M1J1 Summary Notes

JMC Year 1, 2017/2018 syllabus

Fawaz Shah (original notes by Dr Berkshire and Dr Lawn)

The structure of this document is split in two, since the two parts were taught by different lecturers.

Note that the exam will probably require you to PROVE any stated theorems, so you should refer back to the original notes for the proofs.

Boxes cover content in more detail. Titles of some theorems are given in italics.

Contents

Part I.
Applied Methods

1. Definitions

Order (of derivative) An n^{th} derivative has order n.

Order (of ODE) The order of the highest derivative present in an ODE.

Degree (of ODE) The highest power to which a term is raised in an ODE (excluding fractional powers).

Linear (ODE) An ODE which has no terms raised to more than the 1^{st} power, and with no y, x or other derivative terms multiplied by each other.

System of diff. equations A set of simultaneous equations of derivatives, where derivatives of y, x etc. are given w.r.t. a parameter t

Order (of system) The order of the highest derivative present in the system.

Degree (of system) The highest power to which a term is raised in an ODE (excluding fractional powers).

Linear (system) A system which has no terms raised to more than the 1^{st} power, and with no y or other derivative terms multiplied by each other.

Homogeneous (system) A system with no explicit functions of t present.

Order (difference equation) The order of a recurrence relation for u_n is the number of previous u_n terms it relates to.

Linear (difference equation) A recurrence relation for u_n is linear if it only contains u_n terms to the 1^{st} power, and has no u_n terms multiplied by each other.

Homogenous (difference equation) A recurrence relation for u_n is homogenous if there are no explicit functions of n present.

Forward difference operator Δ The forward difference operator is defined as:

$$\Delta U(n) = U(n+1) - U(n) \tag{1}$$

Newton-Raphson method The Newton-Raphson iterative formula to find x such that $f(x) = 0$ is:

$$x_{n+1} = x_n - \frac{f(x_n)}{f'(x_n)} \tag{2}$$

Note that this only works in specific cases.

Taylor series (multivariable) The Taylor series centered around (a, b) for a multivariable function $f(x, y)$ can be written as:

$$f(x,y) = f(a,b) + f'_x(a,b)(x-a) + f'_y(a,b)(y-b) \tag{3}$$

$$+\frac{1}{2!}(f''_{xx}(a,b)(x-a)^2 + 2f''_{xy}(a,b)(x-a)(y-b) \tag{4}$$

$$+f''_{yy}(a,b)(y-b)^2) + \dots \tag{5}$$

where:

$$f'_x = \frac{\partial f}{\partial x}, \quad f''_{xx} = \frac{\partial^2 f}{\partial x^2}, \quad f''_{xy} = \frac{\partial^2 f}{\partial x \partial y} = \frac{\partial^2 f}{\partial y \partial x} \dots \text{ etc.} \tag{6}$$

If we let $h = x - a$ and $k = y - b$ then we can also write the Taylor series as:

$$f(a+h, b+k) = f(a,b) + hf'_x(a,b) + kf'_y(a,b) \tag{7}$$

$$+\frac{1}{2!}(h^2 f''_{xx}(a,b) + 2hkf''_{xy}(a,b) + k^2 f''_{yy}(a,b)) + \dots \tag{8}$$

Exact (total) differential The exact differential of a multivariable function $f(x, y)$ is the equation:

$$df = dx\frac{\partial f}{\partial x} + dy\frac{\partial f}{\partial y} \tag{9}$$

Exact equation A differential equation of the form:

$$P(x,y) + Q(x,y)\frac{dy}{dx} = 0 \tag{10}$$

is called an exact equation iff:

$$\frac{\partial P}{\partial y} \equiv \frac{\partial Q}{\partial x} \tag{11}$$

Note that in this case the LHS of equation 10 forms an exact differential, which can be used to to find the general solution (explained later on).

Scalar field A function that associates a scalar value to every point in (3D) space.

Vector field A function that associates a vector to every point in (3D) space.

Partial differential operator ∇ Imagine we have a 3D vector field with standard basis vectors $\mathbf{i}, \mathbf{j}$ and $\mathbf{k}$. We define the operator ∇ to be:

$$\nabla = \mathbf{i}\frac{\partial}{\partial x} + \mathbf{j}\frac{\partial}{\partial y} + \mathbf{k}\frac{\partial}{\partial z} \tag{12}$$

grad The gradient (grad) of a scalar field produces a vector field. For a scalar field $\phi(x, y, z)$ we denote the grad as $\nabla\phi$. At any point in space, grad is defined as:

$$\nabla\phi = \frac{\partial\phi}{\partial x}\mathbf{i} + \frac{\partial\phi}{\partial y}\mathbf{j} + \frac{\partial\phi}{\partial z}\mathbf{k} \tag{13}$$

div The divergence (div) of a vector field produces a scalar field. For a vector field $\mathbf{u}(x, y, z)$ we denote the div as $\nabla \cdot \mathbf{u}$. At any point in space, div is defined as:

$$\nabla \cdot \mathbf{u} = \frac{\partial u_1}{\partial x} + \frac{\partial u_2}{\partial y} + \frac{\partial u_3}{\partial z} \tag{14}$$

where $\begin{pmatrix} u_1 \\ u_2 \\ u_3 \end{pmatrix}$ is the value of $\mathbf{u}$ at that point.

curl The curl of a vector field produces a vector field. For a vector field $\mathbf{u}(x, y, z)$ we denote the curl as $\nabla \times \mathbf{u}$. At any point in space, curl is defined as:

$$\nabla \times \mathbf{u} = \det\begin{pmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ \frac{\partial}{\partial x} & \frac{\partial}{\partial y} & \frac{\partial}{\partial z} \\ u_1 & u_2 & u_3 \end{pmatrix} \tag{15}$$

$$\tag{16}$$

$$= (\frac{\partial u_3}{\partial y} - \frac{\partial u_2}{\partial z})\mathbf{i} - (\frac{\partial u_3}{\partial x} - \frac{\partial u_1}{\partial z})\mathbf{j} + (\frac{\partial u_2}{\partial x} - \frac{\partial u_1}{\partial y})\mathbf{k} \tag{17}$$

where $\begin{pmatrix} u_1 \\ u_2 \\ u_3 \end{pmatrix}$ is the value of $\mathbf{u}$ at that point.

