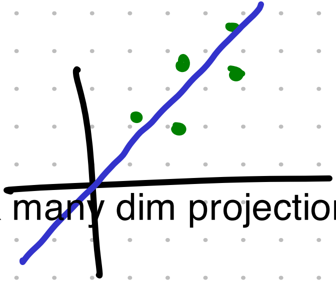
CS 2810 Day 8
Feb 11 2022

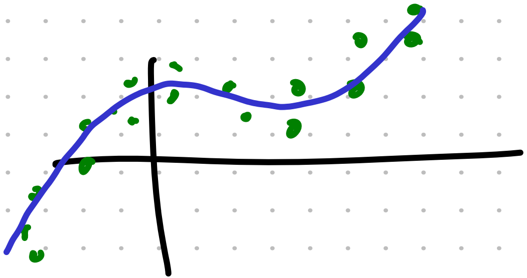
INSTALL

Computing with python / numpy
- build matrices, add/scale/multiply matrices, load matrix from csv, get inverse
get a row/col of matrix

Line of best fit (1d) & 1dim projections



Polynomial of best fit (many dimensions) & many dim projections



Computing with python / numpy

- build matrices
- add/scale/multiply matrices
- get inverse of a matrix:
  (whats an inverse?) A^{-1}
- get a row/col of matrix

- load matrix from csv:
  np.genfromtxt('mystery_matrix.csv', delimiter=',')

ICA 1:
You are welcome to handwrite / screenshot the answers to each question below, without showing the code used, for your ICA submission.

- Compute the matrix multiplication given below:

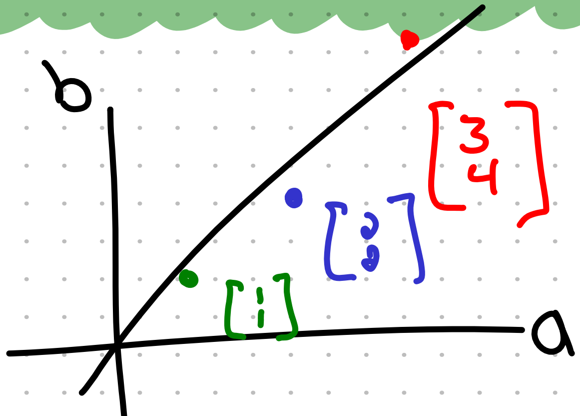
$$\begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{bmatrix} \begin{bmatrix} 10 & 11 \\ 12 & 13 \\ 14 & 15 \end{bmatrix}$$

- Load the data given in 'mystery_matrix.csv' given on website:

```
np.genfromtxt('mystery_matrix.csv', delimiter=',')
```

- What is the dot product of the second and fourth columns of the mystery matrix above?

Line of Best Fit (1 Dim)



Goal: Find a line of the form: b = ma which best fits data

$\vec{b} = m\vec{a}$

↑ SLOPE

$1 = m\,1$

$2 = m\,2$

$4 = m\,3$

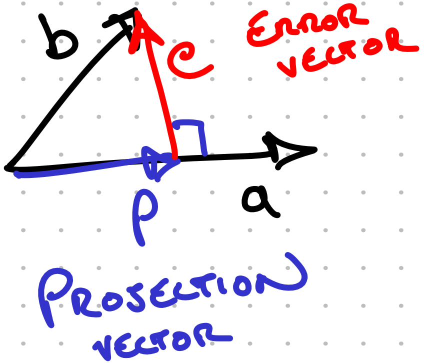
$$\underbrace{\begin{bmatrix} 1 \\ 2 \\ 4 \end{bmatrix}}_{\vec{b}} = m \underbrace{\begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}}_{\vec{a}}$$

no straight line can pass through all 3 points
-> no m exists with b = ma

instead, what if we find the vector p which
- has a solution p = ma
    - p is in the span of {a}
- is closest to b (llb-pll is minimized)

1 dimensional projections (often just called projections)

Among all vectors in the span of {a}, which is closest to b?



