MSc Mas6002 Introductory Material
Background Mathematics

This is a list of mathematical ideas and techniques that are likely to come in useful on the MSc in Statistics. You should have a look through the list and revise topics you feel rusty on.

1. SET THEORY Some of the detail here isn't often needed.

   We usually work with objects which are members of some large set or **space** Ω

   Elements (inclusion): $\omega \in \Omega$ means that ω is a member (or element) of Ω. If ω is not an element of Ω we write $\omega \notin \Omega$.

   Empty set $\emptyset$: a set with no elements

   Subset A $\subset B$ if $\omega \in A \Longrightarrow \omega \in B$

   Union $\omega \in A \cup B$ if $\omega \in A$ or $\omega \in B$; $\cup_{i=1}^{n} A_i = A_1 \cup A_2 \cup \ldots \cup A_n$

   Intersection $\omega \in A \cap B$ if $\omega \in A$ and $\omega \in B$; $\cap_{i=1}^{n} A_i = A_1 \cap A_2 \cap \ldots \cap A_n$

   A, B **mutually exclusive** if $A \cap B = \emptyset$

   A, B **complementary** if $A \cap B = \emptyset$ and $A \cup B = \Omega$ (then B is written as $\overline{A}$ or A^c and is called the **complement** of A)
   Note that $(A^c)^c = A$ (the complement of the complement of A is A itself)

   $A_1, \ldots, A_n$ are **exhaustive** if $\cup_{i=1}^{n} A_i = \Omega$

   $A_1, \ldots, A_n$ form a **partition** of Ω if they are exhaustive and mutually exclusive
   i.e. $\cup_{i=1}^{n} A_i = \Omega$ and $A_i \cap A_j = \emptyset,\ i \neq j$.i.e. $\omega \in$ exactly one A_i.

   Set difference $\omega \in A \setminus B$ if $\omega \in A$ and $\omega \notin B$. i.e. $A \setminus B = A \cup \overline{B}$

   Size of set $|A|$ is no. of elements in A .

Laws $\cup$ and $\cap$	are	commutative	e.g.	$A \cap B = B \cap A$
		associative	e.g.	$(A \cap B) \cap C = A \cap (B \cap C)$
	and	distributive	e.g.	$A \cup (B \cap C) = (A \cup B) \cap (A \cup C)$

   De Morgan's Laws $\overline{\{\cup_{i=1}^{n} A_i\}} = \cap_{i=1}^{n} \overline{A}_i$

   $$\overline{\{\cap_{i=1}^{n} A_i\}} = \cup_{i=1}^{n} \overline{A}_i$$

   Cartesian product $A \times B = \{(\omega_1, \omega_2) : \omega_1 \in A, \omega_2 \in B\}$

Indicator functions $I_A(\omega) = \begin{cases} 1 & \omega \in A \\ 0 & \omega \notin A \end{cases}$

$$\text{e.g. } I_{(0,2]}(x) = \begin{cases} 1 & 0 < x \leq 2 \\ \\ 0 & x \leq 0, x > 2 \end{cases}$$

2. QUADRATIC EQUATIONS

$ax^2 + bx + c = 0$ has roots

$$x = \frac{-b \pm \sqrt{b^2 - 4ac}}{2a}$$

3. RULES OF INDICES AND LOGARITHMS

$$x^a x^b = x^{a+b}$$
$$(x^a)^b = x^{ab}$$

$$\begin{array}{lcl} \log(xy) & = & \log x + \log y \\ \log \frac{x}{y} & = & \log x - \log y \\ \log x^a & = & a \log x \end{array}$$

4. ARITHMETIC AND GEOMETRIC PROGRESSIONS

AP (arithmetic progression): $a, a+d, a+2d, \ldots$
Sum to n terms

$$S_n = na + \frac{1}{2}n(n-1)d.$$

GP (geometric progression): $a, ar, ar^2, \ldots$
Sum to n terms

$$S_n = \frac{a(1-r^n)}{1-r}.$$

Infinite sum (if $|r| < 1$) $S_\infty = \dfrac{a}{1-r}$.

