MATH1853

Notes Fall 2023

Author: Jax

Contact: enhanjax@connect.hku.hk

1 Matrices

Matrices are represented by capitial letters: $A \in R^{r \times c}$. $r \times c$ are the dimensions of a given matrix

1.1 Basic operations

Addition

Matrices can perform per-element addition if their dimensions match.

$\left[\begin{smallmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{smallmatrix}\right] + \left[\begin{smallmatrix} 6 & 5 & 4 \\ 3 & 2 & 1 \end{smallmatrix}\right] = \left[\begin{smallmatrix} 7 & 7 & 7 \\ 7 & 7 & 7 \end{smallmatrix}\right]$

Scalar multiplication

Matrices can perform per-element multiplication with a scalar.

$\left[\begin{smallmatrix} 1 & 2 & 3 \\ 3 & 2 & 1 \end{smallmatrix}\right] \times 2 = \left[\begin{smallmatrix} 2 & 4 & 6 \\ 6 & 4 & 2 \end{smallmatrix}\right]$

Matrix-matrix multiplication

Matrices can only multiply if their <u>inner dimensions</u> match, otherwise there will be no solution.
If a matrix is multiplied by another matrix, each <u>element in the result matrix</u> is the sum of it's corresponding <u>row in the original matrix</u> performing per-element multiplication on it's corresponding <u>col in the applying matrix</u>

$$\begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix} \times \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix} = \begin{bmatrix} 1 \cdot 1 + 2 \cdot 3 & 1 \cdot 2 + 2 \cdot 4 \\ 3 \cdot 1 + 4 \cdot 3 & 3 \cdot 2 + 4 \cdot 4 \end{bmatrix} = \begin{bmatrix} 7 & 10 \\ 15 & 22 \end{bmatrix}$$

Transposition

The function A^T transposes the matrix A. It swaps the rows and columns of a matrix.

$\left[\begin{smallmatrix} 1 & 2 \\ 3 & 4 \end{smallmatrix}\right]^T \equiv \left[\begin{smallmatrix} 1 & 3 \\ 2 & 4 \end{smallmatrix}\right]$ $\quad \left[\begin{smallmatrix} 1 & 2 & 3 \end{smallmatrix}\right]^T \equiv \left[\begin{smallmatrix} 1 \\ 2 \\ 3 \end{smallmatrix}\right]$

1.2 Solving linear systems

Identity matrix

The identity matrix of $I \in \mathbb{R}^{n \times n}$ with square dimensions is the matrix with diagonal elements of 1 and others 0.

$I \in \mathbb{R}^{3 \times 3} = \left[\begin{smallmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{smallmatrix}\right]$

Row-echelon form

A matrix with a bottom-left or top-right triangle of 0's (excluding the diagonal) is considered to be in row-echelon form.

$\begin{bmatrix} 1 & 2 & 3 \\ 0 & 1 & 2 \\ 0 & 0 & 1 \end{bmatrix}$ and $\begin{bmatrix} 1 & 0 & 0 \\ 1 & 2 & 0 \\ 0 & 0 & 1 \end{bmatrix}$ are both considered in row-echelon form

Augmented Matrix and Gaussian Elimination

Augmented matrices are an important tool to solve linear equations. The number of rows is equal to the number of variables in the equation.
GE are operations by <u>rows</u>, performed on augmented matrices. The possible moves are:

1. Addition between rows ($R_1 + R_2$ adds row 2 to row 1, leaving row 2 unchanged)
2. Multiplication of a row by a scalar ($\frac{1}{2}R_1$)
3. Swapping rows ($R_1 \leftrightarrow R_2$)

1.2.1 Strict variables

Consider the following linear system. We can take the coefficients and write it into AM form as followed:

$$\begin{aligned} x_1 + 2x_2 &= 3 \\ 2x_1 + 5x_2 &= 10 \end{aligned} \implies \left[\begin{array}{cc|c} 1 & 2 & 3 \\ \underbrace{2}_{x_1} & \underbrace{5}_{x_2} & 10 \end{array}\right]$$

