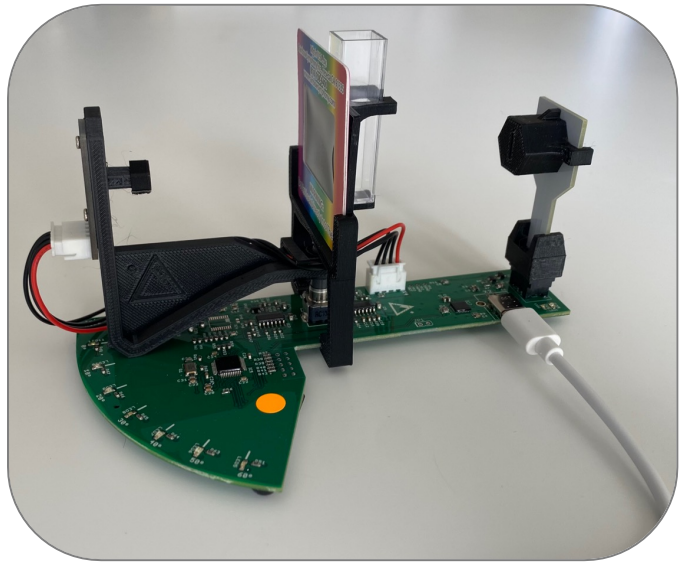
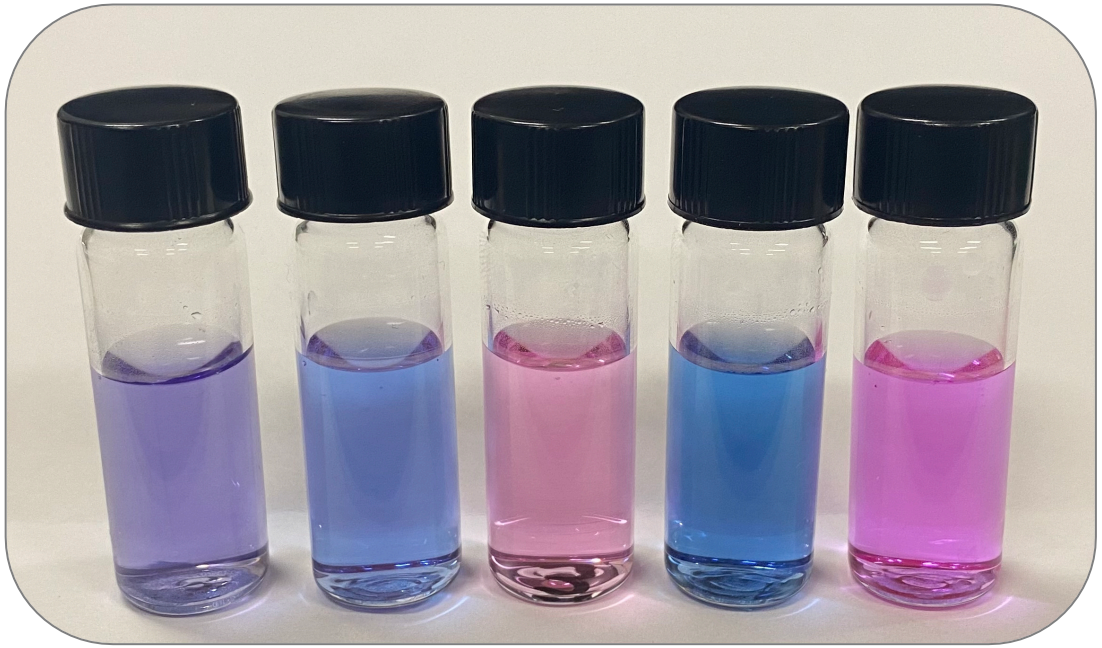
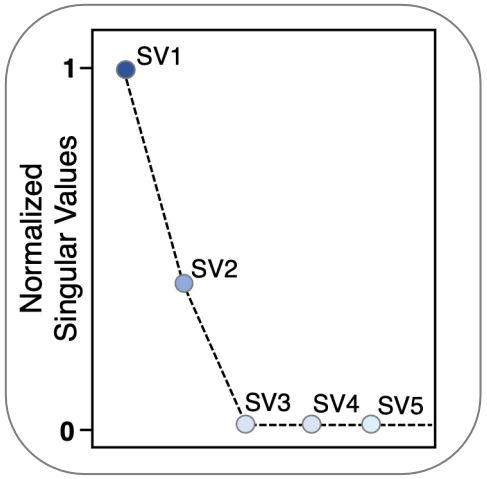
Introduction to advanced data analysis in spectroscopy

Applying singular value decomposition (SVD) to analyze absorbance spectra of chemical mixtures







trimontana

LEARN ANYTHING. EVERYWHERE.

Center for Adapting Flaws into Features

How does light interact with matter?

The Electromagnetic Spectrum

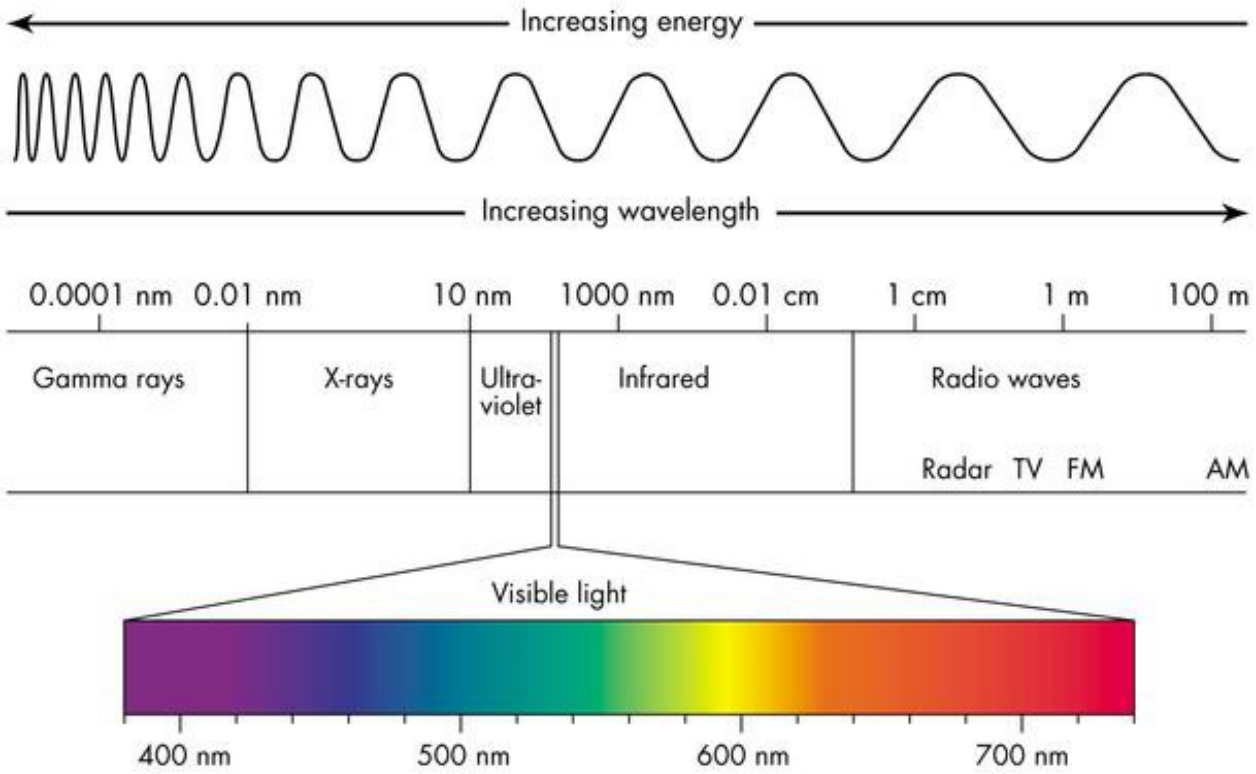
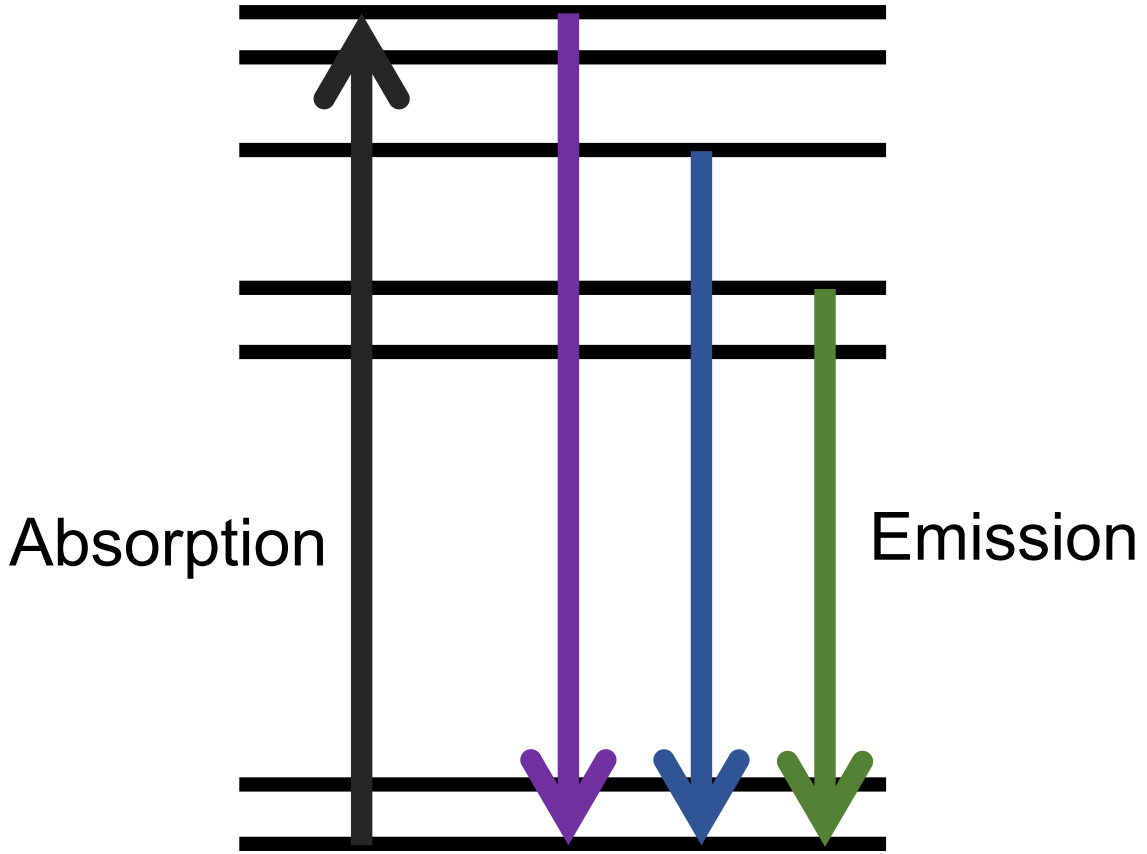


Photo: Cyberphysics.co.uk



Studying chemical systems using absorbance spectroscopy

Review of absorbance spectroscopy:

