VECTOR CALCULUS - Fields, Vector Derivatives and Integrals

CHAIN RULE - for a function $f(x,y)$, $\delta f \approx \frac{\partial f}{\partial x}\delta x + \frac{\partial f}{\partial y}\delta y$

* To change variables to $u(x,y)$ and $v(x,y)$, find $\delta f/\delta u$ and $\delta f/\delta v$
* Total derivative - rate of change along certain curve in x,y plane
  → Take $\delta f/\delta x$, evaluate in general, then put $y = y(x)$

VECTOR FIELDS AND FIELD LINES

* For example in a fluid, field lines tangential to $\underline{v}$ everywhere, called streamlines.
* In general, lines tangential to field everywhere ⇒ field line $y(x)$ satisfies
  $\frac{dy}{dx} = \frac{u_y}{u_x}$ or in 3D $z(x,y)$ satisfies $\frac{\delta x}{u_x} = \frac{\delta y}{u_y} = \frac{\delta z}{u_z} = R$

GRADIENT - $\nabla f = \left[\frac{\partial f}{\partial x}, \frac{\partial f}{\partial y}, \frac{\partial f}{\partial z}\right]$ for scalar field f

* ∇f is a vector, points up steepest slope of f
* Directional Derivative - rate of change in direction $\hat{n}$: $\frac{df}{ds} = \hat{n}\cdot\nabla f$
  → works for vector field $\underline{u}$: $(\hat{n}\cdot\nabla)\underline{u} = (\hat{n}\cdot\nabla)u_x\underline{i} + (\hat{n}\cdot\nabla)u_y\underline{j} + (\hat{n}\cdot\nabla)u_z\underline{k}$
* ∇f is always normal to contours, along which $f = \text{const}$
* 3D Taylor expansion using grad: $f(\underline{r}+\delta\underline{r}) \approx f(\underline{r}) + (\delta\underline{r}\cdot\nabla)f + \frac{1}{2!}(\delta\underline{r}\cdot\nabla)^2 f + \dots$

DIVERGENCE AND CURL - for vector field $\underline{u}$, $\nabla\cdot\underline{u}$ and $\nabla\times\underline{u}$

* Application: fluid flow and mass conservation
  → Consider a cube $\delta V = \delta x\delta y\delta z$. On one face flux $= \underline{u}\cdot\delta\underline{A}$
  → For pair of faces in $\delta z\delta y$ plane, net flux $= \delta y\delta z\left(\delta x\frac{\partial u_1}{\partial x}\right) = \delta V\frac{\partial u_1}{\partial x}$
  → Add in other pairs of faces ⇒ net flow out of volume is: $\delta V\left[\frac{\partial u_1}{\partial x} + \frac{\partial u_2}{\partial y} + \frac{\partial u_3}{\partial z}\right] = \delta V\,\nabla\cdot\underline{u}$
  → But if incompressible, net flux $= 0$ so $\nabla\cdot\underline{u} = 0$
* ~~[illegible]~~ Product Rule: $\nabla\times(\phi\underline{u}) = \phi(\nabla\times\underline{u}) + (\nabla\phi)\times\underline{u}$
  → Loads of other identities on Databook
* Second Derivatives: Remember $\nabla^2 = \nabla\cdot\nabla$
* Solenoidal (or incompressible) if $\nabla\cdot\underline{B} = 0$
* Irrotational (or conservative) if $\nabla\times\underline{E} = 0$ ⇒ $\underline{E} = \nabla\phi$ a scalar potential

VOLUME AND AREA INTEGRALS, JACOBIAN

* For area integrals, just integrate twice with x and y, either order.
* Mean Values: $\bar{\rho} = \text{mass}/\text{volume} = \iiint \rho\, dV / \iiint dV$
* Change of Variable: from x, y to $u(x,y)$, $v(x,y)$
  → Want area of $dxdy$ in terms of du and dv. Use vectors: $A = |\underline{a}\times\underline{b}|$
  → Form $\underline{a}$ by moving short δu in direction $v = \text{const.}$: $\underline{a} = \delta\underline{x} = \frac{\partial \underline{x}}{\partial u}\delta u$
  → Similarly $\underline{b}$ is δv in direction $u = \text{const}$.
  → So $A = \left|\left[\frac{\partial x}{\partial u}\delta u, \frac{\partial y}{\partial u}\delta u, 0\right]\times\left[\frac{\partial x}{\partial v}\delta v, \frac{\partial y}{\partial v}\delta v, 0\right]\right| = \left\|\begin{matrix}\frac{\partial x}{\partial u} & \frac{\partial y}{\partial u}\\ \frac{\partial x}{\partial v} & \frac{\partial y}{\partial v}\end{matrix}\right\|\delta u\delta v = \left|\frac{\partial(x,y)}{\partial(u,v)}\right|\delta u\delta v$
  → Analogous result for 3D
* Useful Jacobians to remember ($dxdy = |J|dudv$)
  → Plane polar coordinates: r
  → Cylindrical polars: ρ
  → Spherical polars: $r^2\sin\theta$

VECTOR CALCULUS - Gauss's and Stoke's Theorems

GAUSS'S THEOREM - $\iint_S \underline{u}\cdot d\underline{A} = \iiint_V \nabla\cdot\underline{u}\, dV$

* Take a small volume element $\delta V = \delta x \delta y \delta z$, then total net flux $= \nabla\cdot\underline{u}\,\delta V$
  → This is the sum as the flux out as defined by surface integral.
  → Now ~~sum~~ ~~all~~ if $\nabla\cdot\underline{u}$ constant over element, $\iint_{\text{element}} \underline{u}\cdot d\underline{A} = \nabla\cdot\underline{u}\,\delta V = \iiint_{\text{element}} \nabla\cdot\underline{u}\, dV$
  → Sum over whole volume, internal fluxes cancel.

