# *Linear Algebra: Applications of Matrices*

Glenn Bruns
CSUMB

Some of the material in these slides comes from Géron's notes:
https://github.com/ageron/handson-ml/blob/master/math_linear_algebra.ipynb

# Learning outcomes

After this lecture you should be able to:

1. Use matrix multiplication to:
    - create derive new features from existing features
    - make predictions from a linear model
    - perform projections

# Review: Matrix multiplication

$$\begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} =$$

$$\begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix} \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} =$$

# Application: creating new features

Suppose we want to create a new feature that is derived from existing features.

feature data

$$\begin{pmatrix} 1 & -5 & 3 \\ 2 & -6 & 4 \end{pmatrix} \begin{pmatrix} 1 \\ .5 \\ .2 \end{pmatrix}$$

take this linear combination of existing features

$$= 1\begin{pmatrix} 1 \\ 2 \end{pmatrix} + .5\begin{pmatrix} -5 \\ -6 \end{pmatrix} + .2\begin{pmatrix} 3 \\ 4 \end{pmatrix}$$

$$= \begin{pmatrix} 1 \\ 2 \end{pmatrix} + \begin{pmatrix} -2.5 \\ -3 \end{pmatrix} + \begin{pmatrix} 0.6 \\ 0.8 \end{pmatrix}$$

$$= \begin{pmatrix} -0.9 \\ -0.2 \end{pmatrix}$$

# Another view of matrix multiplication

Each column of AB is a linear combination of the columns in A.

$$\begin{pmatrix} 1 & 2 \\ 2 & 3 \end{pmatrix} \begin{pmatrix} 1 & 1 \\ 2 & 3 \end{pmatrix} = \begin{pmatrix} 5 & 7 \\ 8 & 11 \end{pmatrix}$$

$$1\begin{pmatrix} 1 \\ 2 \end{pmatrix} + 2\begin{pmatrix} 2 \\ 3 \end{pmatrix} = \begin{pmatrix} 5 \\ 8 \end{pmatrix}$$

$$1\begin{pmatrix} 1 \\ 2 \end{pmatrix} + 3\begin{pmatrix} 2 \\ 3 \end{pmatrix} = \begin{pmatrix} 7 \\ 11 \end{pmatrix}$$

First column of B defines a linear combination of the columns of A

Second column of B defines a linear combination of the columns of A

# Create multiple new features

Suppose we want to create 3 new features derived from existing features.

this describes how to get third new feature

$$\begin{pmatrix} 1 & -5 & 3 \\ 2 & -6 & 4 \end{pmatrix} \begin{pmatrix} 1 & 1/3 & 1/2 \\ .5 & 1/3 & 1/2 \\ .2 & 1/3 & 0 \end{pmatrix}$$

$$= \begin{pmatrix} -0.9 & -\frac{1}{3} & -2 \\ -0.2 & 0 & -2 \end{pmatrix}$$

$$1\binom{1}{2} + 0.5\binom{-5}{-6} + 0.2\binom{3}{4} = ?$$

# Yet another view

You can also think of matrix multiplication AB as:
matrix A times every column of B

$$\begin{pmatrix} 1 & -5 & 3 \\ 2 & -6 & 4 \end{pmatrix} \begin{pmatrix} 1 & 1/3 & 1/2 \\ .5 & 1/3 & 1/2 \\ .2 & 1/3 & 0 \end{pmatrix}$$

$$\begin{pmatrix} 1 & -5 & 3 \\ 2 & -6 & 4 \end{pmatrix} \begin{pmatrix} 1 \\ .5 \\ .2 \end{pmatrix} \qquad \begin{pmatrix} 1 & -5 & 3 \\ 2 & -6 & 4 \end{pmatrix} \begin{pmatrix} 1/3 \\ 1/3 \\ 1/3 \end{pmatrix} \qquad \begin{pmatrix} 1 & -5 & 3 \\ 2 & -6 & 4 \end{pmatrix} \begin{pmatrix} 1/2 \\ 1/2 \\ 0 \end{pmatrix}$$

$$= \begin{pmatrix} -0.9 & -\frac{1}{3} & -2 \\ -0.2 & 0 & -2 \end{pmatrix}$$

# Application: linear regression

We make a prediction using a linear model:

$$\hat{y} = \theta_0 + \theta_1 x_1 + \theta_2 x_2 + \cdots + \theta_n x_n$$

where $x_1, x_2, \ldots, x\_p$ are the predictors and $\theta_0, \theta_1, \ldots, \theta_n$ are the model coefficients.

Example: predict car mileage from engine size and care weight.  After training our model we have:

$$\hat{y} = 50 - 1.5\, weight\ - 5.2\, size$$

If weight (in tons) is 2.5, and engine size (in 100's of cubic inches) is 3, what is predicted mileage?

