Math Review (and Background) for "Stochastic Modeling for Engineers"

Yoni Nazarathy

October 26, 2011

Below is a brief summary of some of the general definitions and mathematical properties needed for the subject. While most of this summary may appear dry, it is designed to help understand the subject material (which is very very exciting).

1 Sets

- A set is a collection of objects, e.g. $A = \{1, -3, 8, a\}$.
- Sets are not regarded as ordered and can have a finite or infinite number of objects.
- $x \in A$ is read as "x is an element of A". Similarly $x \notin A$. E.g. for the set above we have $1 \in A$ and $4 \notin A$.
- Some sets you know well are $\mathbb{R}$ the set of real numbers, $\mathbb{Z}$ the set of integers and $\mathbb{N}$ the set of natural numbers (sometimes taken to be $\{1, 2, 3, \ldots\}$ and sometimes $\{0, 1, 2, 3, \ldots, \}$).
- We say A is a subset of B (denoted by $A \subset B$) if whenever $x \in A$ we also have $x \in B$.
- We say two sets A and B are equal (denoted $A = B$) if $A \subset B$ and $B \subset A$.
- The empty set, denoted $\emptyset$ has no elements ($\emptyset = \{\}$). It is a subset of any other set.
- We often have a universal set, (in probability study it is often denoted Ω). Having such a set allows us to defined the complement of any subset of Ω: A^c (sometimes denoted $\bar{A}$) this is the set of all elements that are not in A but in Ω. This can also be written as,

$$A^c = \{x \in \Omega : x \notin A\}.$$

- Note that $(A^c)^c = A$. Also, $\Omega^c = \emptyset$.
- The union of two sets A and B, denoted $A \cup B$ is the set that contains all elements that are in either A, B or both. E.g. $\{-2, 0, 3\} \cup \{0, 1\} = \{0, -2, 3, 1\}$. Note that $A \cup A^c = \Omega$.
- The intersection of two sets A and B, denoted $A \cap B$ is the set of all elements that are in both A and B. E.g. $\{-2, 0, 3\} \cup \{0, 1\} = \{0\}$. Note that $A \cap A^c = \emptyset$. Note also that $A \cap B \subset A \cup B$.
- Two sets A and B are disjoint (or mutually exclusive) if $A \cap B = \emptyset$.
- The difference of A and B, denoted $A \backslash B$ (sometimes $A - B$) is the set of elements that are in A and not in B. Note that $A \backslash B = A \cap B^c$.

- Commutative properties: $A \cup B = B \cup A$ and $A \cap B = B \cap A$.
- Associative properties: $A \cup (B \cup C) = (A \cup B) \cup C$ and $A \cap (B \cap C) = (A \cap B) \cap C$.
- Distributive properties: $A \cup (B \cap C) = (A \cup B) \cap (A \cup C)$ and $A \cap (B \cup C) = (A \cap B) \cup (A \cap C)$.
- DeMorgan's rules: $(A \cup B)^c = A^c \cap B^c$, $(A \cap B)^c = A^c \cup B^c$.
- We can use the following notation for unions: $\cup_{\gamma \in \Gamma} A_\gamma$, or similarly for intersections $\cap_{\gamma \in \Gamma} A_\gamma$. This means taking the union (or intersection) of A_γ for all γ in Γ. E.g. if $\Gamma = \{1, 2\}$ it implies $A_1 \cup A_2$ (or similarly for intersection).
- In the above, we can have the "index set" Γ as being either finite or infinite. If it is infinite, it can be both countable or uncountable (defined in next bullet).
- A infinite set is countable if there exists a one-to-one mapping between elements of the set to the natural numbers, otherwise it is called uncountable. E.g. $\mathbb{Z}$ is countable (what is the mapping?).
- A more general for of DeMorgan's rules is: $\big(\cup_{\gamma \in \Gamma} A_\gamma\big)^c = \cap_{\gamma \in \Gamma} A_\gamma^c$ and $\big(\cap_{\gamma \in \Gamma} A_\gamma\big)^c = \cup_{\gamma \in \Gamma} A_\gamma^c$.
- The power set of a set A, denoted 2^A is the set of all subsets of A, e.g.,
$$2^{\{a,b\}} = \{\emptyset,\ \{a\},\ \{b\},\ \{a,b\}\}.$$

