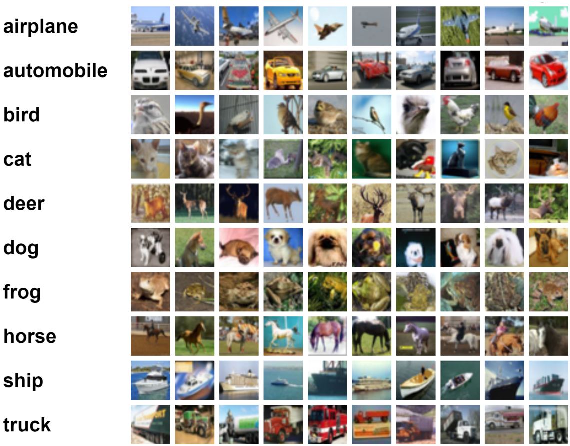
Linear Algebra Basics: Vectors

Foundations of Data Analysis

February 15, 2022

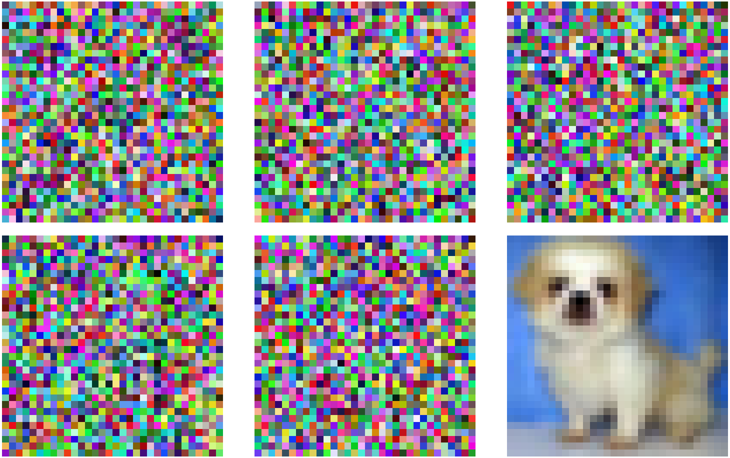
CIFAR-10



$32 \times 32 \times 3 = 3{,}072$ dimensions

10 classes

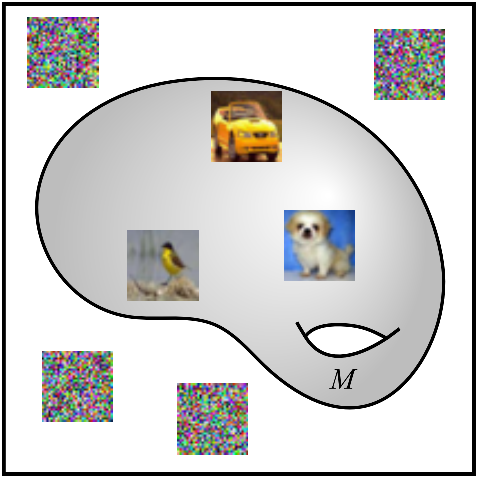
Uniform Random Images



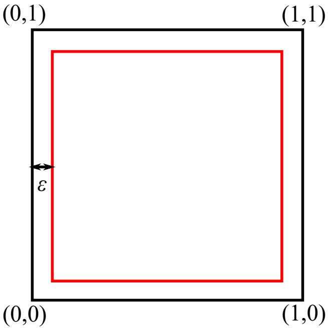
just kidding!

Manifold Hypothesis

Real data lie near lower-dimensional manifolds



Area of a Shrunken Square



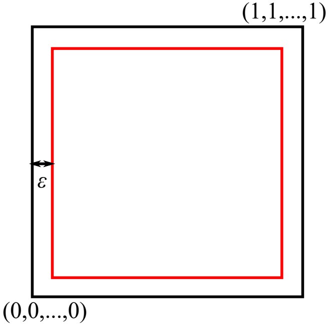
What is the volume of the unit square shrunk by some small amount in each dimension?

$$A = (1 - 2\epsilon)^2$$

Example: $\epsilon = \frac{1}{256}$

$$A \approx 0.9844$$

Volume in High Dimensions



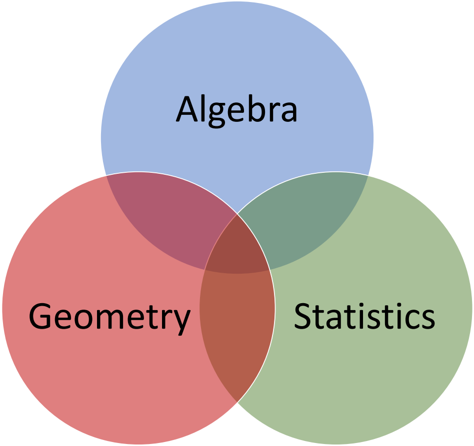
What is the volume of the unit d-cube shrunk by some small amount in each dimension?

$$V = (1 - 2\epsilon)^d$$

Approaches 0 as $d \to \infty$

Example: $256 \times 256 \times 3$ images, $\epsilon = \frac{1}{256}$

$$V \approx 2.0 \times 10^{-670}$$


Algebra
Geometry
Statistics

Types of Data

- **Categorical** (outcomes come from a discrete set)
- **Real-valued** (outcomes come from $\mathbb{R}$)
- **Ordinal** (outcomes have an order, e.g., integers)
- **Vector** (outcomes come from $\mathbb{R}^d$)

Most data is a combination of multiple types!

