Chem190/290, Spring 2024

Math Review

Reference :

Mathematical Methods for Scientists and Engineers, by McQuarrie. This book is available for free as an e-book via the UC
Merced Library. **Please only check out the relevant sections (no need to read everything)**

1. Derivatives

Definition:

$$\frac{dy}{dx} = \lim_{\Delta x \to 0} \frac{\Delta y}{\Delta x}$$

Some notations:

$$\left.\frac{dy}{dx}\right|_{x=a}, y'(x), y''(x), y^{(n)}(x), y^{(n)}(a)$$

Physical meanings: slope, rate, etc.

Differentiation is a linear operation (swap summation and differentiation):

$$[af(x) + bg(x)]' = af'(x) + bf'(x)$$

Product rule:

$$(uv)' = u'v + uv'$$

Example:

$$\frac{d}{d\beta}\left(\frac{1}{Q}\frac{dQ}{d\beta}\right) = \frac{1}{Q}\frac{d^2Q}{d\beta^2} - \left(\frac{1}{Q}\frac{dQ}{d\beta}\right)^2$$

Chain rule:

$$\frac{dy}{dx} = \frac{dy}{du}\frac{du}{dx}$$

Example:

$$\frac{d}{d\beta}e^{-\beta E_i} = -E_i e^{-\beta E_i}$$

$$\frac{d\ln Q}{d\beta} = \frac{1}{Q}\frac{dQ}{d\beta}$$

dy/dx and dx/dy (if x-y is 1-1 mapping)

$$\frac{dx}{dy} = \frac{1}{dy/dx}$$

Application:

$$f'(a)=0 \Rightarrow f(x) \text{ might have a local extremum (min or max) at } x=a, \text{ check } f''(a)$$
$$f'(a)=0, f''(a)>0 \Rightarrow \text{a local min at } x=a$$
$$f'(a)=0, f''(a)<0 \Rightarrow \text{a local max at } x=a$$
$$f'(a)=f''(a)=0 \Rightarrow \text{very likely an inflection point}$$

2. Partial Derivatives

Notation: if $f=f(x,y)$,

$$f_x=\frac{\partial f}{\partial x}=\left(\frac{\partial f}{\partial x}\right)_y$$
$$f_{xy}=\frac{\partial}{\partial y}\left(\frac{\partial f}{\partial x}\right)$$
$$f_{yx}=\frac{\partial}{\partial x}\left(\frac{\partial f}{\partial y}\right)$$

In general, f_{xy} and f_{yx} may not be equal. However, in most case, functions are "well behaved", $f_{xy}=f_{yx}$

Chain rules

If $u=u(x,y),\ x=x(t),\ y=y(t)$

$$\frac{du}{dt}=\frac{\partial u}{\partial x}\frac{dx}{dt}+\frac{\partial u}{\partial y}\frac{dy}{dt}$$

3. Random Variables (r.v.) and Probability

Suppose that an experiment has n possible outcomes E_1,..., E_n, the probability of any particular outcome is

$$P(E_j)=\lim_{N\to\infty}\frac{N_j}{N}$$

where N_j is the number of times the event j occurs in N trials.

Normalization

$$\sum_{j=1}^{N}P(E_j)=1$$

Probability of a union ($\cup$, $+$: union, or; $\cap$ or comma: intersection, and):

$$P(A+B)=P(A)+P(B)-P(A,B)$$

$P(A,B)$: joint probability (the probability that both A and B occur)

$P(A+B)$: the probability of either A or B occurs.

If A and B are mutually exclusive events,

$$P(A+B)=P(A)+P(B)$$

$P(A|B)$: conditional probability (the probability that A occurs, given than B has occurred).

Joint probability = (marginal) probability $\times$ conditional probability:

$$P(A,B) = P(A)P(B|A) = P(B)P(A|B)$$

Bayes's formula or Bayes's theorem:

$$P(B|A) = \frac{P(A|B)P(B)}{P(A)}$$

If A occurs only if one of the set of mutually exclusive and exhaustive events B_1, B_2, ..., B_n occurs.

$$P(B_j|A) = \frac{P(A|B_j)P(B_j)}{P(A)} = \frac{P(A|B_j)P(B_j)}{\sum_{i=1}^{n} P(A|B_i)P(B_i)}$$

If A and B are independent, namely $P(B|A) = P(B)$ and $P(A|B) = P(A)$, we have

$$P(A,B) = P(A)P(B)$$

Example: in a hypothetical country, 40% are men and 60% are women. Among men, 30% are over 6'. Among women, 10% are over 6'.

- How likely to find a man of over 6' in this country?
- How likely to find a woman of over 6' in this country?
- If you find a person of over 6', how likely is it a women?