5. PERMUTATIONS AND COMBINATIONS

No. ways of picking r from n

$$\text{order important} \qquad {}^nP_r = \frac{n!}{(n-r)!}$$

$$\text{order unimportant} \quad {}^nC_r = \begin{pmatrix} n \\ r \end{pmatrix} = \frac{{}^nP_r}{r!} = \frac{n!}{r!(n-r)!}$$

Note relationship $b \begin{pmatrix} a \\ b \end{pmatrix} = a \begin{pmatrix} a-1 \\ b-1 \end{pmatrix}$

6. USE OF SUM AND PRODUCT NOTATION

$$x_1 + x_2 + \ldots + x_n = \sum_{i=1}^n x_i$$

$$ax_1 + ax_2 + \ldots + ax_n = a\sum_{i=1}^n x_i$$

$$a_1x_1 + a_2x_2 + \ldots + a_nx_n = \sum_{i=1}^n a_ix_i$$

$$x_1 \times x_2 \times \ldots \times x_n = \prod_{i=1}^n x_i$$

$$ax_1 \times ax_2 \times \ldots \times ax_n = a^n \prod_{i=1}^n x_i$$

$$a_1x_1 \times a_2x_2 \times \ldots \times a_nx_n = \prod_{i=1}^n a_ix_i$$

$$\text{Also}\quad \log\left(\prod_{i=1}^n x_i\right) = \sum_{i=1}^n \log x_i$$

7. SERIES EXPANSIONS AND LIMITS

Binomial theorem: if n positive integer or $-1 < x < 1$

$$(1+x)^n = 1 + nx + \frac{n(n-1)}{2!}x^2 + \ldots + \frac{n(n-1)\ldots(n-r+1)}{r!}x^r + \ldots$$

Useful series expansions:

$$e^x = 1 + \frac{x}{1!} + \frac{x^2}{2!} + \frac{x^3}{3!} + \ldots$$

$$\ln(1+x) = x - \frac{1}{2}x^2 + \frac{1}{3}x^3 - \ldots \qquad -1 < x \leq 1$$

$$\sin x = x - \frac{x^3}{3!} + \frac{x^5}{5!} - \ldots$$

$$\cos x = 1 - \frac{x^2}{2!} + \frac{x^4}{4!} - \ldots$$

Also:

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

8. MATRICES

multiplication
determinants
inversion —- using $\boldsymbol{A}^{-1} = \dfrac{1}{\det \boldsymbol{A}}\text{adj}\boldsymbol{A}$ not just elementary row/col operations.

General properties:

- $\boldsymbol{A}^T$ (often written $\boldsymbol{A}'$) denotes the transpose of matrix $\boldsymbol{A}$ and is obtained from $\boldsymbol{A}$ by interchanging the rows and columns, that is the columns of $\boldsymbol{A}^T$ are the rows of $\boldsymbol{A}$ and the rows of $\boldsymbol{A}^T$ are the columns of $\boldsymbol{A}$.
- The square matrix $\boldsymbol{A}$ is symmetric if $\boldsymbol{A} = \boldsymbol{A}^T$.
- If the symmetric matrix $\boldsymbol{A}$ is non-singular, then $\boldsymbol{A}^{-1}$ is also symmetric.
- The square matrix $\boldsymbol{A}$ is diagonal if all off-diagonal elements of $\boldsymbol{A}$ are zero.
- Multiplication is distributive over addition and subtraction, so

$$(\boldsymbol{A}-\boldsymbol{B})(\boldsymbol{C}-\boldsymbol{D}) = \boldsymbol{AC} - \boldsymbol{BC} - \boldsymbol{AD} + \boldsymbol{BD}.$$