Vectors

A vector is a list of real numbers:

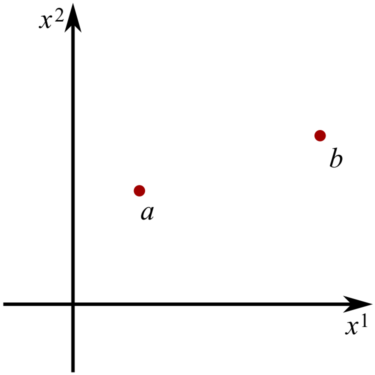
$$x = \begin{bmatrix} x^1 \\ x^2 \\ \vdots \\ x^d \end{bmatrix}$$

Notation: $x \in \mathbb{R}^d$

Notation: We will use superscripts for coordinates, subscripts when talking about a collection of vectors, $x_1, x_2, \ldots, x_n \in \mathbb{R}^d$.

Geometry: Direction and Distance

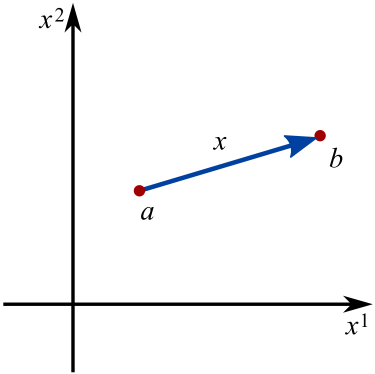
A vector is the difference between two points:



$$a = \begin{pmatrix} a^1 \\ a^2 \\ \vdots \\ a^d \end{pmatrix}, \quad b = \begin{pmatrix} b^1 \\ b^2 \\ \vdots \\ b^d \end{pmatrix},$$

Geometry: Direction and Distance

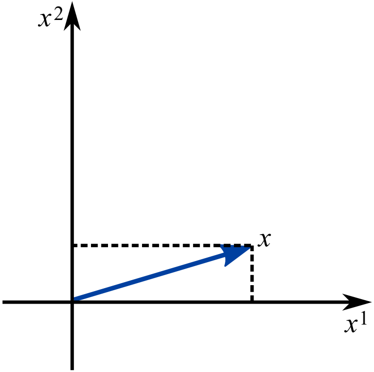
A vector is the difference between two points:



$$a = \begin{pmatrix} a^1 \\ a^2 \\ \vdots \\ a^d \end{pmatrix}, \quad b = \begin{pmatrix} b^1 \\ b^2 \\ \vdots \\ b^d \end{pmatrix},$$

$$x = \begin{bmatrix} x^1 \\ x^2 \\ \vdots \\ x^d \end{bmatrix} = \begin{bmatrix} b^1 - a^1 \\ b^2 - a^2 \\ \vdots \\ b^d - a^d \end{bmatrix}$$

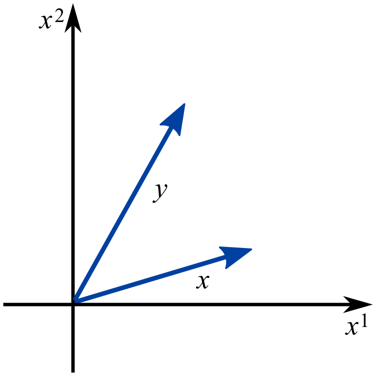
Points as Vectors



We will often treat points as vectors, although they are technically not the same thing.

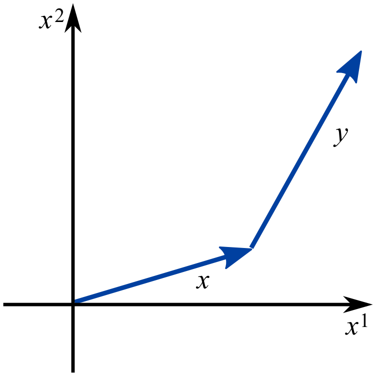
Think of a vector being anchored at the origin: $0 = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix}$

Vector Addition



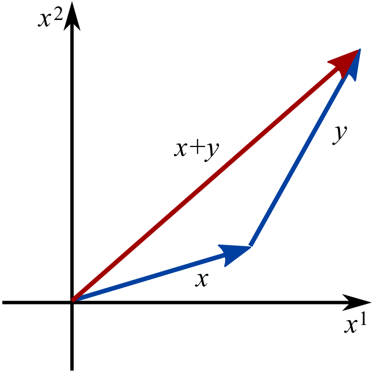
$$x + y = \begin{bmatrix} x^1 + y^1 \\ x^2 + y^2 \\ \vdots \\ x^d + y^d \end{bmatrix}$$