2. 2nd order ODEs

2.1. Special case - y missing

If we can write the 2^{nd} derivative in the form:

$$\frac{d^2y}{dx^2} = f(x, \frac{dy}{dx}) \tag{18}$$

(i.e. no y terms present), then we can make a substitution. Let $P = \frac{dy}{dx}$. This means $\frac{d^2y}{dx^2} = \frac{dP}{dx}$, the refore we have:

$$\frac{dP}{dx} = f(x, P) \tag{19}$$

This is 1st order w.r.t P and can be solved by appropriate 1st order methods.

2.2. Special case - x missing

If we can write the 2^{nd} derivative as:

$$\frac{d^2y}{dx^2} = f(y, \frac{dy}{dx}) \tag{20}$$

(i.e. no x terms present), then we can make the same substitution. Let $P = \frac{dy}{dx}$. This means $\frac{d^2y}{dx^2} = \frac{dP}{dx}$, therefore we have:

$$\frac{dP}{dx} = f(y, P) \tag{21}$$

However, this is not yet a 1st order equation since the derivative is w.r.t. x, but we only have y terms on the RHS.

DIFFERENT TO LAST TIME: we must rewrite $\frac{dP}{dx}$ as a derivative with respect to y. Luckily, we can see that:

$$\frac{dP}{dx} = \frac{dP}{dy}\frac{dy}{dx} = P\frac{dP}{dy} \tag{22}$$

Therefore:

$$P\frac{dP}{dy} = f(y, P) \tag{23}$$

This is 1st order w.r.t P and can be solved by appropriate 1st order methods.

2.3. Finding the CF

The general solution (GS) of a 2nd order ODE can be expressed as the sum of two other functions, called the **complementary function** (CF) and the **particular integral** (PI).

$$y_{GS} = y_{CF} + y_{PI} \tag{24}$$

A 2nd order ODE will usually be presented to us in the form:

$$a\frac{d^2y}{dx^2} + b\frac{dy}{dx} + c = f(x) \tag{25}$$

It can be shown that the CF can be calculated from the LHS of the above equation. We write down the **auxiliary equation**, which is simply the equation:

$$a\lambda^2 + b\lambda + c = 0 \tag{26}$$

using a, b, c from above. Solving this gives us two values, λ_1 and λ_2.

Case 1: $\lambda_1 \neq \lambda_2$, both real

We can express the CF as:

$$y_{CF} = A_1 e^{\lambda_1 x} + A_2 e^{\lambda_2 x} \tag{27}$$

where A_1 and A_2 are arbitrary constants.

Case 2: $\lambda_1 = \lambda_2$, both real

Same as above, but we stick an x in front of one of the clashing parts of the solution.

$$y_{CF} = A_1 e^{\lambda x} + A_2 x e^{\lambda x} \tag{28}$$

Case 3: λ_1, λ_2 are complex

If the auxiliary equation has complex roots, λ_1 and λ_2 will be complex conjugates. The CF can be expressed as:

$$\begin{aligned} y_{CF} &= A_1 e^{(a+bi)x} + A_2 e^{(a-bi)x} \\ &= e^a(A_1 e^{i(bx)} + A_2 e^{-i(bx)}) \\ &= e^a(C_1 cos(bx) + C_2 sin(bx)) \end{aligned} \tag{29}$$

where $C_1 = A_1 + A_2$ and $C_2 = (A_1 - A_2)i$. Note that even though A_1 and A_2 may have been complex, C_1 and C_2 are necessarily real.

2.4. Finding the PI

The particular integral is *any function y_{PI} that satisfies the ENTIRE differential equation.* The particular integral can be calculated depending on the form of the RHS of equation 25. We will refer to the RHS as simply $f(x)$ and the particular integral (as before) as y_{PI}. We can follow some basic rules:

Case 1: $f(x)$ is a polynomial

Try setting y_{PI} as a general polynomial of the same degree. e.g. if $f(x)$ is a quadratic, try setting $y_{PI} = ax^2 + bx + c$ and substituting into the ODE. We will solve for a, b, c, and this will give us y_{PI}.

Case 2: $f(x)$ is a multiple of e^{bx}, e^{bx} NOT in CF

Choose $y_{PI} = Ae^{bx}$ for some real number A.

Case 3: $f(x)$ is a multiple of e^{bx}, e^{bx} IS in CF

We now have a clash between the PI and the CF. We can try $y_{PI} = Axe^{bx}$, i.e. sticking an x in the PI to avoid the clash. If this doesn't work, we can choose $y_{PI} = A(x)e^{bx}$ for some real FUNCTION A (A will be a polynomial of the same degree as the order of the differential equation). Remember to use the CHAIN RULE to differentiate A this time.

At the end remove any clashing terms, i.e. terms of the form $Be^{\lambda x}$ where $e^{\lambda x}$ is already present in the CF. Other terms with more x's included are allowed, e.g. $xe^{\lambda x}$ would not count as a clashing term.

Case 4: $f(x) = A(x)e^{bx}$ where $A(x)$ is a polynomial

Choose $y_{PI} = C(x)e^{bx}$ for some polynomial $C(x)$.

Case 5: $f(x)$ is trigonometric (e.g. sin, cos, sinh etc.)

Look for a pattern in $f(x)$. A good tip for an $f(x)$ with only sines/cosines is to use $y_{PI} = A\cos(x) + B\sin(x)$ and solve for A and B. A similar story for sinh and cosh. CAUTION: sinh, cosh and tanh are actually exponential functions in disguise, so make sure they do not clash with any $e^{\lambda x}$ terms in the CF.

Other cases

If $f(x)$ has a term of the form $e^x\cos(x)$ or $e^x\sin(x)$ then we can rewrite it as the real/imaginary part of a complex function (in this case $e^{(1+i)x}$ would be appropriate, since it expands to $e^x(\cos(x) + i\sin(x))$.