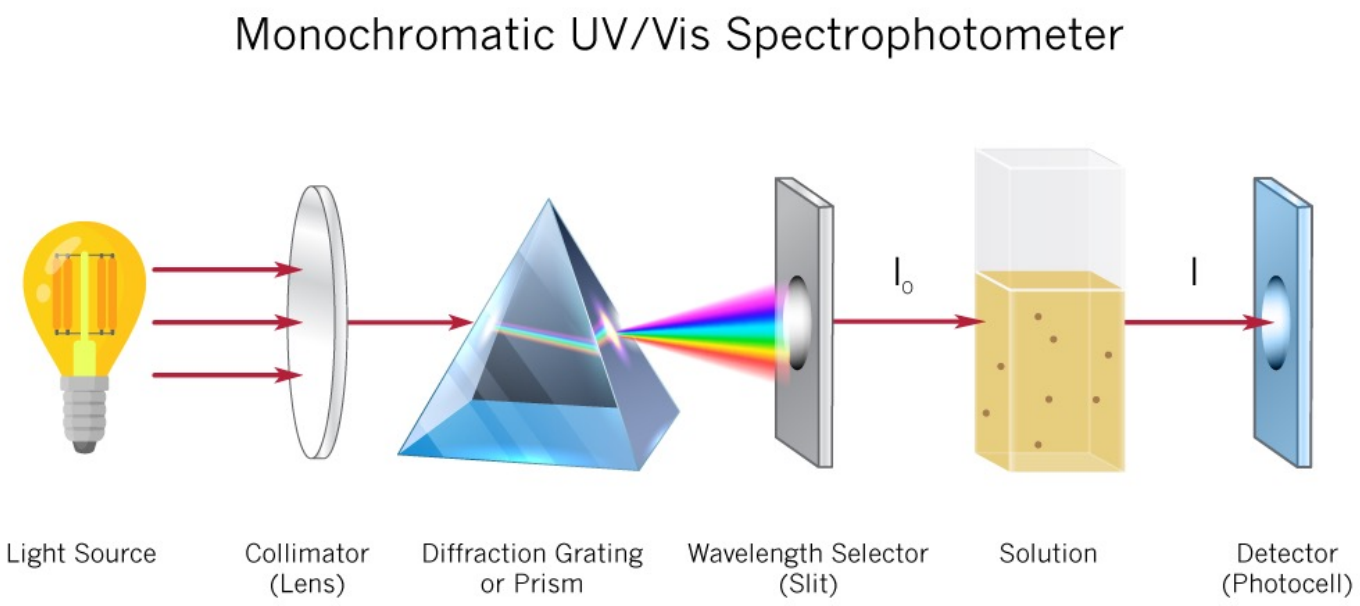
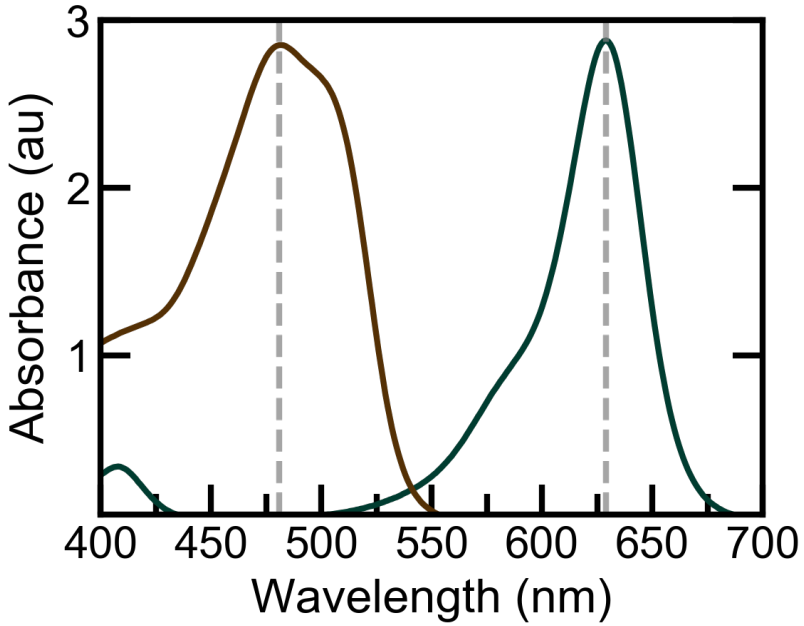


Photo: IMPLEN Spectrophotometers

What properties of a chemical system do you think affect how it will absorb light?

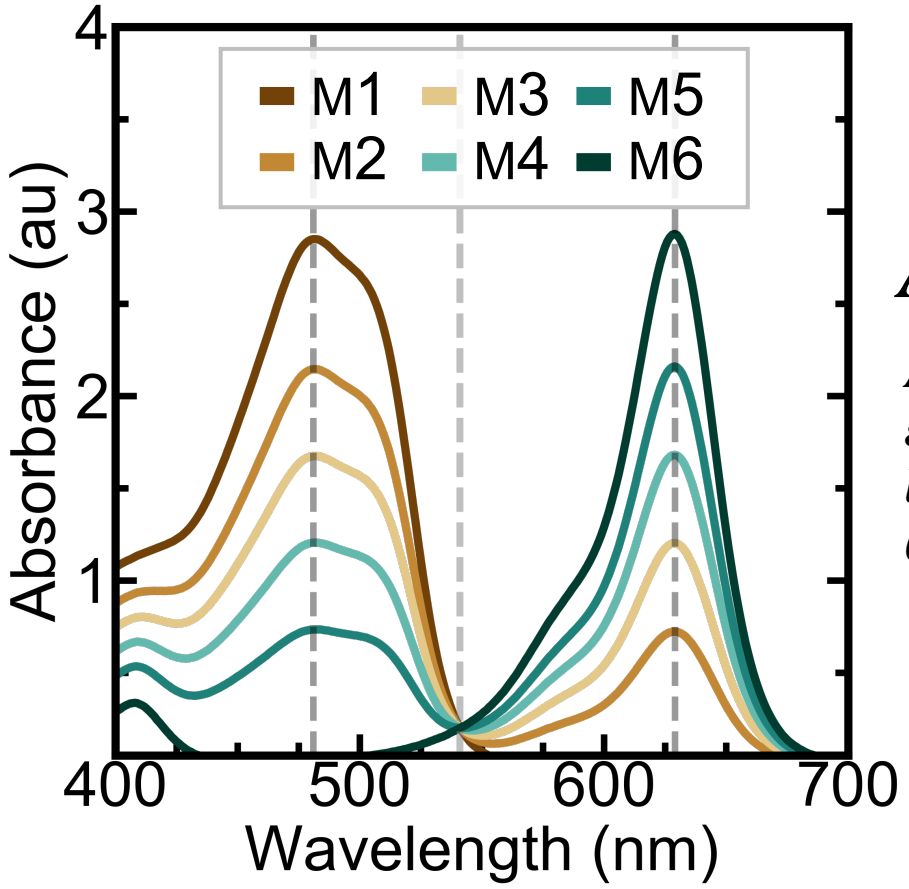
Plotting our data as an absorbance vs. wavelength spectrum:



Two separate spectra plotted above (peaks in each are marked by the dotted line)

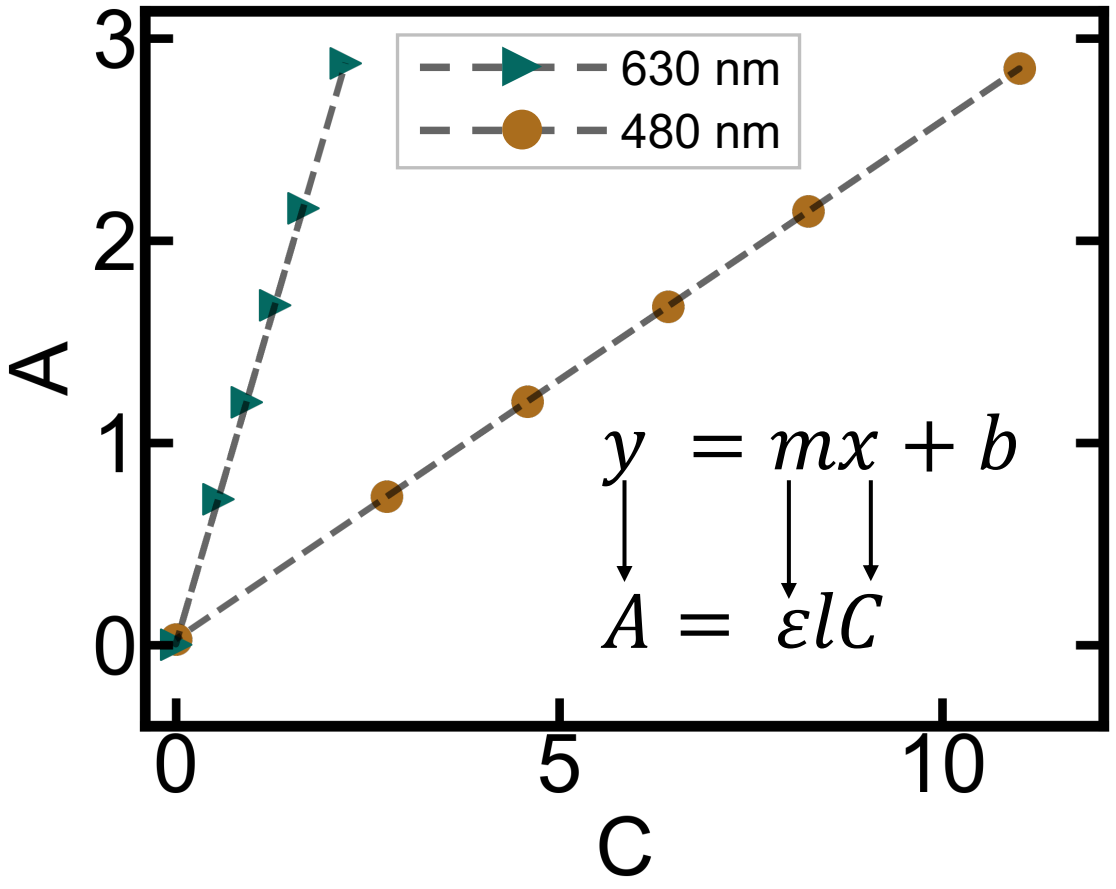
Review of Beer-Lambert Law and analyzing chemical mixtures

How can we use the Beer-Lambert Law to analyze the spectra of the six different mixtures shown below?