STOKE'S THEOREM - $\iint_S \nabla\times\underline{u}\cdot d\underline{A} = \oint_C \underline{u}\cdot d\underline{l}$

* Right hand rule - direction of line integral ⇒ normal direction for $d\underline{A}$
* Take a small element $dx\,dy$, neglect curvature. $d\underline{A} = dx\,dy\,\underline{k}$
* ~~[illegible]~~ With vector field $\underline{u} = [u, v, w]$, $\nabla\times\underline{u} = \underline{k}\left(\frac{\partial v}{\partial x} - \frac{\partial u}{\partial y}\right) + (\underline{i}, \underline{j}$ components$)$
  → So that $\iint_S \nabla\times\underline{u}\cdot d\underline{A} = \int_{x=0}^{X}\int_{y=0}^{Y} \left(\frac{\partial v}{\partial x} - \frac{\partial u}{\partial y}\right) dx\,dy$

$$= \int_0^Y\int_0^X \frac{\partial v}{\partial x}\,dx\,dy - \int_0^X\int_0^Y \frac{\partial u}{\partial y}\,dy\,dx = \int_0^Y (v(X,y) - v(0,y))\,dy - \int_0^X (u(x,Y) - u(x,0))\,dx$$

$$= \int_0^X u(x,0)\,dx + \int_0^Y v(X,y)\,dy + \int_X^0 u(x,Y)\,dx + \int_Y^0 v(0,y)\,dy$$

* This is the line integral $\oint \underline{u}\cdot d\underline{l}$!

COORDINATE-FREE DEFINITIONS - invert above laws

* Gauss's Theorem ⇒ $\nabla\cdot\underline{u} = \lim_{\delta V\to 0} \frac{1}{\delta V}\iint_S \underline{u}\cdot d\underline{A}$
* Stoke's Theorem ⇒ $\nabla\times\underline{u}\cdot d\underline{A} \approx \oint_C \underline{u}\cdot d\underline{l}$ for small dA
  → Then choose different orientations for $\delta\underline{A}$ to find each component

MULTIPLE SURFACES - e.g. 'holes' in surfaces or volumes

* Add two new curves spanning gap between curves, in opposite directions
* The surface is effectively surrounded by a continuous curve
  → The spanning curves cancel each other when Stoke's is applied
  → $\iint_S \nabla\times\underline{u}\cdot d\underline{A} = \oint_{C_1} \underline{u}\cdot d\underline{l} + \oint_{C_2} \underline{u}\cdot d\underline{l}$ but pay attention to sense of integration

VECTOR CALCULUS - Common PDEs

LAPLACE'S EQUATION - $\nabla^2 f = 0$, $\frac{\partial^2 f}{\partial x^2} + \frac{\partial^2 f}{\partial y^2} = 0$

* Separation of Variables - look for solution of form $f(x,y) = X(x)Y(y)$
  → Substitute in ⇒ $\frac{d^2X}{dx^2}Y + X\frac{d^2Y}{dy^2} = 0$, $\frac{1}{X}\frac{d^2X}{dx^2} = \frac{-1}{Y}\frac{d^2Y}{dy^2} = \text{const.}$
  → Only makes sense if both sides are const, independent of x and y
  → This gives us two separate ODEs for X and Y
* Case 1 - +ve separation constant: $\frac{d^2X}{dx^2} = \beta^2 X$, $\frac{d^2Y}{dy^2} = -\beta^2 Y$
  → solve to get $X = Ae^{\beta x} + Be^{-\beta x}$, $Y = C\sin\beta y + D\cos\beta y$
* Case 2 - -ve constant: say $-\gamma^2$
  → solution $X = A\sin\gamma x + B\cos\gamma x$, $Y = Ce^{\gamma y} + De^{-\gamma y}$
* Case 3 - constant $= 0$ ⇒ $\frac{d^2X}{dx^2} = \frac{d^2Y}{dy^2} = 0$
  → Solution $X = Ax + B$, $Y = Cx + D$
* Practically, use boundary conditions to find one of X, Y, then separation constants

POISSON'S EQUATION - $\nabla^2 U = -4\pi G\rho(r)$

* Potential of point mass at origin $\phi = -Gm/r$
* So force on unit mass at distance r is $F = -\nabla\phi = \frac{-Gm}{r^2}e_r$
* Integrate over all of earth's mass: $\phi(r) = \int_{\text{earth}} \frac{-G\rho(r_0)}{|r - r_0|}dV$
* $\nabla^2\phi = 4\pi G\delta(r - r_0)$ satisfied by ϕ

ONE DIMENSIONAL DIFFUSION - $\alpha\nabla^2 T = \frac{\partial T}{\partial t}$ → in 1-D, $\alpha\frac{\partial^2 T}{\partial x^2} = \frac{\partial T}{\partial t}$

* Separate variables - Let $T(x,t) = X(x)F(t)$
  → So $\alpha X''F = XF'$ and $\frac{F'}{F} = \alpha\frac{X''}{X} = \lambda$, say
  → Then $F' = \lambda F$ ⇒ $F \propto e^{\lambda t}$, $X'' = \frac{\lambda}{\alpha}X$ ⇒ $X \propto e^{\pm\sqrt{\frac{\lambda}{\alpha}}x}$
* If λ is +ve, time variation grows exponentially, not realistic
  → Let $\lambda = -\mu$, for example $T = T_0\cos\left(\sqrt{\frac{\mu}{\alpha}}x\right)e^{-\mu t}$
* Can also have complex separation constants because odd-order derivatives.