2 Counting

- For a finite set A, $|A|$ denotes the number of elements in A. E.g. $|\{a, b, c\}| = 3$.
- A k-tuple is simply an ordered list with values $(x_1, \dots, x_k)$.
- The multiplication principle: The number of distinct ordered k-tuples $(x_1, \dots, x_k)$ with components $x_i \in A_i$ is $|A_1| \cdot |A_2| \cdot \ldots \cdot |A_k|$.
- For a finite set A, $|2^A| = 2^{|A|}$. I.e. the number of subsets of a set (the number of elements in the power set) is $2^{|A|}$. To see this, match a binary number with $|A|$ digits to each subset, '1' indicates the element is in the set and '0' indicates that it isn't in the set. The number of values the number can take is $2^{|A|}$ as follows from the multiplication principle.
- The number of ways to choose k objects from a finite set A with $|A| = n$, not requiring the objects to be distinct is: n^k. This is sometimes called sampling with replacement and with ordering. Note that this also corresponds to the number of ways of distributing k distinct balls in n bins where there is no limit on the number of balls that can fit in a bin.
- The number of ways to choose k distinct objects from a finite set A of size n where order matters is
$$n \cdot (n-1) \cdot \ldots \cdot (n-k+1).$$
I.e. this is the number of k-tuples with distinct elements selected from A. This is number also corresponds the number of ways of distributing k distinct balls in n bins where there is a limit of at most one ball per bin.

  Note that if $k = n$ this number is $n!$ (e.g. $5! = 5 \cdot 4 \cdot 3 \cdot 2 \cdot 1 = 120$).

- Each ordering of a finite set of size n is called a permutation. Thus the number of permutations is $n!$.
- It is good to know Stirling's formula:

$$n! \sim \sqrt{2\pi} n^{n+\frac{1}{2}} e^{-n}.$$

The "similar sign" $\sim$ indicates that the ratio of the left hand side and right hand side converges to 1 as $n \to \infty$. For example, try this in Mathematica:

$$\text{ListPlot}\left[\text{Table}\left[\left\{n, \frac{n!}{\sqrt{2\pi} n^{n+\frac{1}{2}} E^{-n}} //N\right\}, \{n, 1, 100\}\right], \text{AxesOrigin} \to \{0, 1\}\right].$$

(Note: We often use $\sim$ to indicate the distribution of a random variable - something completely different!).

- The number of ways of choosing k distinct objects from a finite set A where order doesn't matter is similar to the case where order matters but should be corrected by a factor of $k!$. This number is sometimes called the binomial coefficient:

$$\binom{n}{k} = \frac{n \cdot (n-1) \cdot \ldots \cdot (n-k+1)}{k!} = \frac{n!}{k!(n-k)!}.$$

I.e. this is the number of subsets of size k of a set of size n. It also corresponds to the number of ways of distributing k indistinguishable balls in a n bins with room for at most one ball per bin.

- Some properties of binomial coefficients (all quite easy to see once you know how to look):
  -
  $$\binom{n}{k} = \binom{n}{n-k}.$$
  -
  $$\binom{n}{0} = \binom{n}{n} = 1.$$
  -
  $$\binom{n}{1} = \binom{n}{n-1} = n.$$
  -
  $$\binom{n}{k-1} + \binom{n}{k} = \binom{n+1}{k}.$$
  -
  $$\sum_{k=0}^{n} \binom{n}{k} = 2^n.$$
  - For real numbers a and b and positive integer n,

  $$(a+b)^n = \sum_{k=0}^{n} \binom{n}{k} a^k b^{n-k}.$$

  This is sometimes called the binomial theorem.

