# CS 725: Foundations of Machine Learning

## Overview of linear algebra for ML

August 17, 2020

# Vectors

An $n$-dimensional vector is simply $n$ real numbers, written as a column:

$$\mathbf{u} = \begin{bmatrix} 1 \\ 2 \\ -3 \end{bmatrix}$$

The set of all $n$-dimensional vectors is denoted by $\mathbb{R}^n$

Note: Can consider more general vector spaces over "fields." We'll stick to $\mathbb{R}$.

# Vector Operations

Vector operations: *scalar multiplication* and *vector addition*

*Linear Combination*: Combining a set of vectors in $\mathbb{R}^n$ using these operations yields another vector in $\mathbb{R}^n$

E.g.,

$$\mathbf{u} = \begin{bmatrix} 1 \\ 2 \\ -3 \end{bmatrix} \qquad \mathbf{v} = \begin{bmatrix} -1 \\ -1 \\ 0 \end{bmatrix} \qquad \Rightarrow \qquad 2\mathbf{u} + \mathbf{v} = \begin{bmatrix} 1 \\ 3 \\ -6 \end{bmatrix}$$

# Basis Vectors

Any vector in $\mathbb{R}^n$ can be written as a linear combination of the (standard) basis vectors:

$$\mathbf{u} = \mathbf{e}_1 + 2\mathbf{e}_2 - 3\mathbf{e}_3$$

where

$$\mathbf{e}_1 = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} \qquad \mathbf{e}_2 = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} \qquad \mathbf{e}_3 = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$$

# Linear Transformations

Linear algebra deals with “linear” functions (or linear transformations) that map vectors in $\mathbb{R}^n$ to vectors in $\mathbb{R}^m$

$f$ is said to be linear if

$$f(\mathbf{u} + \mathbf{v}) = f(\mathbf{u}) + f(\mathbf{v})$$

$$f(c\,\mathbf{u}) = c\,f(\mathbf{u})$$

# Linear Transformations

A linear transformation is fully specified by how it acts on the basis vectors:

$$\mathbf{u} = u_1\mathbf{e}_1 + \ldots + u_n\mathbf{e}_n \implies f(\mathbf{u}) = u_1 f(\mathbf{e}_1) + \ldots + u_n f(\mathbf{e}_n)$$

Thus $\big(f(\mathbf{e}_1), \ldots, f(\mathbf{e}_n)\big)$ fully determines $f$

Can denote such a transformation by a matrix, $\mathbf{A} = \begin{bmatrix} \mathbf{a}_1 & \mathbf{a}_2 & \ldots & \mathbf{a}_n \end{bmatrix}$ where each $\mathbf{a}_i = f(\mathbf{e}_i)$ is a column vector.

# Linear Transformation using Matrix Multiplication

Let $\mathbf{A} = \begin{bmatrix} \mathbf{a}_1 & \mathbf{a}_2 & \dots & \mathbf{a}_n \end{bmatrix}$ be an $m \times n$ matrix.

Multiplication of a vector $\mathbf{u}$ by the matrix $\mathbf{A}$ is defined so that $f(\mathbf{u}) = \mathbf{Au}$:

$$\begin{bmatrix} \mathbf{a}_1 & \mathbf{a}_2 & \dots & \mathbf{a}_n \end{bmatrix} \begin{bmatrix} u_1 \\ u_2 \\ \vdots \\ u_n \end{bmatrix} = \sum_i u_i \mathbf{a}_i =: \mathbf{Au}$$

Multiplying two matrices: $\mathbf{BA} = \begin{bmatrix} \mathbf{Ba}_1 & \mathbf{Ba}_2 & \dots & \mathbf{Ba}_n \end{bmatrix}$ where $\mathbf{Ba}_i = f_{\mathbf{B}}(\mathbf{a}_i) = f_{\mathbf{B}} f_{\mathbf{A}}(\mathbf{e}_i)$.

# Matrix Multiplication

If $\mathbf{V}$ is a matrix of size $n \times m$ and $\mathbf{W}$ is an $m \times p$ matrix, then the matrix product $\mathbf{U} = \mathbf{VW}$ is of size $n \times p$ and can be computed as follows:

$$\mathbf{V} = \begin{bmatrix} v_{1,1} & \dots & v_{1,m} \\ \vdots & \vdots & \vdots \\ v_{n,1} & \dots & v_{n,m} \end{bmatrix}, \mathbf{W} = \begin{bmatrix} w_{1,1} & \dots & w_{1,p} \\ \vdots & \vdots & \vdots \\ w_{m,1} & \dots & w_{m,p} \end{bmatrix}, \mathbf{U} = \begin{bmatrix} u_{1,1} & \dots & u_{1,p} \\ \vdots & \vdots & \vdots \\ u_{n,1} & \dots & u_{n,p} \end{bmatrix}$$

where $u_{i,j} = \sum_{k=1}^{m} v_{i,k} w_{k,j}$

Matrix multiplication is not necessarily commutative ($\mathbf{AB} \neq \mathbf{BA}$), but it is associative ($\mathbf{A}(\mathbf{BC}) = (\mathbf{AB})\mathbf{C}$).

# Determinants

Determinants are defined for square matrices $\mathbf{A} \in \mathbb{R}^{n \times n}$, and denoted by $\det(\mathbf{A})$ or $|\mathbf{A}|$.

$$
\begin{aligned}
|a_{11}| &= a_{11} \\
\left| \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} \right| &= a_{11}a_{22} - a_{12}a_{21} \\
\left| \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix} \right| &= \begin{array}{c} a_{11}a_{22}a_{33} - a_{11}a_{23}a_{32} - a_{12}a_{21}a_{33} \\ +a_{12}a_{23}a_{31} + a_{13}a_{21}a_{32} - a_{13}a_{22}a_{31} \end{array}
\end{aligned}
$$