If $f(x)$ is more complicated, we may have to be imaginative with the choice of y_{PI}. e.g. for $f(x) = Ae^{ax} + Be^{bx}$ we could choose $y_{PI} = Ce^{ax} + De^{bx}$ for some constants C, D. Again be careful of terms that clash with the CF.

3. Solving systems of differential equations

A homogeneous 1st order system of equations can be written as:

$$\begin{aligned} \frac{dx}{dt} &= F(x,y) \\ \frac{dy}{dt} &= G(x,y) \end{aligned} \tag{30}$$

Let us choose an example coupled system:

$$\begin{aligned} \frac{dx}{dt} &= ax + by \\ \frac{dy}{dt} &= cx + dy \end{aligned} \tag{31}$$

We can rewrite this in matrix form:

$$\frac{d}{dt}\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} a & b \\ c & d \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix} \tag{32}$$

The system is now of the form

$$\frac{d}{dt}v = Mv \tag{33}$$

If we set $v = Ve^{\lambda t}$, where V is a constant vector independent of x, y or t, then we get

$$\begin{aligned} \lambda V &= MV \\ (M - \lambda I_n)V &= 0_v \\ \det(M - \lambda I_n) &= 0 \end{aligned} \tag{34}$$

Predictably, we find two eigenvalues λ_1, λ_2 and (any) two eigenvectors v_1, v_2. The solution to the system is given by:

$$\begin{pmatrix} x \\ y \end{pmatrix} = A_1 v_1 e^{\lambda_1 t} + A_2 v_2 e^{\lambda_2 t} \tag{35}$$

The dimension of the eigenvectors will always match the number of variables being dealt with, for example a possible scenario is:

$$\begin{pmatrix} x \\ y \end{pmatrix} = A_1 \begin{pmatrix} 3 \\ -5 \end{pmatrix} e^{-3t} + A_2 \begin{pmatrix} 7 \\ -2 \end{pmatrix} e^{2t} \tag{36}$$

The values of the individual derivatives can be found by reading off the rows of the matrices.

$$\begin{aligned} x &= 3A_1e^{-3t} + 7A_2e^{2t} \\ y &= -5A_1e^{-3t} + -2A_2e^{2t} \end{aligned} \tag{37}$$

Complex eigenvalues

If the eigenvalues turn out to be complex conjugates, the solution can be written as:

$$\begin{pmatrix} x \\ y \end{pmatrix} = A_1v_1e^{(a+bi)t} + A_2v_2e^{(a-bi)t} \tag{38}$$

(Note that A_1 and A_2 may be complex). We can do some rearranging like before to tidy up the solution:

$$\begin{aligned} \begin{pmatrix} x \\ y \end{pmatrix} &= A_1v_1e^{(a+bi)t} + A_2v_2e^{(a-bi)t} \\ &= e^a(A_1v_1e^{i(bt)} + A_2v_2e^{-i(bt)}) \\ &= e^a(C_1\cos(bt) + C_2\sin(bt)) \end{aligned} \tag{39}$$

where $C_1 = A_1v_1 + A_2v_2$ and $C_2 = (A_1v_1 - A_2v_2)i$. Note that C_1 and C_2 are vectors.

3.1. Phase portraits

To draw a phase portrait, we can follow some algorithmic steps:

1. Sketch the lines of the eigenvectors in the x-y plane
2. Draw arrows on the eigenvectors which go outwards/inwards, depending on whether the corresponding eigenvalues are positive/negative respectively
3. Sketch out other contour lines on the x-y plane with arrows following the pattern of the existing lines

4. Difference equations

Difference equations are very similar to differential equations, both in their construction and in their solution. Terms are given as $U(n), U(n+1)$ etc. instead of $\frac{dy}{dx}, \frac{d^2y}{dx^2}$. We can reapply definitions such as **auxiliary equation**, **complementary function** and **particular integral**, and once again the general solution is defined by:

$$U(n)_{GS} = U(n)_{CF} + U(n)_{PI} \tag{40}$$

4.1. Finding the CF

For a general recurrence relation:

$$aU(n+2) + bU(n+1) + cU(n) = f(n) \tag{41}$$

we can set up the auxiliary equation:

$$a\lambda^2 + b\lambda + c = 0 \tag{42}$$

where the solutions to the quadratic are given by λ_1 and λ_2.

Case 1: $\lambda_1 \neq \lambda_2$

$$U(n)_{CF} = A_1(\lambda_1)^n + A_2(\lambda_2)^n \tag{43}$$

for arbitrary constants A_1, A_2.

Case 2: $\lambda_1 = \lambda_2$

We stick an n in front of one of the clashing terms:

$$U(n)_{CF} = A_1(\lambda)^n + A_2 n(\lambda)^n \tag{44}$$

4.2. Finding the PI

We refer to the RHS of equation 40 as $f(n)$.

Case 1: $f(n) = Cp^n$ **where** $p \neq \lambda_1, \lambda_2, (C$ **constant)**

Try a solution of the form $U(n)_{PI} = Ap^n$ for some constant A.

Case 2: $f(n) = Cp^n$ **where** $p = \lambda_1$ **or** $p = \lambda_2, (C$ **constant)**

We can try $U(n)_{PI} = Anp^n$ for some constant A. Alternatively, try a solution of the form $U(n)_{PI} = A(n)p^n$ where $A(n)$ is a polynomial with the same degree as the order of the recurrence relation. Remember to remove any clashing terms at the end.

Case 3: $f(n)$ **is a polynomial of degree** n

Choose $U(n)_{PI}$ as a general polynomial of degree n.

Case 4: $f(n)$ **is (polynomial in** $n)p^n$

Choose $U(n)_{PI}$ as (polynomial of degree $n)p^n$.

