Mathematics

A refresher

Per-Olof Widmark

Theoretical Chemistry
Chemistry Department
Lund University

ESQC 2019

Outline

Numbers

Derivatives

Calculus of variation

Vectors

Levi-Civita et. al.

Vector calculus

Vector spaces

Operators

Matrices

Distributions

Decompositions

Diagonalizing matrices

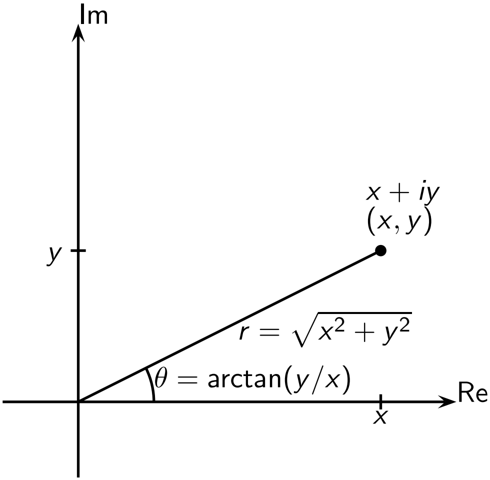
Numbers

- Natural numbers, $\mathbb{N}$: all whole non negative numbers, 0, 1, 2, 3, 4, ...
- Integers, $\mathbb{Z}$: all whole numbers $\ldots, -3, -2, -1, 0, 1, 2, 3, \ldots.$
- Rational numbers, $\mathbb{Q}$: all numbers that can be written as $\frac{p}{q}$ where p and q are integers, $q \neq 0$.
- Irrational number, $\mathbb{P}$: A number that is the limit of a sequence of rational numbers but is not rational. For example, Leibniz formula gives π which is an irrational number

$$\frac{\pi}{4} = 1 - \frac{1}{3} + \frac{1}{5} - \frac{1}{7} + \ldots \quad (1)$$

- Real numbers, $\mathbb{R}$: all rational and irrational numbers.
- Complex numbers, $\mathbb{C}$: all numbers of the form $x + iy$ where x and y are real and i is the imaginary unit defined as $i^2 = -1$.

Complex numbers



Complex numbers

In the formulas below we assume that $z = x + iy$, $z_1 = x_1 + iy_1$ and $z_2 = x_2 + iy_2$.

- From a set theory point of view: $\mathbb{C} = \mathbb{R}^2$, an ordered pair of real numbers (x, y).
- Addition: $z = z_1 + z_2 = (x_1 + x_2) + i(y_1 + y_2)$
- Subtraction: $z = z_1 - z_2 = (x_1 - x_2) + i(y_1 - y_2)$
- Multiplication: $z = z_1 \cdot z_2 = (x_1 x_2 - y_1 y_2) + i(x_1 y_2 + x_2 y_1)$
- Complex conjugate: $z^* = x - iy$
- $(z_1 z_2)^* = z_1^* z_2^*$
- Norm: $|z| = \sqrt{zz^*} = \sqrt{x^2 + y^2}$
- Division: $z = \frac{z_1}{z_2} = \frac{z_1 z_2^*}{z_2 z_2^*} = \frac{z_1 z_2^*}{|z_2|^2}$

Complex numbers — polar representation

Any point in $\mathbb{R}^2$ can be represented in polar coordinates

$$\begin{cases} x = r\cos(\theta) \\ y = r\sin(\theta) \end{cases} \tag{2}$$

and this also holds true for $\mathbb{C}$. Combine this with Eulers formula

$$e^{i\theta} = \cos(\theta) + i\sin(\theta) \tag{3}$$

we obtain

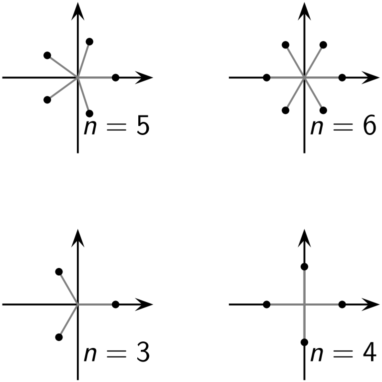
$$z = x + iy = re^{i\theta} = r\cos(\theta) + ir\sin(\theta). \tag{4}$$

Operations on complex numbers

Using the polar representation makes certain operation simpler. In the formulas below we assume that $z = x + iy = re^{i\theta}$, $z_1 = x_1 + iy_1 = r_1 e^{i\theta_1}$ and $z_2 = x_2 + iy_2 = r_2 e^{i\theta_2}$.

- Multiplication: $z = z_1 \cdot z_2 = r_1 r_2 e^{i(\theta_1+\theta_2)}$
- Division: $z = \frac{z_1}{z_2} = \frac{r_1}{r_2} e^{i(\theta_1-\theta_2)}$
- Square root: $\sqrt{z} = \sqrt{re^{i\theta}} = \sqrt{r} e^{i\theta/2}$
- Finding roots: $z^3 = 1$ has roots 1, $e^{2\pi i/3}$ and $e^{4\pi i/3} = e^{-2\pi i/3}$ by realizing that $1 = e^0 = e^{2\pi i} = e^{4\pi i} = e^{6\pi i} \ldots$

Roots of $z^n = 1$



Derivatives

- Ordinary derivative of a function $f(x)$ of one variable x:

$$\frac{df}{dx} = \lim_{h\to 0} \frac{f(x+h) - f(x)}{h} = f'(x) \tag{1}$$

- Partial derivative of a function $f(x, y)$ of two variables x and y with respect to x:

$$\frac{\partial f}{\partial x} = \lim_{h\to 0} \frac{f(x+h, y) - f(x, y)}{h} = f'_x \tag{2}$$

- Total derivative of a function $f(x, y)$ of two variables x and y with respect to x:

$$\frac{\delta f}{\delta x} = \lim_{h\to 0} \frac{f(x+h, y(x+h)) - f(x, y(x))}{h} \tag{3}$$

$$= \frac{\partial f}{\partial x} + \frac{\partial f}{\partial y}\frac{dy}{dx} \tag{4}$$

$$= \frac{df}{dx} \tag{5}$$

Total derivative

$$\begin{aligned}
\frac{\delta f}{\delta x} &= \lim_{h\to 0} \frac{f(x+h, y(x+h)) - f(x, y(x))}{h} && (6)\\
&= \lim_{h\to 0} \frac{\overbrace{f(x+h, y(x))}^{\text{add}} - f(x, y(x))}{h} && (7)\\
&+ \lim_{h\to 0} \frac{f(x+h, y(x+h)) - \overbrace{f(x+h, y(x))}^{\text{subtract}}}{h} && (8)\\
&= \frac{\partial f}{\partial x} + \frac{\partial f}{\partial y}\frac{dy}{dx} && (9)\\
&= \frac{df}{dx} && (10)
\end{aligned}$$

Analytic functions

Consider a complex valued function $f(z)$ of one complex variable $z = x + iy$ which can be written as

$$f(z) = u(x, y) + iv(x, y) \tag{11}$$

where x, y, u and v are real. When is the derivative well defined using $\Delta z = \Delta x + i\Delta y$?

$$\begin{aligned}
\frac{df}{dz} = f'(z) &= \lim_{\Delta x \to 0} \frac{\Delta u + i\Delta v}{\Delta x} && (12)\\
&= \frac{\partial u}{\partial x} + i\frac{\partial v}{\partial x} && (13)\\
&= \lim_{\Delta y \to 0} \frac{\Delta u + i\Delta v}{i\Delta y} && (14)\\
&= \frac{1}{i}\frac{\partial u}{\partial y} + \frac{\partial v}{\partial y} && (15)
\end{aligned}$$

Cauchy-Riemann condition

The condition from above

$$\frac{\partial u}{\partial x} + i\frac{\partial v}{\partial x} = \frac{1}{i}\frac{\partial u}{\partial y} + \frac{\partial v}{\partial y} \tag{16}$$

leads to the Cauchy-Riemann condition

$$\frac{\partial u}{\partial x} = \frac{\partial v}{\partial y}; \quad \frac{\partial u}{\partial y} = -\frac{\partial v}{\partial x} \tag{17}$$

This condition (and all four partial derivatives exist and are continious) leads to a function that have a well defined derivative $f'(z)$, and such a function is called an *analytic function*.

An analytic function

Consider the function

$$\begin{aligned} f(z) &= e^z &&= e^{x+iy} \\ &= e^x e^{iy} &&= e^x\cos(y) + ie^x\sin(y) \\ & &&= u(x,y) + iv(x,y). \end{aligned} \tag{18}$$

Evaluate the partial derivatives

$$\begin{aligned} \frac{\partial u}{\partial x} &= e^x\cos(y) &&= u \\ \frac{\partial v}{\partial x} &= e^x\sin(y) &&= v \\ \frac{\partial u}{\partial y} &= -e^x\sin(y) &&= -v \\ \frac{\partial v}{\partial y} &= e^x\cos(y) &&= u \end{aligned} \tag{19}$$

All conditions are met thus e^z is an analytic function and

$$f'(z) = \frac{\partial u}{\partial x} + i\frac{\partial v}{\partial x} = u + iv = e^z. \tag{20}$$

Why e^z?

The function e^z is analytic for all values $z \neq \infty$ which mean that the Taylor expansion

$$e^z = 1 + z + \frac{1}{2!}z^2 + \frac{1}{3!}z^3 + \frac{1}{4!}z^4 + \ldots = \sum_{k=0} \frac{1}{k!}z^k \quad (21)$$

is convergent for all $z \neq \infty$. We can define

$$e^X = \sum_{k=0} \frac{1}{k!}X^k \quad (22)$$

for basically any X! For example

$$U = e^X; \quad X \text{ is anti-Hermitian matrix, } U \text{ is unitary} \quad (23)$$

$$e^{\hat{T}}; \quad \text{Exponential ansatz in Coupled Cluster theory} \quad (24)$$

Functions

- A function takes a number as input and returns a number: $y = f(x)$.
- A functional takes a function as input and returns a number
  - $E[\Psi] = \int \Psi^* \hat{H} \Psi d\tau$
  - $F[\rho] = \int \rho^2(\mathbf{r}) |\nabla \rho(\mathbf{r})|^2 d\tau$, (GGA LYP)

Functional derivative

Consider a functional $F[\rho] = \int G(x, \rho, \rho', \ldots)dx$ and let us define the functional derivative

$$\frac{\delta F}{\delta \rho} \tag{25}$$

by

$$\int \frac{\delta F}{\delta \rho}\phi dx = \lim_{\epsilon \to 0} \frac{F[\rho + \epsilon\phi] - F[\rho]}{\epsilon} \tag{26}$$

$$= \left[\frac{d}{d\epsilon}F[\rho + \epsilon\phi]\right]_{\epsilon=0} \tag{27}$$

where the functions $\rho(x)$ and $\phi(x)$ are chosen to fulfill the boundary conditions of the problem.