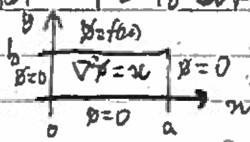
WAVE EQUATION - e.g. small vibrations of stretched string

* Consider equilibrium of small element subject to tension and inertial forces.
  → Define $c = \sqrt{\tau/m}$, get $\frac{\partial^2 y}{\partial t^2} = c^2\frac{\partial^2 y}{\partial x^2}$
* Try $y = X(x)F(t)$ again:
  → $XF'' = c^2X''F$ ⇒ const. $= \frac{F''}{F} = c^2\frac{X''}{X}$, say -ve constant $-\omega^2$
  → then solution can be written $y = e^{i\omega(t \pm x/c)}$ - travelling waves!
* Impose fixed end conditions: $y(0,t) = y(L,t) = 0$, -ve constant again
  → so $X = A\sin\beta x + B\cos\beta x$ but $B = 0$. $X(L) = 0$ ⇒ $A\sin\beta L = 0$
  → then $\beta L = n\pi$ for +ve integers n. Take $A = 1$
* Now $F''/F = -n^2\pi^2c^2/L^2$ ⇒ $F = C\sin\beta ct + D\cos\beta ct$
  → We have found n vibration modes of string.
* Alternatively, try change of variables $\xi = x + ct$, $\eta = x - ct$
  → Turns out $\frac{\partial^2 y}{\partial\xi\partial\eta} = 0$ ⇒ $y = f(x+ct) + g(x-ct)$, d'Alembert solution.

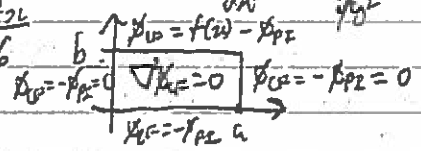
VECTOR CALCULUS – Types of PDE, Advanced PDE Tricks

TYPES OF PDE
* Elliptic – Laplace and Poisson. If we make a perturbation to the problem (e.g. extra point mass) effects all points in space
* Hyperbolic – Wave equation. If we disturb our stretched string introduce perturbation local to x and t. Only felt in certain region of the x–t plane, disturbance travels along $x \pm ct = \text{const.}$
* Parabolic – Diffusion equation. Thermal waves propagate from source but decay exponentially as they travel. So a perturbation influences whole domain but has very small effect on remote parts.

SOLUTION OF PDE BY CHANGE OF VARIABLE – Diffusion equation
* Dimensional Analysis – $T = fn(T_0, x, t, \alpha) \Rightarrow T = T_0 fn(\eta)$ with $\eta = \frac{x}{2\sqrt{\alpha t}}$
* $\frac{\partial T}{\partial x} = \frac{\partial T}{\partial \eta}\frac{\partial \eta}{\partial x} = T'/2\sqrt{\alpha t}$ so $\frac{\partial^2 T}{\partial x^2} = \frac{1}{4\alpha t}T''$, also $\frac{\partial T}{\partial t} = T'\frac{\partial \eta}{\partial t} = \frac{-xT'}{4\sqrt{\alpha}\,t^{3/2}}$
  → Substitute in, simplify to get $-2\eta T' = T''$
* Now let $f(\eta) = T'(\eta)$, $f' = -2\eta f \Rightarrow \int \frac{df}{f} = -\eta^2 + K = \ln f$, $\Rightarrow f = e^K e^{-\eta^2}$
* Then $T = A\int_0^{\eta} e^{-\eta^2}\,d\eta$ ... error function solution $T = T_0\,\mathrm{erf}\left(\frac{x}{2\sqrt{\alpha t}}\right)$



PI + CF for PDEs – for $\nabla^2\phi = \frac{\partial^2\phi}{\partial x^2} + \frac{\partial^2\phi}{\partial y^2} = x$
* Linearity: $\phi = \phi_{CF} + \phi_{PI}$ with $\nabla^2\phi_{CF} = 0$, $\nabla^2\phi_{PI} = x$
* Look for solutions $\phi_{PI} = fn(x)$ only $\Rightarrow \nabla^2\phi_{PI} = \frac{\partial^2\phi_{PI}}{\partial x^2} + \cancel{\frac{\partial^2\phi_{PI}}{\partial y^2}} = x$
  → $\phi_{PI} = x^3/6 + Ax + B$
  → Boundary Conditions $\Rightarrow \phi_{PI} = \frac{x^3}{6} - \frac{a^2 x}{6}$
* Now for complementary function:
  → Split into $\phi_{CF} = \phi_1 + \phi_2$, solve each to satisfy top and bottom boundary conditions.