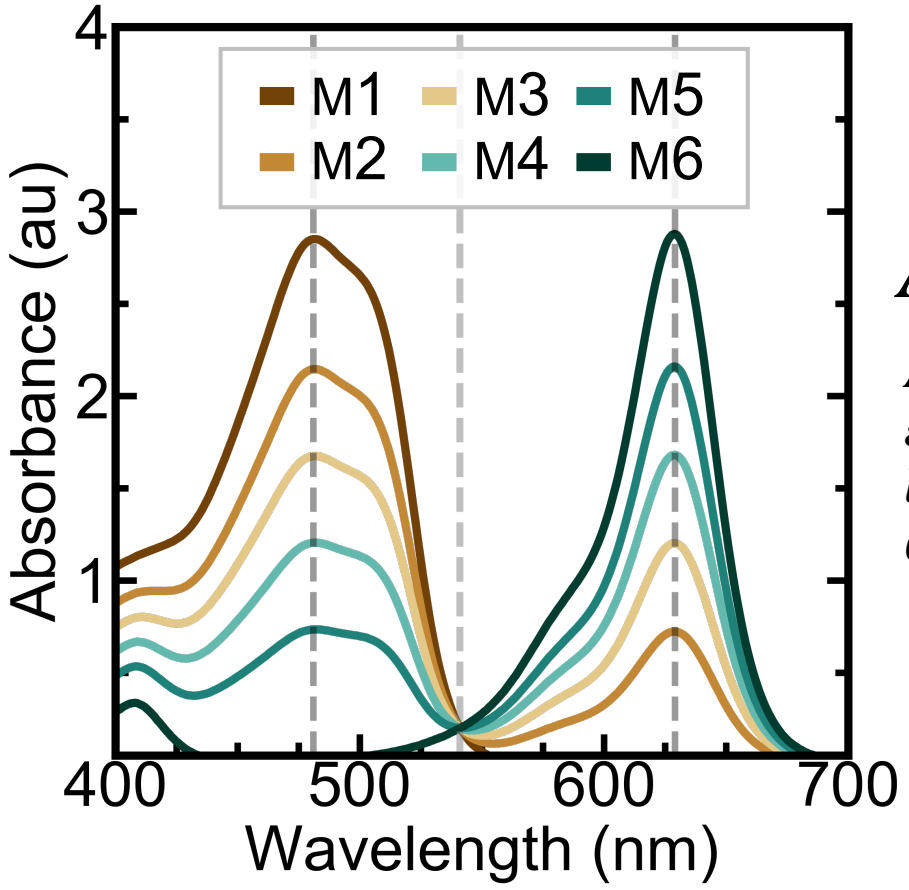
$$A = \varepsilon l C$$

A = Absorbance
ε = molar absorptivity
l = path length
C = concentration



Review of Beer-Lambert Law and analyzing chemical mixtures

How can we use the Beer-Lambert Law to analyze the spectra of the six different mixtures shown below?



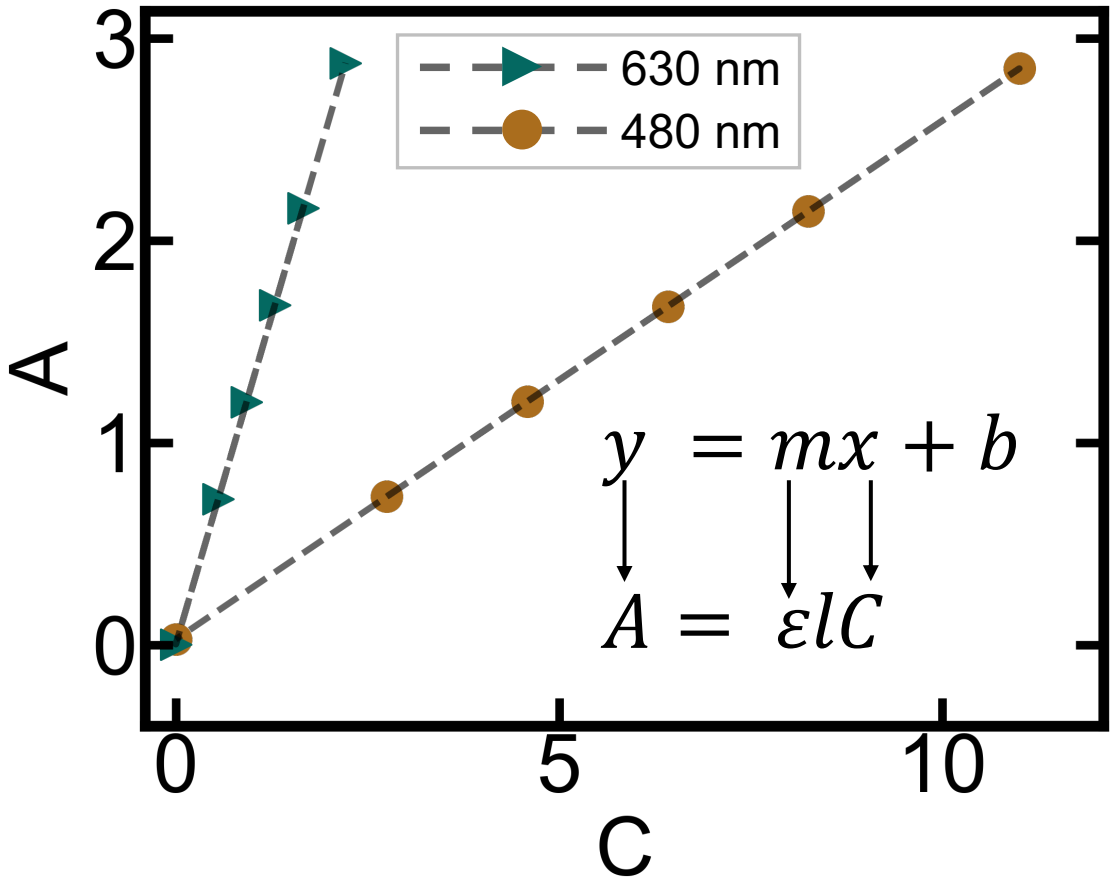
$$A = \varepsilon l C$$

A = Absorbance

ε = molar absorptivity

l = path length

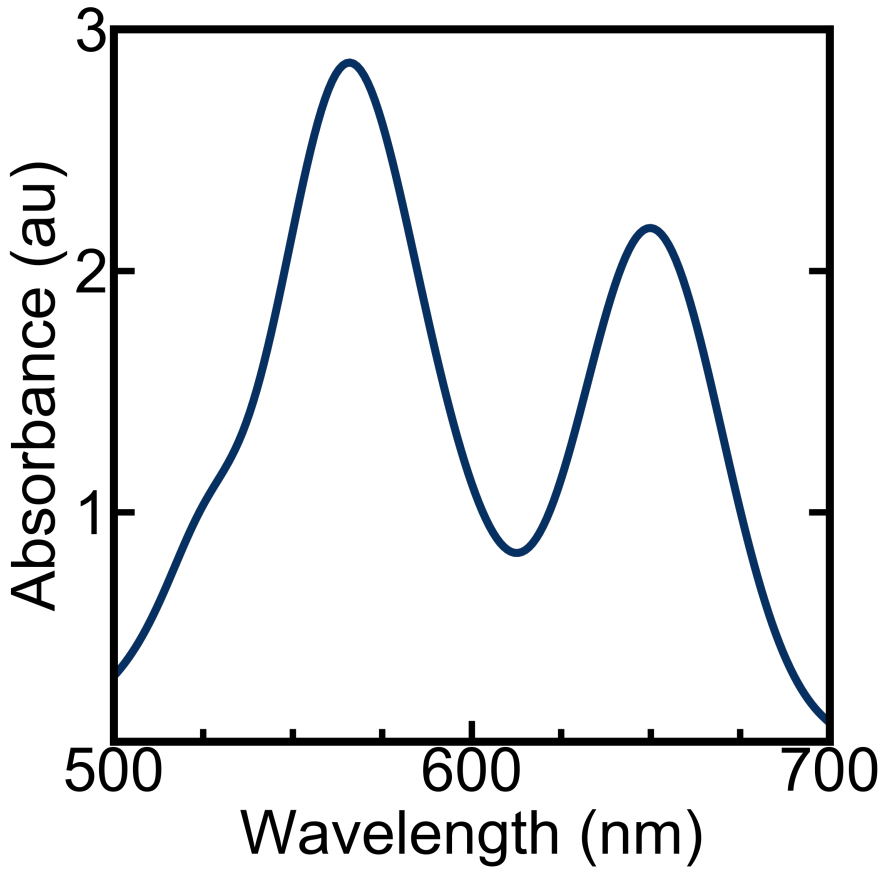
C = concentration

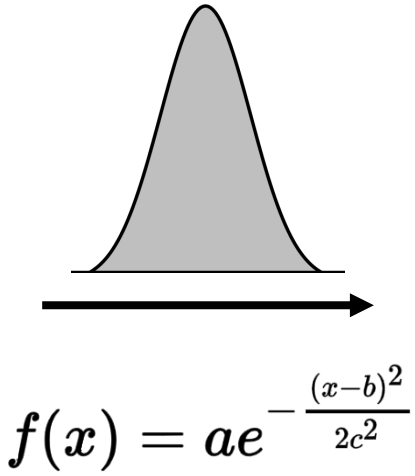


What happens if we start to have peak overlap, or we don't know ε or C?

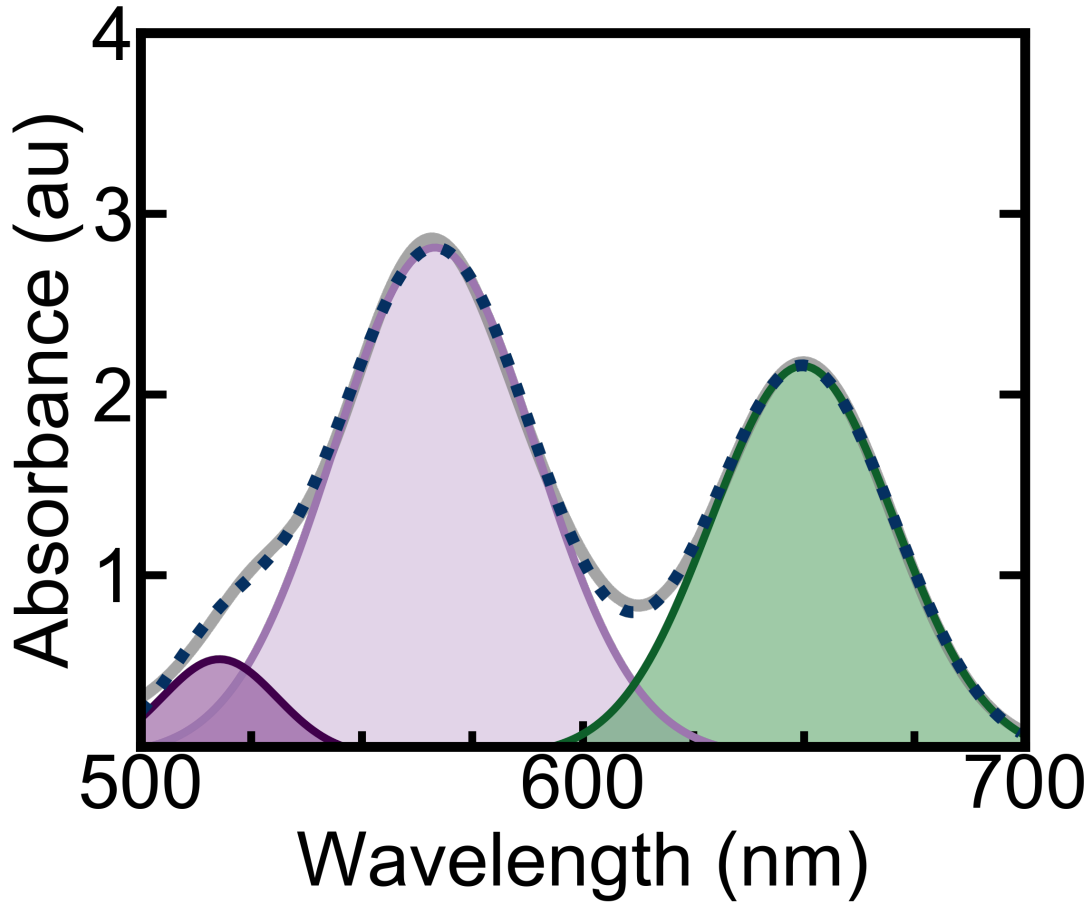
Other common techniques for extracting chemical information