Δ (forward difference operator)

The Δ operator is defined as:

$$
\begin{aligned}
\Delta U(n) &= U(n+1) - U(n) && (45)\\
\Delta^2 U(n) &= (U(n+2) - U(n+1)) - (U(n+1) - U(n)) && (46)\\
&\vdots && (47)\\
&\text{etc.} && (48)
\end{aligned}
$$

Note that:

$$
\begin{aligned}
\Delta n^k &= (n+1)^k - n^k && (49)\\
\Delta(\text{polynomial of degree k}) &= \text{polynomial of degree } k-1 && (50)
\end{aligned}
$$

4.3. Difference tables

We can construct difference tables, where every entry in a row corresponds to the difference between two terms in the row above:

n	1	2	3	4	5	6	7	...
$R(n)$	1	2	4	8	16	31	57	...
$\Delta R(n)$	1	2	4	8	15	26		
$\Delta^2 R(n)$	1	2	4	7	11			
$\Delta^3 R(n)$	1	2	3	4				
$\Delta^4 R(n)$	1	1	1					
$\Delta^5 R(n)$	0	0						

We say the above sequence $R(n)$ is quartic, since the $\Delta^4 R(n)$ terms are constant. To calculate the next number in the sequence $R(n)$, we can add another 0 to the Δ^5 row and follow the pattern upwards.

5. Fixed point iteration

Fixed-point iteration is a method of finding an approximate solution to $f(x) = 0$ for some function f. We rewrite $f(x) = 0$ as an iterative formula:

$$x_{n+1} = F(x_n) \tag{51}$$

for some function F. Let a denote the value that x_n tends towards. The associated **fixed point** of this formula is the point $(a, 0)$. The fixed point is also the point of intersection between the lines $y = x$ and $y = F(x)$. We can find the fixed point by noting that, at the fixed point:

$$a = F(a) \tag{52}$$

We let the error term ϵ_n be defined as:

$$\epsilon_n = a - x_n \tag{53}$$

Case 1: $F'(a) \neq 0$

In this case we say the formula has linear convergence. The error in each iterate changes by a factor of $F'(a)$:

$$\epsilon_{n+1} \approx \epsilon_n F'(a) \tag{54}$$

Upon further expansion we get:

$$\begin{aligned} \epsilon_{n+1} &\approx \epsilon_n F'(a) && (55) \\ &\approx \epsilon_{n-1} F'(a)^2 && (56) \\ &\vdots && (57) \\ &\approx \epsilon_0 F'(a)^{n+1} && (58) \end{aligned}$$

It is easy to see that if $|F'(a)| < 1$, the error terms will get tend to 0 and the iterative formula will indeed converge to a. If $|F'(a)| > 1$, the error terms get bigger and the formula diverges from a.

Case 2: $F'(a) = 0, F''(a) \neq 0$

In this case the formula has quadratic convergence.

$$\epsilon_{n+1} \approx \frac{\epsilon_n^2}{2} F''(a) \tag{59}$$

In this case we can see that for convergence, we must have $|F''(a)| < 1$.

5.1. Newton-Raphson method

The Newton-Raphson method is a special case of fixed-point iteration. The Newton-Raphson iterative formula to find a solution to $f(x) = 0$ is given by:

$$x_{n+1} = x_n - \frac{f(x_n)}{f'(x_n)} \tag{60}$$

Newton-Raphson always converges quadratically, however it will only converge in specific situations (e.g. when there are no stationary points in between x_0 and a).

6. Partial differentation and vector calculus

For a multivariable function $z = f(x, y)$, we can define the partial derivatives (derivatives w.r.t one of the variables at a time). These are written as:

$$\frac{\partial z}{\partial x}, \frac{\partial z}{\partial y} \tag{61}$$

We can define second-order derivatives and so on:

$$\frac{\partial^2 z}{\partial x^2} = \frac{\partial}{\partial x}(\frac{\partial z}{\partial x}) \tag{62}$$

$$\tag{63}$$

$$\frac{\partial^2 z}{\partial x \partial y} = \frac{\partial}{\partial x}(\frac{\partial z}{\partial y}) \tag{64}$$

$$\tag{65}$$

$$\frac{\partial^2 z}{\partial y^2} = \frac{\partial}{\partial y}(\frac{\partial z}{\partial y}) \tag{66}$$

We can write the partial derivatives at (a, b) as $f'_x(a, b)$, $f''_{xx}(a, b)$, $f''_{xy}(a, b)$ etc. (just different notation).

A useful result is that for any continuous function $z = f(x, y)$:

$$\frac{\partial^2 z}{\partial x \partial y} \equiv \frac{\partial^2 z}{\partial y \partial x} \tag{67}$$

Another useful result is that:

$$\frac{\partial x}{\partial z} \equiv \frac{1}{(\frac{dz}{dx})} \tag{68}$$

but ONLY if the same variable is being kept constant in the expressions for $\frac{\partial z}{\partial x}$ and $\frac{\partial x}{\partial z}$.

We can define an expression called the exact (total) differential of z, which is:

$$dz = \frac{\partial z}{\partial x}dx + \frac{\partial z}{\partial y}dy \tag{69}$$

If we divide this equation by dt, we get the formula for a 'full' (i.e. non-partial) derivative of z. This is the multivariable equivalent of the 'chain rule':

$$\frac{dz}{dt} = \frac{\partial z}{\partial x}\frac{dx}{dt} + \frac{\partial z}{\partial y}\frac{dy}{dt} \tag{70}$$

In the above situation x and y are both parametric functions of t.

If we are asked to differentiate an implicit function, i.e. of the form $F(x, y, z) = 0$ where z cannot be isolated on one side, we can perform the equivalent of implicit differentiation:

$$dF = \frac{\partial F}{\partial x}dx + \frac{\partial F}{\partial y}dy + \frac{\partial F}{\partial z}dz = 0 \tag{71}$$

To get, for example, $\frac{\partial z}{\partial x}$, we can set $dy = 0$ and rearrange (since we assume y is constant when finding $\frac{\partial z}{\partial x}$).