MULTIPLE DIMENSIONS – example: rectangular tensioned membrane
→ Governing equation $\rho\frac{\partial^2 z}{\partial t^2} = T\left(\frac{\partial^2 z}{\partial x^2} + \frac{\partial^2 z}{\partial y^2}\right)$, let $c^2 = T/\rho$

* Seek solution of form $z(x, y, t) = X(x)Y(y)F(t)$
  → $XYF'' = c^2(X''YF + XY''F)$
  → $\frac{F''}{F} = c^2\left(\frac{X''}{X} + \frac{Y''}{Y}\right) = \text{constant} = -\omega^2$ and $\frac{X''}{X} = \frac{-\omega^2}{c^2} - \frac{Y''}{Y}$
  → Or $\frac{X''}{X} = A$, $\frac{Y''}{Y} = B$ where $\omega^2 = -c^2(A+B)$

Linear Algebra - Columns, Spaces, LU Factorisation

The Column Picture - Vector Multiplication

* To find general effect of multiplying by a matrix, consider base vectors
  * → $A\underline{i} = \underline{a_1}$, i.e. first column of A, same for other directions
  * → Then for any $\underline{x}$, $A\underline{x} = A(x\underline{i} + y\underline{j} + z\underline{k}) = x\underline{a_1} + y\underline{a_2} + z\underline{a_3}$
* Consider 3 equations in 3 unknowns, or $A\underline{x} = \underline{b}$
  * → In vector form $x\underline{a_1} + y\underline{a_2} + z\underline{a_3} = \underline{b}$
  * → If $\underline{a_i}$ are linearly independent, always have unique solution for $\underline{b}$
  * → If A is singular, columns not independent. i.e. $\underline{a_3} = \alpha\underline{a_1} + \beta\underline{a_2}$
    $A\underline{c}$ lies in plane of $\underline{a_1}$ and $\underline{a_2}$ for any $\underline{c}$. So if $\underline{b}$ doesn't lie in this plane, no solutions. If $\underline{b}$ is in plane, infinite solutions.
  * → $A\underline{x} = x\underline{a_1} + y\underline{a_2} + z\underline{a_3} = x\underline{a_1} + y\underline{a_2} + z(\alpha\underline{a_1} + \beta\underline{a_2}) = \underline{(x+\alpha z)\underline{a_1} + (y+\beta z)\underline{a_2}}$

Vector Spaces and Subspaces

* A vector space in R^m is set of $\underline{x}$ of form $\underline{x} = \lambda\underline{u} + \mu\underline{v} + \nabla\underline{w} + \dots$
  * → where $\underline{u}, \underline{v}, \dots$ are ^{m} fixed vectors, λ, μ are m scalar parameters.
  * → $\underline{u}, \underline{v}, \dots$ are said to span the space.
* Subspaces have fewer than m spanning vectors.
* A set of independent vectors spanning a space form a basis
* The vector space spanned by columns of A is column space of A
  * → The dimension (degrees of freedom or number of ~~[illegible]~~ basis vectors) = rank(A)

The Column/Row Pictures - Matrix Multiplication

* If $A = BC$ $\underline{b_i}$ are columns of B, $\tilde{\underline{c}}_j$ are rows of C
  * → $A = \underline{b_1}\tilde{\underline{c}}_1^T + \underline{b_2}\tilde{\underline{c}}_2^T + \dots + \underline{b_k}\tilde{\underline{c}}_k^T$
    * → i.e. BC is sum of outer products of columns of B with rows of C
* Example - Permutation matrix to swap 2nd and 3rd columns of a 3x3 matrix
  * → If $A = BC$, can write $\underline{a_1} = \underline{b_1}$, $\underline{a_2} = \underline{b_3}$, $\underline{a_3} = \underline{b_2}$
  * → So $A = \underline{b_1}\underline{i}^T + \underline{b_2}\underline{k}^T + \underline{b_3}\underline{j}^T$
* Row Picture: $\tilde{\underline{a}}_i = b_{i1}\tilde{\underline{c}}_1 + b_{i2}\tilde{\underline{c}}_2 + \dots + b_{ik}\tilde{\underline{c}}_k$
* Note that can rearrange columns of B if we rearrange identically rows of C
  * → Can also divide column of B by scalar, if multiply row of C by same scalar.

LU ~~Factor~~isation - $\boxed{A = LU} = \underline{l_1}\tilde{\underline{u}}_1^T + \underline{l_2}\tilde{\underline{u}}_2^T + \dots + \underline{l_m}\tilde{\underline{u}}_m^T$

* L is lower triangular, 1s down leading diagonal, 0s above
* U is upper echelon, non-zero elements on or above leading diagonal.
* Choose $\tilde{\underline{u}}_1$ to be first row of A, $\underline{l_1}$ determined by making $\underline{l_1}\tilde{\underline{u}}_1^T = \underline{a_1}$ for first column
  * → Repeat to choose $\underline{l_2}$ and $\tilde{\underline{u}}_2$ etc.
* Can improve algorithm robustness using partial pivoting.
  * → Choose largest element in each column to pivot on to zero the column.
  * → Produces L which isn't lower triangular. Use permutation matrix $\boxed{PA = LU}$