Then, we perform GE on the created augmented matrix to make the left side row-echelon.

$$\left[\begin{array}{cc|c} 1 & 2 & 3 \\ 2 & 5 & 10 \end{array}\right] \xrightarrow{R_2 - 2R_1} \left[\begin{array}{cc|c} 1 & 2 & 3 \\ 0 & 1 & 4 \end{array}\right]$$

Deriving from the AM, we know:

$$\begin{gathered} x_1 + 2x_2 = 3 \\ x_2 = 4 \end{gathered}$$

So the solution is:

$$\begin{bmatrix} 3 - 2 \cdot 4 \\ 4 \end{bmatrix} \equiv \begin{bmatrix} -5 \\ 4 \end{bmatrix}$$

1.2.2 Inconsistent systems

We say a system is inconsistent if a full row of 0 on the left is equal to a non-zero number. We can also conclude that there are no solutions to the system. The following is an example of a system with no solutions:

$$\begin{aligned} x_1 + 2x_2 &= 3 \\ 2x_1 + 4x_2 &= 10 \end{aligned} \implies \left[\begin{array}{cc|c} 1 & 2 & 3 \\ 2 & 4 & 10 \end{array}\right] \xrightarrow{R_2 - 2R_1} \left[\begin{array}{cc|c} 1 & 2 & 3 \\ 0 & 0 & 4 \end{array}\right]$$

1.2.3 Free variable systems

Free variables are variables with their corresponding columns having **no leading 1s** before 0s in the augmented matrix. If a system has free variables, the solution exists in a space, defined by the general solution.

Consider the following system with free variables, we can derive:

$$\left[\begin{array}{ccccc|c} 1 & 0 & -5 & 0 & 8 & 3 \\ 0 & 1 & -4 & -1 & 0 & 6 \\ 0 & 0 & \underbrace{0}_{free} & \underbrace{0}_{free} & 1 & 0 \end{array}\right] \implies \begin{array}{c} x_1 - 5x_3 + 8x_5 = 3 \\ x_2 - 4x_3 - x_4 = 6 \\ x_5 = 0 \end{array}$$

We can ignore the free variables, and substitute x_5 into both equations, deriving the general solution to be:

$$x = \begin{bmatrix} 3 + 5x_3 \\ 6 + x_4 + 4x_3 \\ x_3 \\ x_4 \\ 0 \end{bmatrix}$$

1.3 Matrix determinant

Determinant is a property of a matrix, represented by $|A|$ or $det(A)$. Only square matrices have a determinant.

1.3.1 Finding the value of determinant

$det(A \in \mathbb{R}^{2\times 2})$

$for \quad A = \left[\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right], \quad |A| = ad - bc$

$det(A \in \mathbb{R}^{n\times n})$

Choose any row or column, then for each element, ignore their corresponding row and column to give the *co-factor* of the element.
$det(A) = \sum(a \times det(\text{co-factor}) \times (-1)^{r_a + c_a})$
Where a is an element from the picked row or column, and $r_a c_a$ are the row and column number of a

---

$|\begin{smallmatrix} 1 & 2 & 3 \\ 1 & 1 & 1 \\ 3 & 2 & 1 \end{smallmatrix}| = 1|\begin{smallmatrix} 1 & 1 \\ 2 & 1 \end{smallmatrix}| - 2|\begin{smallmatrix} 1 & 1 \\ 3 & 1 \end{smallmatrix}| + 3|\begin{smallmatrix} 1 & 1 \\ 3 & 2 \end{smallmatrix}| = 0$

Determinant of row-echelon form matrices

Product of diagonal elements. $|\begin{smallmatrix} 1 & 2 & 3 \\ 0 & 2 & 3 \\ 0 & 0 & 3 \end{smallmatrix}| = 1 \cdot 2 \cdot 3 = 6$