Some observations:

1. obs1: e is at a right angle (orthogonal) to a
2. obs2: b = p + e
3. obs3: p is in the span of a

2

$$e = b - p$$

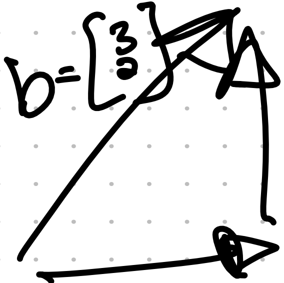
3

SOME CONSTANT

$$p = ca$$

1

$$e^T a = 0 \rightarrow (b-p)^T a = 0 \rightarrow (b - ca)^T a = 0$$



$b = \begin{bmatrix} 3 \\ 2 \end{bmatrix}$

$e = \begin{bmatrix} 0 \\ 2 \end{bmatrix}$

$p = \begin{bmatrix} 3 \\ 0 \end{bmatrix}$

$e + p = b$

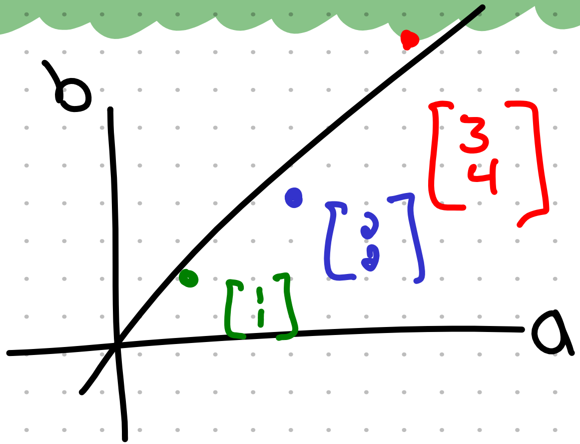
$$0 = (b - ca)^T a = b^T a - c\, a^T a$$

$$c\, a^T a = b^T a$$

$$c = \frac{b^T a}{a^T a}$$

$$ca = p = \frac{b^T a}{a^T a} a$$

LINE OF BEST FIT (1 DIM)



Goal: Find a line of the form: b = ma which best fits data

$$b = ma$$

$$\begin{bmatrix} 1 \\ 2 \\ 4 \end{bmatrix} = m \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}$$

NO SOLUTIONS

SOLVE $p = ma$ INSTEAD

$$p = ma$$

PROJECT b ONTO a

$$p = \frac{b^T a}{a^T a} a = \frac{1\cdot1 + 2\cdot2 + 4\cdot3}{1\cdot1 + 2\cdot2 + 3\cdot3} \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix} = \frac{17}{14} \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}$$

$b = ma$

$$\begin{bmatrix} 1 \\ 2 \\ 4 \end{bmatrix} = m \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}$$

b = ma has no solutions

so we'd like to replace b with p, where p ...
... is as close as possible to b
... is in the span of a -> (p = ma has some solution)

now we solve p = ma

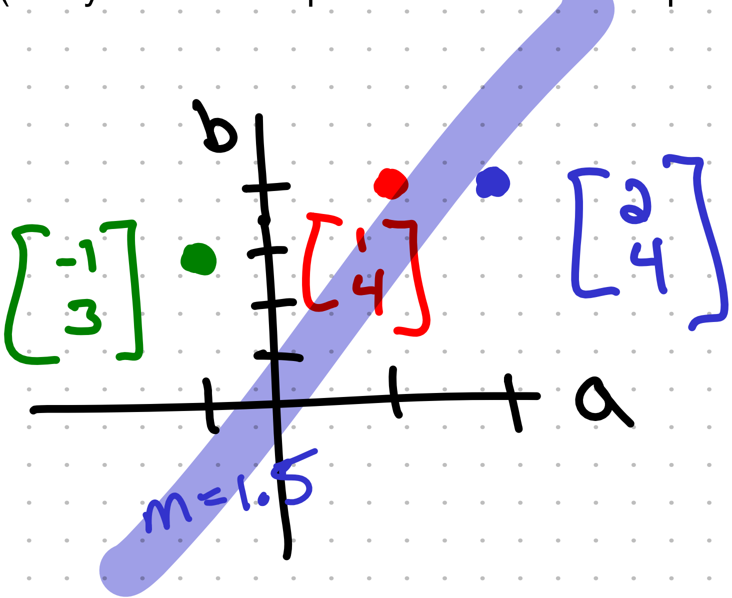
$p = ma$

$$p = \frac{17}{14} \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix} = m \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix} \qquad m = \frac{17}{14}$$