6.1. Change of variables

Often we want to change variables, from $z = f(x, y)$ to $z' = g(r, \theta)$ for some g. We can write down the usual substitution equations:

$$x = r\cos(\theta) \tag{72}$$

$$y = r\sin(\theta) \tag{73}$$

$$r = (x^2 + y^2)^{\frac{1}{2}} \tag{74}$$

$$\theta = \arctan(\frac{y}{x}) \tag{75}$$

Note that different variables are being kept constant in each of the equations, e.g.

$$\frac{\partial x}{\partial r} = \cos(\theta) \quad \text{(keeping } \theta \text{ constant)} \tag{76}$$

$$\frac{\partial r}{\partial x} = \frac{1}{2} \cdot 2x \cdot (x^2 + y^2)^{-\frac{1}{2}} = \cos(\theta) \quad \text{(keeping } y \text{ constant)} \tag{77}$$

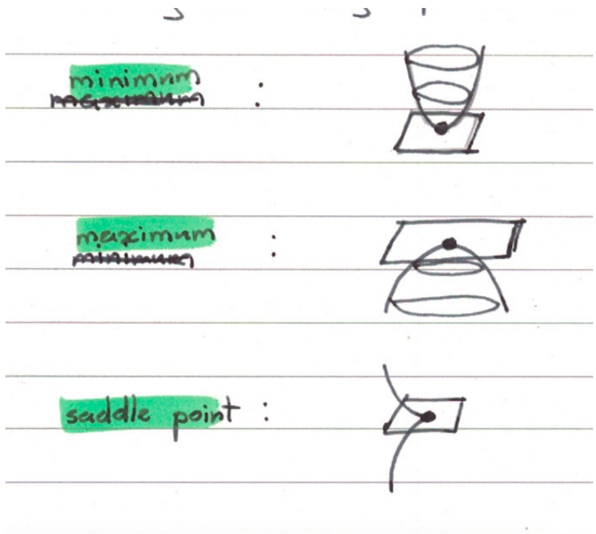
Therefore:

$$\frac{\partial x}{\partial r} \neq \frac{1}{\frac{\partial r}{\partial x}} \tag{78}$$

6.2. Stationary points

We can derive rigorously how to use the Taylor series of $z = f(x, y)$ to find and classify its stationary points (but i wont cos i dont have time lel).

There are three types of stationary points: maxima, minima, and saddle points. Any stationary point lies tangent to the horizontal plane (so $\frac{\partial z}{\partial x} = \frac{\partial z}{\partial y} = 0$).



The algorithmic method for finding and classifying stationary points is as follows:

1. Find the stationary points, by solving the simultaneous equations:

$$\frac{\partial z}{\partial x} = 0 \ , \quad \frac{\partial z}{\partial y} = 0 \tag{79}$$

2. At each stationary point, find the following:

$$A = \frac{\partial^2 z}{\partial x^2}, \ B = \frac{\partial^2 z}{\partial x \partial y}, \ C = \frac{\partial^2 z}{\partial y^2} \tag{80}$$

3. Calculate $B^2 - AC$ for each stationary point:

$$B^2 - AC > 0 \quad \Rightarrow \quad \text{SADDLE POINT} \tag{81}$$
$$B^2 - AC < 0, \ A > 0 \quad \Rightarrow \quad \text{MINIMUM} \tag{82}$$
$$B^2 - AC < 0, \ A < 0 \quad \Rightarrow \quad \text{MAXIMUM} \tag{83}$$
$$\tag{84}$$

6.3. Exact equations

Let us say we have an ODE of the form:

$$P(x,y) + Q(x,y)\frac{dy}{dx} = 0 \tag{85}$$

(note the coefficients are multi-variable functions). This can be rewritten as:

$$P(x,y)dx + Q(x,y)dy = 0 \tag{86}$$

We can try the exact equations method. We say an equation is exact iff:

$$\frac{\partial P}{\partial y} \equiv \frac{\partial Q}{\partial x} \tag{87}$$

This simple condition implies some important results. It can be shown that an exact equation implies the LHS of equation 86 is an exact (total) differential). This means it can be written as df, where f is some function of x and y. But the equation of this total differential is:

$$df = \frac{\partial f}{\partial x}dx + \frac{\partial f}{\partial y}dy \tag{88}$$

Comparing to equation 86 we can note 3 things:

$$\begin{aligned} P(x,y) &= \frac{\partial f}{\partial x} \\ Q(x,y) &= \frac{\partial f}{\partial y} \\ df &= 0 \end{aligned} \tag{89}$$

We integrate $P(x,y)$ w.r.t x and $Q(x,y)$ w.r.t y and 'merge' the two expressions together (i.e. for any matching terms, write them down only once) to give us an expression for $f(x,y)$. Ignore constants of integration. $df = 0$ tells us that $f(x,y) = c$ by integration. Therefore the general solution is given by:

$$f(x,y) = c \tag{90}$$

for some arbitrary constant c.

6.4. grad, div and curl

All scalar fields have an associated vector quantity called grad.
All vector fields have an associated scalar quantity called div, and an associated vector quantity called curl.
The definitions of these quantities are given in the definitions section.

Part II.
Linear Algebra

7. Definitions

$\mathbb{R}^n$ The set of all column vectors with height n.

Zero vector A column vector with all n entries 0. Denoted by $\mathbf{0}_n$.

Standard basis vectors of $\mathbb{R}^n$ The set of column vectors with n entries, with a single entry as 1 and the rest 0.

Linear combination A linear combination of vectors $\mathbf{v}_1 ... \mathbf{v}_n$ is an expression:

$$\lambda_1 \mathbf{v}_1 + \lambda_2 \mathbf{v}_2 + ... + \lambda_n \mathbf{v}_n \tag{91}$$

where $\lambda_1 ... \lambda_n \in \mathbb{R}$.

Span The span of of a set of vectors is the set of all linear combinations of those vectors.

