Math Basics for Machine Learning
Graded Assignment 3

Faran Taimoor Butt

Fall 2024

Instructions

This is the third graded assignment for the Math Basics for Machine Learning course. It contains two tasks. The instructions, as well as links to supplementary material, are given in the task descriptions.

Provide **detailed solutions** to the tasks in this assignment. Then, save your solution document as a .pdf file and submit it by filling in the corresponding Google form.

For some tasks, you might use Python rather than perform computations by hand. If you do so, attach your code as well (e.g., link to the Colab notebook with code).

In total, you can earn 10 points for this assignment. This score will contribute to you final score for this course.

You must submit your answers by **Sunday, October 6, 18:59 Moscow Time**. *Late submissions will **not** be accepted.*

It is the idea that you complete this assignment individually. Do not collaborate or copy answers of somebody else.

Have fun!

1. (5 points) ML researchers work either at universities ("academia") or in private companies ("industry"). Official statistics shows that every year, 20% of academic workers transition to industry, while only 1% moves from industry to academia. How will this distribution evolve over years? In this task, you will answer this question using a simple model.

   We can collect the information we have into a single matrix T, where the first column and the first row correspond to industry, while the second column and the second row are related to academia:

$$T = \begin{bmatrix} 0.99 & 0.2 \\ 0.01 & 0.8 \end{bmatrix}$$

   You can read this matrix as follows:

   - *first column:* 0.99 of ML researchers in industry stay there, while 0.01 transition to academia;
   - *second column:* 0.2 of the academic researchers transition to industry, while the remaining 0.8 stays in academia.

   Suppose that in some year 0, there were $x_0^{industry}$ researchers working in industry and $x_0^{academia}$ academic researchers. Given the statistics above, we can then estimate industry vs. academia research distribution next year as follows:

$$x_1 = Tx_0, \; x_0 = \left[x_0^{industry}, x_0^{academia}\right]^T$$

   (a) (2 points) Imagine that in 2024, 30% of the ML researchers work in industry, while 70% work in academia. How will this distribution be changing over the next 100 years (assuming that the transition matrix stays the same)?

   Make a plot to visualize industry vs. academia researchers distribution for each year from 2023 up to 2124 according to our model. Explain what you see.

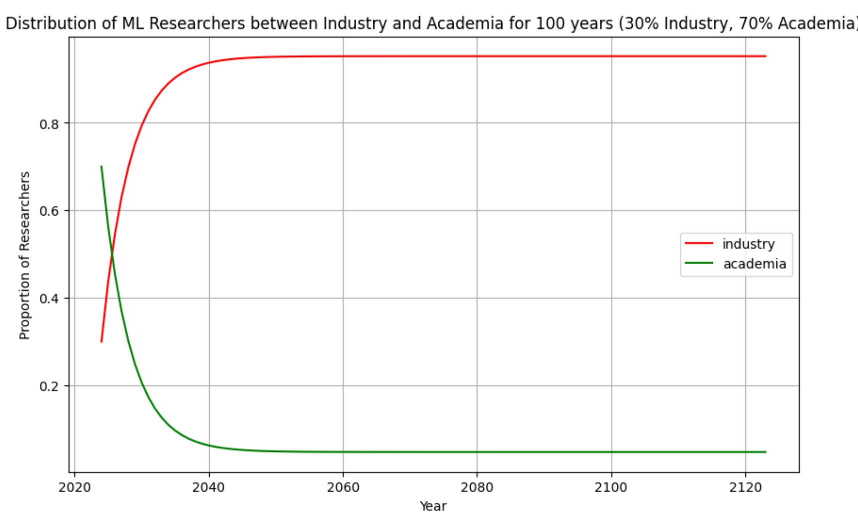

Figure 1: Distribution of ML Researchers between industry and Academia for 100 years

Solution: As the initial distribution for the industry was set to 30% and for the academia, it was set to 70%.It is clearly visible in the above figure that the academia curve starts near 0.7 and the industry curve starts at 0.3.We observed that over time, the proportion of researchers in industry gradually **increases** while the proportion in academia **decreases**.The system tends to a steady state where around 90% of researchers are in industry and 5% are in academia. This is due to the higher retention rate in industry and the lower rate of transition from industry to academia.