ICA 2:

b $\qquad$ a $\qquad$ $p = \frac{b^T a}{a^T a} a$

1. Find the projection of [2, 3, 4]^T in the span of [1, 2, 3]^T
2. Using a projection, find the line of best fit of the form b = ma for the scatter points shown below
3. There's a glaring problem with our line of best fit above, what is it? How might you fix it? (eerily similar to a problem we had with perceptrons having to pass through the origin too...)



$$\begin{bmatrix} -1 \\ 1 \\ 2 \end{bmatrix} m = \begin{bmatrix} 3 \\ 4 \\ 4 \end{bmatrix}$$

$$a\,m = b$$

$$p = \frac{b^T a}{a^T a} a \qquad a = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix} \qquad b = \begin{bmatrix} 2 \\ 3 \\ 4 \end{bmatrix}$$

$$= \frac{1\cdot 2 + 2\cdot 3 + 3\cdot 4}{1\cdot 1 + 2\cdot 2 + 3\cdot 3} \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}$$

$$\begin{bmatrix} -1 \\ 1 \\ 2 \end{bmatrix} m = \begin{bmatrix} 3 \\ 4 \\ 4 \end{bmatrix}$$

$$a \; m = b$$

CAN'T SOLVE

$$b = 1.5a$$

$$m = 1.5$$

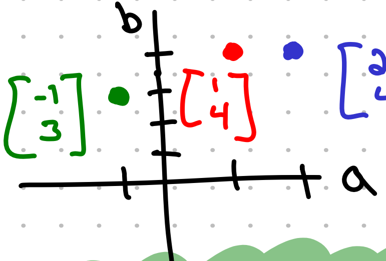
$$\begin{bmatrix} -1 \\ 1 \\ 2 \end{bmatrix} m = am = p = 1.5 \begin{bmatrix} -1 \\ 1 \\ 2 \end{bmatrix}$$

$$p = \frac{a^T b}{a^T a} a$$

$$= \frac{-1 \cdot 3 + 1 \cdot 4 + 2 \cdot 4}{1^2 + 1^2 + 2^2} \begin{bmatrix} -1 \\ 1 \\ 2 \end{bmatrix}$$

$$= 1.5 \begin{bmatrix} -1 \\ 1 \\ 2 \end{bmatrix}$$

LINE OF BEST FIT (MANY DIMENSIONS)



OLD MODEL $b = ma$

$3 = -1m$

$4 = 1m$

$4 = 2m$

$$\begin{bmatrix}3\\4\\4\end{bmatrix} = m\begin{bmatrix}-1\\1\\2\end{bmatrix}$$

NEW MODEL

$$b = m_0 + m_1 a$$

$3 = m_0 + m_1(-1)$

$4 = m_0 + m_1(1)$

$4 = m_0 + m_1(2)$

$$\begin{bmatrix}3\\4\\4\end{bmatrix} = m_0\begin{bmatrix}1\\1\\1\end{bmatrix} + m_1\begin{bmatrix}-1\\1\\2\end{bmatrix}$$

$$\underbrace{\begin{bmatrix}3\\4\\4\end{bmatrix}}_{b} = m_0\begin{bmatrix}1\\1\\1\end{bmatrix} + m_1\begin{bmatrix}-1\\1\\2\end{bmatrix} = \underbrace{\begin{bmatrix}1 & -1\\1 & 1\\1 & 2\end{bmatrix}}_{A}\underbrace{\begin{bmatrix}m_0\\m_1\end{bmatrix}}_{m}$$

b = Am has no solutions ...

what if we solve p = Am instead where p is
- in the span of the columns of A
  - (there is a linear combination of cols of A which equals p ... it's our target vector m!)
- closest to b

... we need a way of projecting a vector into the span of many vectors (previously only one)

Polynomial of Best Fit (Many Dimensions)