Example of functional derivative

Let us try to minimize the functional

$$F[\rho] = \int_0^1 (\rho')^2 dx; \quad \rho(0) = 0; \rho(1) = 1. \tag{28}$$

We then have that $\phi(0) = \phi(1) = 0$ in order to maintain the boundary conditions for $\rho + \epsilon\phi$. Now use the defintion

$$\begin{aligned}
\int \frac{\delta F}{\delta \rho}\phi dx &= \lim_{\epsilon\to 0} \frac{\int_0^1 (\rho' + \epsilon\phi')^2 dx - \int_0^1 (\rho')^2 dx}{\epsilon} && (29)\\
&= \lim_{\epsilon\to 0} \frac{\int_0^1 \epsilon (2\rho'\phi') dx + \int_0^1 \epsilon^2 (\phi')^2 dx}{\epsilon} && (30)\\
&= \int_0^1 2\rho'\phi' dx && (31)\\
&= \left[2\rho'\phi\right]_0^1 - \int_0^1 2\rho''\phi dx && (32)\\
\frac{\delta F}{\delta \rho} &= -2\rho''(=0) && (33)
\end{aligned}$$

Calculus of variation

Consider the problem: minimize the functional F with respect to the function ρ subject to the conditions $\rho(a) = \rho_a$ and $\rho(b) = \rho_b$.

$$F[\rho] = \int_a^b G(x, \rho, \rho')\, dx \qquad (1)$$

Make F stationary with respect to variations of ρ:

$$\delta F[\rho] = F[\rho + \phi] - F[\rho] \qquad (2)$$

where ϕ is "small".

Functional derivative

Replace $\rho(x)$ with $\rho(x) + \phi(x)$ with $\phi(a) = \phi(b) = 0$ to satisfy the boundary conditions.

$$\begin{aligned}
\delta F &= F[\rho + \phi] - F[\rho] && (3)\\
&= \int_a^b G(x, \rho + \phi, \rho' + \phi')\, dx - \int_a^b G(x, \rho, \rho')\, dx && (4)\\
&= \int_a^b \left(\frac{\partial G}{\partial \rho}\phi + \frac{\partial G}{\partial \rho'}\phi' \right) dx + \mathcal{O}(\phi^2) && (5)\\
&= \int_a^b \frac{\partial G}{\partial \rho}\phi\, dx + \left[\frac{\partial G}{\partial \rho'}\phi \right]_a^b - \int_a^b \left(\frac{d}{dx}\frac{\partial G}{\partial \rho'} \right) \phi\, dx && (6)\\
&= \int_a^b \underbrace{\left(\frac{\partial G}{\partial \rho} - \frac{d}{dx}\frac{\partial G}{\partial \rho'} \right)}_{\text{functional derivative}} \phi\, dx = 0; \quad \forall \phi && (7)
\end{aligned}$$

The Euler-Lagrange equation

Minimizing the functional $F[\rho]$ leads to the Euler-Lagrange equation

$$\frac{\delta F}{\delta \rho} = \frac{\partial G}{\partial \rho} - \frac{d}{dx}\frac{\partial G}{\partial \rho'} = 0 \qquad (8)$$

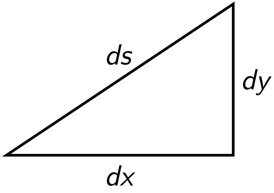
Example

What is the shortest distance between points $(0,0)$ and $(1,1)$? The length of a line segment is given by

$$ds = \sqrt{dx^2 + dy^2} = \sqrt{1 + (y')^2}dx \tag{9}$$

so the functional we want to minimize is

$$S[y] = \int_0^1 \sqrt{1 + (y')^2}\, dx; \quad G(x, y') = \sqrt{1 + (y')^2} \tag{10}$$



Example

The functional does not depend on y directly so

$$\frac{\partial G}{\partial y} - \frac{d}{dx}\frac{\partial G}{\partial y'} = 0 \quad \rightarrow \quad \frac{\partial G}{\partial y'} = c \tag{11}$$

$$\frac{\partial G}{\partial y'} = \frac{1}{2}\left\{1 + (y')^2\right\}^{-1/2}(2y') = \frac{y'}{\sqrt{1+(y')^2}} = c \tag{12}$$

massage this a bit and we get

$$y' = \frac{c}{\sqrt{1-c^2}} = k; \quad \rightarrow \quad y = kx + l; \quad \rightarrow \quad y = x \tag{13}$$

Extensions

Several dependent variables:

$$F = \int G(x, \rho_1, \rho_1', \rho_2, \rho_2', \dots) \tag{14}$$

leads to a set of simultaneous equations

$$\frac{\partial G}{\partial \rho_i} - \frac{d}{dx}\frac{\partial G}{\partial \rho_i'} = 0 \tag{15}$$

Several independent variables:

$$F = \int G(x_1, x_2, \dots, \rho, \frac{\partial \rho}{\partial x_1}, \frac{\partial \rho}{\partial x_2}, \dots) \tag{16}$$

yields the equation

$$\frac{\partial G}{\partial \rho} - \sum_i \frac{\partial}{\partial x_i}\frac{\partial G}{\partial \rho_i'} = \frac{\partial G}{\partial \rho} - \nabla \cdot \frac{\partial G}{\partial \nabla \rho} = 0; \quad \rho_i' = \frac{\partial \rho}{\partial x_i} \tag{17}$$

Extensions

Higher orders:

$$F = \int G(x, \rho, \rho', \rho'', \dots) \tag{18}$$

yields

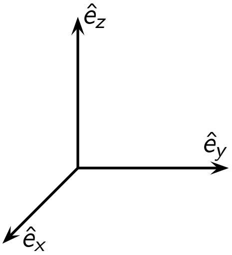
$$\frac{\partial G}{\partial \rho} - \frac{d}{dx}\frac{\partial G}{\partial \rho'} + \frac{d^2}{dx^2}\frac{\partial G}{\partial \rho''} - \dots = 0 \tag{19}$$

Unrestricted upper point:

$$\int_a^b \left(\frac{\partial G}{\partial \rho} - \frac{d}{dx}\frac{\partial G}{\partial \rho'} \right) \phi \, dx + \left[\frac{\partial G}{\partial \rho'} \phi \right]_a^b = 0 \tag{20}$$

$$\frac{\partial G}{\partial \rho} - \frac{d}{dx}\frac{\partial G}{\partial \rho'} = 0 \text{ and } \left. \frac{\partial G}{\partial \rho'} \right|_{x=b} = 0 \tag{21}$$

Vectors

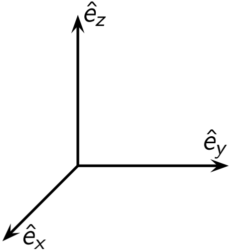


Vectors

- Coordinate system, basis vectors: $\hat{e}_x$, $\hat{e}_y$ and $\hat{e}_z$.
- n-tuples: $\mathbf{r} = (r_x, r_y, r_z)$ vs. $\mathbf{r} = r_x\hat{e}_x + r_y\hat{e}_y + r_z\hat{e}_z$
- Row vector: $\mathbf{r} = (r_x, r_y, r_z)$
- Column vector: $\mathbf{r} = (r_x, r_y, r_z)^T = \begin{pmatrix} r_x \\ r_y \\ r_z \end{pmatrix}$
- Generalisation to n dimensions: $\mathbf{r} = (r_1, r_2, \ldots r_n)^\mathsf{T}$
- Scalar product $\langle \mathbf{a}|\mathbf{b}\rangle = \mathbf{a}^\mathsf{T}\mathbf{b} = \sum_{i=1}^{n} a_i b_i$
- Norm: $||\mathbf{r}|| = \sqrt{\langle \mathbf{r}|\mathbf{r}\rangle}$ (Pythagorean theorem)
- Vector product, cross product: $\hat{e}_x \times \hat{e}_y = \hat{e}_z$ (Right handed coordinate system, only in 3D)

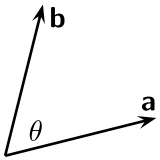
Coordinates, basis vectors

- Right handed coordinate system
- Basis vectors: $\hat{e}_x$, $\hat{e}_y$ and $\hat{e}_z$.



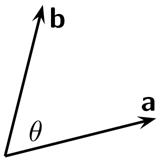
Scalar product

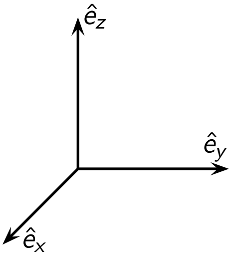
- $\mathbf{a} \cdot \mathbf{b} \stackrel{\text{def}}{=} a_x b_x + a_y b_y + a_z b_z = \mathbf{b} \cdot \mathbf{a}$
- $\mathbf{a} \cdot \mathbf{b} = \langle \mathbf{a} | \mathbf{a} \rangle = \mathbf{a}^t \mathbf{b}$
- $\mathbf{a} \cdot \mathbf{b} = |\mathbf{a}||\mathbf{b}| \cos\theta$
- $||\mathbf{a}||^2 = \mathbf{a} \cdot \mathbf{a} = a_x a_x + a_y a_y + a_z a_z$; Norm, Pythagorean theorem.