3 Summations and Series

- The notation $\sum_{i=1}^{n} a_i$ implies $a_1 + a_2 + \ldots + a_n$. There is similar notation for products: $\prod_{i=1}^{n} a_i = a_1 \cdot a_2 \cdot \ldots \cdot a_n$, yet the rest of this section concentrates on summations.
- $\sum_{i=1}^{n} c = nc$ (the constant c does not depend on i).
- $\sum_{j=k}^{m} a_i = \sum_{j=1}^{m-k+1} a_{j+k-1}$.
- For certain sequences, $a_1, \ldots, a_n$, the summation can be simplified, e.g.:
  - Arithmetic progression: $\sum_{i=1}^{n} i = \frac{n(n+1)}{2}$.
  - Geometric progression: $\sum_{i=0}^{n} q^i = \frac{1-q^{n+1}}{1-q}$ for $q \neq 1$.
  - $\sum_{i=1}^{n} i^2 = \frac{(2n+1)(n+1)n}{6}$.
- An infinite series (or simply series) is a sum of an infinite (countable) number of elements, $\sum_{i=1}^{\infty} a_i$. This series is defined as the limit of $\sum_{i=1}^{n} a_i$ as $n \to \infty$. Note that this limit may converge to a finite value or not.
- A necessary (but not sufficient) condition for convergence of a series is $\lim_{n\to\infty} a_n = 0$. The classic example of the insufficiy of this condition is the case where $a_n = 1/n$ (in this case the series is the harmonic series. Try this for example in Mathematica:

$$\text{ListPlot}\left[\text{Table}\left[\left\{n, \text{Sum}\left[\frac{1}{i}//N, \{i,1,n\}\right] - \text{Log}[n]\right\}, \{n,1,1000\}\right], \text{AxesOrigin} \to \{0, 0.57\}\right].$$

- The Geometric series converges if $|q| < 1$: $\sum_{i=0}^{\infty} q^i = \frac{1}{1-q}$.
- The Taylor series expansion of e^x is very useful:

$$e^x = \sum_{i=0}^{\infty} \frac{x^i}{i!}.$$

  Related Taylor series expansions of $\cos x$, $\sin x$, $\cosh x$ and $\sinh x$ are given in the sections below.
- A useful relation for any sequence, $\{a_k\}$ is $\sum_{i=1}^{n} \sum_{j=i}^{n} a_j = \sum_{i=1}^{n} i a_i$. Note that n can be ∞ if the series converge.

4 Integrals

- $\int f(x)dx$ is the indefinite integral of $f(\cdot)$ (note that we do not specify a domain). It equals $F(x) + C$ for any constant C where $F(\cdot)$ is a function such that $\frac{d}{dx}F(x) = f(x)$. For example $\int x^{\alpha} dx = \frac{1}{\alpha+1} x^{\alpha+1} + C$ for any $\alpha \neq -1$. (If $\alpha = -1$ then it is $Log(x) + C$). Note that when we write *Log* we mean with respect to the natural base e. Here F is called the antiderivative of f.
- A definite integral has a domain of the form $[a, b]$ and is written as $\int_a^b f(x)dx$. One way to think of this integral is as the limit of the sum $\sum_{i=1}^{n} f(a + i\frac{b-a}{n}) \cdot \frac{b-a}{n}$ as $n \to \infty$. Thus when $f(x) \geq 0$ in $[a, b]$ the integral is the area between the function and the horizontal axis. If $f(x)$ is both positive and negative the integral is the difference of the area between the positive part and the axis and the area between the negative part of the axis. For example:

- $\int_{-1}^{1} \sqrt{1-x^2}dx = \pi/2$, why?
- $\int_{-1}^{1} x dx = 0$.
- The above is the special case of the following: For an odd function, $f(x)$, $\int_{-a}^{a} f(x)dx = 0$. An odd function is a function such that $f(-x) = -f(x)$. ($\sin(x)$ is an odd function).
- An even function is a function such that $f(-x) = f(x)$. ($\cos(x)$ is an even function). For an even function, $f(x)$, $\int_{-a}^{a} f(x)dx = 2\int_{0}^{a} f(x)dx$.
- The so called fundamental theorem of integral calculus is that if F is the antiderivative of a continuous function f then,

$$\int_a^b f(t)dt = G(a) - G(b).$$

- The standard methods of finding the antiderivative of a function (solving an integral) are:
  - Using the linearity of the integral, for example (for $n, m \neq -1$):

$$\int ax^n + bx^m dx = a\int x^n dx + b\int x^m dx = \frac{a}{n+1}x^{n+1} + \frac{b}{m+1}x^{m+1} + C.$$

  - Using the method of substitution: $\int f\big(g(x)\big)g'(x)dx$ can be written as $\int f(u)du$ by setting $u = g(x)$ and $du = g'(x)dx$. Then, $\int f\big(g(x)\big)g'(x)dx = F\big(g(x)\big) + C$, where F is the anti-derivative of f (this is an application of the chain rule for differentiation: $\frac{d}{dx} f\big(g(x)\big) = f'\big(g(x)\big)g'(x)$).
  - Using the product rule (integration by parts).
  - Using partial fraction expansion (for integration of rational functions).
- Some integrals can not be evaluated in closed form, for example, there is no simple expression for $\int e^{-x^2/2}dx$. For such cases if we want to calculate definite integrals over some range $[a, b]$ it is done numerically.

5 Cosine, Sine etc..

- Points in the plane can be represented in polar form, (θ, r) with the angle $\theta \in [0, 2\pi]$ and the radius $r \in [0, \infty)$.
- A point represented in polar form, (θ, r), has cartesian coordinates $(r\cos(\theta), r\sin(\theta))$.
- The above may be viewed as a definition of the cosine and sine functions over $\theta \in [0, 2\pi]$. For this, extend θ to the whole real line (not just in $[0, 2\pi]$) and require that,

$$\cos(\theta + 2\pi k) = \cos(\theta), \quad , \sin(\theta + 2\pi k) = sin(\theta), \quad \text{for} \quad k = \dots, -2, -1, 0, 1, 2, \dots.$$

- The above requires that cosine and sine be periodic functions with a period of 2π. Recal that the real function $f(x)$ is periodic with period $d > 0$ if,

$$f(x) = f(x + d), \ \forall x.$$

  Observe that if a function is periodic with period d then it is periodic with period $k\,d$, for any integer k. Obviously the period of cosine and sine is 2π.

- The Pythagorean theorem implies that,

$$\cos^2\theta + \sin^2\theta = 1.$$

Note: We sometimes omit the "()" in these functions and also mean $cos^2\theta$ to be $(\cos\theta)^2$.

- $\frac{d}{dx}\sin x = \cos x$, $\frac{d}{dx}\cos x = -\sin x$
- Taylor's series expansion of cosine and sine, yields,

$$\begin{aligned} \cos x &= 1 - \frac{x^2}{2!} + \frac{x^4}{4!} - \frac{x^6}{6!} + \ldots = \sum_{k=0}^{\infty} \frac{(-1)^k x^{2k}}{(2k)!} \\ \sin x &= x - \frac{x^3}{3!} + \frac{x^5}{5!} - \frac{x^7}{7!} + \ldots = \sum_{k=0}^{\infty} \frac{(-1)^k x^{2k+1}}{(2k+1)!} \end{aligned}$$

- Many useful identities can be obtained from Euler's formula: $e^{ix} = \cos x + i\sin x$ (see also next section). For example:

$$\big(\cos\alpha + i\sin\alpha\big)\big(\cos\beta + i\sin\beta\big) = e^{i\alpha}e^{i\beta} = e^{i(\alpha+\beta)} = \cos(\alpha+\beta) + i\sin(\alpha+\beta).$$