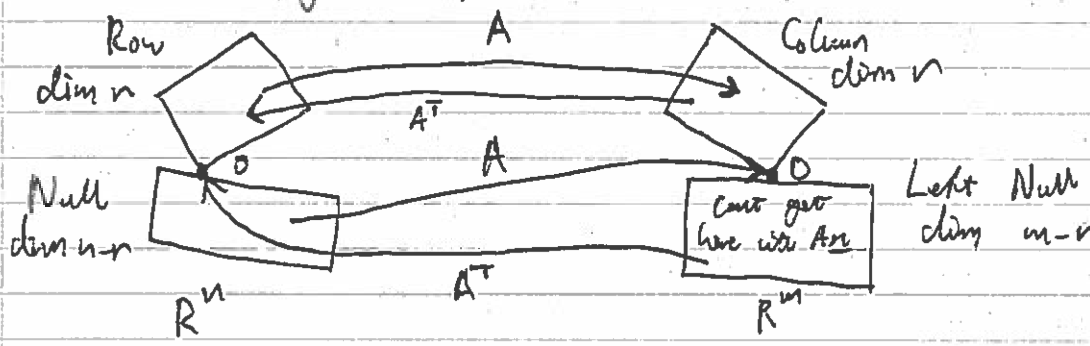
~~[illegible]~~ LINEAR ALGEBRA – $A\underline{x}=\underline{b}$; Fundamental Subspaces

SOLUTION OF $A\underline{x}=\underline{b}$

* No solution of $\underline{b}$ outside column space of A
  → If we express $\underline{b}$ in terms of colums of L, should only need those in column space
* To solve $A\underline{x}=\underline{b}$, use LU decomposition $LU\underline{x}=\underline{b}$
  → First find $\underline{c}$ such that $L\underline{c}=\underline{b}$
  → Then find $U\underline{x}=\underline{c}$ (finding Null Space of A)
  → Find variables in terms of free variables
* We find, in general, $\underline{x}=\underline{x}_0+t_1\underline{n}_1+t_2\underline{n}_2+\dots$
  → These $\underline{n}_i$ are general solution of $A\underline{n}=0$ – basis for null space
  → Because L^{-1} exists, null space of U = null space of A

FOUR FUNDAMENTAL SUBSPACES – of $A=LU$

* Column Space – columns of L corresponding to non-zero rows of U
* Null Space – set free variables to 1 in turn, solve $U\underline{x}=0$
* Row Space – All non-zero rows of U, $\perp$ to Null Space
* Left Null Space – perpendicular to all vectors in Column Space
* Useful properties
  → Dimension of Row Space = r = rank(A) = dimension of Column Space
  → Dimension of Null Space = number of free variables = $n-r$
  → Everything also applies to A^T
  → For orthogonal complements, sum of subspace dimensions = total for space



ROW MANIPULATION – Because $A=BC=\underline{b}_1\underline{c}_1^T+\underline{b}_2\underline{c}_2^T+\dots$

* Add an extra column at end of C: ~~[illegible]~~, then $C'=[C\ \underline{d}]$, $A'=[A\ \dots$
* $[A\ \underline{b}]=B[C\ \underline{d}]$. Using LU factorisation $[A\underline{b}]=L[U\ \underline{c}]$
  → Then $[A\ \underline{b}]=[LU\ \ L\underline{c}]$ gives us $\underline{c}$ with no extra work
* if we extend a matrix A with I: $[A\ I]=L[U\ D]$
* Manipulate (linearly) rows and colums of $[AI]$ to make first three columns identity matrix. Ignore coefficients stored in L
  → So afterwards we have $[A\ I]=L[I\ D]$, let $C=L^{-1}$
* Take $C[AI]=[CA\ \ CI]=[I\ \ D]$
  → now if $CA=I$ and $CI=D \Rightarrow \underline{C=D=A^{-1}}$

LINEAR ALGEBRA – Least Squares, Gram-Schmidt, QR Factorisation

LEAST SQUARES SOLUTION OF $A\underline{x} = \underline{b}$

* Measure b at different times t, want to fit $b = C + Dt$
* In general $A\bar{x} = \underline{b}$, $\begin{bmatrix} 1 & t_1 \\ 1 & t_2 \\ \dots \\ 1 & t_m \end{bmatrix} \begin{bmatrix} C \\ D \end{bmatrix} = \begin{bmatrix} b_1 \\ b_2 \\ \dots \\ b_m \end{bmatrix}$
  → This equation inconsistent, look for best fit $\bar{x}$
* Least squares solution for x minimises $|A\underline{x} - \underline{b}|^2 = (A\underline{x} - \underline{b}) \cdot (A\underline{x} - \underline{b})$
  → $A\bar{x}$ must be in column space, so for any $\underline{b}$ want point at end of perpendicular dropped onto column space.
* In general $\underline{b} = b_{col} + b_{left}$ with $b_{col} \cdot b_{left} = 0$
* To get rid of b_{left} multiply original problem by A^t
  → $A^t A \bar{x} = A^t \underline{b} = A^t b_{col} + \cancel{A^t b_{left}}^{\,0}$, given $\underline{b}$ can solve for $\bar{x}$
* Note: if columns of A independent $A^t A$ is always invertible
  → rank$(A) = n$ if columns independent ⇒ dim of row space is n, null space dimension $= n - r = 0$ so $A\underline{x} = 0$ only solution $\underline{x} = 0$
  → So dimension of null space of $A^t A = 0$, column space of $A^t A = n$
  → i.e. Column space of $A^t A$ is all of R^n ⇒ invertible.