- $(\boldsymbol{A}+\boldsymbol{B})^T = (\boldsymbol{A}^T + \boldsymbol{B}^T)$ and $(\boldsymbol{AB})^T = \boldsymbol{B}^T\boldsymbol{A}^T$.
- If $\boldsymbol{a}$ is a (column) vector of length n then $\boldsymbol{a}^T\boldsymbol{a} = {a_1}^2 + {a_2}^2 + \ldots + {a_n}^2$.
- If $\boldsymbol{A}$ is a $n \times p$ matrix, then $\boldsymbol{AA}^T$ is a $n \times n$ matrix obtained by taking products of the rows of $\boldsymbol{A}$, whilst $\boldsymbol{A}^T\boldsymbol{A}$ is a $p \times p$ matrix obtained by taking products of the columns of $\boldsymbol{A}$ and thus both are symmetric.

- A square matrix $\boldsymbol{A} = (\boldsymbol{a}_1, \boldsymbol{a}_2, \ldots, \boldsymbol{a}_p)$ whose columns are normalised (that is ${\boldsymbol{a}_i}^T\boldsymbol{a}_i = 0$ for all $i \neq j$) is called an **orthogonal** matrix. Hence for an orthogonal matrix $\boldsymbol{A}^T\boldsymbol{A} = \boldsymbol{AA}^T = \boldsymbol{I}$ and $\boldsymbol{A}^T = \boldsymbol{A}^{-1}$.
- If $\boldsymbol{A}$ is square then the *trace* of $\boldsymbol{A}$ is the sum of its diagonal entries and is written as trace($\boldsymbol{A}$) or tr($\boldsymbol{A}$). Thus $\text{tr}(\boldsymbol{A}) = \sum_i \boldsymbol{A}_{ii}$; now

$$\text{tr}(\boldsymbol{AB}) = \sum_i (\boldsymbol{AB})_{ii} = \sum_i \sum_k \boldsymbol{A}_{ik}\boldsymbol{B}_{ki} = \sum_k \sum_i \boldsymbol{B}_{ki}\boldsymbol{A}_{ik} = \text{tr}(\boldsymbol{BA}).$$

  A special case of this is that if $\boldsymbol{M}$ is an $r \times r$ matrix and $\boldsymbol{S}$ is an $r \times p$ matrix then (with $\boldsymbol{A} = \boldsymbol{MS}$ and $\boldsymbol{B} = \boldsymbol{S}^T$)

$$\text{tr}\left(\boldsymbol{S}^T\boldsymbol{MS}\right) = \text{tr}\left(\boldsymbol{MSS}^T\right)$$

Basic Properties of Eigenvalues and Eigenvectors

Let $\boldsymbol{A}$ be a real $p \times p$ matrix.
The eigenvalues of $\boldsymbol{A}$ are the roots of the p-degree polynomial in λ:

$$q(\lambda) = \det(\boldsymbol{A} - \lambda\boldsymbol{I}_p) = 0.$$

Let these be $\lambda_1, \lambda_2, \ldots, \lambda_p$.
Since the matrices $\boldsymbol{A} - \lambda_i\boldsymbol{I}_p$ are singular (i.e. have zero determinant) there exist vectors x_i called the **eigenvectors** of $\boldsymbol{A}$ such that

$$(\boldsymbol{A} - \lambda_i\boldsymbol{I}_p) = 0, \text{ i.e. } \boldsymbol{A}x_i - \lambda_i x_i = 0.$$

(Strictly, if $\boldsymbol{A}$ is non-symmetric, the x_i are right-eigenvectors and we can define left-eigenvectors y_i such that $y_i\boldsymbol{A} - \lambda_i y_i = 0$)

- $\sum_{i=1}^p \lambda_i =$ trace($\boldsymbol{A}$).
- $\prod_{i=1}^p \lambda_i = \det(\boldsymbol{A})$, often written $|\boldsymbol{A}|$.
- $\boldsymbol{A}$ and $\boldsymbol{CAC}^{-1}$ have identical eigenvalues for $\boldsymbol{C}$ non-singular.
- Eigenvectors of $\boldsymbol{CAC}^{-1}$ are $\boldsymbol{C}x_i$.
- $\boldsymbol{AB}$ and $\boldsymbol{BA}$ have identical non-zero eigenvalues.
- Eigenvectors of $\boldsymbol{BA} = \boldsymbol{B} \times$ those of $\boldsymbol{AB}$.
- $\boldsymbol{A}$ symmetric $\Rightarrow$ eigenvalues real.
- $\boldsymbol{A}$ symmetric $\Rightarrow$ eigenvectors corresponding to distinct eigenvalues are orthogonal.