Taking the real part of the above we obtain:

$$\cos\alpha\cos\beta - \sin\alpha\sin\beta = \cos(\alpha+\beta).$$

Taking the imaginary part of the above we obtain:

$$\sin\alpha\cos\beta + \cos\alpha\sin\beta = \sin(\alpha+\beta).$$

Many other identities arise from this, e.g. by taking $\beta = \alpha$ we have (from the $\cos(\alpha+\beta)$ law above),

$$\cos^2\alpha - \sin^2\alpha = \cos(2\alpha),$$

but from Pythagorean theorem we can 'get rid of the $\sin^2\alpha$ and get,

$$\cos^2\alpha - (1 - \cos^2\alpha) = \cos(2\alpha),$$

which rearranges to,

$$\cos^2\alpha = \frac{1 + \cos(2\alpha)}{2}.$$

There are many other such useful identities.

- The hyperbolic functions, cosh and sinh are defined as follows:
  - $\cosh x = \frac{1}{2}(e^x + e^{-x})$.
  - $\sinh x = \frac{1}{2}(e^x - e^{-x})$.

Their name perhaps comes from the fact that,

$$\cosh^2 t - \sinh^2 t = 1,$$

and hence the coordinates of the curve $P(t) = (\cosh^2 t, \sinh^2 t)$ lie on the hyperbola $x^2 - y^2 = 1$. The series expansions of these functions are:

  - $\cosh x = \sum_{k=0}^{\infty} \frac{x^{2k}}{(2k)!}$.
  - $\sinh x = \sum_{k=0}^{\infty} \frac{x^{2k+1}}{(2k+1)!}$.

6 Complex Numbers

Think of i as $\sqrt{-1}$: There is no real number x such that $x^2 = -1$, hence there is no real solution to the equation, $x^2 + 1 = 0$, thus the imaginary number i introduced. Electrical engineers sometimes use the notation j for i, reserving the latter for current.

Think of a complex number z as,

$$z = a + b\, i,$$

where the real part, $\Re(z) = a$ and the imaginary part, $\Im(z) = b$.

Denote the complex numbers, $z = a + b\, i$ and $w = c + d\, i$. They can be,

- Added (or subtracted) as follows,

$$z \pm w = (a \pm c) + i(b \pm d).$$

- Multiplied as follows,

$$zw = (a + b\, i)(c + d\, i) = ac + i^2 bd + i(ad + bc) = (ac - bd) + i(ad + bc).$$

The conjugate of a complex number $z = a + b\, i$ is,

$$\bar{z} = a - b\, i.$$

The absolute value of z is,

$$|z| = \sqrt{z\bar{z}} = \sqrt{a^2 + b^2}.$$

Thus multiplying a complex number by it's conjugate yields a real number $(a^2 + b^2)$, this is useful for dividing. Assume $w \neq 0$, then,

$$\frac{z}{w} = \frac{z\bar{w}}{c^2 + d^2} = \frac{ac + bd}{c^2 + d^2} + \frac{bc - ad}{c^2 + d^2} i.$$

We have, $\overline{z \pm w} = \bar{z} \pm \bar{w}$, $\overline{zw} = \bar{z}\bar{w}$, $\overline{z/w} = \bar{z}/\bar{w}$. A special example is,

$$\frac{1}{z} = \frac{\bar{z}}{|z|^2}.$$

Polar form representation of complex numbers is very useful for multiplication/division. For $z = a + bi$, the magnitude (called modulus) of z is $r = |z|$ and the argument of z, φ is the angle between the x-axis and z expressed in radians. For positive, a this is simply $\arctan(b/a)$, observe that this is a value in the range $(-\frac{\pi}{2}, \frac{\pi}{2})$. For other values, more care is needed.