Vector product

- $\mathbf{a} \times \mathbf{b} \overset{\text{def}}{=} (a_y b_z - a_z b_y, a_z b_x - a_x b_z, a_x b_y - a_y b_x)^\mathsf{T}$
- $\mathbf{a} \times \mathbf{b} = -\mathbf{b} \times \mathbf{a} \quad \to \quad \mathbf{a} \times \mathbf{a} = 0$
- $\mathbf{a} \times \mathbf{b} \perp \mathbf{a}$ and $\mathbf{a} \times \mathbf{b} \perp \mathbf{b}$
- $|\mathbf{a} \times \mathbf{b}| = |\mathbf{a}||\mathbf{b}| \sin\theta$
- $\hat{e}_x \times \hat{e}_y = \hat{e}_z; \quad \hat{e}_y \times \hat{e}_z = \hat{e}_x; \quad \hat{e}_z \times \hat{e}_x = \hat{e}_y$





Useful relations

- $\alpha(\mathbf{a}+\mathbf{b}) = \alpha\mathbf{a} + \alpha\mathbf{b}$; distributive wrt mult. by scalar
- $\mathbf{a}+\mathbf{b} = \mathbf{b}+\mathbf{a}$; commutative addition
- $(\mathbf{a}+\mathbf{b})+\mathbf{c} = \mathbf{a}+(\mathbf{b}+\mathbf{c})$; associative addition
- $\mathbf{a}\cdot\mathbf{b} = \mathbf{b}\cdot\mathbf{a}$; commutative scalar product
- $\mathbf{a}\times\mathbf{b} = -\mathbf{b}\times\mathbf{a}$; anticommutative vector product
- $(\mathbf{a}+\mathbf{b})\cdot\mathbf{c} = \mathbf{a}\cdot\mathbf{c}+\mathbf{b}\cdot\mathbf{c}$; distributive scalar product
- $(\mathbf{a}+\mathbf{b})\times\mathbf{c} = \mathbf{a}\times\mathbf{c}+\mathbf{b}\times\mathbf{c}$; distributive vector product
- $\mathbf{a}\cdot(\mathbf{b}\times\mathbf{c}) = \mathbf{b}\cdot(\mathbf{c}\times\mathbf{a}) = \mathbf{c}\cdot(\mathbf{a}\times\mathbf{b}) = [\mathbf{a},\mathbf{b},\mathbf{c}]$; scalar triple product, CAB rule.
- $\mathbf{a}\times(\mathbf{b}\times\mathbf{c}) \neq (\mathbf{a}\times\mathbf{b})\times\mathbf{c}$; **not** associative
- $\mathbf{a}\times(\mathbf{b}\times\mathbf{c}) = (\mathbf{a}\cdot\mathbf{c})\mathbf{b} - (\mathbf{a}\cdot\mathbf{b})\mathbf{c}$; vector triple product.

Tensor notation

- a_i denote all components of a rank 1 tensor.
- a_{ij} denote all components of a rank 2 tensor.
- $a_i b_i = a_1 b_1 + a_2 b_2 + \ldots + a_n b_n$; repeated indices imply contraction (summation).
- $a_{ij} b_{jk} c_{kl} \equiv \sum_{j=1}^{n} \sum_{k=1}^{n} a_{ij} b_{jk} c_{kl} = t_{il}$
- $a_{ij} \delta_{jk} = a_{ik}$
- $\delta_{ii} = n$

Levi-Civita

The Levi-Civita tensor can be defined as

$$\begin{cases} \epsilon_{xyz} = \epsilon_{yzx} = \epsilon_{zxy} = & 1 \\ \epsilon_{xzy} = \epsilon_{zyx} = \epsilon_{yxz} = & -1 \\ \epsilon_{ijk} = & 0 \quad \text{otherwise} \end{cases} \tag{1}$$

- $\epsilon_{ijk} = -\epsilon_{jik}$; for any pair of indices.
- In general: $\epsilon_{i_1,i_2,\ldots,i_n} = (-1)^p$ where p is the number of pairwise permutations.

Useful relations

- $\mathbf{a} \cdot \mathbf{b} = a_i b_i = \delta_{ij} a_i b_j$; scalar product.
- $\mathbf{a} \times \mathbf{b} = \epsilon_{ijk} \hat{e}_i a_j b_k$; vector product.
- $(\mathbf{a} \times \mathbf{b})_i = \epsilon_{ijk} a_j b_k$; vector product.
- $(\mathbf{a} \times \mathbf{b})_z = \epsilon_{zxy} a_x b_y + \epsilon_{zyx} a_y b_x = a_x b_y - a_y b_x$
- $(\mathbf{a} \times \mathbf{b}) \cdot \mathbf{c} = (\mathbf{a} \times \mathbf{b})_i c_i = \epsilon_{ijk} a_j b_k c_i$; scalar triple product.
- $\nabla^2 \Phi = \frac{\partial^2 \Phi}{\partial x_i \partial x_i} = \delta_{ij} \frac{\partial^2 \Phi}{\partial x_i \partial x_j}$
- $\epsilon_{ijk} = -\epsilon_{jik}$; for any pair of indices.
- $\epsilon_{ijk} \epsilon_{klm} = \delta_{il} \delta_{jm} - \delta_{im} \delta_{jl}$; contracting k.
- $\epsilon_{kji} \epsilon_{klm} = -\epsilon_{ijk} \epsilon_{klm}$; etc.
- $\epsilon_{ijk} \epsilon_{ijm} = 2\delta_{km}$; contracting i, j.
- $\epsilon_{ijk} \epsilon_{ijk} = 6$; contracting i, j, k.

Differentiation of a vector

Consider a vector in Cartesian coordinates which depends on a variable u

$$\mathbf{a}(u) = a_x(u)\hat{e}_x + a_y(u)\hat{e}_y + a_z(u)\hat{e}_z \tag{1}$$

the derivative is then

$$\frac{d\mathbf{a}}{du} = \lim_{h\to 0} \frac{\mathbf{a}(u+h) - \mathbf{a}(u)}{h} \tag{2}$$

$$= \frac{da_x}{du}\hat{e}_x + \frac{da_y}{du}\hat{e}_y + \frac{da_z}{du}\hat{e}_z \tag{3}$$

Example

Consider the position of a particle as a function of time and its time derivatives

$$\mathbf{r}(t) = x(t)\hat{e}_x + y(t)\hat{e}_y + z(t)\hat{e}_z \tag{4}$$

$$\frac{d\mathbf{r}}{dt}(t) = \frac{dx}{dt}\hat{e}_x + \frac{dy}{dt}\hat{e}_y + \frac{dz}{dt}\hat{e}_z \tag{5}$$

$$\frac{d^2\mathbf{r}}{dt^2}(t) = \frac{d^2x}{dt^2}\hat{e}_x + \frac{d^2y}{dt^2}\hat{e}_y + \frac{d^2z}{dt^2}\hat{e}_z \tag{6}$$

where we have

- Position: $\mathbf{r}(t)$
- Velocity: $\mathbf{v}(t) = \frac{d\mathbf{r}}{dt} = \dot{\mathbf{r}}(t)$
- Acceleration: $\mathbf{a}(t) = \frac{d^2\mathbf{r}}{dt^2} = \ddot{\mathbf{r}}(t)$

Composite expressions

$$\frac{d}{du}(\phi \mathbf{a}) = \phi \frac{d\mathbf{a}}{du} + \frac{d\phi}{du}\mathbf{a} \tag{7}$$

$$\frac{d}{du}(\mathbf{a} \cdot \mathbf{b}) = \mathbf{a} \cdot \frac{d\mathbf{b}}{du} + \frac{d\mathbf{a}}{du} \cdot \mathbf{b} \tag{8}$$

$$\frac{d}{du}(\mathbf{a} \times \mathbf{b}) = \mathbf{a} \times \frac{d\mathbf{b}}{du} + \frac{d\mathbf{a}}{du} \times \mathbf{b} \tag{9}$$

$$\frac{d}{du}(\mathbf{a} \cdot \mathbf{b}) = \frac{d}{du}(a_x b_x + a_y b_y + a_z b_z) \tag{10}$$

$$= a_x \frac{db_x}{du} + \frac{da_x}{du} b_x + \ldots \tag{11}$$

$$= \mathbf{a} \cdot \frac{d\mathbf{b}}{du} + \frac{d\mathbf{a}}{du} \cdot \mathbf{b} \tag{12}$$

Vector operators

- $\phi(x, y, z)$; scalar field (f.x. electrostatic potential)
- $\mathbf{a}(x, y, z)$; vector field (f.x. electric field)
- $\nabla \equiv \hat{e}_x \frac{\partial}{\partial x} + \hat{e}_y \frac{\partial}{\partial y} + \hat{e}_z \frac{\partial}{\partial z}$; vector operator (nabla).

$$\nabla = \begin{pmatrix} \frac{\partial}{\partial x} \\ \frac{\partial}{\partial y} \\ \frac{\partial}{\partial z} \end{pmatrix}; \quad \nabla \phi = \begin{pmatrix} \frac{\partial \phi}{\partial x} \\ \frac{\partial \phi}{\partial y} \\ \frac{\partial \phi}{\partial z} \end{pmatrix} \tag{13}$$

Grad, div, curl

$$\begin{aligned}
\operatorname{grad}\phi &= \nabla\phi &&= \frac{\partial\phi}{\partial x}\hat{e}_x + \frac{\partial\phi}{\partial y}\hat{e}_y + \frac{\partial\phi}{\partial z}\hat{e}_z \\
\operatorname{div}\mathbf{a} &= \nabla\cdot\mathbf{a} &&= \frac{\partial a_x}{\partial x} + \frac{\partial a_y}{\partial y} + \frac{\partial a_z}{\partial z} \\
\operatorname{curl}\mathbf{a} &= \nabla\times\mathbf{a} &&= \left(\frac{\partial a_z}{\partial y} - \frac{\partial a_y}{\partial z}\right)\hat{e}_x \\
& && + \left(\frac{\partial a_x}{\partial z} - \frac{\partial a_z}{\partial x}\right)\hat{e}_y \\
& && + \left(\frac{\partial a_y}{\partial x} - \frac{\partial a_x}{\partial y}\right)\hat{e}_z
\end{aligned} \tag{14}$$

Combining grad, div, curl

$$
\begin{array}{llcl}
\text{div grad} & \nabla\cdot(\nabla\phi) & = & \frac{\partial^2\phi}{\partial x^2}+\frac{\partial^2\phi}{\partial y^2}+\frac{\partial^2\phi}{\partial z^2} \\
\text{curl grad} & \nabla\times(\nabla\phi) & = & 0 \\
\text{grad div} & \nabla(\nabla\cdot\mathbf{a}) & = & (\frac{\partial^2 a_x}{\partial x^2}+\frac{\partial^2 a_y}{\partial x\partial y}+\frac{\partial^2 a_z}{\partial x\partial z})\hat{e}_x \\
 & & + & (\frac{\partial^2 a_y}{\partial y^2}+\frac{\partial^2 a_z}{\partial y\partial z}+\frac{\partial^2 a_x}{\partial y\partial x})\hat{e}_y \\
 & & + & (\frac{\partial^2 a_z}{\partial z^2}+\frac{\partial^2 a_x}{\partial z\partial x}+\frac{\partial^2 a_y}{\partial z\partial y})\hat{e}_z \\
\text{div curl} & \nabla\cdot(\nabla\times\mathbf{a}) & = & 0 \\
\text{curl curl} & \nabla\times(\nabla\times\mathbf{a}) & = & \nabla(\nabla\cdot\mathbf{a})-\nabla^2\mathbf{a}
\end{array}
\tag{15}
$$

