LINEAR ALGEBRA FOR DATA SCIENCE IN R

Motivation

Eric Eager
Data Scientist at Pro Football Focus

Data - The Atom of Data Science

```
  height weight forty vertical bench broad_jump three_cone shuttle
1     71    192  4.38     35.0    14        127       6.71    3.98
2     73    298  5.34     26.5    27         99       7.81    4.71
3     77    256  4.67     31.0    17        113       7.34    4.38
4     74    198  4.34     41.0    16        131       6.56    4.03
5     76    257  4.87     30.0    20        118       7.12    4.23
6     78    262  4.60     38.5    18        128       7.53    4.48
```

Vectors - Storing Univariate Data

$$\vec{x} = \begin{pmatrix} 1 \\ 5 \\ -2 \\ 4 \end{pmatrix}$$

$$\vec{x}^T = \begin{pmatrix} 1 & 5 & -2 & 4 \end{pmatrix}$$

$$\vec{y} = \begin{pmatrix} 11 & -7 & 12 & 14 & 21 \end{pmatrix}$$

$$\vec{y}^T = \begin{pmatrix} 11 \\ -7 \\ 12 \\ 14 \\ 21 \end{pmatrix}$$

Vectors - Storing Univariate Data

```
> x <- rep(1, 4)
> x
[1]  1  1 1 1
```

```
> y <- seq(2, 8, by = 2)
> y
[1] 2 4 6 8
```

```
> z <- c(1, 5, -2, 4)
> z
[1]  1  5 -2  4
```

```
> z[3] <- 7
> z
[1]  1  5 7  4
```

Matrices - Storing Tables of Data

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

Matrices - Storing Tables of Data

Case	Variable 1	Variable 2
A	-2	-4
B	-1	-2
C	0	0
D	1	2
F	2	4

Matrices - Storing Tables of Data

```
> matrix(2, 3, 2)
     [,1] [,2]
[1,]    2    2
[2,]    2    2
[3,]    2    2
```

```
> matrix(c(1, -1, 2, 3, 2, -2), nrow = 2, ncol = 3, byrow = TRUE)
     [,1] [,2] [,3]
[1,]    1   -1    2
[2,]    3    2   -2
```

```
> matrix(c(1, -1, 2, 3, 2, -2), nrow = 2, ncol = 3, byrow = FALSE)
     [,1] [,2] [,3]
[1,]    1    2    2
[2,]   -1    3   -2
```

```
> A[2, 1] <- 100
> print(A)
     [,1] [,2] [,3]
[1,]    1    2    2
[2,]  100    3   -2
```

LINEAR ALGEBRA FOR DATA SCIENCE IN R

Let's practice!

LINEAR ALGEBRA FOR DATA SCIENCE IN R

Matrix-Vector Operations

Eric Eager
Data Scientist at Pro Football Focus

How Matrix-Vector Multiplication Works

$$A\vec{x} =$$

$$\underbrace{\begin{pmatrix} 1 & -1 \\ 2 & 1 \\ 4 & -2 \end{pmatrix}}_{3x2} \times \underbrace{\begin{pmatrix} 1 \\ 2 \end{pmatrix}}_{2x1} = \begin{pmatrix} 1 \times 1 & + & (-1) \times 2 \\ 2 \times 1 & + & 1 \times 2 \\ 4 \times 1 & + & (-2) \times 2 \end{pmatrix} = \underbrace{\begin{pmatrix} -1 \\ 4 \\ 0 \end{pmatrix}}_{3x1}$$

How Matrix-Vector Multiplication Works

$$A\vec{x} = \begin{pmatrix} 1 & -1 \\ 2 & 1 \\ 4 & -2 \end{pmatrix} \times \begin{pmatrix} 1 \\ 2 \end{pmatrix} = \begin{pmatrix} 1 \times 1 & + & (-1) \times 2 \\ 2 \times 1 & + & 1 \times 2 \\ 4 \times 1 & + & (-2) \times 2 \end{pmatrix} = \begin{pmatrix} -1 \\ 4 \\ 0 \end{pmatrix}$$

3x2 2x1 3x1

How Matrix-Vector Multiplication Works

```
> A
     [,1] [,2]
[1,]    1   -1
[2,]    2    1
[3,]    4   -2

> b
[1] 1 2
```

```
> A%*%b
     [,1]
[1,]   -1
[2,]    4
[3,]    0
```

How Matrix-Vector Multiplication Works

$$A\vec{x} =$$

$$\begin{pmatrix} 1 & 3 & 2 \\ 1 & 1 & 4 \end{pmatrix} \times \begin{pmatrix} 0 \\ 1 \\ 2 \end{pmatrix} = \begin{pmatrix} 1 \times 0 + 3 \times 1 + 2 \times 2 \\ 1 \times 0 + 1 \times 1 + 4 \times 2 \end{pmatrix} = \begin{pmatrix} 7 \\ 9 \end{pmatrix}$$

How Matrix-Vector Multiplication Works

```
> A[1,]%*%b
     [,1]
[1,]    7

> A[2,]%*%b
     [,1]
[1,]    9

> A%*%b
     [,1]
[1,]    7
[2,]    9
```

Matrix-Vector Multiplication Motivation

Teams	Johns Hopkins	F & M	Gettysburg	Dickinson	McDaniel
Johns Hopkins	-	Loss, 12 - 14	Win 49-35	Win 49-0	Win 49-7
F & M	Win, 14 - 12	-	Loss, 31-38	Win 36-28	Win 35-10
Gettsyburg	Loss 35-49	Win, 38-31	-	Loss 13-23	Win 35-3
Dickinson	Loss 0-49	Loss 28-36	Win 23-13	-	Win 38-31
McDaniel	Loss 7-49	Loss 10-35	Loss 3-35	Loss 31-38	-

