Class 6 overview: 29 Feb

Announcements & Carry-forwards

1. We start a new lane is Statistics today,
2. And we begin a programming assignment that evolves into a class project
3. Office hours will now start 6:30 PM on Tuesday.

Class survey: Should we organize into small (3 person) groups for projects?

First hour

Midterm, multiple short answer questions

Second hour

- Continuation of *Probability 4*
- Statistics Unit 1 - Variance, inference and sampling.

Third hour

(Continuation of Stats)

- Linear Algebra 5, Gram-Schmidt orthogonalization

Fourth hour

Session on data simulation - discussion on class project contents

Assignments

Variance:

- Derivation of variance formulas. Using the definition of variance, $\sigma^2(X) = E[(X-\mu)^2]$.show the algebra to derive these;

a. $\sigma^2(X) = E[X^2] - E[X]^2$.

b. $c^2\sigma^2(X) = \sigma^2(cX)$,

c. $\sigma^2(X+Y) = \sigma^2(X) + \sigma^2(Y) \text{ when } X \perp Y$.

d. $\sigma^2(\overline{x}) = \sigma^2(X)/n$, e.g. the sample mean variance scales as 1/n.

Orthonormal projections

- Simplify the projection operator $A(A^TA)^{-1}A^T$ when the columns of ***A*** are orthonormal.

Programming:

- Pick a non-linear function on a fixed interval
- Simulate your function y = f(x) + “noise”
- create a “polynomial” design matrix 1, x, x^2, x^3
- Gram-Schmidt normalization of your polynomial design matrix
- Projection — see what you get? Visualization & notebook to show it.

Files

- Probability Unit 4
- Statistics_Unit1
- Linear_algebra_unit5a
- Normal_simulation_w_variance.ipynb

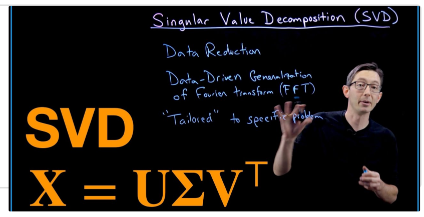
Suggested video

Steve Brunton has a really awesome lecture series on this with a lot of visual examples if anyone wants to try and skip the equation heavy explanations lol

Singular Value Decomposition (SVD): Overview

This video presents an overview of the singular value decomposition (SVD), which is one of the most widely used algorithms for data processing, reduced-order modeling, and high-dimensional statistics.

https://www.youtube.com/watch?v=gXbThCXjZFM&list=PLMrJAkhIeNNSVjnsviglFoY2nXildDCcv



In general, Brunton's lectures on Physics and Machine Learning are worth viewing:

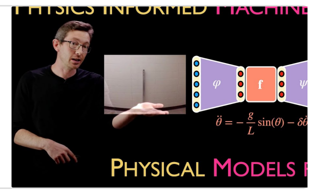
raw notebooks

In Google drive: notebooks/visualizing_dependent_data.ipynb

Physics Informed Machine Learning: High Level Overview of AI and ML in Science and Engineering

This video describes how to incorporate physics into the machine learning process. The process of machine learning is broken down into five stages: (1) formulating a problem to model, (2) collecting and curating training data to inform the model, (3) choosing an architecture with which to represent the model, (4) designing a loss

https://youtu.be/JoFW2uSd3Uo?si=ylUK2WE-VJwIF3MC



Additional Graphics

Class teamwork

1. Models, Theories, and Data

How is ML different than conventional statistics?