Useful relations

$$
\begin{aligned}
\nabla(\phi + \psi) &= \nabla\phi + \nabla\psi \\
\nabla \cdot (\mathbf{a} + \mathbf{b}) &= \nabla \cdot \mathbf{a} + \nabla \cdot \mathbf{b} \\
\nabla \times (\mathbf{a} + \mathbf{b}) &= \nabla \times \mathbf{a} + \nabla \times \mathbf{b} \\
\nabla(\phi\psi) &= \psi\nabla\phi + \phi\nabla\psi \\
\nabla \cdot (\phi\mathbf{a}) &= \phi\nabla \cdot \mathbf{a} + \mathbf{a} \cdot \nabla\phi \\
\nabla \cdot (\mathbf{a} \times \mathbf{b}) &= \mathbf{b} \cdot (\nabla \times \mathbf{a}) - \mathbf{a} \cdot (\nabla \times \mathbf{b})
\end{aligned} \tag{16}
$$

Line integrals

Consider an integral from point A to point B along the curve C:

- $\int_C \phi \, d\mathbf{r}$
- $\int_C \mathbf{a} \cdot d\mathbf{r}$
- $\int_C \mathbf{a} \times d\mathbf{r}$

Introduce a parametrization

$$C = \{\mathbf{r}(u); u_0 \leq u \leq u_1\} \tag{17}$$

$$d\mathbf{r} = (dx, dy, dz) = \left(\frac{dx}{du}, \frac{dy}{du}, \frac{dz}{du}\right) du \tag{18}$$

$$\int_C \mathbf{a} \cdot d\mathbf{r} = \int_C a_x dx + a_y dy + a_z dz = \tag{19}$$

$$\int_{u_0}^{u_1} \left[a_x(u)\frac{dx}{du} + a_y(u)\frac{dy}{du} + a_z(u)\frac{dz}{du}\right] du \tag{20}$$

Green's theorem in a plane

Consider the functions $P(x,y)$ and $Q(x,y)$ that are continuous with continuous partial derivatives in a simply connected region R (no holes). The curve C is the boundary of this region, then

$$\oint_C (P(x,y)dx + Q(x,y)dy) = \int\int_R \left(\frac{\partial Q}{\partial x} - \frac{\partial P}{\partial y}\right) dx\, dy \tag{21}$$

Surface/volume integrals

Consider a surface S in three dimensions. We can define the surface integrals

- $\int_S \phi \, dS$
- $\int_S \phi \, d\mathbf{S}$
- $\int_S \mathbf{a} \cdot d\mathbf{S}$
- $\int_S \mathbf{a} \times d\mathbf{S}$

where $d\mathbf{S}$ is a vector with the magnitude of the area element and the direction perpendiculer to the surface,

$$d\mathbf{S} = \hat{n} dS \tag{22}$$

Introduce a parametrization

$$S = \{\mathbf{r}(u, v); u_0 \leq u \leq u_1; v_0 \leq v \leq v_1\} \tag{23}$$

Consider the volume V. We can define the volume integrals

- $\int_V \phi \, dV$
- $\int_V \mathbf{a} \, dV$

Some theorems

The divergence theorem:

$$\int_V \nabla \cdot \mathbf{a} \, dV = \oint_S \mathbf{a} \cdot d\mathbf{S} \tag{24}$$

Stokes' theorem:

$$\int_S (\nabla \times \mathbf{a}) \cdot d\mathbf{S} = \oint_C \mathbf{a} \cdot d\mathbf{r} \tag{25}$$

Example — what is $\nabla^2 \frac{1}{r}$

$$
\begin{aligned}
\nabla \frac{1}{r} &= \left(\hat{e}_x \frac{\partial}{\partial x} + \hat{e}_y \frac{\partial}{\partial y} + \hat{e}_z \frac{\partial}{\partial x}\right) \{x^2 + y^2 + z^2\}^{-1/2} \\
&= -\tfrac{1}{2}\{x^2 + y^2 + z^2\}^{-3/2}(2x\hat{e}_x + 2y\hat{e}_y + 2z\hat{e}_z) \\
&= -\frac{\mathbf{r}}{r^3}
\end{aligned}
\tag{26}
$$

$$
\begin{aligned}
\nabla^2 \frac{1}{r} &= -\nabla \cdot \frac{\mathbf{r}}{r^3} \\
&= -\left(\frac{\partial}{\partial x}\frac{x}{r^3} + \frac{\partial}{\partial y}\frac{y}{r^3} + \frac{\partial}{\partial z}\frac{z}{r^3}\right) \\
&= -\frac{3}{r^3} + \frac{3}{2}\left(\frac{2x^2}{r^5} + \frac{2y^2}{r^5} + \frac{2z^2}{r^5}\right) \\
&= -\frac{3}{r^3} + \frac{3}{r^3} \\
&= 0
\end{aligned}
\tag{27}
$$

What about $\nabla^2 \frac{1}{r}$ at $r = 0$?

Example — what is $\nabla^2 \frac{1}{r}$

What about $\nabla^2 \frac{1}{r}$ at $r = 0$? Use divergence theorem with sphere of radius R.

$$\int_V \nabla \cdot \mathbf{a} \, dV = \oint_S \mathbf{a} \cdot d\mathbf{S} \tag{28}$$

$$\begin{aligned} \int_V (\nabla \cdot \nabla \tfrac{1}{r}) \, dV &= \oint_S \nabla \tfrac{1}{r} \cdot d\mathbf{S} \\ &= -\oint_S \tfrac{\mathbf{r}}{r^3} \cdot \tfrac{\mathbf{r}}{r} \, dS \\ &= -\oint_S \tfrac{1}{r^2} \, dS \\ &= -\tfrac{1}{R^2} 4\pi R^2 \\ &= -4\pi \end{aligned} \tag{29}$$

so we get

$$\int_V \nabla^2 \frac{1}{r} \, dV = -4\pi \tag{30}$$

If this is true then we must have that

$$\nabla^2 \frac{1}{r} = -4\pi \delta(\mathbf{r}) \tag{31}$$

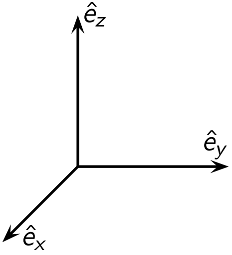
Vector spaces

Vector space: A set of objects called vectors, ($\mathbf{a}$, $\mathbf{b}$, $\mathbf{c}$) with an addition and a multiplication with scalars, real or complex, (α, β) subject to the conditions:

1. $\mathbf{a} + \mathbf{b} = \mathbf{b} + \mathbf{a}$; commutative addition.
2. $(\mathbf{a} + \mathbf{b}) + \mathbf{c} = \mathbf{a} + (\mathbf{b} + \mathbf{c})$; associative addition.
3. $\mathbf{a} + \mathbf{0} = \mathbf{a}$; existence of null vector (identity).
4. $\mathbf{a} + (-\mathbf{a}) = \mathbf{0}$; existence of inverse.
5. $(\alpha + \beta)\mathbf{a} = \alpha\mathbf{a} + \beta\mathbf{a}$; distributive.
6. $\alpha(\mathbf{a} + \mathbf{b}) = \alpha\mathbf{a} + \alpha\mathbf{b}$; distributive.
7. $\alpha(\beta\mathbf{a}) = (\alpha\beta)\mathbf{a}$; compatibility.
8. $1\,\mathbf{a} = \mathbf{a}$; multiplication with one.

Basis vectors and dimensionallity

- If $c_1\mathbf{x_1} + c_2\mathbf{x_2} + \ldots + c_n\mathbf{x_n} = \mathbf{0}$ is fulfilled only if all $c_i = 0$, then the vectors are linearly independent.
- N linearly independent vectors **span** a vector space of dimension N.
- If there are N linearly independent vectors but not $N+1$, the vector space is said to be N-dimensional.
- There needs to be N linearly independent basis vectors to **span** a N-dimensional vector space.



Examples

- Vectors in two or three dimensions such as forces or velocities.
- The space of ordered pairs of numbers such that
  - $(x_1, y_1) + (x_2, y_2) = (x_1 + x_2, y_1 + y_2)$ and
  - $\alpha(x, y) = (\alpha x, \alpha y)$.
- Complex numbers, basically the same as above.
- The space of all functions $f(x) = \sum_{k=1}^{\infty} c_k \sin(kx)$ on the interval $[0, \pi]$.

Normed vector spaces

We can add a norm to a vector space, $||\mathbf{u}||$ such that:

- $||\mathbf{u}|| \geq 0$, and $||\mathbf{u}|| = 0$ if and only if $\mathbf{u} = \mathbf{0}$
- $||\alpha \mathbf{u}|| = |\alpha| \, ||\mathbf{u}||$
- $||\mathbf{u} + \mathbf{v}|| \leq ||\mathbf{u}|| + ||\mathbf{v}||$ (triangle inequality)

A few norms for $\mathbb{R}^3$:

- $||\mathbf{u}||_1 = |u_1| + |u_2| + |u_3|$
- $||\mathbf{u}||_\infty = \max\{|u_1|, |u_2|, |u_3|\}$
- $||\mathbf{u}||_2 = \sqrt{u_1^2 + u_2^2 + u_3^2}$ (Euclidian norm)
- $||\mathbf{u}||_p = (|u_1|^p + |u_2|^p + |u_3|^p)^{1/p}$ (ℓ^p norm)

Inner product

Define an **inner product** that takes two vectors and form a scalar such that

- $\langle \mathbf{u}|\mathbf{v}\rangle = \langle \mathbf{v}|\mathbf{u}\rangle^*$
- $\langle \mathbf{u}|\alpha\mathbf{v} + \beta\mathbf{x}\rangle = \alpha\langle \mathbf{u}|\mathbf{v}\rangle + \beta\langle \mathbf{u}|\mathbf{x}\rangle$

which implies

- $\langle \alpha\mathbf{u} + \beta\mathbf{v}|\mathbf{x}\rangle = \alpha^*\langle \mathbf{u}|\mathbf{x}\rangle + \beta^*\langle \mathbf{v}|\mathbf{x}\rangle$
- $\langle \alpha\mathbf{u}|\beta\mathbf{v}\rangle = \alpha^*\beta\langle \mathbf{u}|\mathbf{v}\rangle$.

