# Preliminaries

Garrett Thomas

May 29, 2018

## 1 About

This document is part of a series of notes about math and machine learning. You are free to distribute it as you wish. The latest version can be found at http://gwthomas.github.io/notes. Please report any errors to gwthomas@stanford.edu.

## 2 Sets

A **set** is a collection of objects which is determined entirely by its contents. The elements of a set have no ordering or multiplicity within the set. If $S$ is a set, we write $x \in S$ to denote that the object $x$ is a member of the set $S$, and $x \notin S$ to denote that $x$ is not a member of $S$. The set which contains no elements is referred to as **the empty set** and denoted $\varnothing$.

If $A$ and $B$ are sets, and every element of $A$ is also an element of $B$, we say that $A$ is a **subset** of $B$ (or equivalently, that $B$ is a **superset** of $A$) and write $A \subseteq B$. Two sets $A$ and $B$ are equal if both $A \subseteq B$ and $B \subseteq A$, which is to say that they contain exactly the same elements.

### 2.1 Specifying sets

Sets can be written down in a number of ways. If there are only a few elements, one can simply list them out, e.g. $\{\text{cat}, 3, a\}$. For sets too large to enumerate, ellipsis notation $(\dots)$ indicates that the reader should be able to infer the pattern, e.g. $\{2, 4, 6, 8, \dots\}$ or $\{2, 4, 6, \dots, 100\}$. The trailing ellipses in the first of these indicates that the pattern continues infinitely.

It is also common to use **set-builder** notation to specify a set by means of one or more conditions. For example, to specify the set formed by keeping only the positive elements of some set $A$, we might write $\{x \in A : x \text{ is positive}\}$.

One can also describe sets with natural language, as long as the description is unambiguous, e.g. "the set of all humans who were born before the year 2000".

### 2.2 Constructing sets from existing sets

The **intersection** of $A$ and $B$, denoted $A \cap B$, is the set of elements which lie in both $A$ and $B$:

$$A \cap B \triangleq \{x : x \in A, x \in B\}$$

$A$ and $B$ are said to be **disjoint** if they have no elements in common, i.e. $A \cap B = \varnothing$.

The **union** of $A$ and $B$, denoted $A \cup B$, is the set of elements which lie in $A$ or $B$ (or both):

$$A \cup B \triangleq \{x : x \in A \text{ or } x \in B\}$$

Clearly $A \cap B \subseteq A, B \subseteq A \cup B$ for any sets $A$ and $B$.

These binary operations have natural generalizations to combinations of many sets. Let $\{A_i\}_{i \in I}$ be a collection of sets indexed by some set $I$. Then the intersection and union of this collection are

$$\bigcap_{i \in I} A_i \triangleq \{x : x \in A_i \text{ for all } i \in I\}$$

$$\bigcup_{i \in I} A_i \triangleq \{x : x \in A_i \text{ for at least one } i \in I\}$$

The **difference** of $A$ and $B$ is the set of elements which lie in $A$ but not in $B$:

$$A \setminus B \triangleq \{x : x \in A, x \notin B\}$$

If $A \subseteq B$, the **complement** of $A$ (in $B$) is defined as $A^{\mathrm{c}} \triangleq B \setminus A$. Note that the notation for complements does not explicitly include the containing set $B$; this set should be clear from context.

The interplay between complements, intersections, and unions is well-captured by **De Morgan's laws**:

$$(A \cap B)^{\mathrm{c}} = A^{\mathrm{c}} \cup B^{\mathrm{c}} \qquad\qquad (A \cup B)^{\mathrm{c}} = A^{\mathrm{c}} \cap B^{\mathrm{c}}$$

or, more generally,

$$\left(\bigcap_{i \in I} A_i\right)^{\mathrm{c}} = \bigcup_{i \in I} A_i^{\mathrm{c}} \qquad\qquad \left(\bigcup_{i \in I} A_i\right)^{\mathrm{c}} = \bigcap_{i \in I} A_i^{\mathrm{c}}$$

# 3 Common sets of numbers

The **natural numbers** are the set $\mathbb{N} \triangleq \{0, 1, 2, \dots\}$.