1.3.2 Determinant related properties

1. Vector dependency
2. Matrix inverse
3. Matrix eigenpairs

1.4 Vectors

Vectors are matrices with a single column, represented by letters with an arrow above them. $\overrightarrow{v} = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}$

Dot product

The dot product of two vectors $\vec{a}$ and $\vec{b}$ is essentially $\vec{a}^T \cdot \vec{b}$

Vector length

$||\vec{v}||$ gives the length of a vector. $||\begin{bmatrix} a & b & \dots \end{bmatrix}^T|| = \sqrt{a^2 + b^2 + \dots}$

Unit vectors

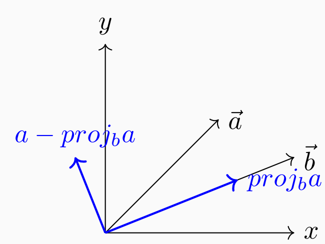
They are vectors with the length of one. $\frac{\vec{v}}{||\vec{v}||}$ transforms the vector to a unit vector.

Vector dependency

If $\begin{vmatrix} v_1 & v_2 & \dots \end{vmatrix} = 0$, the vectors are dependent of each other

Orthogonal projection

The *orthogonal projection* of a onto b is $proj_b a = (\frac{a^T b}{b^T b})b$.
The component of a perpendicular to b is given by $a - proj_b a$



Orthonormal matrices

Orthonormal matrices have vector columns length 1, and which any product of two vectors is 0. For an orthonormal matrix A, $AA^T = I$

Expressing a vector as a linear combination

We can express a vector $\vec{v}$ as the *linear combination* of other vectors $a, b, \dots$ by:
$v = proj_a v + proj_b v + \dots$

Vector spans

Vectors have the same span if they are *linearly dependent*
The dimension of $span(A)$ is given by the number of non-zero rows after GE

1.5 Matrix inverse

Inverse is a property and function of a matrix, represented by A^{-1}

Matrix inversibility

The matrix has an inverse if $|A| \neq 0$, and that it is a square matrix.

$A^{-1} \in \mathbb{R}^{2\times 2}$

$$A^{-1} = \frac{1}{|A|}\begin{bmatrix} d & -b \\ -c & a \end{bmatrix}$$

$A^{-1} \in \mathbb{R}^{n\times n}$

$$\begin{bmatrix} A \mid I \end{bmatrix} \xrightarrow{GE} \begin{bmatrix} I \mid A^{-1} \end{bmatrix}$$

1.6 Matrix eigenpairs

Matrices have properties *eigenvalues* λ_n and *eigenvectors* v_n. They must satisfy the condition $Av = \lambda v$, and has the following properties:

1. λ_n and v_n come in *pairs* (each value correspond to a vector)
2. Matrices can only have $\lambda_1, \lambda_2 \dots \lambda_n$ where n is the smallest dimension of the matrix
3. An v_n includes all vectors of its multiple, and is non-zero

By convention, $\lambda_1 > \lambda_2 > \cdots > \lambda_n$

Solving for eigenvalues

Note that we can rearrange the condition to $(A - I\lambda)v = 0$, and into $|A - I\lambda| = 0$. Rearrange into polynomial equation and solve for λ by first trial and error (if degree > 2) and then factorization.

The eigenvalues of $\left[\begin{smallmatrix} 1 & 1 \\ 4 & 1 \end{smallmatrix}\right]$ can be found by $\left|\begin{smallmatrix} 1-\lambda & 1 \\ 4 & 1-\lambda \end{smallmatrix}\right| = 0 : \quad (1-\lambda)^2 - 4 = 0 \quad \rightarrow \quad \lambda_1 = 3,\ \lambda_2 = -1$

Algebraic and Geometric Multiplicity

AM is the number of times the value of λ occur (e.g. $(1-\lambda)^2 = 0, \lambda_1 = 1$ has $AM = 2$)
GM is the size of nullspace (line of 0s) when plugging λ into $(A - I\lambda)v = 0$

Solving for eigenvectors

Plug each λ into $(A - I\lambda)v = 0$, perform GE, then let the deterministic variable(s) as 1.