# Application: linear regression

Making a prediction:

$$\hat{y} = \theta_0 \times \theta_1 x_1 + \theta_2 x_2 + \cdots + \theta_n x_n$$

vectorized form:

$$\hat{y} = \theta^T \cdot \mathbf{x}$$

$\theta, \mathbf{x}$ are column vectors

$$(\theta_0, \theta_1, \theta_2) \begin{pmatrix} 1 \\ x_1 \\ x_2 \end{pmatrix} = \theta_0 + \theta_1 x_1 + \theta_2 x_2$$

Mileage model example:

$$\hat{y} = 50 - 1.5\, weight\ - 5.2\, size$$

If $x = \begin{pmatrix} 1 \\ 2.5 \\ 3 \end{pmatrix}$, then $\hat{y} = (50, -1.5, -5.2) \begin{pmatrix} 1 \\ 2.5 \\ 3 \end{pmatrix} = 30.65$

# Matrices as vector transformers

Here's a function that maps a vector to another vector:

$$f\left(\begin{pmatrix} x \\ y \\ z \end{pmatrix}\right) = \begin{pmatrix} 2x \\ x + y \\ 3z \end{pmatrix}$$

Q: what is $f\left(\begin{pmatrix} -1 \\ 2 \\ 3 \end{pmatrix}\right)$ ?

Another way to write it:

$$f(u) = \begin{pmatrix} 2 & 0 & 0 \\ 1 & 1 & 0 \\ 0 & 0 & 3 \end{pmatrix} u$$

3 x 3   3 x 1 → 3 x 1

Lots of functions on vectors could be written like this:

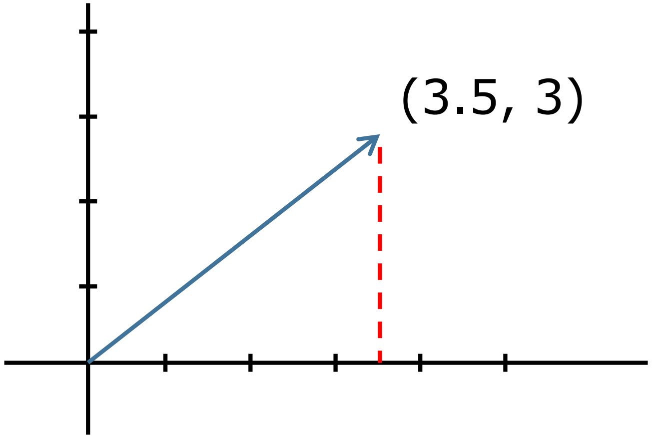
$$f(u) = Au$$

(A is $n \times n$, where $u$ has length $n$)

You can think of u being "transformed" by A.

# Projection matrices

What is the projection of the vector onto the x axis?

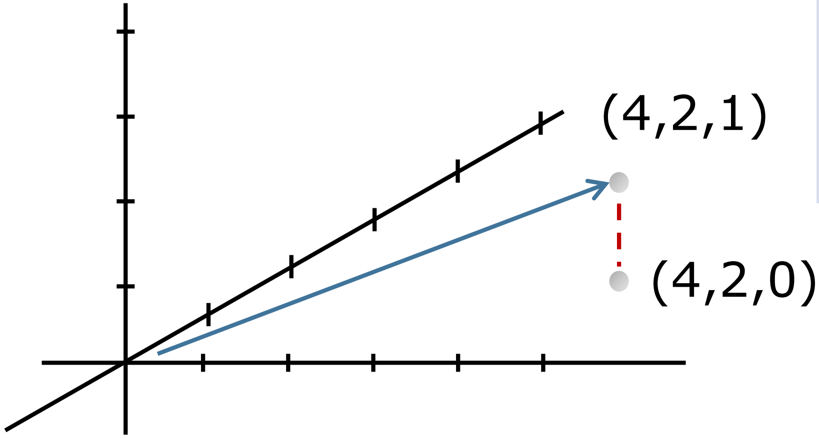



$$\begin{pmatrix} & \\ & \end{pmatrix}\begin{pmatrix}3.5\\3\end{pmatrix} = \begin{pmatrix}3.5\\0\end{pmatrix}$$

What values should be in the 2 by 2 projection matrix to give the desired projection?

# Projection matrices

What is the projection of the vector onto the x,y plane?




$$\begin{pmatrix} & & \\ & & \\ & & \end{pmatrix}\begin{pmatrix}4\\2\\1\end{pmatrix}=\begin{pmatrix}4\\2\\0\end{pmatrix}$$

What values should be in the 3 by 3 projection matrix to give the desired projection?

And to project the same vector onto the z axis?

# Summary

Some applications of matrix multiplication to machine learning

- derive new features as the linear combination of existing features
- compute predictions from a linear model
- project vectors