Code:

Click here to see code

(b) (1 point) In some alternative universe, in 2024, the statistics look different: 90% of the researchers work in academia and only 10% are in industry. What will the distribution between industry and academia research be by 2124 in that case? How does it compare to the previous result?

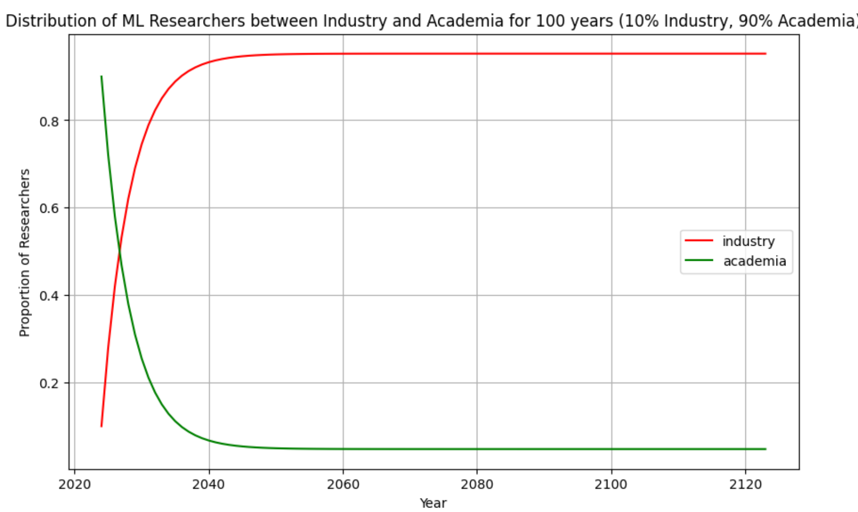

Figure 2: Distribution of ML Researchers between industry and Academia for 100 years

Solution: Although the initial distribution of ML researchers is 90% of researchers working in Academia and 10% of researchers working in Industry now. Over time, the system again converges to the same steady state, where about 90% of researchers will be in industry and 5% in academia. This demonstrates that the initial conditions affect the early years, but the system eventually reaches the same equilibrium as in the previous question where initial conditions were different but the Transition Matrix forces the initial state to converge to equilibrium values.

Code:

Click here to see code

(c) (2 points) Compute eigenvalues and eigenvectors of the transition matrix T and note that one of the eigenvalues is equal to 1. How is this (and the corresponding eigenvector) related to the evolution of the industry vs. academia research distribution that you have observed in the previous task? Explain.

Solution: Given matrix T:

$$\text{T} = \begin{bmatrix} 0.99 & 0.2 \\ 0.01 & 0.8 \end{bmatrix}$$

The eigenvalue equation can be written as:

$$\text{Tv} = \lambda \text{ v}$$

As eigenvalue equation can be written as

$$\det(\text{T-}\lambda \text{ E}) = 0$$

Solving for T-λ E

$$\begin{bmatrix} 0.99 & 0.2 \\ 0.01 & 0.8 \end{bmatrix} - \begin{bmatrix} \lambda & 0 \\ 0 & \lambda \end{bmatrix} = \begin{bmatrix} 0.99-\lambda & 0.2 \\ 0.01 & 0.8-\lambda \end{bmatrix} = 0$$