For $\lambda_1 = 3$ and $A = \left[\begin{smallmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{smallmatrix}\right]$:

$$(A - 3\lambda)v = 0 \Longrightarrow \left[\begin{array}{ccc|c} -2 & 1 & 1 & 0 \\ 1 & -2 & 1 & 0 \\ 1 & 1 & -2 & 0 \end{array}\right] \xrightarrow{GE} \left[\begin{array}{ccc|c} -2 & 1 & 1 & 0 \\ 0 & -1 & 1 & 0 \\ 0 & 0 & 0 & 0 \end{array}\right] GM = 1$$

From the augmented matrix: $-2x_1 + x_2 + x_3 = 0, \quad -x_2 + x_3 = 0.$
As x_3 is in both equations, let $x_3 = 1$: $x_2 = 1, x_1 = 1 \rightarrow v_1 = \left[\begin{smallmatrix} 1 \\ 1 \\ 1 \end{smallmatrix}\right]$

Solving for eigenvectors with GM 2

Find the other eigenvector with $v = z_2 - proj_{z_1} z_2$

For $v = \begin{bmatrix} -x_2-2x_3 \\ x_2 \\ x_3 \end{bmatrix}$, find the two solution forms as $z_1 = \begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix}$, $z_2 \begin{bmatrix} -2 \\ 0 \\ 1 \end{bmatrix}$

Then let $z_1 = v_1$, and solve for v_2 with $v_2 = z_2 - proj_{z_1} z_2 = \begin{bmatrix} -2 \\ 0 \\ 1 \end{bmatrix} - \frac{2+0+0}{1+1+0} \begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix}$

Related properties:

1. Matrix diagonalization
2. Quadratic form
3. Differential equations

1.7 Matrix diagonalization

Condition

A matrix can only be diagonalized if all of it's eigenpairs has their $AM = GM$.

Eigendecomposition

$$A = VDV^{-1}, \quad \text{where } V = \begin{bmatrix} v_1 \\ \dots \\ v_n \end{bmatrix}, \; D = \begin{bmatrix} \lambda_1 & 0 & 0 \\ 0 & \ddots & 0 \\ 0 & 0 & \lambda_n \end{bmatrix}$$

We can also derive $A^k = VD^kV^{-1}$

1.8 Differential equations

Given $x(0)$, solve for $x(t)$ with the following steps:

1. Solve for eigenpairs
2. $xk = cve^{\lambda t} + \cdots + cve^{\lambda t}$
3. Solve c by augmented matrix $\left[\begin{array}{ccc|c} v_1 & \dots & v_n & x(0) \end{array}\right]$

1.9 Quadratic form

A quadratic equation can be written into the form $Q(x) = xAx^T$, where A is a symmetric matrix.

Symmetric matrices

A matrix is symmetric if $A = A^T$

Linear to matrix quadratic

$$\begin{aligned} Q(x) &= ax_1^2 + bx_2^2 + cx_3^2 + dx_1x_2 + ex_2x_3 + fx_1x_3 \\ &= \begin{bmatrix} x_1 & x_2 & x_3 \end{bmatrix} \begin{bmatrix} a & d/2 & f/2 \\ d/2 & b & e/2 \\ f/2 & e/2 & c \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} \end{aligned}$$

Definitiveness of a quadratic

For $Q(x) = xAx^T$, considering the eigenvalues of A:

- $Q(x)$ is *positive definitive* if $\forall\lambda > 0$
- $Q(x)$ is *negative definitive* if $\forall\lambda < 0$
- $Q(x)$ is *indefinite* if otherwise
- $Q(x)$ is *p/n semi-definitive* if the conditions are inclusive (i.e. $\lambda \geq or \leq 0$)

Minimum and maximum values

$\lambda_{max} = max\{xAx^T : ||x|| = 1\}$
$\lambda_{min} = min\{xAx^T : ||x|| = 1\}$
Apply the corresponding $v \to x$ to find the min/max of $Q(x)$. Note that $||x|| = 1$ is a required restriction on $Q(x)$, otherwise its size would be infinite.