For example

- $\langle \mathbf{u}|\mathbf{v}\rangle = u_1 v_1 + u_2 v_2 + u_3 v_3$ (for $\mathbb{R}^3$)
- $\langle \mathbf{u}|\mathbf{v}\rangle = u_1^* v_1 + u_2^* v_2 + u_3^* v_3$ (for $\mathbb{C}^3$)
- $\langle \Psi_1|\Psi_2\rangle = \int \Psi_1^*\Psi_2\, dV$ (overlap of wavefunctions)

Inner product spaces

Define a compatible norm as

- $||\mathbf{u}|| = \langle \mathbf{u}|\mathbf{u}\rangle^{1/2}$

for example

- $||\mathbf{u}|| = \sqrt{u_1 u_1 + u_2 u_2 + u_3 u_3}$ (for $\mathbb{R}^3$)
- $||\mathbf{u}|| = \sqrt{u_1^* u_1 + u_2^* u_2 + u_3^* u_3}$ (for $\mathbb{C}^3$)
- $||\Psi|| = \sqrt{\int \Psi^* \Psi \, dV}$ (norm of wavefunction)

Define orthogonality as

- $\langle \mathbf{u}|\mathbf{v}\rangle = 0$

A set of base vectors that fulfill

- $\langle \mathbf{u}_i|\mathbf{u}_j\rangle = \delta_{ij}$

is called an *orthonormal* basis.

Some useful relations

- Schwartz's inequality: $|\langle \mathbf{u}|\mathbf{v}\rangle| \leq ||\mathbf{u}||\,||\mathbf{v}||$ and equality only if $\mathbf{u} = \alpha \mathbf{v}$.
- The triangle inequality: $||\mathbf{u}+\mathbf{v}|| \leq ||\mathbf{u}|| + ||\mathbf{v}||$
- Bessels inequality: $||\mathbf{u}||^2 \geq \sum_i |\langle \hat{\mathbf{e}}_i|\mathbf{u}\rangle|^2$ where $\hat{\mathbf{e}}_i$ is a set of orthonormal basis vectors.
- The parallelogram equality: $||\mathbf{u}+\mathbf{v}||^2 + ||\mathbf{u}-\mathbf{v}||^2 = 2(||\mathbf{u}||^2 + ||\mathbf{v}||^2)$.

Hilbert spaces

Cauchy sequence: A sequence of vectors, $\{x_i\}_{i=1}^{\infty}$, is called a Cauchy sequence if for every small ϵ there is a finite integer N such that $||x_m - x_n|| < \epsilon$ for $n > N$ and $m > N$.

Complete space: A vector space is complete if any Cauchy sequence converges to an element in the vector space.

Hilbert space: A complete inner product space is called a Hilbert space.

Sobolev spaces

L^2 space: A vector space with all functions that satisfy $\int |f|^2 \, d\tau < \infty$.

Sobolev spaces: A vector space where the function and the derivative up to a given order lie in the L^2 space.

Some vector spaces

- $\mathbb{R}^3$, all points in 3-dimensional space.
- All infinite sequences of real or complex numbers, $\{c_i\}_{i=1}^{\infty}$, such that $\sum_i |c_i|^2 < \infty$.
- All functions f such that $\int f^* f \, d\tau < \infty$ with the inner product $\langle f|g\rangle = \int f^* g \, d\tau$.
- All functions (orbitals) ϕ that can be formed from a basis set $\{\chi_i\}_{i=1}^{n}$, $\phi = \sum_i c_i \chi_i$ with the inner product above.
- The space of coefficients c_i above. Note that $\{\chi_i\}_{i=1}^{n}$ is normally a nonorthogonal basis so we get the inner product $\langle c^{(1)}|c^{(2)}\rangle = \sum_{ij} c_i^{(1)} \langle \chi_i|\chi_j\rangle c_j^{(2)} = \sum_{ij} c_i^{(1)} S_{ij} c_j^{(2)}$

Linear operators

A linear operator $\hat{A}$ is a mapping that maps an element x in vector space V to an element $z = \hat{A}x$ in vector space V' in such a way that

$$\hat{A}(\alpha\mathbf{x} + \beta\mathbf{y}) = \alpha\hat{A}\mathbf{x} + \beta\hat{A}\mathbf{y} \tag{1}$$

where both $\hat{A}\mathbf{x}$ and $\hat{A}\mathbf{y}$ are members in V'. V and V' can be the same vector space.

Hermitian adjoint operator

Consider a linear operator $\hat{A}$ mapping vectors from one Hilbert space H_1 to another Hilbert space H_2. The adjoint operator $\hat{A}^\dagger$ then maps from H_2 to H_1 in such a way that

$$\langle h_2|\hat{A}h_1\rangle_{H_2} = \langle \hat{A}^\dagger h_2|h_1\rangle_{H_1} \tag{2}$$

Let $H_1 = H_2 = H$ be the Hilbert space of all functions $\int f^* f \, d\tau < \infty$ and the inner product $\langle h_2|h_1\rangle = \int h_2^* h_1 \, d\tau$

$$\langle h_2|\hat{A}h_1\rangle = \langle \hat{A}^\dagger h_2|h_1\rangle \tag{3}$$

$\hat{A}^\dagger$ is called the Hermitian adjoint operator.

Adjoint operator

Consider the operator $\hat{A} = \frac{d}{dx}$ for functions on the interval $(-\infty, \infty)$ such that $\int f^* f \, dx < \infty$.

$$\langle f | \hat{A} g \rangle = \int f^* g' \, dx \tag{4}$$

$$= [f^* g]_{-\infty}^{\infty} - \int (f')^* g \, dx \tag{5}$$

$$= \langle (-\hat{A}) f | g \rangle \tag{6}$$

Thus $\hat{A}^\dagger = -\hat{A}$ is the adjoint operator.

Hermitian operator

Consider the operator $\hat{A} = i\frac{d}{dx}$ for functions on the interval $(-\infty, \infty)$ such that $\int f^* f \, dx < \infty$.

$$\begin{aligned}
\langle f|\hat{A}g\rangle &= \int f^*(ig')\,dx && (7)\\
&= i\,[f^*g]_{-\infty}^{\infty} - i\int (f')^* g\,dx && (8)\\
&= \int (if')^* g\,dx && (9)\\
&= \langle \hat{A}f|g\rangle && (10)
\end{aligned}$$

Thus $\hat{A}^\dagger = \hat{A}$ is a self-adjoint or Hermitian operator.

Self adjoint?

Consider the operator $\hat{A} = i\frac{d}{dx}$ for functions on the interval $[0, 1]$.

$$\begin{aligned}
\langle f|\hat{A}g\rangle &= \int f^*(ig')\,dx & (11)\\
&= i\,[f^*g]_0^1 - i\int (f')^* g\,dx & (12)\\
&= i\,[f^*g]_0^1 + \int (if')^* g\,dx & (13)\\
&= \langle \hat{A}f|g\rangle + i\,\{f^*(1)g(1) - f^*(0)g(0)\} & (14)
\end{aligned}$$

Self-adjoint/Hermitian? Depends on boundary conditions.

Some Hermitian operators

With proper boundary conditions

Linear momentum: $\hat{p}_x = -i\hbar\frac{d}{dx}$.

Angular momentum: $\hat{L}^2$

Spin: $\hat{S}^2$

Position: $\hat{x} = x\cdot$

Potential energy: $\hat{V} = V(x)\cdot$

Hamiltonian: $\hat{H} = \frac{\hat{p}\cdot\hat{p}}{2m} + \hat{V}$

An n by m matrix

A rectangular array of numbers with n rows and m columns, dimensions $n \times m$.

$$A = \begin{pmatrix} a_{11} & a_{12} & a_{13} & \cdots & a_{1m} \\ a_{21} & a_{22} & a_{23} & \cdots & a_{2m} \\ a_{31} & a_{32} & a_{33} & \cdots & a_{3m} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & a_{n3} & \cdots & a_{nm} \end{pmatrix} \tag{1}$$

A square n by n matrix

A rectangular array with equal number of rows and columns is called a *square* matrix.

$$A = \begin{pmatrix} a_{11} & a_{12} & a_{13} & \cdots & a_{1n} \\ a_{21} & a_{22} & a_{23} & \cdots & a_{2n} \\ a_{31} & a_{32} & a_{33} & \cdots & a_{3n} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & a_{n3} & \cdots & a_{nn} \end{pmatrix} \tag{2}$$

Basic algebra: addition

Addition: $A = B + C$

$$\begin{pmatrix} b_{11} & b_{12} & b_{13} \\ b_{21} & b_{22} & b_{23} \\ b_{31} & b_{32} & b_{33} \end{pmatrix} + \begin{pmatrix} c_{11} & c_{12} & c_{13} \\ c_{21} & c_{22} & c_{23} \\ c_{31} & c_{32} & c_{33} \end{pmatrix} = \tag{3}$$

$$\begin{pmatrix} a_{11} = b_{11} + c_{11} & a_{12} = b_{12} + c_{12} & a_{13} = b_{13} + c_{13} \\ a_{21} = b_{21} + c_{21} & a_{22} = b_{22} + c_{22} & a_{23} = b_{23} + c_{23} \\ a_{31} = b_{31} + c_{31} & a_{32} = b_{32} + c_{32} & a_{33} = b_{33} + c_{33} \end{pmatrix} \tag{4}$$

- Associative: $(A + B) + C = A + (B + C)$
- Commutative: $A + B = B + A$
- Matching dimensions: A and B are $n \times m$.

Basic algebra: multiplication

Multiplication: $A = BC$

$$\begin{pmatrix} \dots & \dots & \dots \\ \dots & \dots & \dots \\ b_{31} & b_{32} & b_{33} \end{pmatrix} \begin{pmatrix} \dots & c_{12} & \dots \\ \dots & c_{22} & \dots \\ \dots & c_{32} & \dots \end{pmatrix} = \tag{5}$$

$$\begin{pmatrix} \dots & \dots & \dots \\ \dots & \dots & \dots \\ \dots & b_{31}c_{12} + b_{32}c_{22} + b_{33}c_{32} & \dots \end{pmatrix} \tag{6}$$

- $a_{ij} = \sum_k b_{ik} c_{kj}$
- Associative: $(AB)C = A(BC)$
- NOT commutative: $AB \neq BA$, but may be under certain conditions.
- Matching dimensions: A is $n \times m$ thus B is $n \times l$ and C is $l \times m$.