Fitting absorbance peaks to known functions





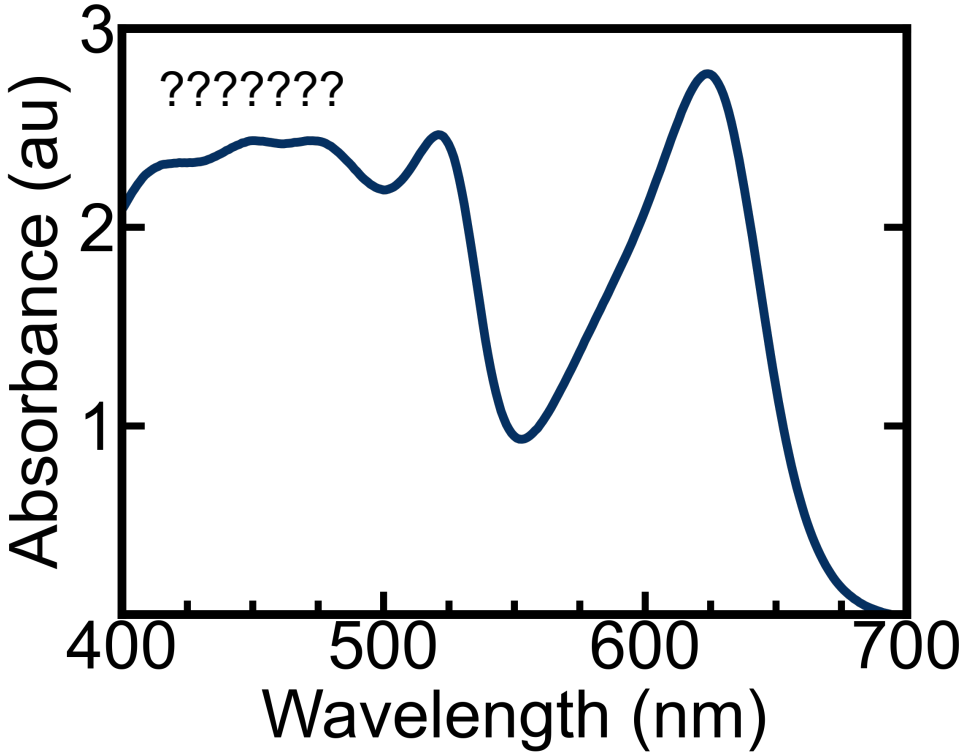
$$f(x) = ae^{-\frac{(x-b)^2}{2c^2}}$$



What happens if we don't know anything about the number of components or their lineshapes?

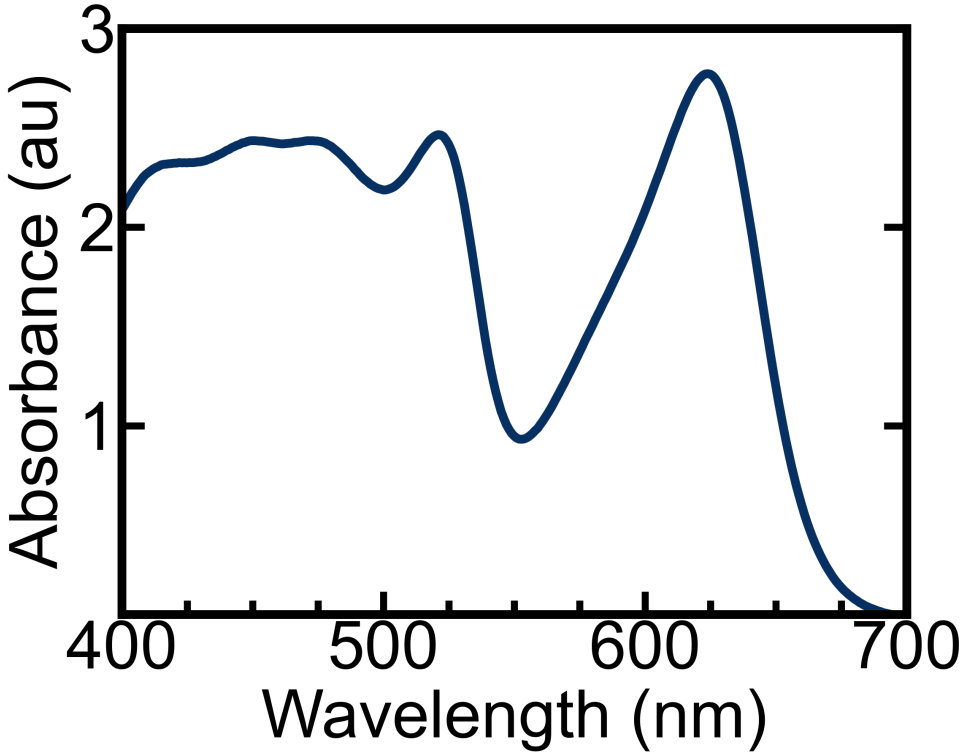
Dimensionality Reduction Techniques

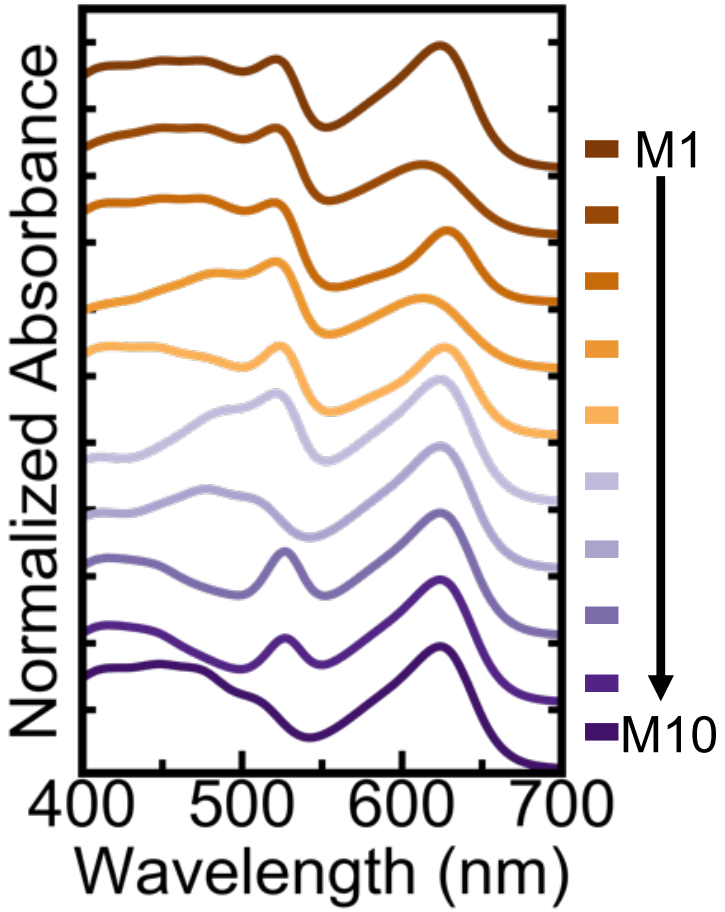
How many unique dyes do you have in the chemical mixture below?



Dimensionality Reduction Techniques

How many unique dyes do you have in the chemical mixture below?

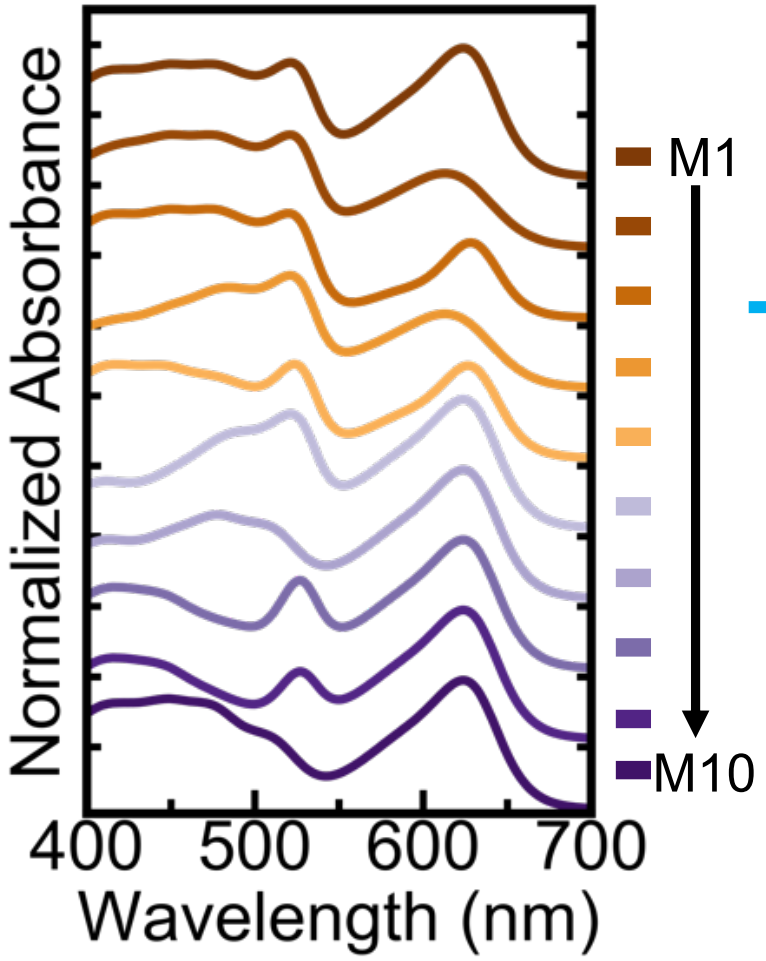




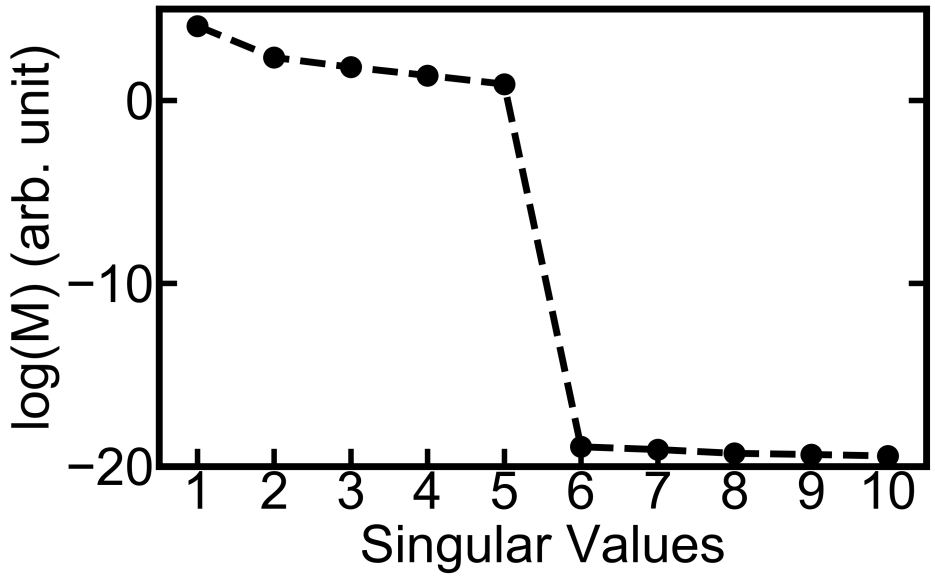
If we have enough mixtures containing varying concentrations of the same dyes, we can use singular value decomposition (SVD) to answer this question!