Newer Model

$$b = m_0 + m_1 a + m_2 a^2$$

$$3 = m_0 + m_1(-1) + m_2(-1)^2$$

$$4 = m_0 + m_1(1) + m_2(1)^2$$

$$4 = m_0 + m(2) + m_2(2)^2$$

$$\begin{bmatrix} 3 \\ 4 \\ 4 \end{bmatrix} = m_0 \begin{bmatrix} -1^0 \\ 1^0 \\ 2^0 \end{bmatrix} + m_1 \begin{bmatrix} -1^1 \\ 1^1 \\ 2^1 \end{bmatrix} + m_2 \begin{bmatrix} -1^2 \\ 1^2 \\ 2^2 \end{bmatrix}$$

New Model

$$b = m_0 + m_1 a$$

$$3 = m_0 + -1 m_1$$

$$4 = m_0 + 1 m_1$$

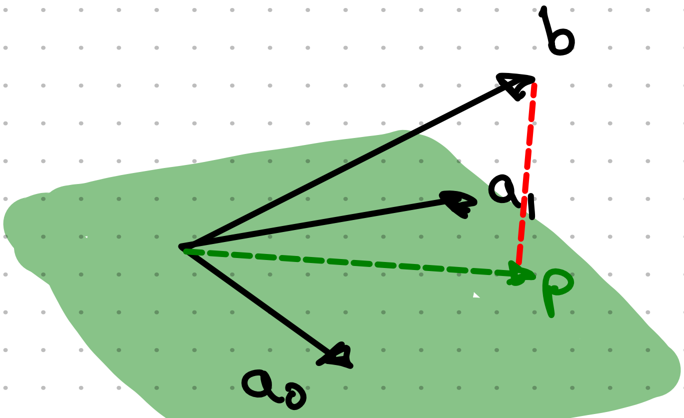
$$4 = m_0 + 2 m_1$$

$$\underbrace{\begin{bmatrix} 3 \\ 4 \\ 4 \end{bmatrix}}_{b} = m_0 \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} + m_1 \begin{bmatrix} -1 \\ 1 \\ 2 \end{bmatrix} = \underbrace{\begin{bmatrix} 1 & -1 \\ 1 & 1 \\ 1 & 2 \end{bmatrix}}_{A} \underbrace{\begin{bmatrix} m_0 \\ m_1 \end{bmatrix}}_{m}$$

$$\begin{bmatrix}3\\4\\4\end{bmatrix} = m_0\begin{bmatrix}-1^0\\1^0\\2^0\end{bmatrix} + m_1\begin{bmatrix}-1^1\\1^1\\2^1\end{bmatrix} + m_2\begin{bmatrix}-1^2\\1^2\\2^2\end{bmatrix} = \begin{bmatrix}-1^0 & -1^1 & -1^2\\1^0 & 1^1 & 1^2\\2^0 & 2^1 & 2^2\end{bmatrix}\begin{bmatrix}m_0\\m_1\\m_2\end{bmatrix}$$

PROJECTIONS (MULTI DIMENSIONAL)

Goal: find the vector p, in the span of {a_0, a_1, ... } which is closest to b



WITHOUT EXPLANATION

$$A = \begin{bmatrix} | & | \\ a_0 & a_1 \\ | & | \end{bmatrix}$$

$$p = A(A^TA)^{-1}A^Tb$$

Two questions:
- whats a matrix to the -1 power? ... next slide
- do you expect students to compute that mess by hand?

SINCE $Am = p = A(A^TA)^{-1}A^Tb$

$$m = (A^TA)^{-1}A^Tb$$

COEFFICIENTS OF LINE OF BEST FIT

IPYND DEMO

ICA3:
Load the matrices A and b from A.csv and b.csv respectively. (See zip next to today's notes on site)

A is the matrix with the polynomial manipulation shown here applied:

- column 0 is a to the 0th power
    - also known as: bias term
- column 1 is a to the 1st power
- column 2 is a to the 2nd power

b is a vector which contains

- where are the scatter points within A and b we're building the line of best fit to?
- Using numpy and python, find the m vector which defines the polynomial of best fit between a and b
- What polynomial is represented by this particular m vector? (Please round to the nearest integer)