GRAM-SCHMIDT PROCESS

* Least-squares solution would be easier if could project b directly onto column space: $\underline{b}_{col} = \lambda_1 \underline{a}_1 + \lambda_2 a_2 + \dots$
  → But $\underline{a}$s are not orthogonal, finding the λs is difficult.
* If instead column space aligned with coordinate directions $\underline{i}, \underline{j}, \underline{k} \dots$ and other base vectors $\underline{m}, \underline{n}, \underline{o}, \dots$ lay in left-null space
  → Could then write $\underline{b} = b_1 \underline{i} + b_2 \underline{j} + \dots$, strip off $\underline{m}, \underline{n}$ to give b_{col}
  → Because basis vectors orthogonal coefficients $b_1 = \underline{i} \cdot \underline{b}$, $b_2 = \underline{j} \cdot \underline{b}$ etc.
* Gram Schmidt generates mutually orthogonal unit vectors from arbitrary set
  → Firstly $\underline{q}_1 = \underline{a}_1 / |\underline{a}_1|$. Form $\underline{q}_2$ by subtracting part which is parallel to $\underline{q}_1$ and then normalising. $\bar{a}_2 = a_2 - (\underline{q}_1 \cdot \underline{a}_2) \underline{q}_1$, $\bar{q}_2 = \bar{a}_2 / |\bar{a}_2|$
  → Repeat: $\bar{a}_3 = a_3 - (q_1 \cdot a_3) q_1 - (q_2 \cdot a_3) q_2$, $q_3 = \bar{a}_3 / |\bar{a}_3|$ etc.

QR FACTORISATION – $A = QR$

→ Q has mutually orthogonal unit vectors, R is upper triangular
* Can write Gram-Schmidt process as relationship between matrices:
  → $A = QR$, $\begin{bmatrix} \uparrow & \uparrow & & \uparrow \\ a_1 & a_2 & \dots & a_n \\ \downarrow & \downarrow & & \downarrow \end{bmatrix} = \begin{bmatrix} \uparrow & \uparrow & & \uparrow \\ q_1 & q_2 & \dots & q_n \\ \downarrow & \downarrow & & \downarrow \end{bmatrix} \begin{bmatrix} q_1 \cdot a_1 & q_1 \cdot a_2 & \dots & q_1 \cdot a_n \\ 0 & q_2 \cdot a_2 & \dots & q_2 \cdot a_n \\ 0 & 0 & \dots & \dots \\ 0 & 0 & \dots & q_n \cdot a_n \end{bmatrix}$
* Least Squares much easier. If $A = QR$, $A^t A \bar{x} = A^t b$
  ⇒ $(QR)^t QR \bar{x} = (QR)^t b$, $R^t Q^t Q R \bar{x} = R^t Q^t b$ but because qs orthogonal, $Q^t Q = I$
  → then $R^t R \bar{x} = R^t Q^t b$. Also R^t is invertible because R is
  → Finally $\boxed{R\bar{x} = Q^t b}$, can solve by back-substitution.

LINEAR ALGEBRA – Eigenvalues, Eigenvectors, Diagonalisation

EIGENVALUES AND EIGENVECTORS FOR SYMMETRIC MATRICES

* Always real eigenvalues, λ and orthogonal eigenvectors $\underline{u}$. $Au_1 = \lambda u_1$ etc.
  → Then $AU = U\Lambda \Rightarrow A = U\Lambda U^T$ since $U^{-1} = U^T$
* In its coordinate system A is a simple diagonal matrix.
  → $A^k = U\Lambda^k U^T$ or in vector form $A^k \underline{x} \simeq \alpha_1 \lambda_1^k u_1$ where λ_1 is largest magnitude eigenvalue, with $\underline{x} = \alpha_1 u_1 + \alpha_2 u_2 + \alpha_3 u_3 \dots$

NON-SYMMETRIC EXTENSION

* Characteristic equation $A\underline{x} = \lambda \underline{x} \Rightarrow (A - \lambda I)\underline{x} = 0$
  → Note A must be square.
* For a non-zero solution for x, $A - \lambda I$ is singular
  → so $(A - \lambda I)^{-1}$ must not exist $\Rightarrow \det(A - \lambda I) = 0$
  → x must be in null space of $A - \lambda I$, which must not just be 0
* Let $P(\lambda) = \det(A - \lambda I) = (-\lambda)^n + \alpha_1(-\lambda)^{n-1} + \alpha_2(-\lambda)^{n-2} + \dots + \alpha_n = 0$
  → n eigenvalues, either real or complex conjugate pairs
  → Factorise $P(\lambda)$, set $= 0 \Rightarrow \det(A) =$ product of eigenvalues.
* Defective matrices – less eigenvectors than eigenvalues
  → can only happen if an eigenvalue is repeated.

DIAGONALISING A MATRIX – $AX = X\Lambda$, eigenvalues λ, eigenvectors x

* In general eigenvectors are not orthogonal.
* However columns of X are eigenvectors ⇒ independent
  → Column space of X has dimension n, must be all of R^n ⇒ X invertible, $\boxed{A = X\Lambda X^{-1}}$
* This is a similarity relationship – A and Λ are similar
  → Note X is not unique – can scale arbitrarily any of eigenvectors
* In physical problems often a choice of variables to set up problem in.
  → E.g. loop or element currents in an electric circuit.
  → The matrices that result from this ambiguity are similar
  → The two sets of variables can be related by $M\underline{x} = \underline{y}$
  → Assume M is invertible – no information lost re-casting the problem
* Can set up problem in $\underline{x}$ with A or in y with A'
  → $A\underline{x} = M^{-1}A'y = M^{-1}A'M\underline{x}$ i.e. $A = M^{-1}A'M$, A and A' are similar.
* Similar matrices have same set of eigenvalues: $A\underline{x} = \lambda \underline{x} \Rightarrow M^{-1}A'M\underline{x} = \lambda \underline{x}$
  → so $A'M\underline{x} = \lambda M\underline{x}$ i.e. $A'y = \lambda y$. Eigenvectors related by $y = M\underline{x}$
* A matrix similar to a diagonal matrix is diagonalisable
* Any matrix with distinct eigenvalues can be diagonalised
* Non-symmetric matrices with repeated eigenvalues may be diagonalisable.