2 Complex numbers

The imaginary number $i = \sqrt{-1}$. A complex number is usually denoted as $z = a + bi$. In operations, the imaginary number can be considered as an unknown, though note that $i^2 = -1$ (and so forth).

2.1 Useful items

Common properties

The *modulus*: $|z| = \sqrt{a^2 + b^2}$
The *conjugate*: $conj(z) = \hat{z} = a - bi$
The *argument*: $arg(z) = \theta = 2tan^{-1}(\frac{b}{a+|z|})$

Polar form

We can write a complex number in polar form: $z = r(\cos\theta + i\sin\theta), \quad$ where $r = |z|$.
This represents the point on a circle with radius r, at the angle θ.

Euler's formula

$re^{i\theta} = r(\cos\theta + i\sin\theta)$.

Using the formula, we can easily derive:

- $(\cos\theta + i\sin\theta)^n = (\cos n\theta + i\sin n\theta)$
- $(\cos\theta + i\sin\theta)^{-1} = (\cos\theta - i\sin\theta)$
- $(\sin\theta + i\cos\theta) = i(\cos\theta - i\sin\theta) = ie^{-i\theta}$

Complex roots

Consider $nz = k$, the solutions to the equation is:

$\underbrace{kw^0 + kw^1 + \cdots + kw^{n-1}}_{n \text{ solutions}} \quad for\ w = e^{(\frac{2\pi i}{n})}$,

Set notation

Elements in a set are *unique*.

$$A = \{\underbrace{\underbrace{1}_{\text{An element}} \quad ,2,3}_{\text{Elements}}\} \qquad B = \{\underbrace{x}_{\text{For all}} \mid \underbrace{0 < x < 2}_{\text{Such that}} \underbrace{,}_{\text{And}} \quad x \in \mathbb{R}\} = \{1\}$$

2.2 Image under complex function

Given a complex function $f(z)$ and a set of complex numbers $D = \{g(z) \in \mathbb{C} : h(z)\}$, our goal is to solve for the function $f(D)$ in a form that we can identify the shape / image.

2.2.1 General steps

1. Find $f(D)$ by finding $f(g(z))$
2. Let $f(g(z))$ be $x' + y'i$, then solve for x and y in terms of x' and y'
   (Note that $g(z)$ and $h(z)$ are functions of $x + yi$)
3. Substitute your x and y to $f(D)$, so that all unknowns are in terms of x' and y'
4. Rearrange $h(z)$ to solve the shape of the image

2.2.2 Simple example

Let $D = \{x + iy \in \mathbb{C}\ ||x + iy| > 1\}$, with complex function $f(z) = -1/z$. Find $f(D)$, then sketch the picture of $f(D)$ on the complex plane. 23 Dec, Part 2 Assignment 2 Question 3

We first find $f(D)$ by finding $f(x + yi)$:

$$f(x + yi) = \frac{-1}{x + yi} \to f(D) = \{\frac{-1}{x + yi} \in \mathbb{C}||x + yi| > 1\}$$

Then let $\frac{-1}{x+yi}$ be $x' + y'i$, then solve for x and y in terms of x' and y'

$$x' + y'i = \frac{-1}{x + yi} \to x + yi = \frac{-1}{x' + y'i}$$

Substitute the values we found into $f(D)$

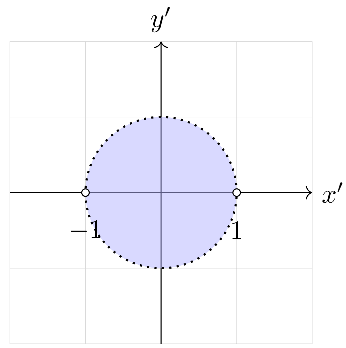
$$f(D) = \{x' + y'i \in \mathbb{C} | |\frac{-1}{x' + y'i}| > 1\}$$

