CS166 Handout 02

Spring 2021 March 30, 2021

Mathematical Terms and Identities

Thanks to Andy Nguyen and Julie Tibshirani for their advice on this handout.

This handout covers mathematical concepts, notation, and identities that may be useful over the course of CS166. The topics included here are usually taught over a combination of CS103, CS107, CS109, and CS161, though because the content in those classes shifts from quarter to quarter there's a chance that you may not have seen everything here. If that's the case, no worries! We'll try to refresh the concepts as they arise in the quarter. On the other hand, if the majority of the topics here seem unfamiliar, you might want to reach out to us to make sure that you have the right prerequisites for the course.

Set Theory

The set of all natural numbers is denoted $\mathbb{N}$. We include 0, so $\mathbb{N} = \{ 0, 1, 2, 3, \dots \}$.

The set $\mathbb{Z}$ consists of all integers: $\mathbb{Z} = \{ \dots, -3, -2, -1, 0, 1, 2, 3, \dots \}$

The set $\mathbb{R}$ consists of all real numbers.

The set $\emptyset$ is the empty set consisting of no elements.

If x belongs to set S, we write $x \in S$. If x does not belong to S, we write $x \notin S$.

The ***union*** of two sets S_1 and S_2 is denoted $S_1 \cup S_2$. Their ***intersection*** is denoted $S_1 \cap S_2$, ***difference*** is denoted $S_1 - S_2$ or $S_1 \setminus S_2$, and ***symmetric difference*** is denoted $S_1 \Delta S_2$.

If S_1 is a ***subset*** of S_2, we write $S_1 \subseteq S_2$. If S_1 is a ***strict subset*** of S_2, we denote this by $S_1 \subsetneq S_2$.

The ***power set*** of a set S (denoted $\wp(S)$) is the set of all subsets of S.

The ***Cartesian product*** of two sets S_1 and S_2 is the set $S_1 \times S_2 = \{ (a, b) \mid a \in S_1 \text{ and } b \in S_2 \}$

First-Order Logic

The negations of the basic propositional connectives are as follows:

$$\neg(\neg p) \equiv p$$

$$\neg(p \land q) \equiv \neg p \lor \neg q$$

$$\neg(p \lor q) \equiv \neg p \land \neg q$$

$$\neg(p \to q) \equiv p \land \neg q$$

$$\neg(p \leftrightarrow q) \equiv p \leftrightarrow \neg q$$

The negations of the $\exists$ and $\forall$ quantifiers are as follows:

$$\neg \forall x.\ \varphi \equiv \exists x.\ \neg\varphi$$

$$\neg \exists x.\ \varphi \equiv \forall x.\ \neg\varphi$$

The statement "iff" abbreviates "if and only if."

Summations

The sum of the first n natural numbers ($0 + 1 + 2 + \ldots + n - 1$) is given by

$$\sum_{i=0}^{n-1} i = \frac{n(n-1)}{2}$$

The sum of the first n terms of the arithmetic series a, $a + b$, $a + 2b$, …, $a + (n - 1)b$ is

$$\sum_{i=0}^{n-1} (a+ib) = a\sum_{i=0}^{n-1} 1 + b\sum_{i=0}^{n-1} i = an + \frac{bn(n-1)}{2}$$

The sum of the first n terms of the geometric series $1, r, r^2, r^3, \ldots, r^{n-1}$ is given by

$$\sum_{i=0}^{n-1} r^i = \frac{r^n-1}{r-1}$$

As a useful special case, when $r = 2$, we have

$$\sum_{i=0}^{n-1} 2^i = 2^n-1$$

In the case that $|r| < 1$, the sum of all infinite terms of the geometric series is given by

$$\sum_{i=0}^{\infty} r^i = \frac{1}{1-r}$$

The following summation often arises in the analysis of algorithms: when $|r| < 1$, we have

$$\sum_{i=0}^{\infty} i r^i = \frac{r}{(1-r)^2}$$

Inequalities

The following identities are useful for manipulating inequalities:

If $A \leq B$ and $B \leq C$, then $A \leq C$

If $A \leq B$ and $C \geq 0$, then $CA \leq CB$

If $A \leq B$ and $C \leq 0$, then $CA \geq CB$

If $A \leq B$ and $C \leq D$, then $A + C \leq B + D$

If $A, B \in \mathbb{Z}$, then $A \leq B$ iff $A < B + 1$

If f is any monotonically increasing function and $A \leq B$, then $f(A) \leq f(B)$