COMPUTATION OF EIGENVALUES

* Power Method – start with a guess, v_0 and keep multiplying by A
  → Eventually component corresponding to largest eigenvalue dominates
* Inverse Power – Use A^{-1} in power method. Converges to eigenvector corresponding to largest λ of A^{-1} i.e. smallest λ of A.
* Shifted Inverse Power – replace A with $A - \alpha I$. Has same eigenvectors but all eigenvalues reduced by α. Vary α to pick out each.

LINEAR ALGEBRA - Singular Value Decomposition

EIGENVALUE PROBLEM FOR A^tA

* A^tA is always square, has eigenvalues and eigenvectors
  - → It is also symmetric ∴ diagonalisable.
  - → Choose eigenvectors to be orthonormal ⇒ $A^tAQ = Q\Delta$ ⇒ $A^tA = Q\Delta Q^t$ where Q columns are the eigenvectors and eigenvalues in diagonal of Δ
* The eigenvalues of A^tA cannot be negative: $A^tAq = \lambda q$
  - → $q^tA^tAq = \lambda q^tq$, $(Aq)^tAq = \lambda q^tq$ ⇒ $\lambda = |Aq|^2/|q|^2 \geq 0$
* Will be rank(A) = r non-zero eigenvalues and n-r which are zero
  - → Because for a zero eigenvalue $Aq = 0$ ⇒ q in null space
  - → For non-zero eigenvalue q not in null space. But qs orthogonal ⇒ in row space.
  - → $q_1 \cdots q_r$ are orthogonal basis for Row Space
  - → $q_{r+1} \cdots q_n$ are orthogonal basis for Column Space.

SINGULAR VALUES

* Define for non-zero eigenvalues of A^tA $\sigma_i = \sqrt{\lambda_i}$ (singular values)
  - → $A^tA\, q_i = \sigma_i^2 q_i$ for $1 \leq i \leq r$
* To generalise the concept of eigenvalues. Demand $Aq_i = \sigma_i \hat{q}_i$
* $\hat{q}_1, \hat{q}_2 \cdots \hat{q}_r$ are orthogonal unit eigenvectors of AA^t
  - → add $\hat{q}_{r+1} \cdots \hat{q}_m$, eigenvectors corresponding to zero eigenvalues of AA^t
* We then have $AA^t = \hat{Q}\hat{\Delta}\hat{Q}^t$
* ~~Because Q is an orthogonal matrix $Q^{-1} = Q^t$ so that~~
  ~~→ $AQ = \hat{Q}\Sigma$~~
* So $Aq_i = \sigma_i \hat{q}_i$ and $A^t\hat{q}_i = \sigma_i q_i$ ⇒ $Aq_i = \sigma_i \hat{q}_i$
  - → In matrix form: $AQ = \hat{Q}\Sigma$ ⇒ $\boxed{A = \hat{Q}\Sigma Q^t}$ (because Q is orthogonal $Q^{-1} = Q^t$)

→ this is singular value decomposition.

PROBABILITY – Basics and Probability Distributions

BASICS – probability quantifies likelihood of an uncertain event

* 3 Axioms – $p(E) \geq 0$, $p(\Omega) = 1$, $P(E_1 \cup E_2) = P(E_1) + P(E_2)$
  → Where Ω is sample space, E_1 and E_2 disjoint events. ($\cup \equiv$ OR)
* General Addition – $p(E_1 \cup E_2) = p(E_1) + p(E_2) + P(E_1 \cap E_2)$ ($\cap \equiv$ AND)
* Conditional Probability – $p(A$ given $B) = p(A|B) = p(A \cap B)/p(B)$
* Bayes Theorem – $p(A|B)\,p(B) = p(A \cap B) = p(B|A)\,p(A) \Rightarrow P(A|B) = p(B|A)\,p(A)/p(B)$
* Random Variable – associates a numerical value with outcome of random experiment
  → Probability Function – $P(X=x)$ is probability random variable X has value x
* Mean = Average = Expectation – $E[X] = \sum_i p(x_i) x_i$
* Surprise$(x_i) = -\log(p(x_i))$, entropy$(p) = E[-\log(p)] = -\sum_i p(x_i)\log(p(x_i))$