Random variable (r.v.): a rule or a formula assigns numerical values to each of the events in an experiment.

Example: toss a die, and the number on each face is a r.v.. Let's denote it as X, so $X = 1, 2, \ldots, 6$. In this example, X is a discrete r.v. Intuition tells us that $P(X = 1) = \ldots = P(X = 6) = 1/6$.

$P(X = X_j)$ or simply $P(X_j)$ is the probability distribution.

Expectation value (or mean value, average) of X:

$$\mu = E[X] = \langle X \rangle = \sum_{j=1}^{n} X_j P(X_j)$$

Continuous r.v.

r.v. takes continuous values. We use x.

Define probability density function (PDF), $f(x)$:

Normalization:

$$\int_{-\infty}^{\infty} f(x)dx = 1$$

Probability of having x between a and b:

$$P(a \le x \le b) = \int_a^b f(x)dx$$

Expectation value ($g(x)$ is a function of x):

$$\mu = \langle x \rangle = \int_{-\infty}^{\infty} x f(x)dx$$

$$\langle x^2 \rangle = \int_{-\infty}^{\infty} x^2 f(x)dx$$

Variance:

$$\sigma^2 = \left\langle (x-\mu)^2 \right\rangle = \left\langle x^2 \right\rangle - \langle x \rangle^2$$

If r.v. x and y are independent (or uncorrelated),

$$\langle xy \rangle = \langle x \rangle \langle y \rangle$$

(However, $\langle xy \rangle = \langle x \rangle \langle y \rangle$ does not necessarily imply the independence of x and y, e.g., $y = x^2$; but for jointly Gaussian r.v., it does.)

Gaussian distribution (or normal distribution):

Gaussian PDF:

$$f(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-(x-\mu)^2/2\sigma^2}$$

Notation: $x \sim N(\mu, \sigma^2)$.

Standard normal distribution: $N(0,1)$

4. Regression and Correlation

x: independent variable, may be measured without appreciable error.

Y: dependent variable, which is a function of X.

Regression analysis: consider Y to be a r.v.; y is the mean value of Y; goal: $y = y(x)$;

Linear regression: $y = a + bx$, where b is called the regression coefficient.

Method of least squares (least squares fit)

$$a = \langle y \rangle - b \langle x \rangle$$
$$b = \frac{\langle xy \rangle - \langle x \rangle \langle y \rangle}{\left\langle x^2 \right\rangle - \langle x \rangle^2}$$

where

$$\langle x \rangle = \frac{1}{n} \sum_{i=1}^{n} x_i, \ etc$$

Define variance of x, s_x^2, and covariance of the sample, s_{xy}

$$s_x^2 = \frac{1}{n-1} \sum_{i=1}^{n} (x_i - \langle x \rangle)^2$$
$$s_{xy} = \frac{1}{n-1} \sum_{i=1}^{n} (x_i - \langle x \rangle)(y_i - \langle y \rangle)$$

Therefore,

$$b = \frac{s_{xy}}{s_x^2}$$

The values of a and b obtained from the equations above turn out to be the maximum likelihood estimators of the linear regression parameters.

We assume that variable x can be measured without appreciable error, and we call it the independent variable. The other variable, Y, which is a function of x, is subject to some imprecision or uncertainty, and is considered to be a r.v.

Correlation analysis: both X and Y are r.v.'s.

Sample **(linear)** correlation coefficient of the sample:

$$r = \frac{s_{xy}}{s_x s_y}$$

Note that $s_x > 0$, $s_y > 0$, but s_{xy} can be positive or negative.

Often we report:

$$r^2 = \frac{s_{xy}^2}{s_x^2 s_y^2} \in [0, 1]$$

r is a measure of the **linear** correlation between x_i and y_i.

If $r = 0$, x_i and y_i are uncorrelated.

Note that population correlation is defined in a similar way.

Uncorrelatedness and Independence

Independence: $h(x, y) = f(x)g(y) \Rightarrow \langle xy \rangle = \langle x \rangle \langle y \rangle$

Uncorrelatedness: $\langle xy \rangle = \langle x \rangle \langle y \rangle$

Independence implies uncorrelatedness; however, uncorrelatedness does not necessarily imply independence. (e.g., $Y = X^2$)

If X and Y are jointly normally distributed, independence = uncorrelatedness.

Note: If x_i and y_i do not indicate there's at least somewhat of a linear relationship, the correlation coefficient doesn't mean much. (e.g., $Y = X^2$)

5. Matrix

Check sections 9.3, 10.1, 10.2 and 10.5 of McQuarrie's book on your own