Vector Addition



$$x + y = \begin{bmatrix} x^1 + y^1 \\ x^2 + y^2 \\ \vdots \\ x^d + y^d \end{bmatrix}$$

Vector Addition



$$x + y = \begin{bmatrix} x^1 + y^1 \\ x^2 + y^2 \\ \vdots \\ x^d + y^d \end{bmatrix}$$

Scalar Multiplication

Multiplication between a vector $x \in \mathbb{R}^d$ and a scalar $s \in \mathbb{R}$:

$$sx = s \begin{bmatrix} x^1 \\ x^2 \\ \vdots \\ x^d \end{bmatrix} = \begin{bmatrix} sx^1 \\ sx^2 \\ \vdots \\ sx^d \end{bmatrix}$$

Statistics: Vector Mean

Given vector data $x_1, x_2, \ldots, x_n \in \mathbb{R}^d$, the mean is

$$\bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i = \begin{bmatrix} \frac{1}{n}\sum_{i=1}^{n} x_i^1 \\ \frac{1}{n}\sum_{i=1}^{n} x_i^2 \\ \vdots \\ \frac{1}{n}\sum_{i=1}^{n} x_i^d \end{bmatrix}$$

Notice that this is a vector of means in each dimension.

Vector Norm

The norm of a vector is its length:

$$\|x\| = \sqrt{\sum_{i=1}^{d} (x^i)^2}$$

Statistics: Total Variance

Remember, the equation for the variance of scalar data, $y_1, \ldots, y_n \in \mathbb{R}$:

$$\mathrm{var}(y) = \frac{1}{n-1} \sum_{i=1}^{n} (y_i - \bar{y})^2.$$

For **total variance** for vector data, $x_1, \ldots, x_n \in \mathbb{R}^d$, is

$$\mathrm{var}(x) = \frac{1}{n-1} \sum_{i=1}^{n} \|x_i - \bar{x}\|^2.$$

Dot Product

Given two vectors, $x, y \in \mathbb{R}^d$, their dot product is

$$\langle x, y \rangle = x^1 y^1 + x^2 y^2 + \cdots + x^d y^d = \sum_{i=1}^{d} x^i y^i.$$

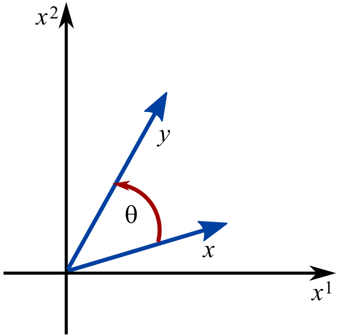
Also known as the **inner product**.

Relation to norm:

$$\|x\| = \sqrt{\langle x, x \rangle}$$

Geometry: Angles and Lengths

The dot product tells us the angle θ between two vectors, $x, y \in \mathbb{R}^d$:



$$\langle x, y \rangle = \|x\| \|y\| \cos \theta.$$

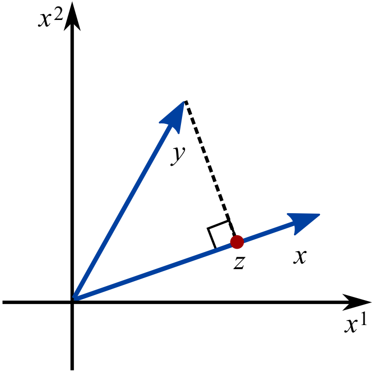
Or, rewriting to solve for θ: $\theta = \arccos \frac{\langle x,y \rangle}{\|x\| \|y\|}$.

Geometry: Orthogonality

Two vectors at a 90 degree angle ($\pi/2$ radians) are called orthogonal.
There dot product is zero:

$$\langle x, y\rangle = \|x\|\|y\|\cos\frac{\pi}{2} = \|x\|\|y\|0 = 0$$

Geometry: Projection

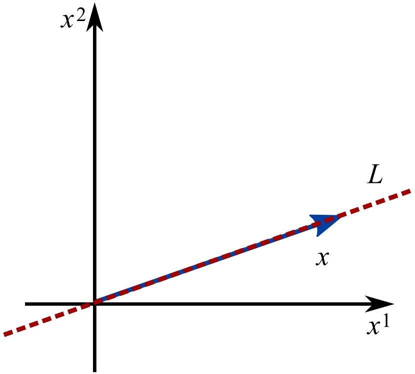


$$z = \frac{x}{\|x\|^2}\langle x, y\rangle$$

Equation for a Line

Line passing through the origin along vector $x \in \mathbb{R}^d$

$$L = \{tx : t \in \mathbb{R}\}$$



Linear Independence

Two vectors, $x_1, x_2 \in \mathbb{R}^d$, are linearly independent if they aren't scaled versions of each other:

$$sx_1 \neq x_2, \quad \text{for all } s \in \mathbb{R}.$$

Equation for a Plane

Two linearly independent vectors, $x, y \in \mathbb{R}^d$, span a plane:

$$H = \{sx + ty : s \in \mathbb{R}, t \in \mathbb{R}\}$$