Chapter 2. Mathematical Methods

Before discussing probability and statistics, we first review basic mathematical methods that we shall use in this course.

2.1 Notation

The following conventions and set notation will be used throughout the course:

Notation	Set	Description
$\mathbb{R}$	$(-\infty, \infty)$	The real numbers
$\mathbb{R}^+$	$(0, \infty)$	The positive real numbers
$\mathbb{Z}$	$\{\ldots, -2, -1, 0, 1, 2, \ldots\}$	The integers
$\mathbb{Z}^+$	$\{1, 2, 3, \ldots\}$	The positive integers
$\mathbb{N}$	$\{0, 1, 2, 3, \ldots\}$	The natural numbers

2.2 Log and Exponential

Where there is any ambiguity, by "log" we will mean the natural logarithm, sometimes written "ln" elsewhere. For any other base b (e.g. $b = 10$), we will write $\log_b$. So

$$\log \equiv \log_e \equiv \ln .$$

Rules:

- $\log(xy) = \log(x) + \log(y) \implies \log\left(\prod_i x_i\right) = \sum_i \log(x_i)$
- $\log(x^y) = y \log(x)$
- $\log(e^x) = x$
- $\lim_{x \to 0} \log(x) = -\infty$

For exponential, we will use the notations "e" or "exp" interchangeably. So

$$e^x \equiv \exp(x).$$

Rules:

- $\exp(x + y) = \exp(x)\exp(y) \implies \exp\left(\sum_i x_i\right) = \prod_i \exp(x_i)$
- $\exp(x)^y = \exp(xy)$
- $\exp\{\log(x)\} = x$
- $\exp(0) = 1$

2.3 Arithmetic and Geometric Progressions

Consider the infinite sequence of numbers

$$a, a+d, a+2d, a+3d, \ldots$$

The first term in this sequence is a, and then each subsequent term is equal to the previous term plus d, the common difference. Any such sequence is known as an arithmetic progression.

Formulae:

- n^{th} term = $a + (n-1)d$
- Sum of first n terms, $S_n = \dfrac{n}{2}\{2a + (n-1)d\}$
- (Infinite sum, $S_\infty = \pm\infty$, unless $a = d = 0$)

Consider the infinite sequence of numbers

$$a, ar, ar^2, ar^3, \ldots$$

The first term in this sequence is a, and then each subsequent term is equal to the previous term multiplied by r, the common ratio. Any such sequence is known as a geometric progression.

Formulae:

- n^{th} term = ar^{n-1}
- Sum of first n terms, $S_n = \dfrac{a(1-r^n)}{1-r}$, if $r \neq 1$
- Infinite sum, $S_\infty = \dfrac{a}{1-r}$, if $|r| < 1$; (S_∞ diverges otherwise for $a \neq 0$)

2.4 Calculus

Let f, g be functions of a variable x. Then the derivative of f with respect to x, denoted as $\dfrac{df}{dx}$ or $f'(x)$, is

$$\frac{df}{dx} \equiv f'(x) \equiv \lim_{h \to 0} \frac{f(x+h) - f(x)}{h}.$$

Formulae:

- Chain Rule: $\dfrac{d}{dx} f(g(x)) = f'(g(x))g'(x)$
- Product Rule: $\dfrac{d}{dx} f(x)g(x) = f'(x)g(x) + f(x)g'(x)$
- Quotient Rule: for $g(x) \neq 0$, $\dfrac{d}{dx}\dfrac{f(x)}{g(x)} = \dfrac{f'(x)g(x) - f(x)g'(x)}{\{g(x)\}^2}$

Fundamental Theorem of Calculus:

$$\frac{d}{dx}\int_{u=a}^{x} f(u)du = f(x),$$

so differentiation and integration are inverse operations.

Formulae:

- Change of variable: if $y = g(x)$, $\displaystyle\int_a^b f(x)dx = \int_{g(a)}^{g(b)} f(g^{-1}(y))g^{-1\prime}(y)dy$
- By parts: $\displaystyle\int_a^b f(x)g'(x)dx = [f(x)g(x)]_a^b - \int_a^b f'(x)g(x)dx$

Both differentiation and integration are additive. That is, for functions f, g,

$$\frac{d}{dx}\{f(x) + g(x)\} = \frac{df(x)}{dx} + \frac{dg(x)}{dx},$$
$$\int\{f(x) + g(x)\}dx = \int f(x)dx + \int g(x)dx.$$

And for any constant $c \in \mathbb{R}$,

$$\frac{d}{dx}\{c \cdot f(x)\} = c \cdot \frac{df(x)}{dx},$$
$$\int\{c \cdot f(x)\}dx = c \cdot \int f(x)dx.$$

2.5 Function images and inverses

Suppose f is a function $f : X \to Y$. For $A \subseteq X$, the image of A under f, written $f(A)$ is the subset of Y given by

$$f(A) = \{y \in Y | f(x) = y \text{ for some } x \in A\}.$$

The image of X under f can be referred to simply as the *image of* f. Recall the inverse of f, should it exist, is denoted f^{-1} and has the property

$$f^{-1}(f(x)) = x$$

for any value $x \in X$. The inverse image of f could therefore be defined as the image of Y under f^{-1}.

More generally, for $B \subseteq Y$, the inverse image of B under (possibly non-invertible) f is given by

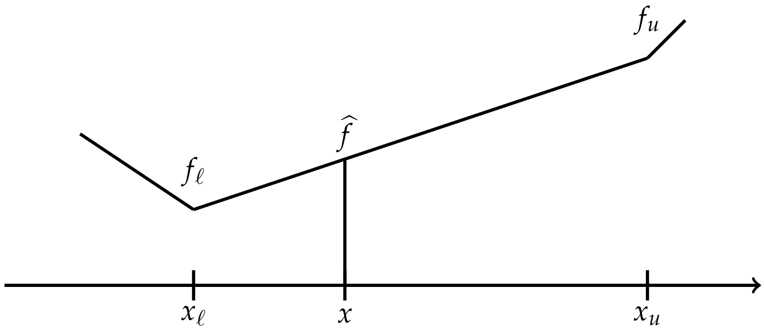
$$f^{-1}(B) = \{x \in X | f(x) \in B\}.$$

2.6 Interpolation

Consider a function $f : \mathbb{R} \to \mathbb{R}$ of unknown form, where all that is known about f is the value it takes at each of a pre-determined discrete set of points $\mathcal{X} = \{a = x_1 < x_2 < \ldots < x_k = b\}$; for each of these values x, denote the corresponding function value $f_x = f(x)$.

Interpolation is the task of finding an approximate value of f for a general point x in the interval $[a, b]$, say $\hat{f}(x)$. (Extrapolation would be the task of finding an approximate value of f for x outside the interval $[a, b]$.)

The most commonly used approximation is linear interpolation, which assumes the underlying function f can be considered approximately piecewise linear between the set of known points.



Let x_ℓ, x_u be nearest pair of points in $\mathcal{X}$ on either side of x. Then $f(x)$ is linearly approximated by

$$\widehat{f}(x) = f_{x_\ell} + (x - x_\ell)\frac{(f_{x_u} - f_{x_\ell})}{(x_u - x_\ell)}.$$