Example — matrix multiply

$$\begin{pmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} \cos\theta\, x_1 + \sin\theta\, x_2 \\ -\sin\theta\, x_1 + \cos\theta\, x_2 \end{pmatrix}$$

$$\begin{pmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{pmatrix} \begin{pmatrix} \cos\phi & \sin\phi \\ -\sin\phi & \cos\phi \end{pmatrix} =$$

$$\begin{pmatrix} \cos\theta\cos\phi - \sin\theta\sin\phi & \cos\theta\sin\phi + \sin\theta\cos\phi \\ -\sin\theta\cos\phi - \cos\theta\sin\phi & -\sin\theta\sin\phi + \cos\theta\cos\phi \end{pmatrix} =$$

$$\begin{pmatrix} \cos(\theta+\phi) & \sin(\theta+\phi) \\ -\sin(\theta+\phi) & \cos(\theta+\phi) \end{pmatrix}$$

Basic algebra: scaling

$$\lambda \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix} = \begin{pmatrix} \lambda a_{11} & \lambda a_{12} & \lambda a_{13} \\ \lambda a_{21} & \lambda a_{22} & \lambda a_{23} \\ \lambda a_{31} & \lambda a_{32} & \lambda a_{33} \end{pmatrix} \quad (7)$$

- Distributive: $\lambda(A+B) = \lambda A + \lambda B$, $(\lambda+\mu)A = \lambda A + \mu A$.
- Associative: $(\lambda\mu)A = \lambda(\mu A)$

The null and identity matrices

The null matrix: $A0 = 0A = 0$ and $A + 0 = 0 + A = A$.

$$0 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \tag{8}$$

The identity matrix: $AI = IA = A$.

$$I = E = 1 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \tag{9}$$

Functions of matrices

Use Taylor expansion as definition for *square* matrices, for example

$$e^A = \sum_k \frac{1}{k!} A^k \tag{10}$$

Transpose of a matrix

Transpose is the interchange of rows and columns

$$A = \begin{pmatrix} a_{11} & a_{12} & a_{13} & \cdots & a_{1m} \\ a_{21} & a_{22} & a_{23} & \cdots & a_{2m} \\ a_{31} & a_{32} & a_{33} & \cdots & a_{3m} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & a_{n3} & \cdots & a_{nm} \end{pmatrix} \tag{11}$$

$$A^T = \begin{pmatrix} a_{11} & a_{21} & a_{31} & \cdots & a_{n1} \\ a_{12} & a_{22} & a_{32} & \cdots & a_{n2} \\ a_{13} & a_{23} & a_{33} & \cdots & a_{n2} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ a_{1m} & a_{2m} & a_{3m} & \cdots & a_{mn} \end{pmatrix} \tag{12}$$

$(AB)^T = B^T A^T$

Complex and Hermitian conjugate

Complex conjugate

$$A^* = \begin{pmatrix} a_{11}^* & a_{12}^* & a_{13}^* & \cdots & a_{1m}^* \\ a_{21}^* & a_{22}^* & a_{23}^* & \cdots & a_{2m}^* \\ a_{31}^* & a_{32}^* & a_{33}^* & \cdots & a_{3m}^* \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ a_{n1}^* & a_{n2}^* & a_{n3}^* & \cdots & a_{nm}^* \end{pmatrix} \qquad (13)$$

Hermitian conjugate

$$A^\dagger = (A^T)^* = (A^*)^T = \begin{pmatrix} a_{11}^* & a_{21}^* & a_{31}^* & \cdots & a_{n1}^* \\ a_{12}^* & a_{22}^* & a_{32}^* & \cdots & a_{n2}^* \\ a_{13}^* & a_{23}^* & a_{33}^* & \cdots & a_{n2}^* \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ a_{1m}^* & a_{2m}^* & a_{3m}^* & \cdots & a_{mn}^* \end{pmatrix} \qquad (14)$$

$(AB)^\dagger = B^\dagger A^\dagger$

Example — transpose and Hermitian conjugate

$$\begin{pmatrix} 1+i & 1+2i \\ 2+i & 2+2i \\ 3+i & 3+2i \end{pmatrix}^T = \begin{pmatrix} 1+i & 2+i & 3+i \\ 1+2i & 2+2i & 3+2i \end{pmatrix}$$

$$\begin{pmatrix} 1+i & 1+2i \\ 2+i & 2+2i \\ 3+i & 3+2i \end{pmatrix}^\dagger = \begin{pmatrix} 1-i & 2-i & 3-i \\ 1-2i & 2-2i & 3-2i \end{pmatrix}$$

Determinants

$$|A| = \begin{vmatrix} a_{11} & a_{12} & a_{13} & \cdots & a_{1n} \\ a_{21} & a_{22} & a_{23} & \cdots & a_{2n} \\ a_{31} & a_{32} & a_{33} & \cdots & a_{3n} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & a_{n3} & \cdots & a_{nn} \end{vmatrix} \tag{15}$$

$$|A| = \sum_{\sigma \in S_n} \mathrm{sgn}(\sigma) a_{1,\sigma_1} a_{2,\sigma_2} \ldots a_{n,\sigma_n} \tag{16}$$

where S_n is the set of all permutations of the numbers $\{1, 2, \ldots, n\}$ and $\mathrm{sgn}(\sigma) = -1$ for an odd number of pairwise permutations while $\mathrm{sgn}(\sigma) = +1$ for an even number of pairwise numbers. For $n = 2$ we have

$$|A| = \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} = a_{11}a_{22} - a_{12}a_{21} \tag{17}$$

Determinants — recursive definition

$$\begin{vmatrix} a_{11} & a_{12} & a_{13} & \cdots & a_{1n} \\ a_{21} & a_{22} & a_{23} & \cdots & a_{2n} \\ a_{31} & a_{32} & a_{33} & \cdots & a_{3n} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & a_{n3} & \cdots & a_{nn} \end{vmatrix} = a_{11} \overbrace{\begin{vmatrix} a_{22} & a_{23} & \cdots & a_{2n} \\ a_{32} & a_{33} & \cdots & a_{3n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n2} & a_{n3} & \cdots & a_{nn} \end{vmatrix}}^{\text{remove row 1 and col 1}}$$

$$-a_{12} \overbrace{\begin{vmatrix} a_{21} & a_{23} & \cdots & a_{2n} \\ a_{31} & a_{33} & \cdots & a_{3n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n3} & \cdots & a_{nn} \end{vmatrix}}^{\text{remove row 1 and col 2}} + a_{13} \overbrace{\begin{vmatrix} a_{21} & a_{22} & \cdots & a_{2n} \\ a_{31} & a_{32} & \cdots & a_{3n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{vmatrix}}^{\text{remove row 1 and col 3}} - \cdots \tag{18}$$

Determinant 3×3

$$\begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} = a_{11} \begin{vmatrix} a_{22} & a_{23} \\ a_{32} & a_{33} \end{vmatrix} - a_{12} \begin{vmatrix} a_{21} & a_{23} \\ a_{31} & a_{33} \end{vmatrix}$$

$$+a_{13} \begin{vmatrix} a_{21} & a_{22} \\ a_{31} & a_{32} \end{vmatrix} = \tag{19}$$

$$a_{11}a_{22}a_{33} - a_{11}a_{23}a_{32} - a_{12}a_{21}a_{33} +$$

$$a_{12}a_{23}a_{31} + a_{13}a_{21}a_{32} - a_{13}a_{22}a_{31} \tag{20}$$

Determinants — some properties

- $|A|$ is the product of the eigenvalues
- $|A^T| = |A|$, $|A^*| = |A|^*$, $|A^\dagger| = |(A^*)^T| = |A^*| = |A|^*$
- Interchange of two rows or columns will change the sign.
- $|\lambda A| = \lambda^n |A|$
- Linear dependence in rows or columns $\rightarrow \quad |A| = 0$.
- Identical rows or columns $\rightarrow \quad |A| = 0$.
- Add one row to another row does not change the value of the determinant.
- $|AB| = |A|\,|B|$

Example

- The determinant is the product of the eigenvalues.
- The trace of a matrix (the sum of the diagonal elements) is the sum of the eigenvalues.

$$\begin{vmatrix} 1 & 1 \\ 1 & 1 \end{vmatrix} = 0; \quad \mathrm{tr}\begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix} = 2$$

$$\lambda_1 + \lambda_2 = 2; \quad \lambda_1 \times \lambda_2 = 0; \quad \lambda_{1,2} = 0, 2$$

$$\Psi_{\mathrm{HF}} = \frac{1}{\sqrt{n!}} \begin{vmatrix} \phi_1(1) & \phi_2(1) & \phi_3(1) & \cdots & \phi_n(1) \\ \phi_1(2) & \phi_2(2) & \phi_3(2) & \cdots & \phi_n(2) \\ \phi_1(3) & \phi_2(3) & \phi_3(3) & \cdots & \phi_n(3) \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \phi_1(n) & \phi_2(n) & \phi_3(n) & \cdots & \phi_n(n) \end{vmatrix}$$

Inverse

A *square* matrix may have an inverse such that

$$A^{-1}A = AA^{-1} = I \tag{21}$$

Using the properties of a determinant

$$1 = |I| = |A^{-1}A| = |A|\,|A^{-1}| \tag{22}$$

so $|A| \neq 0$ is a necessary and also sufficient condition for an inverse to exist. If $|A| = 0$ the matrix is called *singular*.

Rank

The rank of a $n \times m$ matrix is given by the number of linearly independent vectors v_i in A. It is also given by the number of linearly independent vectors w_k in A.

$$A = \begin{pmatrix} \uparrow & \uparrow & & \uparrow \\ v_1 & v_2 & \dots & v_m \\ \downarrow & \downarrow & & \downarrow \end{pmatrix} \tag{23}$$

$$A = \begin{pmatrix} \leftarrow & w_1 & \rightarrow \\ \leftarrow & w_2 & \rightarrow \\ & \dots & \\ \leftarrow & w_n & \rightarrow \end{pmatrix} \tag{24}$$

Example — a rank $m < n$ matrix

A rank 1 matrix.

$$c\,c^{\mathsf{T}} = \begin{pmatrix} c_1 \\ c_2 \\ c_3 \end{pmatrix} \begin{pmatrix} c_1 & c_2 & c_3 \end{pmatrix} = \begin{pmatrix} c_1 c_1 & c_1 c_2 & c_1 c_3 \\ c_2 c_1 & c_2 c_2 & c_2 c_3 \\ c_3 c_1 & c_3 c_2 & c_3 c_3 \end{pmatrix}$$

A HF density matrix for m occupied orbitals has rank m

$$D = \sum_{i=1}^{m} \eta c_i\, c_i^{\mathsf{T}}$$

Special matrices

Diagonal:

$$\begin{pmatrix} a_{11} & 0 & 0 & 0 \\ 0 & a_{22} & 0 & 0 \\ 0 & 0 & a_{33} & 0 \\ 0 & 0 & 0 & a_{44} \end{pmatrix} \tag{25}$$

Tridiagonal

$$\begin{pmatrix} a_{11} & a_{12} & 0 & 0 \\ a_{21} & a_{22} & a_{23} & 0 \\ 0 & a_{32} & a_{33} & a_{34} \\ 0 & 0 & a_{43} & a_{44} \end{pmatrix} \tag{26}$$

Special matrices

Lower/upper triangular

$$\begin{pmatrix} a_{11} & 0 & 0 & 0 \\ a_{21} & a_{22} & 0 & 0 \\ a_{31} & a_{32} & a_{33} & 0 \\ a_{41} & a_{42} & a_{43} & a_{44} \end{pmatrix} \begin{pmatrix} a_{11} & a_{12} & a_{13} & a_{14} \\ 0 & a_{22} & a_{23} & a_{24} \\ 0 & 0 & a_{33} & a_{34} \\ 0 & 0 & 0 & a_{44} \end{pmatrix} \tag{27}$$

Special matrices

- Symmetric and antisymmetric: $S^T = S$ and $A^T = -A$. Any matrix $X = S + A$.
- Hermitian and antihermitian: $H^\dagger = H$ and $A^\dagger = -A$. Any matrix $X = H + A$.
- Orthogonal: $O^{-1} = O^T$
- Unitary: $U^{-1} = U^\dagger$.