The **integers** are the set $\mathbb{Z} \triangleq \{\dots, -2, -1, 0, 1, 2, \dots\}$.

The **rational numbers** are the set $\mathbb{Q} \triangleq \{\frac{p}{q} : p \in \mathbb{Z}, q \in \mathbb{Z}, q \neq 0\}$.

The **real numbers** are denoted $\mathbb{R}$. We will not define them here but trust that the reader is familiar with their basic properties.

# 4 Tuples

An $n$**-tuple** (where $n$ is a positive integer) is an ordered list of $n$ objects, denoted with parentheses, e.g. $(x_1, x_2, \dots, x_n)$. If the number of items is unimportant, we will simply say **tuple**. A 2-tuple is referred to as an **ordered pair**, and a 3-tuple as a **triple**.

The **Cartesian product** of $n$ sets $A_1, \dots, A_n$ is the set of all $n$-tuples whose entries come from those sets:

$$A_1 \times \cdots \times A_n \triangleq \{(x_1, \dots, x_n) : x_1 \in A_1, \dots, x_n \in A_n\}$$

For example, $\{1, 2, 3\} \times \{4, 5\} = \{(1, 4), (2, 4), (3, 4), (1, 5), (2, 5), (3, 5)\}$. If $A$ is a set, the notation $A^n$ means the Cartesian product of $n$ copies of $A$, i.e.

$$A^n \triangleq \underbrace{A \times \cdots \times A}_{n \text{ times}}$$

We will often have cause to refer to $\mathbb{R}^n$, the set of $n$-tuples of real numbers.

## 5 Functions

A **function** is a rule which associates to each element of a certain set, known as the **domain** of the function, exactly one element of another set, known as the **codomain** of the function. We write $f : X \to Y$ to denote that $f$ is a function from $X$ to $Y$; here $X$ is the domain (sometimes denote $\text{dom}(f)$), and $Y$ the codomain. The element of $Y$ associated to $x \in X$ is written $f(x)$.

Perhaps the simplest example of a function is the **identity** function on a set $X$, which maps every element to itself:

$$\begin{aligned} \text{id}_X : X &\to X \\ x &\mapsto x \end{aligned}$$

Functions can be chained together if their domains and codomains are appropriately matched up. If $f : X \to Y$ and $g : Y \to Z$, then their **composition** $g \circ f$ is defined as

$$\begin{aligned} g \circ f : X &\to Z \\ x &\mapsto g(f(x)) \end{aligned}$$

The **image** of $S \subseteq X$ under $f$, denoted $f(S)$, is the set of elements in $Y$ that are mapped to by elements of $S$:

$$f(S) \triangleq \{f(x) : x \in S\}$$

Conversely, the **preimage** of $S \subseteq Y$ under $f$, denoted $f^{-1}(S)$, is the set of elements of $X$ which map to some element of $S$:

$$f^{-1}(S) \triangleq \{x \in X : f(x) \in S\}$$

The **range** of $f$ is the image of the entire domain of $f$: $\text{range}(f) = f(X)$.

A function is said to be **surjective** if its range is the entirety of its codomain, i.e. $f(X) = Y$. It is **injective** if $f(x) = f(y)$ implies $x = y$ for all $x, y \in X$. A function that is both surjective and injective is said to be **bijective**. If $f$ is bijective, it possesses an **inverse** $f^{-1} : Y \to X$ such that $f^{-1} \circ f = \text{id}_X$ (i.e. $f^{-1}(f(x)) = x$ for all $x \in X$) and $f \circ f^{-1} = \text{id}_Y$.[1]

[1] Note the overloading of the notation $f^{-1}$! It is unfortunate but standard.