Applying singular value decomposition

How many unique dyes do you have in the chemical mixtures below?



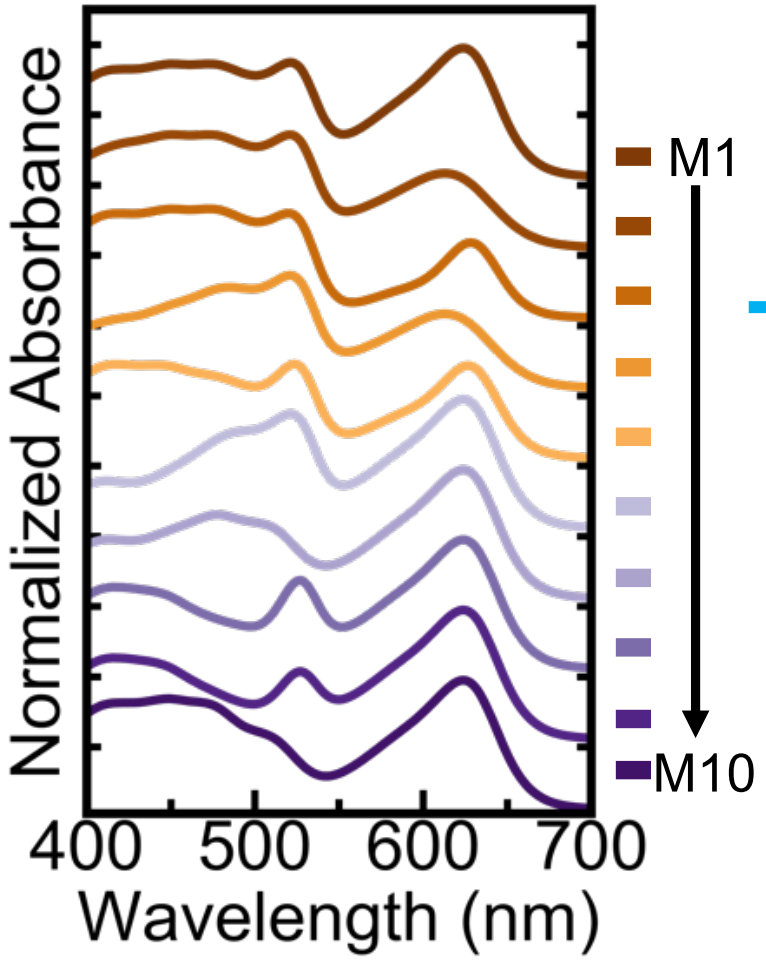
SVD



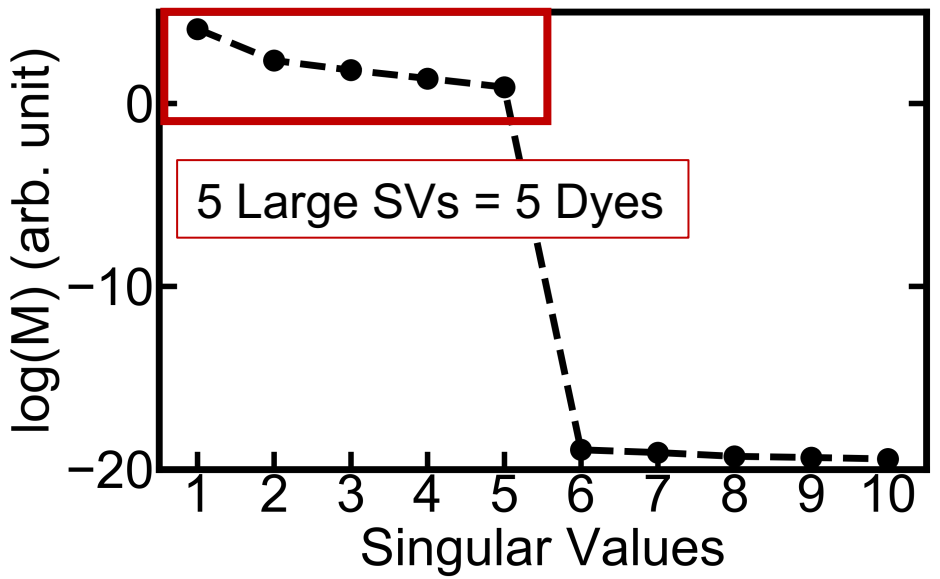
SVD can help tell us how many unique components we have in a collection of mixtures

Applying singular value decomposition

How many unique dyes do you have in the chemical mixtures below?



SVD

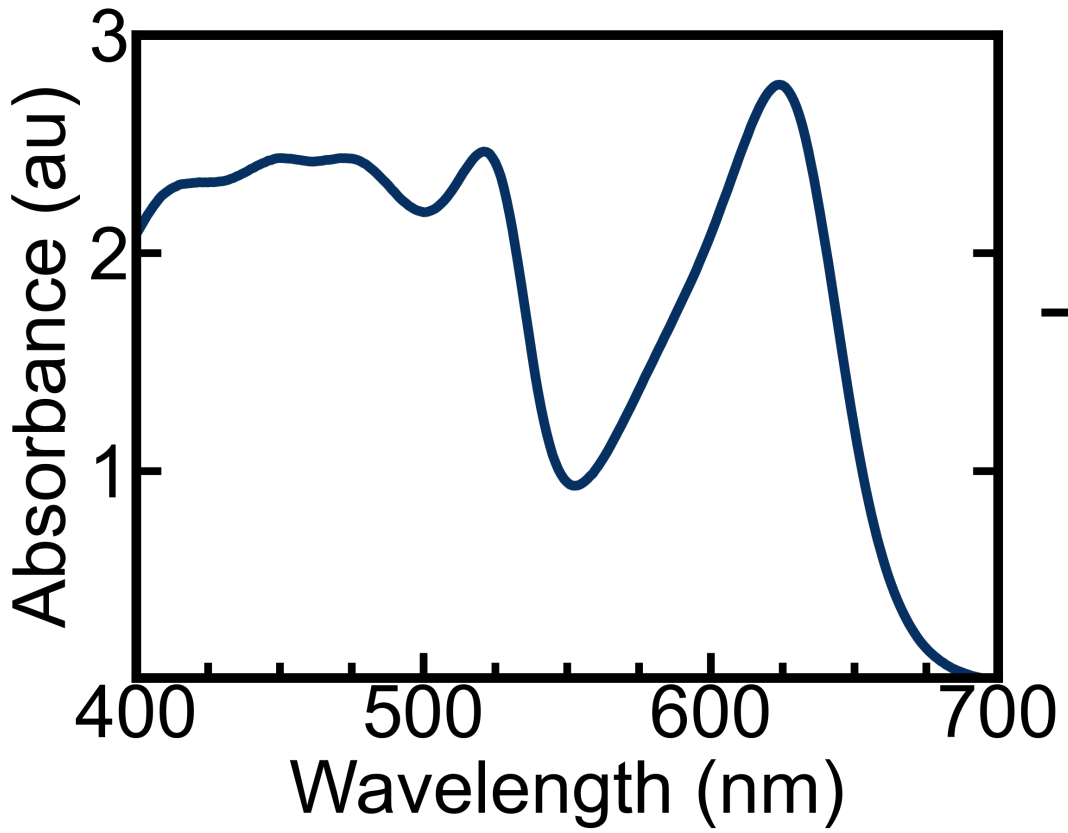


SVD can help tell us how many unique components we have in a collection of mixtures

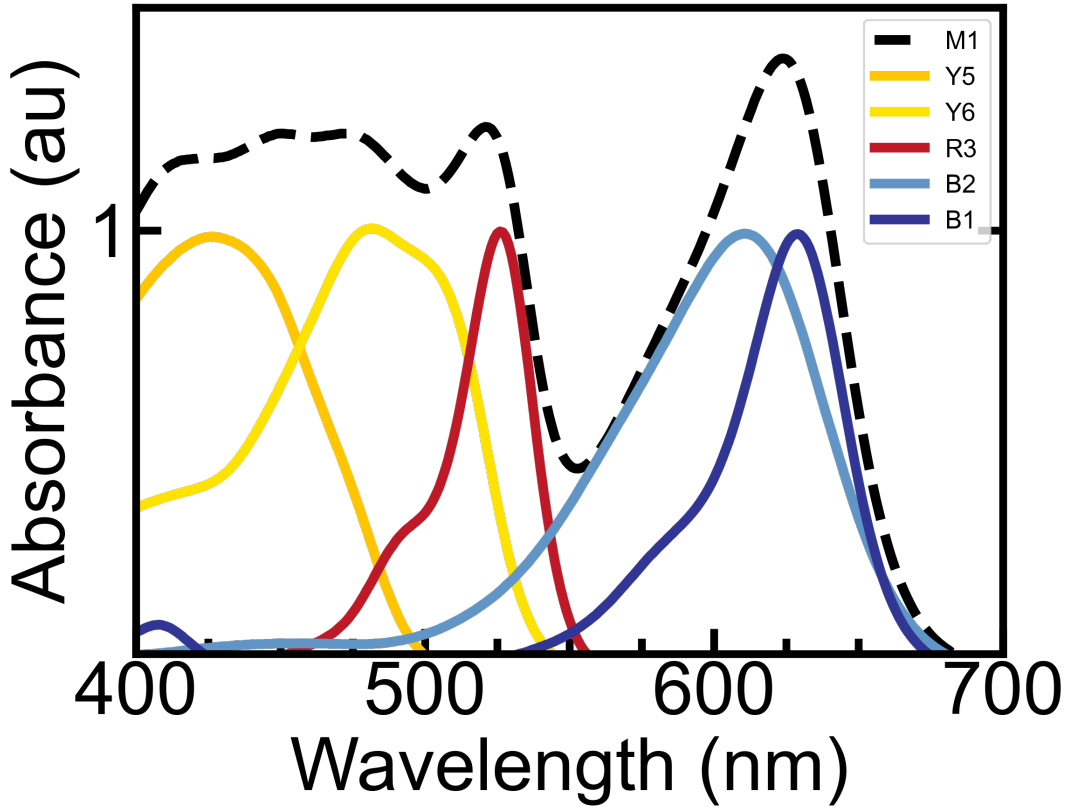
- Allows for determination of number of unique components in complex chemical mixtures
- Requires multiple sample mixtures with varying concentrations of each component
- Does not directly give you chemical parameters (e.g., ε)

Applying singular value decomposition

How many unique dyes do you have in the chemical mixture below?