Then finally we rearrange the right side of the equation as:

$$\frac{1}{\sqrt{x'^2 + y'^2}} > 1$$

$$1 > x'^2 + y'^2$$

Therefore, we can conclude that the image is a filled circle with radius 1 (excluding edges):



2.2.3 A more elegant way

We can use Euler's formula to solve questions involving circles. Using the above example:

We let $x + yi$ as $re^{i\theta}$ and substitute:

$$D = \{re^{i\theta} \in \mathbb{C} | \underbrace{r > 1}_{\text{As } |z| \text{ is } r}\}$$

Then following the usual steps:

$$f(D) = \{-r^{-1}e^{-i\theta} \in \mathbb{C} | r > 1\}$$

$$f(D) = \{r^{-1}\underbrace{e^{-i\theta+i\pi}}_{-1=e^{i\pi}} \in \mathbb{C} | r > 1\}$$

$$Let\ r' = r^{-1}, \theta' = \pi - \theta \to r = r'^{-1}$$

$$f(D) = \{r'e^{i\theta'} \in \mathbb{C} | r^{-1} > 1\}$$

$$f(D) = \{r'e^{i\theta'} \in \mathbb{C} | r < 1\}$$

3 Hyperbolic functions

The scope of this course only focuses on these three functions:

$$\sinh x = \frac{e^{ix} + e^{-ix}}{2} \qquad \cosh x = \frac{e^{ix} - e^{-ix}}{2} \qquad \tanh x = \frac{\sinh x}{\cosh x}$$

4 Probabilities

Combinations

The formula for combinations, also known as the binomial coefficient, is given by:

$$\binom{n}{r} = \frac{n!}{r!(n-r)!}$$

Which represents the number of ways to choose r items from a set of n distinct items without regard to their order.

Pascal's rule

Pascal's rule is given by $\binom{n+1}{r} = \binom{n}{r} + \binom{n}{r-1}$

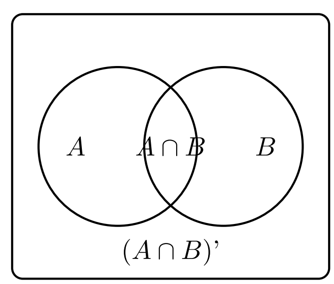
Probability measures the likelihood of an event happening. It assigns a numerical value between 0 and 1 to an event, where 0 represents an impossible event and 1 indicates a certain event.
For example, the probability of flipping a fair coin and getting heads is 0.5, as there are two equally likely outcomes (heads or tails).
The key to understanding probabilities (or anything else) is practice.

4.1 Infinite sets

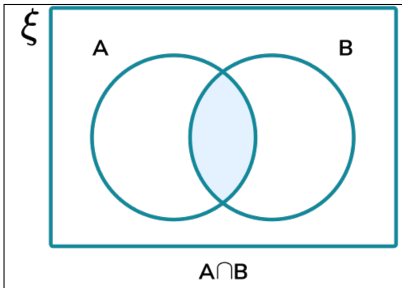
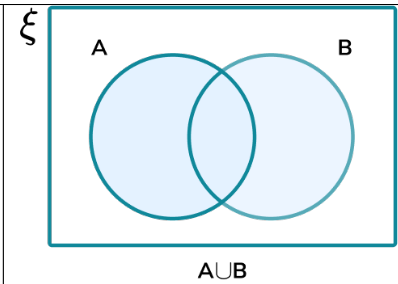
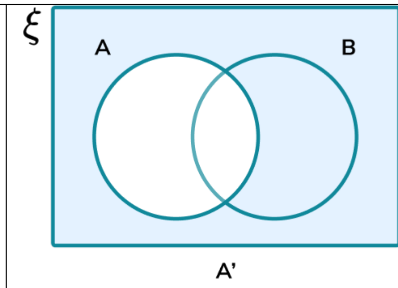
Set	**Description**	**Example**
$\mathbb{Z}$	Integers	$-1, 0, 1$
$\mathbb{N}$	Natural	$0, 1, 2$
$\mathbb{R}$	Real	Any number $\frac{4}{7}, 0.1, 1, \pi$
$\mathbb{Q}$	Rational	Any number that can be expressed as a fraction $\frac{22}{7}$
$\mathbb{C}$	Complex	$a + bi$