Dot product The dot product of $\mathbf{v} = \begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix}$ and $\mathbf{u} = \begin{pmatrix} u_1 \\ \vdots \\ u_n \end{pmatrix}$ is defined as:

$$\mathbf{v} \cdot \mathbf{u} = v_1 u_1 + v_2 u_2 + ... + v_n u_n \tag{92}$$

Norm The norm (aka length) of a vector $\mathbf{v} = \begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix}$ in $\mathbb{R}^n$ is defined by:

$$||\mathbf{v}|| = \sqrt{\sum_{i=1}^{n} v_i^2} \tag{93}$$

Unit vector Any $\mathbf{v}$ such that $||\mathbf{v}|| = 1$.

Zero matrix An $m \times n$ matrix with all entries 0. Denoted by $\mathbf{0}_{m \times n}$.

Transpose The transpose of a matrix $\mathbf{A}$ is the result of flipping all rows with columns and vice versa. Denoted by $\mathbf{A}^T$.

(a_{ij}) **notation** For an $m \times n$ matrix $\mathbf{A}$, we can say $\mathbf{A} = (a_{ij})$ if we represent the matrix as:

$$\begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1j} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2j} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots & & \vdots \\ a_{i1} & a_{i2} & \cdots & a_{ij} & \cdots & a_{in} \\ \vdots & \vdots & & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mj} & \cdots & a_{mn} \end{pmatrix} \tag{94}$$

i denotes the row number and j denotes the column number. Hence we have $1 \leq i \leq m$ and $1 \leq j \leq n$.

Matrix multiplication Let $\mathbf{A} = (a_{ij})$ be an $m \times n$ matrix. Let $\mathbf{B} = (b_{jk})$ be an $n \times l$ matrix. We can define the matrix product $\mathbf{C} = \mathbf{AB}$ as an $m \times l$ matrix:

$$(c_{ik}) = \sum_{j=1}^{n} a_{ij} b_{jk} \tag{95}$$

Note that a condition on matrix multiplication being possible is:

$$\text{no. of columns in } \mathbf{A} = \text{no. of rows in } \mathbf{B} \tag{96}$$

Note that in general, matrix multiplication is NOT commutative.

Leading diagonal Let $\mathbf{A} = (a_{ij})$. The leading diagonal is made up of all entries of the form a_{ii}, e.g. a_{11}, a_{22}, etc.

Identity matrix A square $n \times n$ matrix where the leading diagonal is all 1s and the rest of the entries are 0. Denoted by I_n.

Linear equation A linear equation of variables $x_1 ... x_n \in \mathbb{R}$ is given by:

$$\lambda_1 x_1 + \lambda_2 x_2 + ... + \lambda_n x_n = b \tag{97}$$

where $\lambda_1 ... \lambda_n, b \in R$.

System of linear equations A system of linear equations is a set of simultaneous linear equations.

Free variable A variable in a system of linear equations that can be set to any value, without affecting the validity of the solution.

Basic variable Any variable that is not a free variable (i.e. a solution only exists for particular values of this variable).

Augmented matrix For a system of linear equations represented as $\mathbf{Ax} = \mathbf{b}$, the augmented matrix is $(\mathbf{A} \mid \mathbf{b})$. This is the matrix $\mathbf{A}$ but with $\mathbf{b}$ added in as an extra column on the right.

Row operation There are 3 types of row operation:

- Swapping rows
- Multiplying a row by a scalar value
- Adding a multiple of one row to another

Leading entry The leading entry of a row is the first non-zero entry in that row.

Row echelon form (REF) A matrix is in row echelon form if:

- All non-zero rows are above rows of all zeroes
- Each leading entry is 1, and is in a column to the right of the leading entry in the row above
- All entries in the column below a leading entry are 0

Note that any square matrix in row echelon form is upper triangular.

Reduced row echelon form (RREF) A matrix is in reduced row echelon form if:

- The matrix is in reduced echelon form
- Each leading entry is the only non-zero entry in its column

Any square matrix in RREF is upper AND lower triangular (and hence a diagonal matrix).

Gaussian Elimination An algorithm for putting a matrix into REF or RREF form via row operations (explained in theorems section).

Upper triangular A matrix is upper triangular if $a_{ij} = 0$ for $i > j$.
i.e. only the top-right half contains non-zero entries.

Strictly upper triangular A matrix is strictly upper triangular if $a_{ij} = 0$ for $i \geq j$.

Lower triangular A matrix is lower triangular if $a_{ij} = 0$ for $i < j$.
i.e. only the bottom-left half contains non-zero entries.

Strictly lower triangular A matrix is strictly lower triangular if $a_{ij} = 0$ for $i \leq j$.

Diagonal matrix A diagonal matrix is such that $a_{ij} = 0$ if $i \neq j$.
i.e. Only the leading diagonal contains non-zero entries.

(i, j) **minor** The (i, j) minor of a matrix $\mathbf{A}$ is a submatrix obtained by removing a row i and a column j. Denoted by $\mathbf{A}_{ij}$.

Inverse The inverse of an $n \times n$ matrix $\mathbf{A}$ is such that:

$$\mathbf{A}^{-1}\mathbf{A} = \mathbf{A}\mathbf{A}^{-1} = \mathbf{I}_n \tag{98}$$

Note that this is one of the only occasions where matrix multiplication is commutative. Not all matrices have inverses.

Singular A matrix is called singular if it is non-invertible.

Determinant A specific number calculated from the entries in a matrix (see theorems). Used in computing inverses. Denoted by $\det(\mathbf{A})$.

Elementary matrix A matrix that differs from the identity matrix by one row operation.

Eigenvalue An eigenvalue of a matrix $\mathbf{A}$ is some $\lambda \in R$ that satisfies:

$$\mathbf{A}\mathbf{v} = \lambda\mathbf{v} \tag{99}$$

for some $\mathbf{v} \in \mathbb{R}^n$, called the eigenvector.

Eigenvalue An eigenvector of a matrix $\mathbf{A}$ is some $\mathbf{v} \in R^n$ that satisfies:

$$\mathbf{A}\mathbf{v} = \lambda\mathbf{v} \tag{100}$$

for some $\lambda \in \mathbb{R}$, called the eigenvalue.