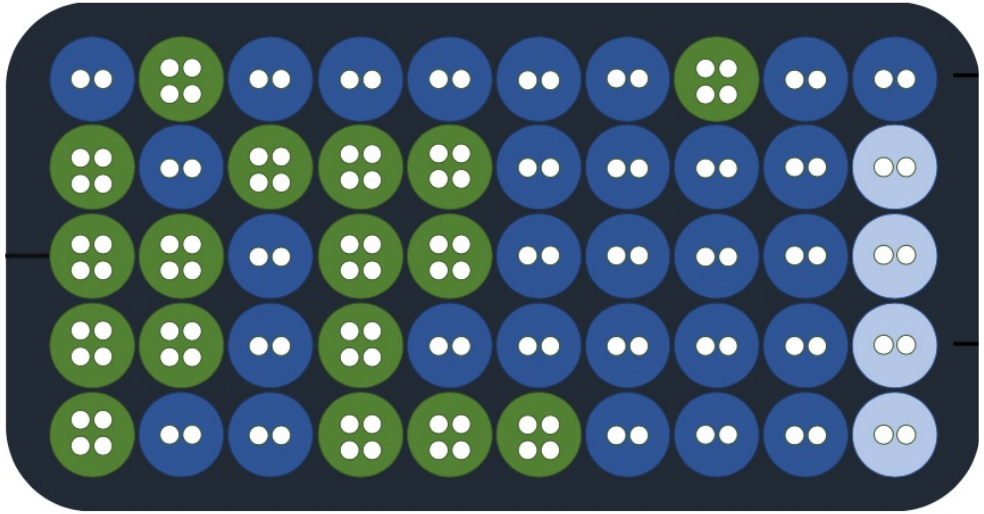


5 Large SVs = 5 Dyes!

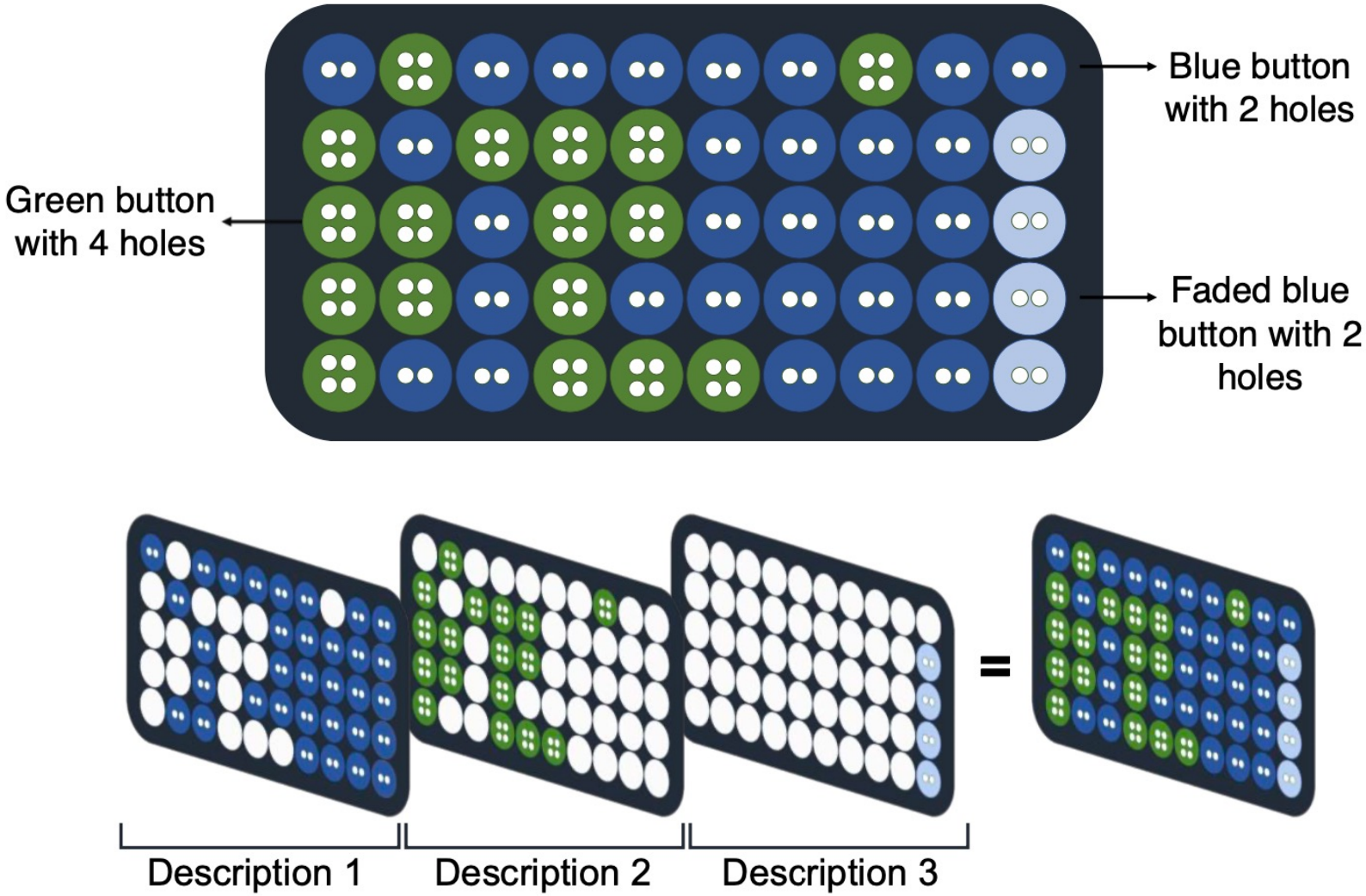


How does SVD work?

How many unique features would you need to list to completely describe all buttons in this collection?

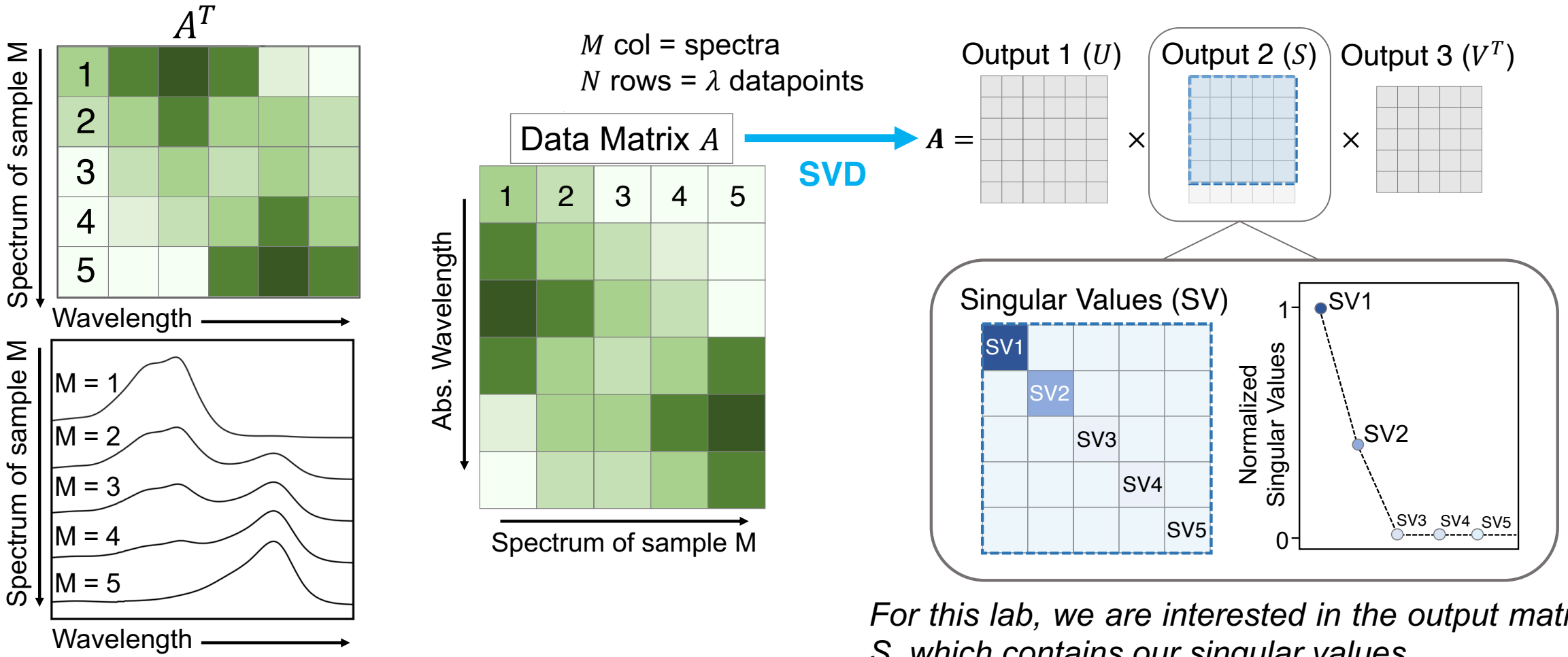


How does SVD work?



How does SVD work?

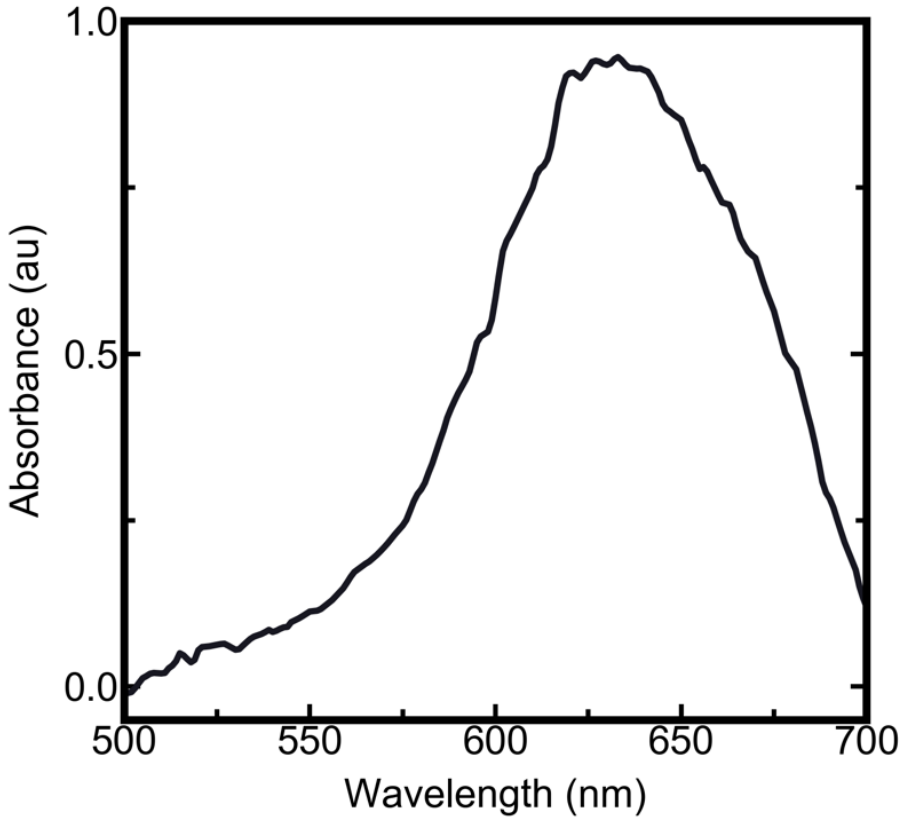
Given an N x M input matrix *A* of data, SVD returns three outputs that can be used to redescribe your data:



For this lab, we are interested in the output matrix S, which contains our singular values

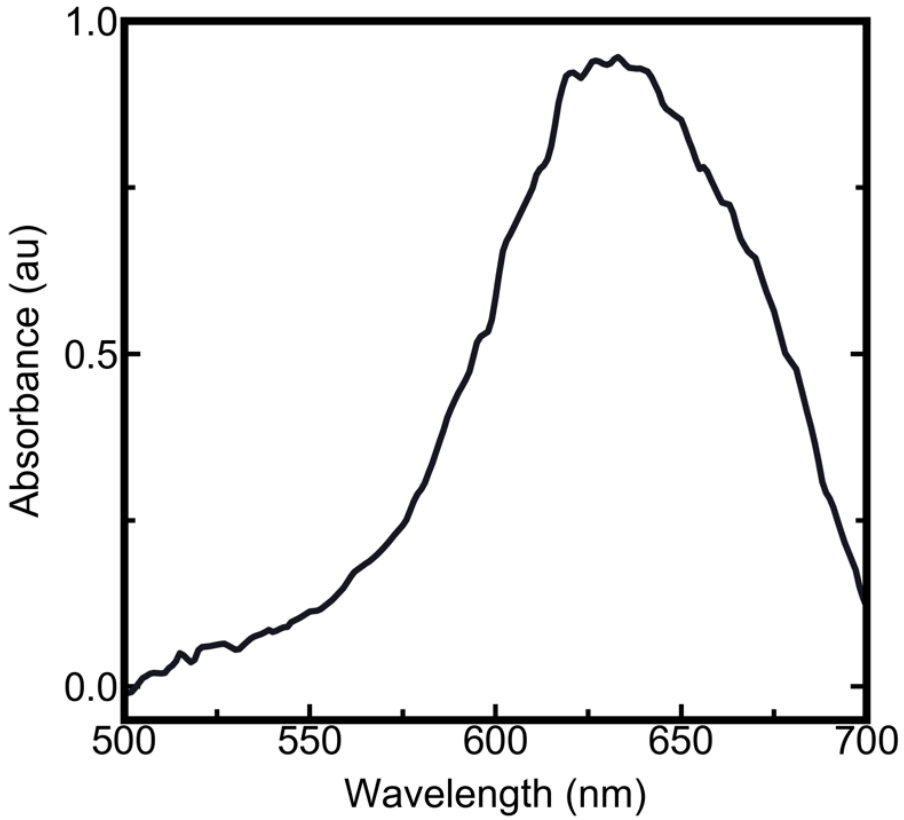
How does noise affect our results?

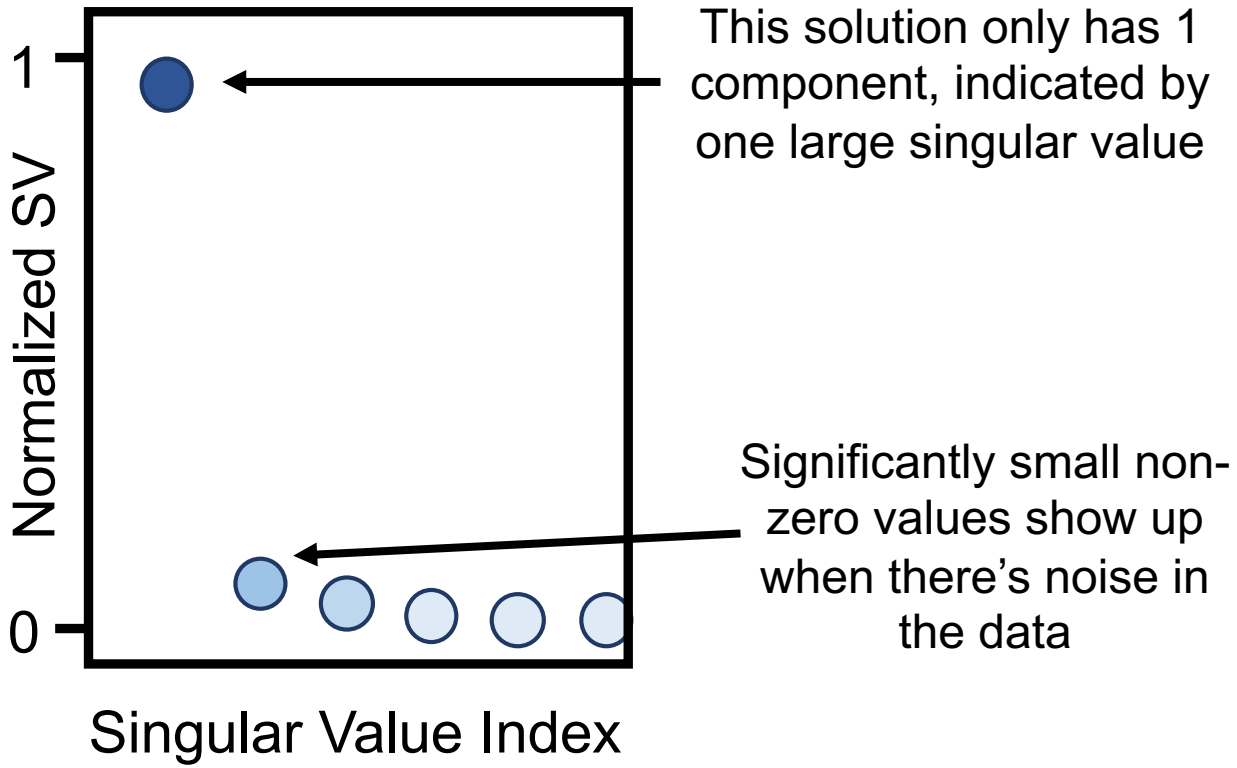
Will noise show up in our results? If so, how will it affect our results?



How does noise affect our results?

Will noise show up in our results? If so, how will it affect our results?

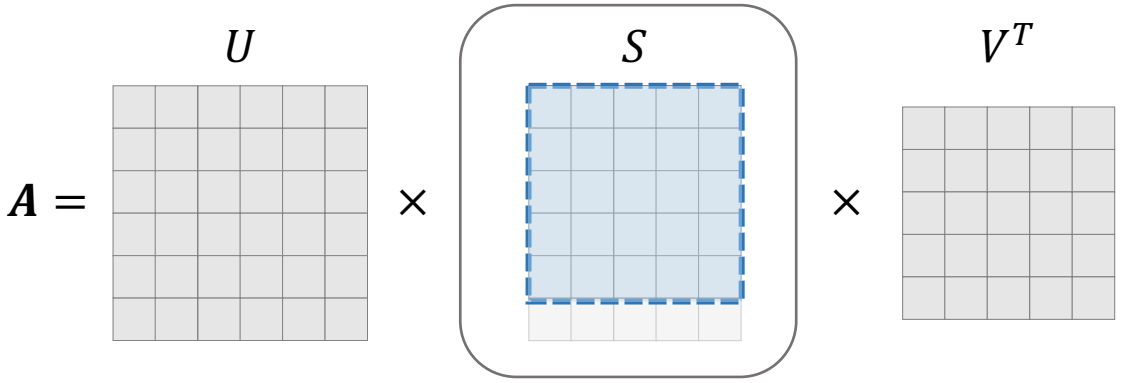


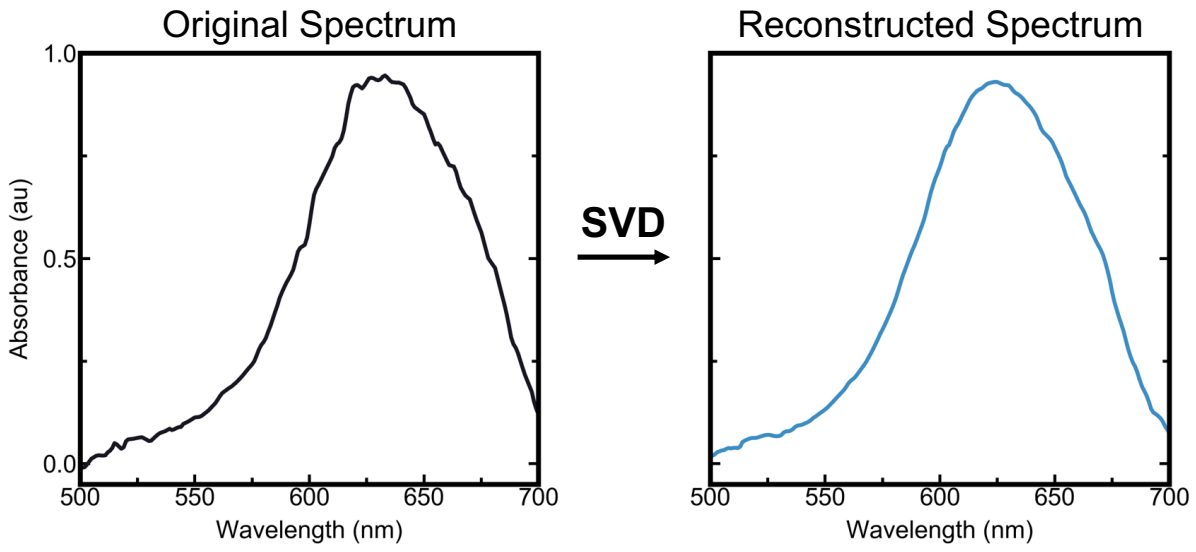


On day 2 of this lab, you'll learn how to decide when values are small enough to be attributed to noise