4.2 Venn diagrams

Venn diagrams illustrate concepts like intersections, unions, which is a good way to visualize probabilities. The overlapping regions indicate elements that belong to multiple sets, while non-overlapping regions represent elements unique to specific sets.
In this example, two circles are drawn to represent sets A and B. The overlapping region is labeled as the intersection of sets A and B, denoted by A ∩ B.



4.3 Probability notation and event types

		
A and B	A or B	Not A
Intersection	Union	Compliment

Conditional probabilities

The probability of A given that B occurs: $P(A \mid B) = \frac{P(A \cup B)}{P(B)}$

Independent events

A and B are said to be *independent* if $P(A) \times P(B) = P(A \cap B)$.
Being independent means that the probability of an event *has no influence on the other*

4.4 Random variables

Notation

Random variables are denoted with capital letters X
The possible outcomes are denoted with regular letters x
Probability that the outcome of X is x is denoted by $P(X = x)$

Expected value and Variance

$E(X^n) = \sum(x^n \cdot P(X = x))$ gives the expected value, which represents the mean value (outcome) of the random variable.
$Var(X) = E(X^2) - E(X)^2$ gives the variance, which is a measure of the variability of the random variable's outcomes.

Operations of $E(X)$ and $Var(X)$

$E(X + Y) = E(X) + E(Y)$, which any addition / subtraction function within $E()$ can be expanded.

If X and Y are independent: $Var(X + Y) = Var(X) + Var(Y), \qquad E(XY) = E(X) \cdot E(Y)$

$for\ Y = aX + b : E(Y) = aE(X) + b, \qquad Var(Y) = a^2 Var(X)$

4.5 Joint random variables

Joint random variables are in the form $P(X = x, Y = y)$. We can visualize the joint distribution in the following way:

$X \backslash Y$	x_1	x_2	
y_1	$P(X = x_1, Y = y_1)$	$P(X = x_2, Y = y_1)$	$P(Y = y_1)$
y_2	$P(X = x_1, Y = y_2)$	$P(X = x_2, Y = y_2)$	$P(Y = y_2)$
	$P(X = x_1)$	$P(X = x_2)$	

Note that the sum of a column or row results in the corresponding variable's probability of outcome.

Expected value

$E(X + Y) = \sum((x + y)P(X = x, Y = y))$
$E((XY)^n) = \sum((xy)^n P(X = x, Y = y))$

4.6 Random variables of random variables (outcomes)

For random variables modelled in the following way:

$$\bar{X} = X_1 + X_2 + \cdots + X_n$$

We can deduce that:

$$E(\bar{X}) = \frac{E(X_1) + \cdots + E(X_n)}{n} = \frac{nE(X)}{n}$$
$$Var(\bar{X}) = \frac{Var(X_1) + \cdots + Var(X_n)}{n^2} = \frac{nVar(X)}{n^2}$$

Expected value and Variance

$E(\bar{X}) = E(X)$
$Var(\bar{X}) = \frac{Var(X)}{n}$

5 Probability distributions

A *continuous* random variable has probabilities the area under its curve. Hence, $P(X = n)$ for any outcome n is 0. A *discrete* random variables have specific probabilities assigned to an outcome.