If f is any monotonically decreasing function and $A \leq B$, then $f(A) \geq f(B)$

The following inequalities are often useful in algorithmic analysis:

$$e^x \geq 1 + x$$

$$\sqrt[n]{x_1 x_2 \ldots x_n} \leq \frac{x_1+x_2+\ldots+x_n}{n}$$

Floors and Ceilings

The ***floor function*** $\lfloor x \rfloor$ denotes the largest integer less than or equal to x. The ***ceiling function*** $\lceil x \rceil$ denotes the smallest integer greater than or equal to x. These functions obey the rules

$$\lfloor x \rfloor \leq x < \lfloor x \rfloor + 1 \quad \text{and} \quad \lfloor x \rfloor \in \mathbb{Z}$$

$$\lceil x \rceil - 1 < x \leq \lceil x \rceil \quad \text{and} \quad \lceil x \rceil \in \mathbb{Z}$$

Additionally, $\lfloor x + n \rfloor = \lfloor x \rfloor + n$ and $\lceil x + n \rceil = \lceil x \rceil + n$ for any $n \in \mathbb{Z}$.

Asymptotic Notation

Let $f, g : \mathbb{N} \to \mathbb{N}$. Then

$$f(n) = \mathrm{O}(g(n)) \quad \text{if} \quad \exists n_0 \in \mathbb{N}.\ \exists c \in \mathbb{R}.\ \forall n \in \mathbb{N}.\ (n \geq n_0 \to f(n) \leq cg(n))$$

$$f(n) = \Omega(g(n)) \quad \text{if} \quad \exists n_0 \in \mathbb{N}.\ \exists c > 0 \in \mathbb{R}.\ \forall n \in \mathbb{N}.\ (n \geq n_0 \to f(n) \geq cg(n))$$

$$f(n) = \Theta(g(n)) \quad \text{if} \quad f(n) = \mathrm{O}(g(n)) \wedge f(n) = \Omega(g(n))$$

When multiple variables are involved in an expression, big-O notation generalizes as follows: we say that $f(x_1, \ldots, x_n) = \mathrm{O}(g(x_1, \ldots, x_n))$ if there are constants N and c such that for any $x_1 \geq N$, $x_2 \geq N, \ldots, x_n \geq N$, we have $f(x_1, \ldots, x_n) \leq c \cdot \mathrm{g}(x_1, \ldots, x_n)$.

The following rules apply for O notation:

$$\text{If } f(n) = \mathrm{O}(g(n)) \text{ and } g(n) = \mathrm{O}(h(n)), \text{ then } f(n) = \mathrm{O}(h(n)) \quad (\text{also } \Omega, \Theta, o, \omega)$$

$$\text{If } f_1(n) = \mathrm{O}(g(n)) \text{ and } f_2(n) = \mathrm{O}(g(n)), \text{ then } f_1(n) + f_2(n) = \mathrm{O}(g(n)) \quad (\text{also } \Omega, \Theta, o, \omega)$$

$$\text{If } f_1(n) = \mathrm{O}(g_1(n)) \text{ and } f_2(n) = \mathrm{O}(g_2(n)), \text{ then } f_1(n)f_2(n) = \mathrm{O}(g_1(n)g_2(n)) \quad (\text{also } \Omega, \Theta, o, \omega)$$

We can use o and ω notations to denote strict bounds on growth rates:

$$f(n) = o(g(n)) \quad \text{if} \quad \lim_{n \to \infty} \frac{f(n)}{g(n)} = 0 \qquad f(n) = \omega(g(n)) \quad \text{if} \quad \lim_{n \to \infty} \frac{f(n)}{g(n)} = \infty$$

Polynomials, exponents, and logarithms are related as follows:

$$\log_a n = \Theta(\log_b n) \text{ for any fixed constants } a, b > 1$$

$$\text{Any polynomial of degree } k \text{ with positive leading coefficient is } \Theta(n^k)$$

$$\log_b n = o(n^k) \text{ for any } k > 0$$

$$n^k = o(b^n) \text{ for any } b > 1$$

$$b^n = o(c^n) \text{ for any } 1 < b < c$$

In a graph, n denotes the number of nodes ($|V|$) and m denotes the number of edges ($|E|$). In any graph, $m = \mathrm{O}(n^2)$. In a dense graph, $m = \Theta(n^2)$; a sparse graph is one where $m = o(n^2)$.