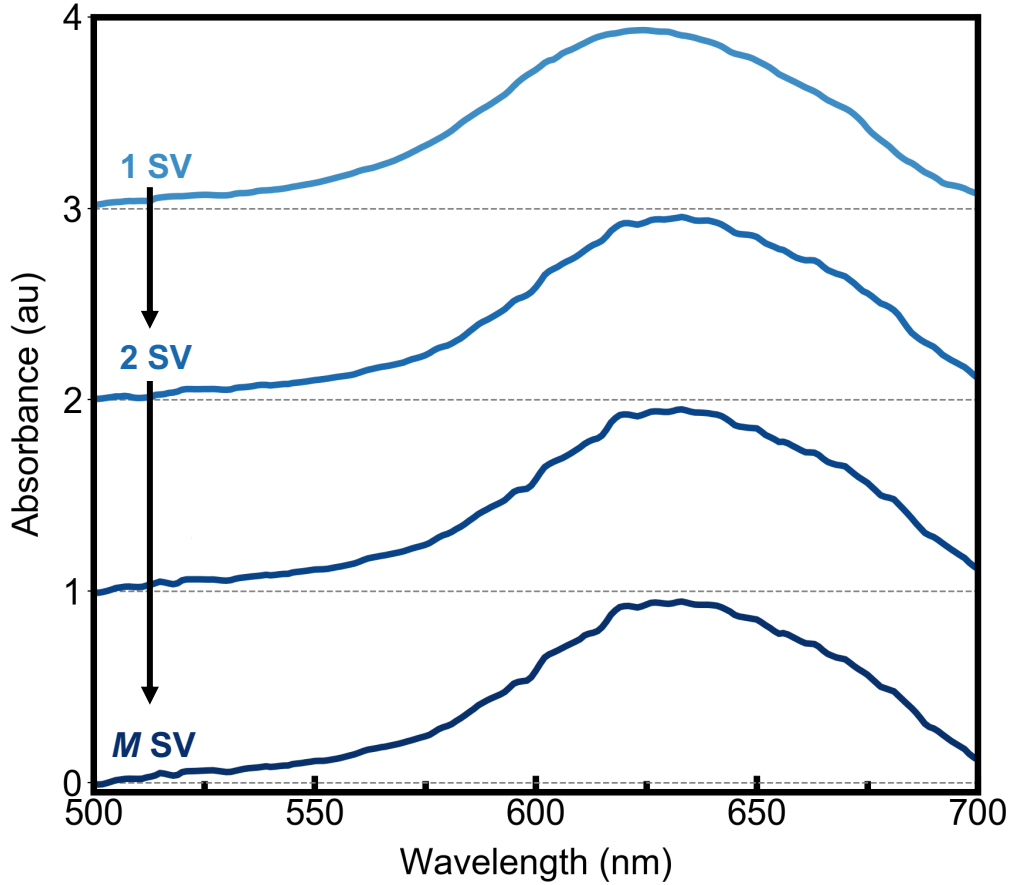
How does noise affect our results?

SVD can also be used as a useful tool for processing data with noise!



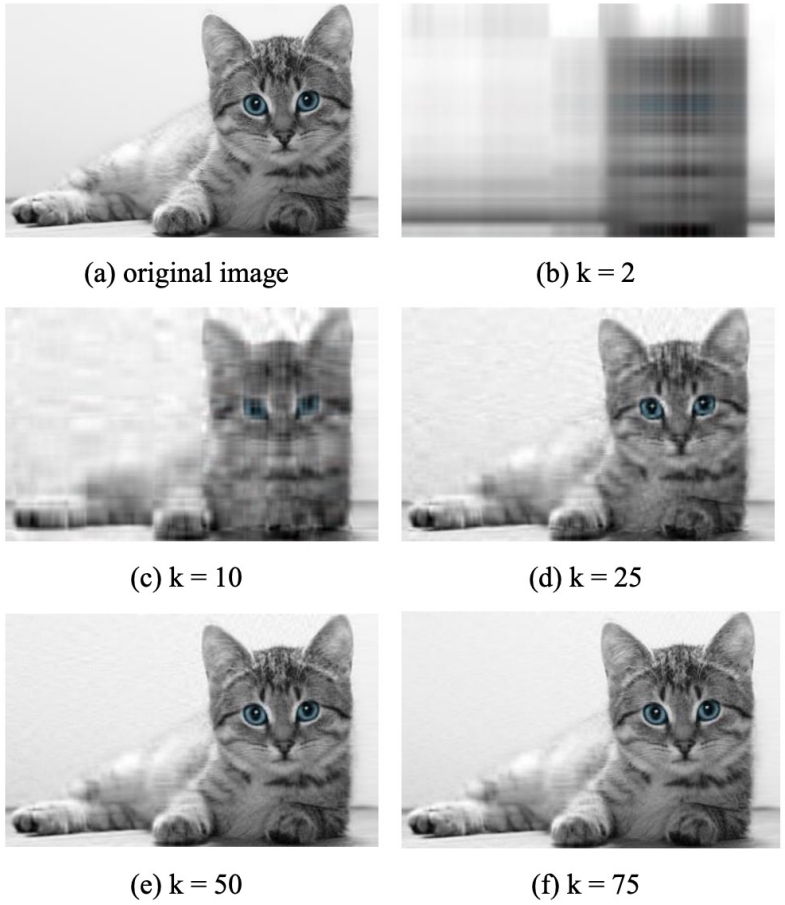


b



Applications of SVD across fields of science and engineering

Image Compression

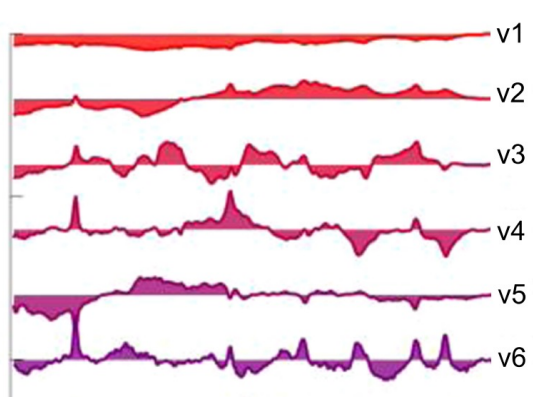


k = number of components used in image reconstruction

Recommender Systems

NETFLIX

Helping analyze chemical mixtures (what you'll be doing today)



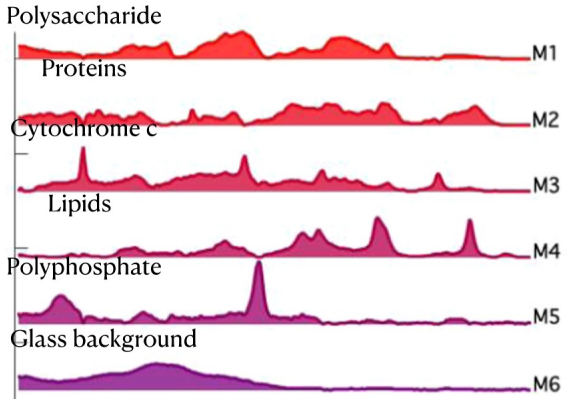
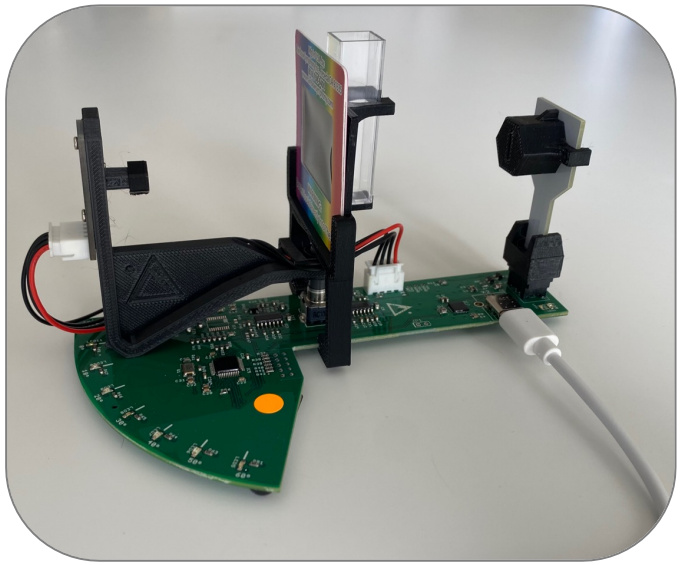
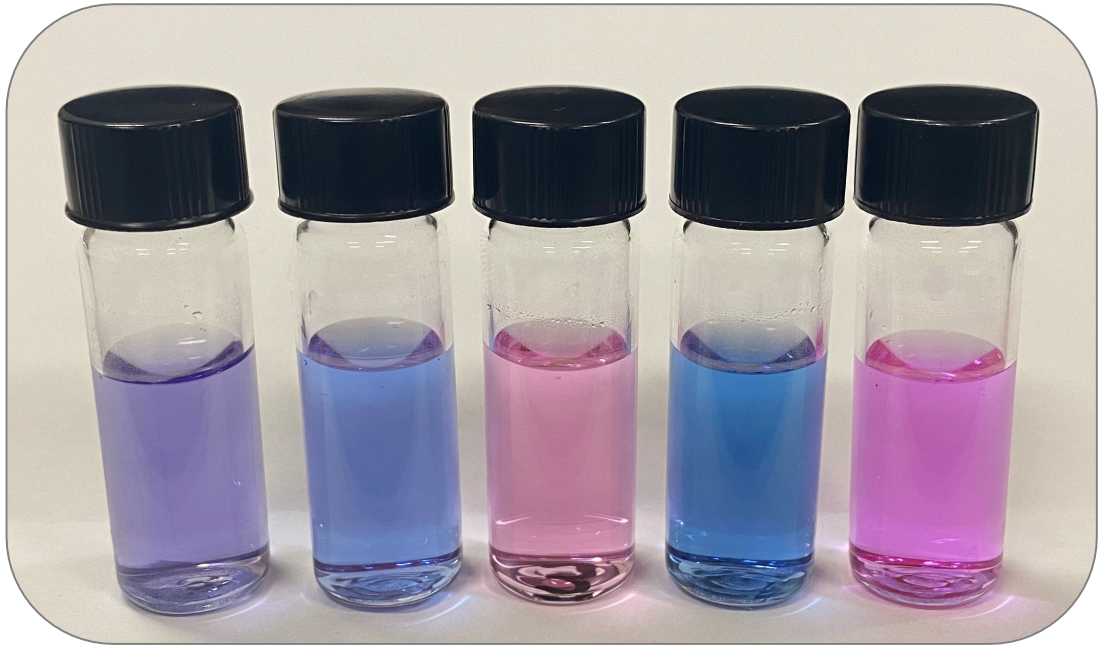


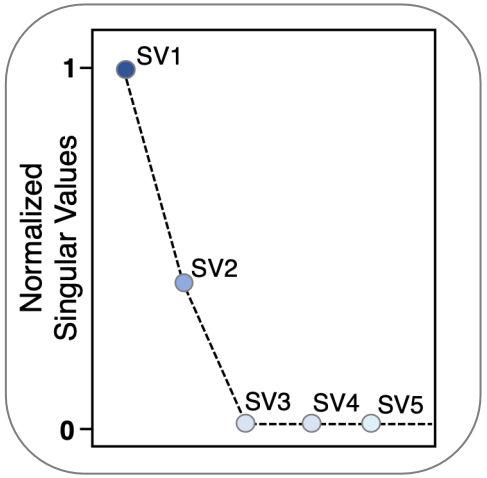
Image 1: (image compression) https://doi.org/10.1145/3274250.3274261
Image 2: (recommender systems) https://developers.google.com/machine-learning/recommendation/collaborative/matrix
Image 3 (spectra): Samuel, A. Z.; Horii, S.; *et al. Anal. Chem.* **2021**, *93* (35), 12139–12146.

In this lab:

You'll be taking on the roll of a food scientist trying to answer the question:
How many unique dyes do I have in a collection of samples?







1: Build a spectrometer and collect absorbance data on known samples → 2: Collect data on your unknown mixtures → 3: Use SVD to analyze your data and compare the results

In the final portion of your lab, you'll be comparing the results from your hand-built spectrometer to data from a commercial instrument and ideal/artificial data