Differentiation with respect to vectors:

If $x = (x_1, x_2, \ldots, x_p)^T$ is a p-vector and $f = f(x)$ is a scalar function of x, we define $\dfrac{\partial f}{\partial x}$ to be the vector $\left(\dfrac{\partial f}{\partial x_1}, \dfrac{\partial f}{\partial x_2}, \ldots, \dfrac{\partial f}{\partial x_p}\right)^T$.

- Quadratic forms: if $f(x) = x^T \boldsymbol{S} x$, where $\boldsymbol{S}$ is a symmetric $p \times p$ matrix, then $\dfrac{\partial f}{\partial x} = 2\boldsymbol{S}x$.
- If $\boldsymbol{S}$ is not symmetric, then $\dfrac{\partial (x^T S x)}{\partial x} = (\boldsymbol{S} + \boldsymbol{S}^T)x$
- Special case: $\boldsymbol{S} = \boldsymbol{I}_p$, $\dfrac{\partial x^T x}{\partial x} = 2x$.
- Inner-products with scalars: if $f(x) = a^T x$, then $\dfrac{\partial f}{\partial x} = a$

9. GEOMETRY

    equation of straight line $y = mx + c$

    equation of circle centre (a, b), radius c $(x-a)^2 + (y-b)^2 = c^2$

    equation of ellipse $\frac{x^2}{a^2} + \frac{y^2}{b^2} = 1$

    plotting simple functions and inequalities

10. CALCULUS

    standard differentials and integrals, especially

$$\frac{d}{dx}(x^n) = nx^{n-1}$$

$$\frac{d}{dx}(\sin x) = \cos x$$

$$\frac{d}{dx}(\cos x) = -\sin x$$

$$\frac{d}{dx}(\tan x) = \sec^2 x$$

$$\frac{d}{dx}(e^x) = e^x$$

$$\frac{d}{dx}(\ln x) = \frac{1}{x}$$

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

and general rules

products $\dfrac{d}{dx}(uv) = u\dfrac{dv}{dx} + v\ \dfrac{du}{dx}$

quotients $\dfrac{d}{dx}\left(\dfrac{u}{v}\right) = \dfrac{v\ du/dx - u\ dv/dx}{v^2}$

function of a function $\dfrac{d}{dx}\ u(v) = \dfrac{du}{dv} \times \dfrac{dv}{dx}$

integrating by parts $\displaystyle\int uv\ dx = \left(\int u\ dx\right) v - \int \left(\int u\ dx\right) \frac{dv}{dx} dx$

double integration.

11. LAGRANGE MULTIPLIERS

Suppose $x = (x_1, x_2, \ldots, x_n)^T$.

To maximize/minimize $f(x)$ (a scalar function of x)
subject to k scalar constraints $g_1(x) = 0, g_2(x) = 0, \ldots, g_k(x) = 0$ where $k < n$,
define $L = f(x) + \sum_{j=1}^{k} \lambda_j g_j(x)$ (the Lagrangian; the λ_j are the Lagrange multipliers)
and max/minimize L with respect to the $n + k$ variables $x_1, x_2, \ldots, x_n, \lambda_1, \lambda_2, \ldots, \lambda_k$.

12. OTHER TOPICS (general ideas required)

Methods of proof : induction and contradiction
Polar coordinates
Convergence
Newton-Raphson method for numerical solution of equations
Complex numbers
Completing the square