Taking det(T-λE):

$$\begin{aligned} &= (0.99-\lambda)(0.8-\lambda) - (0.2)(0.01) \\ &= 0.792 \text{ - } 0.99\lambda - 0.8\lambda + \lambda^2 - 0.002 \\ &= \lambda^2 - 1.79\lambda + 0.79 \\ &= \lambda^2 - \lambda - 0.79\lambda + 0.79 \\ &= \lambda(\lambda-1) - 0.79(\lambda-1) \\ &= (\lambda-1)(\lambda-0.79) \end{aligned}$$

$$\lambda_1 = \mathbf{1} \text{ , } \lambda_2 = \mathbf{0.79}$$

For λ_1

$$\begin{bmatrix} 0.99-\lambda_1 & 0.2 \\ -0.1 & 0.8-\lambda_1 \end{bmatrix} \begin{bmatrix} v_1 \\ v_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

$$\begin{bmatrix} -0.01 & 0.2 \\ 0.01 & -0.2 \end{bmatrix} \begin{bmatrix} v_1 \\ v_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

$$\begin{gathered} -0.01v_1 + 0.2v_2 = 0 \\ 0.2\text{v}_2 = 0.01v_1 \\ \text{v}_2 = 0.05v_1 \end{gathered}$$

For λ_2

$$\begin{bmatrix} 0.99-\lambda_2 & 0.2 \\ -0.1 & 0.8-\lambda_2 \end{bmatrix} \begin{bmatrix} v_1 \\ v_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

$$\begin{bmatrix} 0.2 & 0.2 \\ 0.01 & 0.01 \end{bmatrix} \begin{bmatrix} v_1 \\ v_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

$$\begin{gathered} 0.2v_1 + 0.2v_2 = 0 \\ 0.2\text{v}_2 = -0.2v_1 \\ \text{v}_2 = -v_1 \end{gathered}$$

So eigenvectors and eigenvalues are:

$$\lambda_1 = 1, X_1 = \begin{bmatrix} 1 \\ 0.05 \end{bmatrix}$$

$$\lambda_2 = 0.79, X_2 = \begin{bmatrix} 1 \\ -1 \end{bmatrix}$$

Explaination:

An eigenvalue of 1 indicates a steady-state or equilibrium. This means that as time progresses , the distribution of researchers between industry and academia will stabilize.

In matrix transition models like this an eigenvalue of 1 signifies that this corresponding eigenvector represents the stable or final distribution of researchers between the two sectors (industry and academia) in the long run.

As $V_1 = \begin{bmatrix} 1 \\ 0.05 \end{bmatrix}$ It represents long-term distribution of researchers in industry and academia. When normalized, this eigenvector tells us that about **95% of researchers will work in industry**, and **5% will work in academia** in the long run, regardless of the initial distribution of researchers. This is the equilibrium that the system will reach after a sufficient amount of time, based on the given transition probabilities.

Initial distribution 10% in industry, 90% in academia: Even though we start at a very different distribution with initial conditions of 10% in industry and 90% in academia, the system will converge to the same **steady state distribution** of 95% in industry and 5% in academia. The highlight of the **initial condition doesn't affect the long term outcome**, but only influences how fast our system will reach the equilibrium.

Eigenvalue of 0.79 and its eigenvector:
The eigenvalue of $\lambda_2 = 0.79$ is lesser that $\lambda_1 = 1$ which means any influence of eigenvector associated with λ_2 diminshes over time.The eigen vector $\begin{bmatrix} -1 \\ 1 \end{bmatrix}$ represents the deviation from the steady state and affects the systems short term dynamics.As years progress towards 100 years of completion the eigen value λ_2 influence becomes smaller due to being less than 1 and the system is increasingly dominated by the steady state distribution given by the eigenvalue $\lambda_1 = 1$

2. (5 points) This is a programming task dedicated to SVD. You can find the assignment in the Google Colab notebook.

   First, make your own copy of the notebook (*File* $\rightarrow$ *Save a copy in Drive*) or download the notebook to your machine if you prefer to work locally (*File* $\rightarrow$ *Download*).

   Then, implement your solutions to the tasks formulated in the notebook. You can add **code cells** to write some code and **text cells** in case you want to include additional explanations to your answers in plain English.

   Finally, attach the link to you notebook to the submission form. **Make sure that all the cells are executed and all relevant outputs are being printed out**.

   Solution: Click here to go to code file.