The Master Theorem

If a, b, and d are constants, then the recurrence relation

$$\mathrm{T}(n) = a\mathrm{T}(n / b) + \mathrm{O}(n^d)$$

solves as follows:

$$\mathrm{T}(n)=\begin{cases} \mathrm{O}(n^d) & \text{if } \log_b a < d \\ \mathrm{O}(n^d \log n) & \text{if } \log_b a = d \\ \mathrm{O}(n^{\log_b a}) & \text{if } \log_b a > d \end{cases}$$

Logarithms and Exponents

Logarithms and exponents are inverses of one another: $b^{\log_b x}=\log_b b^x=x$

The ***change-of-base formula*** for logarithms states that

$$\log_b a=\frac{\log_c a}{\log_c b}$$

Sums and differences of logarithms translate into logarithms of products and quotients:

$$\log_b xy=\log_b x+\log_b y \qquad \log_b(x/y)=\log_b x-\log_b y$$

The ***power rule*** for logarithms states

$$\log_b x^y=y\log_b x$$

In some cases, exponents may be interchanged:

$$(a^b)^c=a^{bc}=(a^c)^b$$

We can change the base of an exponent using the fact that logarithms and exponents are inverses:

$$a^c=b^{c\log_b a}$$

We will sometimes make use of the fact that

$$\lim_{n\to\infty}\left(1-\frac{1}{n}\right)^n = \frac{1}{e}.$$

In particular, note that for any $n \geq 1$ that

$$0 \leq \left(1-\frac{1}{n}\right)^n \leq \frac{1}{e}.$$

Probability

If E_1 and E_2 are mutually exclusive events, then

$$\Pr[E_1] + \Pr[E_2] = \Pr[E_1 \cup E_2]$$

For any events E_1, E_2, E_3, …, including overlapping events, the ***union bound*** states that

$$\Pr\left[\bigcup_{i=1}^{\infty} E_i\right] \le \sum_{i=1}^{\infty} \Pr[E_i]$$

The probability of E given F is denoted $\Pr[E \mid F]$ and is given by

$$\Pr[E|F] = \frac{\Pr[E \cap F]}{\Pr[F]}$$

The ***chain rule*** for conditional probability is

$$\Pr[E_n \cap E_{n-1} \cap \ldots \cap E_1] = \Pr[E_n \mid E_{n-1} \cap \ldots \cap E_1] \cdot \Pr[E_{n-1} \mid E_{n-2} \cap \ldots \cap E_1] \cdot \ldots \cdot \Pr[E_1]$$

Two events E_1 and E_2 are called ***independent*** if

$$\Pr[E_1 \cap E_2] = \Pr[E_1] \cdot \Pr[E_2]$$

For any event E, the ***complement*** of that event (denoted $\overline{E}$) represents the event that E does not occur. E and $\overline{E}$ are mutually exclusive, and

$$\Pr[E] + \Pr[\overline{E}] = 1$$

Expected Value

The ***expected value*** of a discrete random variable X is defined as

$$\mathrm{E}[X] = \sum_{i=0}^{\infty} (i \cdot \Pr[X = i])$$

The expected value operator is linear: for any $a, b \in \mathbb{R}$ and any random variable X:

$$\mathrm{E}[aX + b] = a\mathrm{E}[X] + b$$

If X_1, X_2, X_3, … X_n is a finite collection of random variables, ***linearity of expectation*** tells us that

$$\mathrm{E}[\sum_{i=1}^{n} X_i] = \sum_{i=1}^{n} \mathrm{E}[X_i].$$

Variance and Covariance

The ***variance*** of a random variable X is defined as

$$\text{Var}[X] = \text{E}[(X - \text{E}[X])^2] = \text{E}[X^2] - \text{E}[X]^2$$

Accordingly, for any random variable X, notice that

$$\text{Var}[X] \leq \text{E}[X^2].$$

Given two random variables X and Y, the ***covariance*** of X and Y is defined as

$$\text{Cov}[X, Y] = \text{E}[(X - \text{E}[X])(Y - \text{E}[Y])] = \text{E}[XY] - \text{E}[X]\text{E}[Y]$$