Operation on continous ranges

$P(X < x) = P(X \leq x)$
$P(X > x) = 1 - P(X < x)$
$P(a \leq X \leq b) = P(X < b) - P(X < a)$

Operation on discrete ranges

$P(X < x) = P(X \leq x - 1)$
$P(X > x) = 1 - P(X \leq x)$
$P(a \leq X \leq b) = P(X \leq b) - P(X \leq a - 1)$

5.1 Probability distribution functions

A *p.d.f* is a continuous function that returns the probability of the given outcome. The following is an example of a p.d.f:

$$X \sim f(x) = \begin{cases} Kx^2 & 0 \leq x \leq 1 \\ 0 & \text{otherwise} \end{cases}$$

Using the function

$P(X = x) = 0$
$P(a < X < b) = P(a \leq X \leq b) = \int_a^b f(x)dx$

Note that $\int_{\infty}^{-\infty} f(x)dx = 1$, as the total probability of any event is 1. This condition *must be true* for $f(x)$ to be a valid p.d.f. Hence for the example $K = 3$

Expected value

$E(X^n) = \int_{\infty}^{-\infty} x^n f(x)dx$

5.2 Cumulative distribution function

To obtain a *c.d.f* $X' \sim F(x)$where $P(X' = x) = P(X < x)$, all we have to do is integrate the p.d.f:

$$F(X) = \int f(x)dx$$

Using our example:

$$F(X) = \begin{cases} 1 & x > 1 \\ x^3 & 0 < x < 1 \\ 0 & x < 0 \end{cases}$$

And to convert a c.d.f back to it's p.d.f, all we have to do is differentiate $F(x)$:

$$f(x) = \frac{d}{dx}F(x)$$

5.3 Statistical distributions

5.3.1 Common statistical distribution

Continuous	Discrete	Discrete	Discrete	Discrete	Discrete
Exponential	Bernoulli	Binomial	Geometric	Negative Binomial	Possion
$Expo(\lambda)$	$Ber(p)$	$B(n,p)$	$Geo(p)$	$NB(n,p)$	$Po(\lambda)$
$\lambda e^{-\lambda x}$	p if true, else $(1-p)$	$\binom{n}{x}p^x(1-p)^{n-x}$	$p(1-p)^{x-1}$	$\binom{x-1}{n-1}p^n(1-p)^{x-n}$	$\frac{\lambda^x}{x!}e^{-\lambda}$
$E = \lambda^{-1}$	$E = p$	$E = np$	$E = \frac{1}{p}$	$E = \frac{n}{p}$	$E = \lambda$
$Var = \lambda^{-2}$	$Var = p(1-p)$	$Var = np(1-p)$	$Var = \frac{1-p}{p2}$	$Var = \frac{n(1-p)}{p2}$	$Var = \lambda$

5.3.2 Usage cases

- Ber: Outcomes only $True$ or $False$
- B: x successes in n
- Geo: 1st success at x tries
- NB: nth success at x tries
- Po: x successes in interval λ

5.3.3 Normal distribution

The Normal distribution is given by the following formula (which you don't have to memorize):

$$X \sim N(\mu, \sigma^2) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}, \qquad \text{where } \underbrace{\mu}_{\text{mean}}, \underbrace{\sigma^2}_{\text{variance}}$$

To calculate $P(X < x)$, we have to *standardize* our N by $Z \sim N(0,1)$, $P(X < x) = P(Z < \frac{x-\mu}{\sigma})$. Here Z is the *standard normal variable*.

Reading the Z-table

For $P(Z < z) = p$, to find p, locate the header and leftmost column in the z-table such that their *sum* is z. The corresponding intersecting cell is p.
Z_p gives the value z in $P(Z < z) = p$

Critical interval

The critical interval is given by:

$$C.I. = [\bar{X} \pm Z_{\frac{C.L.+1}{2}} \times \frac{\sigma}{\sqrt{n}}], \qquad \text{where } \underbrace{C.L.}_{\text{Confidence level}}, \underbrace{\bar{X}}_{\text{Mean value}}$$

Hence, we can derive that the critical interval width is:

$$2 \times Z_{\frac{C.L.+1}{2}} \times \frac{\sigma}{\sqrt{n}}$$