Similar matrices Two matrices $\mathbf{A}$ and $\mathbf{B}$ are similar if there exists a matrix $\mathbf{P}$ such that:

$$\mathbf{A} = \mathbf{P}^{-1}\mathbf{B}\mathbf{P} \tag{101}$$

Diagonalizable A matrix is diagonalizable if it is similar to a diagonal matrix.

8. Theorems

Matrices can only be added if they have the same dimension.

The matrix product $\mathbf{AB}$ is only defined if number of columns in $\mathbf{A}$ = number of rows in $\mathbf{B}$.

If $\mathbf{A}$ is an $m \times n$ matrix, and $\mathbf{e_k}$ is the k^{th} standard basis vector of $\mathbb{R}^n$, then $\mathbf{Ae_k}$ gives the column vector which is the k^{th} column of A.

Matrix multiplication is associative, distributive but NOT commutative.
There are special cases however when matrix multiplication is commutative:

- Multiplying two diagonal matrices
- Multiplying any matrix by the identity ($\mathbf{AI}_n = \mathbf{I}_n\mathbf{A} = \mathbf{A}$)
- Multiplying any matrix by its inverse ($\mathbf{AA}^{-1} = \mathbf{A}^{-1}\mathbf{A} = \mathbf{I}_n$)

8.1. Systems of linear equations

Let $\mathbf{Ax} = \mathbf{b}$ be a sytem of linear equations, with $(\mathbf{A} \mid \mathbf{b})$ the augmented matrix. Let $(\mathbf{A}' \mid \mathbf{b}')$ be the result of applying a single row operation to $(\mathbf{A} \mid \mathbf{b})$. $\mathbf{x}$ is a solution to $\mathbf{A}'\mathbf{x} = \mathbf{b}'$ iff it is a solution to $\mathbf{Ax} = \mathbf{b}$.

In other words, we can apply row operations to a system of linear equations without affecting the solution set.

Any system of linear equations either has 0, 1 or infinitely many solutions.

Gaussian Elimination
The process of Gaussian elimination can be used to reduce a matrix to REF or RREF. It is applied to augmented matrices to solve their corresponding systems of equations. For a system $\mathbf{Ax} = \mathbf{b}$ let the augmented matrix be $(\mathbf{A} \mid \mathbf{b}) = (a_{ij} \mid b_i)$:

1. Look at row 1, column 1 of $\mathbf{A}$. If $a_{11} = 0$, find a row i such that $a_{i1} \neq 0$ and swap row i with row 1.
2. Multiply row 1 by $\frac{1}{a_{11}}$. We have produced a leading 1 in that column.

3. Subtract multiples of row 1 from all other rows such that all other entries in column 1 are 0.
4. Repeat the process for row 2, column 2, and continue in general for row i, column i until all of matrix $\mathbf{A}$ has been spanned.
5. The matrix is now in REF form.

To get the matrix into RREF form, we must apply a few more steps:

6. Starting from the right-hand side column of $\mathbf{A}$, look at the row which contains the column's leading 1 (say row k). For every non-zero entry above the 1, subtract multiples of row k from the rows above to make all the entries 0.
7. Repeat for all columns of $\mathbf{A}$.

Each row of A should now contain a single 1 with the rest of the entries as 0. Reading off the modified values of column $\mathbf{b}$ gives us the values of the solution vector $\mathbf{x}$.

8.2. Determinants

Computing determinants
Calculating the determinant of a 2×2 matrix is as follows:

$$\det \begin{pmatrix} a & b \\ c & d \end{pmatrix} = ad - bc \tag{102}$$

Calculating the determinant of a 3×3 matrix $\mathbf{A}$ can be done by manipulating the minors of the first row:

$$\det \begin{pmatrix} a_{11} & a_{22} & a_{33} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix} = a_{11} \det(\mathbf{A}_{11}) - a_{12} \det(\mathbf{A}_{12}) + a_{13} \det(\mathbf{A}_{13}) \tag{103}$$

Note the alternating signs present in the above equation.
In general, for any $n \times n$ matrix $\mathbf{A}$ we have:

$$\det \mathbf{A} = \sum_{j=1}^{n} (-1)^{i+j} \det(\mathbf{A}_{ij}) \tag{104}$$

by expanding about any row i.

If two rows are swapped, the determinant is negated.
If a row is multiplied by a scalar, the determinant is also multiplied by the scalar.
If a multiple of one row is added to another, the determinant is unchanged.

For any matrices $\mathbf{A}, \mathbf{B}$:

$$\det(\mathbf{A}) = \det(\mathbf{A}^T) \tag{105}$$

$$\det(\mathbf{AB}) = \det(\mathbf{A})\det(\mathbf{B}) \tag{106}$$

8.3. Inverses

Computing inverses
For any invertible 2×2 matrix, its inverse is calculated as:

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix}^{-1} = \frac{1}{ad - bc}\begin{pmatrix} d & -b \\ -c & a \end{pmatrix} \tag{107}$$

For any invertible $n \times n$ matrix $\mathbf{A}$, calculating its inverse can be done with the following algorithm:

1. Write the augmented matrix $(\mathbf{A} \mid \mathbf{I}_n)$
2. Perform Gaussian elimination such that the left matrix $\mathbf{A}$ reduces into RREF. Since it is a square matrix, its RREF will be $\mathbf{I}_n$.
3. The right matrix (which started as $\mathbf{I}_n$) will transform into $\mathbf{A}^{-1}$

Any matrix $\mathbf{A}$ is singular iff $\det(\mathbf{A}) = 0$.

Any invertible matrix only has one inverse.

If $\mathbf{A}$ and $\mathbf{B}$ are invertible, then $\mathbf{AB}$ is invertible.
If either of $\mathbf{A}$ or $\mathbf{B}$ are singular, then $\mathbf{AB}$ is singular.