Accordingly:

$$\text{Var}[X] = \text{Cov}[X, X]$$

Variance is not a linear operator:

$$\text{Var}[aX + bY] = a^2\text{Var}[X] + 2ab\,\text{Cov}[X, Y] + b^2\text{Var}[Y]$$

The variance of a summation of random variables, including dependent variables, can be simplified using the following rule:

$$\text{Var}[\sum_{i=1}^{n} X_i] = \sum_{i=1}^{n} \text{Var}[X_i] + \sum_{i \neq j} \text{Cov}[X_i, X_j]$$

Two random variables X and Y are called ***uncorrelated*** if

$$\text{E}[XY] = \text{E}[X]\text{E}[Y].$$

Equivalently, the random variables X and Y are uncorrelated if $\text{Cov}[X, Y] = 0$.

Any two independent random variables are uncorrelated, but uncorrelated random variables may not be independent of one another.

If X_1, X_2, …, and X_n are uncorrelated random variables, then

$$\text{Var}[\sum_{i=1}^{n} X_i] = \sum_{i=1}^{n} \text{Var}[X_i].$$

Concentration Inequalities

Markov's inequality says that for any nonnegative random variable X with finite expected value and any $c > 0$, we have both

$$\Pr\left[X \geq c\,\text{E}[X]\right] \leq \frac{1}{c} \quad \text{and} \quad \Pr\left[X \geq c\right] \leq \frac{\text{E}[X]}{c}.$$

Chebyshev's inequality states that for any random variable X with finite expected value that

$$\Pr\left[\ |X - \text{E}[X]| \geq c\sqrt{\text{Var}[X]}\ \right] \leq \frac{1}{c^2} \quad \text{and} \quad \Pr\left[\ |X - \text{E}[X]| \geq c\ \right] \leq \frac{\text{Var}[X]}{c^2}.$$

The ***Chernoff bound*** says that if $X \sim \text{Binom}(n, p)$ for $p < ½$, that

$$\Pr[X \geq \frac{n}{2}] \leq e^{\frac{-n(1/2 - p)^2}{2p}}$$

In the case where p is a fixed constant, notice that the right-hand side is $e^{-\text{O}(1) \cdot n}$.

Useful Probability Equalities and Inequalities

An ***indicator random variable*** is a random variable X where

$$X = \begin{cases} 1 & \text{if event } F \text{ occurs} \\ 0 & \text{otherwise} \end{cases}$$

For any indicator variable, $E[X] = \Pr[F]$. Notice that if X is an indicator, then $X = X^2$, so

$$\mathrm{Var}[X] = E[X^2] - E[X]^2 = E[X] - E[X]^2 = E[X](1 - E[X]) \leq E[X] = \Pr[F].$$

If $X_1, X_2, \ldots, X_n$ are random variables, then

$$\Pr[\max\{X_{1,} X_{2,} \ldots, X_n\} \leq k] = \Pr[X_1 \leq k \cap X_2 \leq k \cap \ldots X_n \leq k]$$

$$\Pr[\min\{X_{1,} X_{2,} \ldots, X_n\} \geq k] = \Pr[X_1 \geq k \cap X_2 \geq k \cap \ldots X_n \geq k]$$

On expectation, repeatedly flipping a biased coin that comes up heads with probability p requires $^1/_p$ trials before the coin will come up heads.

Harmonic Numbers

The nth ***harmonic number***, denoted H_n, is given by

$$H_n = \sum_{i=1}^{n} \frac{1}{i}$$

The harmonic numbers are close in value to ln n: for any $n \geq 1$, we have

$$\ln(n + 1) \leq H_n \leq \ln n + 1,$$

so $H_n = \Theta(\log n)$.

Binary XOR

The ***xor*** ("exclusive or") operation, denoted $\boldsymbol{x \oplus y}$, takes in two bits x and y and evaluates to 1 if x and y are different and 0 if x and y are the same.

The xor operator has 0 as an identity:

$$x \oplus 0 = 0 \oplus x = x.$$

The xor operator is self-inverting:

$$x \oplus x = 0.$$

Xor is also associative and commutative:

$$x \oplus y = y \oplus x$$

$$(x \oplus y) \oplus z = x \oplus (y \oplus z)$$

While xor is nominally defined on pairs of bits, it can be extended to work on pairs of integers as well. That operation works by xoring the corresponding pairs of bits of the numbers in question, and has the same algebraic properties described above (where, in this case, 0 would be interpreted as "a number whose bits are all zeroes").