Change of basis — similarity transform

Consider a new basis

$$\mathbf{e}'_j = \sum_i S_{ij}\mathbf{e}_i \tag{28}$$

How does the representation of vector **u** change?

$$\mathbf{u} = \sum_i x_i\mathbf{e}_i = \sum_j x'_j\mathbf{e}'_j = \sum_i \Big(\sum_j S_{ij}x'_j\Big)\mathbf{e}_i \tag{29}$$

$$x_i = \sum_j S_{ij}x'_j; \quad x = Sx'; \quad x' = S^{-1}x \tag{30}$$

Consider a matrix vector multiplication in original coordinates $y = Ax$ and in transformed coordinates $y' = A'x'$.

$$y = Ax; \quad Sy' = ASx'; \quad y' = S^{-1}ASx' \tag{31}$$

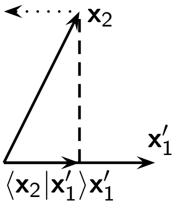
thus

$$A' = S^{-1}AS \tag{32}$$

Gram-Schmidt orthonormalization

Start with a set of vectors, $\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_n$

1. Normalize: $\mathbf{x}'_1 = \mathbf{x}_1/||\mathbf{x}_1||$.
2. Orthogonalize: $\mathbf{x}'_2 = \mathbf{x}_2 - \langle \mathbf{x}_2|\mathbf{x}'_1\rangle \mathbf{x}'_1$
3. Normalize: $\mathbf{x}''_2 = \mathbf{x}'_2/||\mathbf{x}'_2||$.
4. Orthogonalize: $\mathbf{x}'_3 = \mathbf{x}_3 - \langle \mathbf{x}_3|\mathbf{x}'_1\rangle \mathbf{x}'_1 - \langle \mathbf{x}_3|\mathbf{x}''_2\rangle \mathbf{x}''_2$
5. Normalize: $\mathbf{x}''_3 = \mathbf{x}'_3/||\mathbf{x}'_3||$.
6. Orthogonalize: $\mathbf{x}'_4 = \mathbf{x}_4 - \ldots$
7. Etc.



Example — Gram-Schmidt

$\mathbf{x}_1 = (1, 1, 1); \mathbf{x}_2 = (2, 0, 1); \mathbf{x}_3 = (3, 1, -1)$

1. $\mathbf{x}'_1 = \frac{1}{\sqrt{3}}(1, 1, 1)$
2. $\mathbf{x}'_2 = (2, 0, 1) - \frac{1}{3}(2 \cdot 1 + 0 \cdot 1 + 1 \cdot 1)(1, 1, 1)$
   $= (1, -1, 0)$
3. $\mathbf{x}''_2 = \frac{1}{\sqrt{2}}(1, -1, 0)$
4. $\mathbf{x}'_3 = (3, 1, -1) - \frac{1}{3}(3 \cdot 1 + 1 \cdot 1 + (-1) \cdot 1)(1, 1, 1)$
   $- \frac{1}{2}(3 \cdot 1 + 1 \cdot (-1) + (-1) \cdot 0)(1, -1, 0)$
   $= (1, 1, -2)$
5. $\mathbf{x}''_3 = \frac{1}{\sqrt{6}}(1, 1, -2)$

Distributions

Test function: A function $\phi(x)$, that is wellbehaved and fulfills

- $\phi(x)$ is *smooth* (infinitely differentiable)
- $\phi(x)$ has *compact support* (identically zero outside a finite intervall)

Let us define a **distribution** $f(x)$ by the values of the functional

$$\langle f, \phi \rangle = \int_{-\infty}^{\infty} f(x)\phi(x)\, dx \tag{1}$$

for all possible test functions $\phi(x)$.

Example — the Dirac δ function

The Dirac δ function can be defined as

$$\langle \delta, \phi \rangle = \int_{-\infty}^{\infty} \delta(x)\phi(x)\,dx = \phi(0) \tag{2}$$

for any test function $\phi(x)$. We can shift the origin of the test function

$$\phi(a) = \int_{-\infty}^{\infty} \delta(x)\phi(x+a)\,dx = \int_{-\infty}^{\infty} \delta(z-a)\phi(z)\,dz \tag{3}$$

If we let $\phi_t(x) = 1$ on the interval $[-t, t]$ we get

$$\int_{-\infty}^{\infty} \delta(x)\,dx = \lim_{t\to\infty} \int_{-\infty}^{\infty} \delta(x)\phi_t(x)\,dx = 1 \tag{4}$$

- Only defined under an integral sign.

Example — the Heaviside step function

Let us define the Heaviside step function as

$$H(x) = \begin{cases} 0 & x < 0 \\ \frac{1}{2} & x = 0 \\ 1 & x > 0 \end{cases} \tag{5}$$

or in the form of a distribution

$$\langle H, \phi \rangle = \int_{-\infty}^{\infty} H(x)\phi(x)\,dx = \int_{0}^{\infty} \phi(x)\,dx \tag{6}$$

The derivative of a distribution

$$\begin{aligned}
\langle f', \phi \rangle &= \int_{-\infty}^{\infty} f'(x)\phi(x)\,dx && (7)\\
&= [f(x)\phi(x)]_{-\infty}^{\infty} - \int_{-\infty}^{\infty} f(x)\phi'(x)\,dx && (8)\\
&= -\int_{-\infty}^{\infty} f(x)\phi'(x)\,dx && (9)\\
&= -\langle f, \phi' \rangle && (10)
\end{aligned}$$

Connection between $H(x)$ and $\delta(x)$.

Let us take the derivative of $H(x)$,

$$\begin{aligned}
\langle H', \phi \rangle &= \int_{-\infty}^{\infty} H'(x)\phi(x)\,dx && (11)\\
&= -\int_{-\infty}^{\infty} H(x)\phi'(x)\,dx && (12)\\
&= -\int_{0}^{\infty} \phi'(x)\,dx && (13)\\
&= -\left[\phi(x)\right]_0^{\infty} && (14)\\
&= -\{\phi(\infty) - \phi(0)\} && (15)\\
&= \phi(0) && (16)
\end{aligned}$$

thus $H'(x) = \delta(x)$.

Connection between $\delta(x)$ and ...

Fourier back transform with equal amplitudes for all frequencies

$$\delta(x) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \tilde{f}(\omega) e^{i\omega x} d\omega; \quad \tilde{f}(\omega) = 1 \tag{17}$$

Electrostatics for a point charge

$$\nabla^2 \frac{1}{r} = -4\pi\delta(\mathbf{r}) = -4\pi\delta(x)\delta(y)\delta(z) \tag{18}$$

LU decomposition

For a square matrix A we want to rewrite it as

$$A = LU \tag{1}$$

where L is a lower triangular matrix and U is an upper triangular matrix.

$$\begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ l_{21} & 1 & 0 \\ l_{31} & l_{32} & 1 \end{pmatrix} \begin{pmatrix} u_{11} & u_{12} & u_{13} \\ 0 & u_{22} & u_{23} \\ 0 & 0 & u_{33} \end{pmatrix} \tag{2}$$

$$= \begin{pmatrix} u_{11} & u_{12} & u_{13} \\ l_{21}u_{11} & l_{21}u_{12} + u_{22} & l_{21}u_{13} + u_{23} \\ l_{31}u_{11} & l_{31}u_{12} + l_{32}u_{22} & l_{31}u_{13} + l_{32}u_{23} + u_{33} \end{pmatrix} \tag{3}$$

LU decomposition

Now it is simpler to solve a linear equation system

$$Ax = c; \quad LUx = c; \quad Lb = c; \quad Ux = b \tag{4}$$

$$\begin{pmatrix} 1 & 0 & 0 \\ l_{21} & 1 & 0 \\ l_{31} & l_{32} & 1 \end{pmatrix} \begin{pmatrix} b_1 \\ b_2 \\ b_3 \end{pmatrix} = \begin{pmatrix} c_1 \\ c_2 \\ c_3 \end{pmatrix}$$

$$\begin{pmatrix} u_{11} & u_{12} & u_{13} \\ 0 & u_{22} & u_{23} \\ 0 & 0 & u_{33} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} b_1 \\ b_2 \\ b_3 \end{pmatrix}$$

LL decomposition, Cholesky decomposition

Assume that A is symmetric and positive semi-definite matrix, then

$$A = LL^{\dagger} \tag{5}$$

Can be used for data reduction:

1. $A_1 = l^{(1)}(l^{(1)})^{\dagger}; A = A_1 + R_1$. Is R_1 small? Then stop.
2. $A_2 = A_1 + l^{(2)}(l^{(2)})^{\dagger}; A = A_2 + R_2$. Is R_2 small? Then stop.
3. ...