Matrix-Vector Multiplication Motivation

$$
\begin{array}{c|ccccc}
 & \text{JH} & \text{F\&M} & \text{G} & \text{D} & \text{McD} \\
\hline
\text{JH} & 4 & -1 & -1 & -1 & -1 \\
\text{F\&M} & -1 & 4 & -1 & -1 & -1 \\
\text{G} & -1 & -1 & 4 & -1 & -1 \\
\text{D} & -1 & -1 & -1 & 4 & -1 \\
\text{McD} & -1 & -1 & -1 & -1 & 4
\end{array}
\times \begin{pmatrix} r_{\text{JH}} \\ r_{\text{F\&M}} \\ r_{\text{G}} \\ r_{\text{D}} \\ r_{\text{McD}} \end{pmatrix} = \begin{pmatrix} 103 \\ 28 \\ 15 \\ -40 \\ -106 \end{pmatrix}
$$

LINEAR ALGEBRA FOR DATA SCIENCE IN R

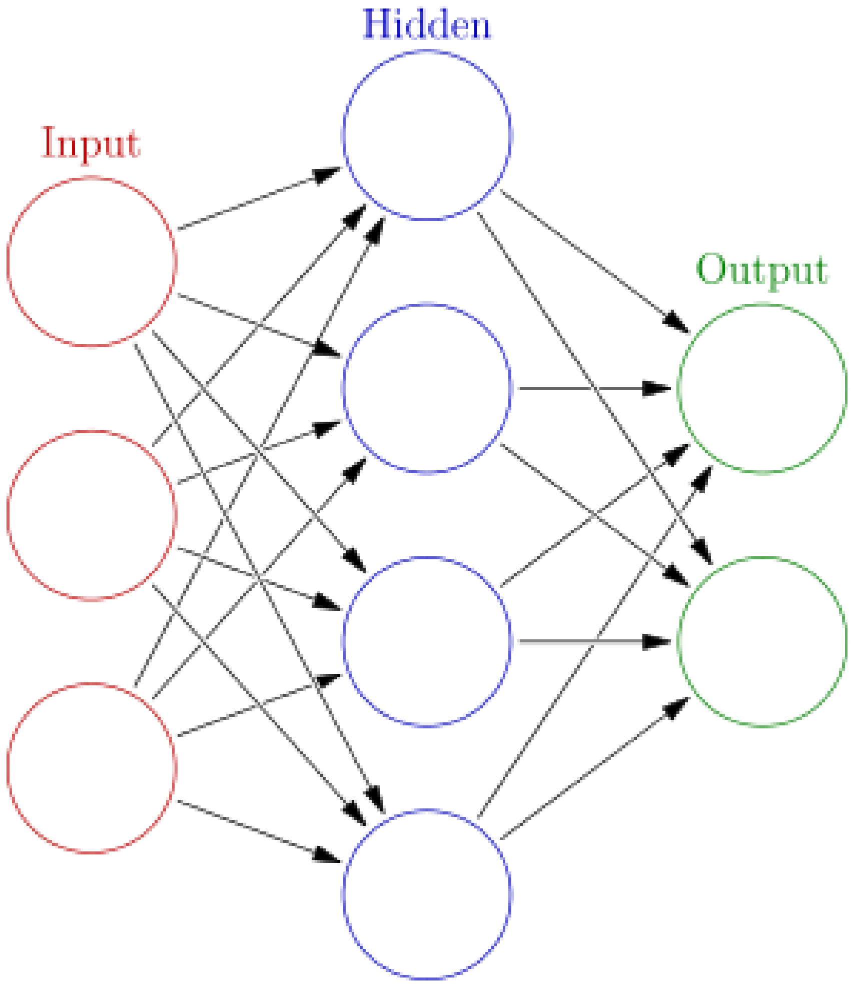
Let's practice!

LINEAR ALGEBRA FOR DATA SCIENCE IN R

Matrix-Matrix Calculations

Data Scientist at Pro Football Focus
Instructor

Matrix-Matrix Multiplication Motivations



How Matrix Multiplication Works

$$AB =$$

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

$$\begin{pmatrix} 1\times 0 \;+\; 2\times 1 & 1\times 1 \;+\; 2\times 2 \\ 2\times 0 \;+\; 1\times 1 & 2\times 1 \;+\; 1\times 2 \end{pmatrix} =$$

$$\begin{pmatrix} 2 & 5 \\ 1 & 4 \end{pmatrix}$$

How Matrix Multiplication Works

```
> A%*%B
     [,1] [,2]
[1,]    2    5
[2,]    1    4
```

```
> B%*%A
     [,1] [,2]
[1,]    2    1
[2,]    5    4
```

```
> A*B
     [,1] [,2]
[1,]    0    2
[2,]    2    2
```

The Identity Matrix

$$I_3 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

The Identity Matrix

```
> A
     [,1] [,2]
[1,]    1    2
[2,]    2    1
> I <- diag(2)
> I
     [,1] [,2]
[1,]    1    0
[2,]    0    1

> I%*%A
     [,1] [,2]
[1,]    1    2
[2,]    2    1
> A%*%I
     [,1] [,2]
[1,]    1    2
[2,]    2    1
```

Additional Importance Concepts for Matrices

1. Square Matrices
2. The Matrix Inverse
3. Singular Matrices
4. Diagonal and Triangular Matrices

LINEAR ALGEBRA FOR DATA SCIENCE IN R

Let's practice!