For any $m \times n$ matrix $\mathbf{A}$:

- If there exists a non-zero vector $\mathbf{v} \in \mathbb{R}^n$ such that $\mathbf{Av} = \mathbf{0}_m$ then $\mathbf{A}$ is singular
- If there exists a non-zero $n \times k$ matrix $\mathbf{B}$ such that $\mathbf{AB} = \mathbf{0}_{m\times k}$ then $\mathbf{A}$ is singular
- If $\mathbf{A}$ contains an all-zero column or an all-zero row, it is singular

If $\mathbf{A}$ is invertible and $\mathbf{A}'$ is obtained from $\mathbf{A}$ by applying a single row operation:

$$\mathbf{A} \text{ invertible} \Leftrightarrow \mathbf{A}' \text{ invertible} \tag{108}$$

For any matrix $\mathbf{A}$:

$$\mathbf{A} \text{ invertible} \Leftrightarrow \text{RREF of } \mathbf{A} \text{ invertible} \tag{109}$$

For any inverible matrices $\mathbf{A}, \mathbf{B}$:

$$(\mathbf{AB})^{-1} = \mathbf{B}^{-1}\mathbf{A}^{-1} \tag{110}$$

8.4. Eigenvalues and eigenvectors

Computing eigenvalues/eigenvectors
An $n \times n$ matrix $\mathbf{A}$ has an eigenvalue λ (and a corresponding eigenvector $\mathbf{v} \in \mathbb{R}^n$) if:

$$\mathbf{A}\mathbf{v} = \lambda\mathbf{v} \tag{111}$$

where $\mathbf{v} \neq \mathbf{0}_n$. By rearranging, we get:

$$(\mathbf{A} - \lambda\mathbf{I}_n)\mathbf{v} = \mathbf{0}_n \tag{112}$$

Therefore:

$$\lambda \text{ is an eigenvector} \Leftrightarrow \mathbf{A} - \lambda\mathbf{I}_n \text{ is singular} \tag{113}$$
$$\Leftrightarrow \det(\mathbf{A} - \lambda\mathbf{I}_n) = 0 \tag{114}$$

Equation 114 is called the characteristic equation. Solving it gives us all the eigenvalues. (Note there will be n solutions.)
To find eigenvectors, we plug the known eigenvalues into equation 112 and use the Gaussian elimination algorithm for solving $\mathbf{A}\mathbf{x} = \mathbf{b}$ to find $\mathbf{v}$.

HOWEVER the resulting augmented matrix will always seem to give us infinite solutions. Therefore there are infinitely many eigenvectors for a single eigenvalue. We can find the general form of the eigenvector and simply state an eigenvector of that form. e.g. if our augmented matrix can be reduced to:

$$\left(\begin{array}{cc|c} 1 & 2 & 0 \\ 1 & 2 & 0 \end{array}\right) \tag{115}$$

then if we let our solution $\mathbf{v} = \begin{pmatrix} v_1 \\ v_2 \end{pmatrix}$ we can see that the only constraint we have is $v_1 + 2v_2 = 0$. Therefore some example eigenvectors are:

$$\begin{pmatrix} 2 \\ -1 \end{pmatrix}, \begin{pmatrix} 4 \\ -2 \end{pmatrix}, \begin{pmatrix} 6 \\ -3 \end{pmatrix} \text{ etc...} \tag{116}$$

since they all fit this constraint. We need only state one of these eigenvectors.

Any $n \times n$ matrix has at most n eigenvalues.

8.5. Diagonalization

We say two matrices $\mathbf{A}$ and $\mathbf{B}$ are similar if there exists a matrix $\mathbf{P}$ such that:

$$\mathbf{A} = \mathbf{P}^{-1}\mathbf{B}\mathbf{P} \tag{117}$$

If $\mathbf{A}$ is similar to $\mathbf{B}$ then $\mathbf{B}$ is similar to $\mathbf{A}$.

Similar matrices have the same eigenvalues.

A matrix is called diagonalizable if it is similar to a diagonal matrix.

Let $\mathbf{v}_1 ... \mathbf{v}_n$ be eigenvectors of $\mathbf{A}$, with corresponding eigenvalues $\lambda_1 ... \lambda_n$.
Let $\mathbf{P} = (\mathbf{v}_1 \mid \mathbf{v}_2 \mid ... \mid \mathbf{v}_n)$.
If $\mathbf{P}$ is invertible then $\mathbf{A}$ is diagonalizable, with:

$$\mathbf{P}^{-1}\mathbf{A}\mathbf{P} = \mathrm{diag}(\lambda_1, \lambda_2 ... \lambda_n) \tag{118}$$

In other words, if we let $\mathbf{D} = \mathrm{diag}(\lambda_1, \lambda_2 ... \lambda_n)$:

$$\mathbf{A} = \mathbf{P}\mathbf{D}\mathbf{P}^{-1} \tag{119}$$

A matrix is only diagonalizable if all its eigenvectors are linearly independent (since only then will $\mathbf{P}$ be invertible).

Computing powers of a diagonalizable matrix $\mathbf{A}$ is easy, since:

$$\begin{aligned}
\mathbf{A}^n &= (\mathbf{P}\mathbf{D}\mathbf{P}^{-1})^n && (120)\\
&= (\mathbf{P}\mathbf{D}\mathbf{P}^{-1})(\mathbf{P}\mathbf{D}\mathbf{P}^{-1}) \; ... \; (\mathbf{P}\mathbf{D}\mathbf{P}^{-1}) && (121)\\
&= \mathbf{P}\mathbf{D}(\mathbf{P}^{-1}\mathbf{P})\mathbf{D}(\mathbf{P}^{-1} \; ... \; \mathbf{P})\mathbf{D}\mathbf{P}^{-1} && (122)\\
&= \mathbf{P}\mathbf{D}^n\mathbf{P}^{-1} && (123)
\end{aligned}$$

since all $\mathbf{P}^{-1}\mathbf{P}$ terms cancel out.

Also note that calculating $\mathbf{D}^n$ is easy, since it is a diagonal matrix:

$$\mathrm{diag}(\lambda_1, \lambda_2 ... \lambda_n)^k = \mathrm{diag}(\lambda_1^k, \lambda_2^k ... \lambda_n^k) \tag{124}$$