DISCRETE DISTRIBUTIONS

* Bernoulli – $X \sim Ber(p)$ where $0 \leq p \leq 1$, X is binary variable
  → $P(X=1) = p$, $P(X=0) = 1-p$, $P(X=x) = p^x(1-p)^{(1-x)}$
  → $E[X] = 0 \times (1-p) + 1 \times p = p$, $Var[X] = E[(X-E[X])^2] = p(1-p)$
* Useful Note: $Var[X] = E[X^2] - E[X]^2$
* Binomial – $X \sim B(n,p)$ where $n = 0, 1, 2 \dots$ and $0 \leq p \leq 1$
  → $P(X=r) = {}_nC_r\, p^r (1-p)^{n-r}$ $E[X] = np$, $Var[X] = np(1-p)$
* Poisson – consider random independent events, average rate/intensity λ
  → Want probability distribution over rate of events.
  → Look at r arrivals in a Binomial $B(n,p)$, let $n \to \infty$, $p = \lambda/n$
  → $$P(X=r) = \lim_{n\to\infty} \frac{n!}{(n-r)!\,r!}\left(\frac{\lambda}{n}\right)^r\left(1-\frac{\lambda}{n}\right)^{n-r}$$
  $$= \lim_{n\to\infty} \underbrace{\frac{n}{n}\frac{n-1}{n}\cdots\frac{n-r+1}{n}}_{1}\left(\frac{\lambda^r}{r!}\right)\underbrace{\left(1-\frac{\lambda}{n}\right)^n}_{\exp(-\lambda)}\underbrace{\left(1-\frac{\lambda}{n}\right)^{-r}}_{1}$$
* So $P(X=r) = \lambda^r \exp(-\lambda)/r!$, a Poisson distribution!
  → $E[X] = Var[X] = \lambda$, derive from limiting Binomial.

CONTINUOUS DISTRIBUTIONS

* Probability density function $p(x)$ defined so $p(x)\,dx$ is the probability that X takes a value between x and $x+dx$.
  → Total probability $= 1$ so $\int_x p(x)\,dx = 1$
* $E[X] = \int_x x\,p(x)\,dx$, $Var[X] = \int_x (x - E[X])^2\,dx = \int_x x^2 p(x)\,dx - E[X]^2$
* Cumulative Probability function $F(x) = \int_0^x p(x')\,dx'$
* Let T represent time between Poisson events. $P(T=t)\,dt = P(t \leq T \leq t+dt)$
  → i.e. $p(\text{none in } t) \times p(\text{one in } dt) = \exp(-\lambda t) \times \lambda dt \exp(-\lambda dt) = \lambda\exp(-\lambda t)\,dt$
  → $T \sim Ex(\lambda)$ where $p(t) = \lambda\exp(-\lambda t)$ Exponential distribution
  → $E[T] = 1/\lambda$, $Var[T] = 1/\lambda^2$
* Gaussian or Normal – Large numbers of small influences add up.
  → All in Databook, integrals not available in closed form.
* Beta Distribution – ~~dist~~ over probabilities i.e. interval $0 \leq x \leq 1$
  → $X \sim Beta(\alpha, \beta)$, $p(x) = \dfrac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)}$ with $\Gamma(\gamma) = (\gamma-1)!$

PROBABILITY – Combining, Manipulating Distributions, Moment Generating Functions

COMBINING PROBABILITIES
* Assume X and Y independent, uniformly distributed $\Rightarrow p(x) = p(y) = 1$
→ Let $S = X + Y$. Consider a 2D map of x and y
→ Joint probability is uniform = 1 over this square.
Take given box $dx\,dy$. $p(x \leq X \leq x+dx)$ and $p(y \leq Y \leq y+dy)$ is $p(x,y)\,dx\,dy$
* For $0 \leq s \leq 1$ $F(S=s)$ = (area under line $s = x+y$) $= s^2/2$
→ So $p(s) = dF(s)/ds = s$ where $0 \leq s \leq 1$
* For $1 \leq s \leq 2$ we have $F(S=s) = 1 - (2-s^2)/2 \Rightarrow p(s) = \frac{dF(s)}{ds} = 2-s$

SUMMING INDEPENDENT VARIABLES
* In general mean of sum is sum of means
→ Variance of sum is sum of variances.
* In continuous case, of $S = X+Y$, $P(S=s) = \int P(X=x)P(Y=s-x)\,dx$
* Sum of two independent Gaussians is another Gaussian – closed under addition

TRANSFORMATIONS OF CONTINUOUS VARIABLES
* If distribution of X is $p(x)$, what is distribution of $Y = f(X)$?
→ We want $p(x)\,dx = p(y)\,dy$, $p(y) = p(x)\left|\frac{dx}{dy}\right| = p(x)\,|J|$
J = the Jacobian
* Linear Transformations don't change distribution shape.
→ Adding a constant shifts the mean
→ Multiply by constant α, $\mu \to \alpha\mu$, $\sigma^2 \to \alpha^2\sigma^2$

MOMENT GENERATING FUNCTIONS
* For discrete random variable define $g(z) = \sum_r z^r p(r)$
→ Useful because $E[R] = g'(1)$ and $E[R^2] = g''(1) + g'(1)$
→ Moment generating functions for common distributions in Databook
* The sum of independent random variables has a moment generating function which is the product of original moment generating functions.
* In continuous case, $g(s) = \int_x \exp(-sx)\,p(x)\,dx$, two-sided Laplace Transform

→ Turns out $E[X] = -g'(0)$ and $E[X^2] = g''(0)$
→ If $Y = X - \beta$, $g_y(s) = \exp(\beta s)\,g_x(s)$
→ If $Y = \alpha X$, $g_y(s) = g_x(\alpha s)$

CENTRAL LIMIT THEOREM
* If X_i are all identically independently distributed random variables, mean μ and variance σ^2, in the limit $\sum_0^n X_i \sim N(n\mu, n\sigma^2)$ for large n.
→ Prove with moment generating function.

HYPOTHESIS TESTING – Assume null hypothesis H_0 is true.
* Find probability of observed outcome, or something more extreme
→ If this is less than desired significance level, reject H_0
* May need double sided test. Or Bayesian comparing two H_0.