We can now express,

$$z = a + bi = re^{\varphi i}.$$

The nice thing is that rules of exponentials work, so for example,

$$zw = r_z e^{\varphi_z i} r_w e^{\varphi_w i} = r_z r_w e^{(\varphi_z + \varphi_w) i)}.$$

7 Fourier Transforms

The functions $R(\tau)$ and $S(f)$ are fourier transform pairs if,

$$S(f) = \mathcal{F}\{R(\tau)\} = \int_{-\infty}^{\infty} R(\tau)e^{-i2\pi f\tau}d\tau,$$

and,

$$R(\tau) = \mathcal{F}^{-1}\{S(f)\} = \int_{-\infty}^{\infty} S(f)e^{i2\pi f\tau}df.$$

Here $\mathcal{F}$ symbolizes the Fourier Transform taking a function from the so-called "time-domain" and generating a function in the "frequency-domain". Similarly $\mathcal{F}^{-1}$ is the inverse Fourier Transform.

The existence of a fourier transform can be stated as follows: If the function $R(\tau)$ is absolutly integrable, i.e.,

$$\int_{-\infty}^{\infty} |R(\tau)|d\tau < \infty,$$

and it is it is "not too crazy": It has at most a finite number of maxima and minima and a finite number of discontinuities in any finite interval, then,

$$R(\tau) = \int_{-\infty}^{\infty} \left[e^{i2\pi f\tau} \int_{-\infty}^{\infty} R(t)e^{-i2\pi ft}dt\right]df.$$

Example: Consider $R(\tau) = e^{-a|\tau|}$, with $a > 0$.

$$\begin{aligned} S(f) &= \int_{-\infty}^{\infty} e^{-a|\tau|}e^{-i2\pi f\tau}d\tau = \int_{-\infty}^{0} e^{(a-i2\pi f)\tau}d\tau + \int_{0}^{\infty} e^{(-a-i2\pi f)\tau}d\tau \\ &= \frac{1}{a-i2\pi f} + \frac{1}{a+i2\pi f} = \frac{2a}{a^2+4\pi^2f^2}. \end{aligned}$$

As can be seen from the example above, the fourier transform turned out to be real valued for every frequency f. Observe that it is also an even function. This is always the case when $R(\tau)$ is an even function as the fourier transform simplifies to,

$$S(f) = \int_{-\infty}^{\infty} R(\tau)\cos(2\pi f\tau)d\tau.$$

Tables of fourier transforms and many basic properties can be found every where.

8 Convolutions

9 Linear Time Invariant Systems

10 Frequency Response

11 Linear Algebra

- For an $N \times M$ matrix A and an $M \times L$ matrix B, the matrix product, AB is the $N \times L$ matrix with element (i, j) (in row i and collum j) being $\sum_{k=1}^{M} a_{i,k}b_{k,j}$. Each element of the resulting matrix is thus the inner product of the i'th collum of A and the j'th collum of B.

- Often we apply the matrix A to an M dimensional (always taken by default as collum) vector x, resulting in an N dimensional vector, y. This is the linear transformation, $y = Ax$.

- Quite often we take $N = M$ (the matrix A is square). And thus the linear transformation $y = Ax$ maps vectors from $\mathbb{R}^N$ to $\mathbb{R}^N$. In this case it is often important to identify the cases where the transformation is a bijection (one to one and onto) and in these cases there is an inverse transformation $x = A^{-1}y$, where A^{-1} is called the inverse matrix of A. When A^{-1} exists we say that A is non-singular, otherwise it is singular. One test for this is $\det(A) \neq 0$ (i.e. if the determinant is 0 then the matrix is singular).

- Finding A^{-1} is a technical matter, i.e. finding the matrix A^{-1} such that $AA^{-1} = I$, or alternatively $A^{-1}A = 1$, (I is the identity matrix). This can be done using row operations or using the so-called Cramer's rule in terms of determinant. By the way, a determinant is a mapping from an $N \times N$ matrix onto the real numbers - further discussed below. A useful case is for a $N = 2$, in this case,

$$A^{-1} = \frac{1}{ad - bc} \begin{bmatrix} d & -b \\ -c & a \end{bmatrix}, \text{when} \quad A = \begin{bmatrix} a & b \\ c & d \end{bmatrix},$$

  here the denominator $ad - bc$ is $\det(A)$.