Cholesky decomposition

$$LL^\dagger = \begin{pmatrix} l_{11} & 0 & 0 \\ l_{21} & l_{22} & 0 \\ l_{31} & l_{32} & l_{33} \end{pmatrix} \begin{pmatrix} l_{11}^\dagger & l_{12}^\dagger & l_{13}^\dagger \\ 0 & l_{22}^\dagger & l_{23}^\dagger \\ 0 & 0 & l_{33}^\dagger \end{pmatrix} \tag{6}$$

$$A_1 = \begin{pmatrix} l_{11} l_{11}^\dagger & l_{11} l_{12}^\dagger & l_{11} l_{13}^\dagger \\ l_{21} l_{11}^\dagger & l_{21} l_{12}^\dagger & l_{21} l_{13}^\dagger \\ l_{31} l_{11}^\dagger & l_{31} l_{12}^\dagger & l_{31} l_{13}^\dagger \end{pmatrix} \tag{7}$$

$$A_2 = \begin{pmatrix} l_{11} l_{11}^\dagger & l_{11} l_{12}^\dagger & l_{11} l_{13}^\dagger \\ l_{21} l_{11}^\dagger & l_{21} l_{12}^\dagger + l_{22} l_{22}^\dagger & l_{21} l_{13}^\dagger + l_{22} l_{23}^\dagger \\ l_{31} l_{11}^\dagger & l_{31} l_{12}^\dagger + l_{32} l_{22}^\dagger & l_{31} l_{13}^\dagger + l_{32} l_{23}^\dagger \end{pmatrix} \tag{8}$$

$$A_3 = \begin{pmatrix} l_{11} l_{11}^\dagger & l_{11} l_{12}^\dagger & l_{11} l_{13}^\dagger \\ l_{21} l_{11}^\dagger & l_{21} l_{12}^\dagger + l_{22} l_{22}^\dagger & l_{21} l_{13}^\dagger + l_{22} l_{23}^\dagger \\ l_{31} l_{11}^\dagger & l_{31} l_{12}^\dagger + l_{32} l_{22}^\dagger & l_{31} l_{13}^\dagger + l_{32} l_{23}^\dagger + l_{33} l_{33}^\dagger \end{pmatrix} \tag{9}$$

Singular value decomposition

Let A be a matrix of dimension $m \times n$. Assume that we can do the rewrite

$$A = USV^\dagger \tag{10}$$

where

- U has dimension $m \times m$ and is *unitary*
- S has dimension $m \times n$ and is *diagonal*, that is $s_{ij} = 0$ unless $i = j$.
- V has dimension $n \times n$ and is *unitary*

Now form the two Hermitian matrices

$$\begin{aligned} A^\dagger A &= VS^\dagger U^\dagger USV^\dagger &= VS^\dagger SV^\dagger \\ AA^\dagger &= USV^\dagger VS^\dagger U^\dagger &= USS^\dagger U^\dagger \end{aligned} \tag{11}$$

where $S^\dagger S$ and $SS^\dagger$ are diagonal matrices of dimensions $n \times n$ and $m \times m$ respectively.

Singular value decomposition

Rewrite

$$\begin{aligned} A^\dagger A &= V S^\dagger S V^\dagger \\ A A^\dagger &= U S S^\dagger U^\dagger \end{aligned} \quad (12)$$

into

$$\begin{aligned} V^\dagger A^\dagger A V &= S^\dagger S \\ U^\dagger A A^\dagger U &= S S^\dagger \end{aligned} \quad (13)$$

which is a diagonalization of $AA^\dagger$ and $A^\dagger A$. The smaller of $SS^\dagger$ and $S^\dagger S$ have eigenvalues λ_i. Thus S is diagonal with $s_{ii}^2 = \lambda_i$. We can now write A as

$$A = \sum_i s_i u^{(i)} (v^{(i)})^\dagger \quad (14)$$

If we discard vectors for small eigenvalues we get a data reduction.

Example — SVD

Start with matrix A

$$A = \begin{pmatrix} 0 & -1 \\ 1 & 1 \\ -1 & 0 \end{pmatrix} \tag{15}$$

and form

$$A^\dagger A = \begin{pmatrix} 0 & 1 & -1 \\ -1 & 1 & 0 \end{pmatrix} \begin{pmatrix} 0 & -1 \\ 1 & 1 \\ -1 & 0 \end{pmatrix} = \begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix} \tag{16}$$

and

$$AA^\dagger = \begin{pmatrix} 0 & -1 \\ 1 & 1 \\ -1 & 0 \end{pmatrix} \begin{pmatrix} 0 & 1 & -1 \\ -1 & 1 & 0 \end{pmatrix} = \begin{pmatrix} 1 & -1 & 0 \\ -1 & 2 & -1 \\ 0 & -1 & 1 \end{pmatrix} \tag{17}$$

Example — SVD

Diagonalize $A^\dagger A$

$$\begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix} \begin{pmatrix} 1 \\ 1 \end{pmatrix} = \lambda_1 \times \begin{pmatrix} 1 \\ 1 \end{pmatrix} = 3 \times \begin{pmatrix} 1 \\ 1 \end{pmatrix} \tag{18}$$

$$\begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix} \begin{pmatrix} 1 \\ -1 \end{pmatrix} = \lambda_2 \times \begin{pmatrix} 1 \\ -1 \end{pmatrix} = 1 \times \begin{pmatrix} 1 \\ -1 \end{pmatrix} \tag{19}$$

so

$$V = \begin{pmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \end{pmatrix} \tag{20}$$

and $S_{11} = \sqrt{3}$ and $S_{22} = 1$

Example — SVD

Diagonalize $AA^\dagger$

$$\begin{pmatrix} 1 & -1 & 0 \\ -1 & 2 & -1 \\ 0 & -1 & 1 \end{pmatrix} \begin{pmatrix} -1 \\ 2 \\ -1 \end{pmatrix} = 3 \times \begin{pmatrix} -1 \\ 2 \\ -1 \end{pmatrix} \tag{21}$$

$$\begin{pmatrix} 1 & -1 & 0 \\ -1 & 2 & -1 \\ 0 & -1 & 1 \end{pmatrix} \begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix} = 1 \times \begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix} \tag{22}$$

$$\begin{pmatrix} 1 & -1 & 0 \\ -1 & 2 & -1 \\ 0 & -1 & 1 \end{pmatrix} \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} = 0 \times \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} \tag{23}$$

Example — SVD

Finally

$$A = USV^\dagger \tag{24}$$

where

$$U = \begin{pmatrix} -\frac{1}{\sqrt{6}} & \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{3}} \\ \frac{2}{\sqrt{6}} & 0 & \frac{1}{\sqrt{3}} \\ -\frac{1}{\sqrt{6}} & -\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{3}} \end{pmatrix} \tag{25}$$

$$S = \begin{pmatrix} \sqrt{3} & 0 \\ 0 & 1 \\ 0 & 0 \end{pmatrix} \tag{26}$$

$$V = \begin{pmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \end{pmatrix} \tag{27}$$

Diagonalization of Hermitian matrix

Any Hermitian matrix, H, can be diagonalized by some unitary matrix, U

$$U^{\dagger} H U = D; \quad D_{ij} = \delta_{ij}\lambda_i \tag{1}$$

For the real symmetric case it can be diagonalized by an orthogonal matrix.

$$O^T H O = D; \quad D_{ij} = \delta_{ij}\lambda_i \tag{2}$$

where λ_i are so called eigenvalues of H.

Eigenvectors

Multiply both sides by U,

$$U^{\dagger}HU = D \quad \rightarrow \quad HU = UD \tag{3}$$

For each column in U, $u^{(k)}$, we have

$$Hu^{(k)} = u^{(k)}\lambda_k \tag{4}$$

where $u^{(k)}$ is an eigenvector with corresponding eigenvalue λ_k.

How to diagonalize

We need to find U or O.

- We need all (most) eigenvalues and eigenvectors
  - Jacobi, Givens, Householder, QR, MRRR, ...
- We need a few eigenvalues and eigenvectors
  - Lanzcos, Davidson, ...

Power method

Consider a trial vector x that is assumed to be an approximation to one eigenvector. Expand this in the (unknown) eigenvectors

$$x = c_1 u^{(1)} + c_2 u^{(2)} + \ldots + c_n u^{(n)} \tag{5}$$

Multiply this vector by H m times to get

$$x' = c_1 \lambda_1^m u^{(1)} + c_2 \lambda_2^m u^{(2)} + \ldots + c_n \lambda_n^m u^{(n)} \tag{6}$$

which will eventually be dominated one eigenvalue. This method is mostly of theoretical interest.

Jacobi's method

Loop through all offdiagonal element, H_{ij}, and perform a 2×2 rotation such that $H_{ij} = H_{ji} = 0$. For example $i = 2$ and $j = 3$.

$$\begin{pmatrix} x & a'_{12} & a'_{13} & x & x \\ a_{21} & a'_{22} & a'_{23} = 0 & a'_{24} & a'_{25} \\ a_{31} & a'_{32} = 0 & a'_{33} & a'_{34} & a'_{35} \\ x & a'_{42} & a'_{43} & x & x \\ x & a'_{52} & a'_{53} & x & x \end{pmatrix} \tag{7}$$

The square sum of the offdiagonal elements is reduced by $2H_{ij}^2$. Very robust method, a bit slow, suitable for smallish matrices.

Givens rotations

Here we reduce the matrix to a tridiagonal form: perform 2×2 rotations with $(i,j) = (2,3)$ in order to make $a_{13} = 0$.

$$\begin{pmatrix} x & a'_{12} & a'_{13} = 0 & x & x \\ a_{21} & a'_{22} & a'_{23} & a'_{24} & a'_{25} \\ a_{31} = 0 & a'_{32} & a'_{33} & a'_{34} & a'_{35} \\ x & a'_{42} & a'_{43} & x & x \\ x & a'_{52} & a'_{53} & x & x \end{pmatrix} \tag{8}$$

followed by rotating $(2,4)$ to make $a_{14} = 0$, etc. Followed by some method to find eigenvalues and eigenvectors, such as MRRR.

Inverse iteration

Consider a normalized trial vector x that is assumed to be an approximation to one eigenvector. Expand this in the (unknown) eigenvectors

$$x = c_1 u^{(1)} + c_2 u^{(2)} + \ldots + c_n u^{(n)} \tag{9}$$

Find an approximate eigenvalue by $\lambda = x^\dagger H x$. Solve the equation

$$(H - \lambda I)x' = x; \quad x' = (H - \lambda I)^{-1} x \tag{10}$$

$$x' = c_1(\lambda_1 - \lambda)^{-1} u^{(1)} + c_2(\lambda_2 - \lambda)^{-1} u^{(2)} + \ldots + c_n(\lambda_n - \lambda)^{-1} u^{(n)} \tag{11}$$

and the procedure is repeated until convergence. The procedure has cubic convergence.

Krylov subspace methods

A Krylov subspace is defined as
$K_n(A,b) = \mathrm{span}\left\{\mathbf{b}, A\mathbf{b}, A^2\mathbf{b}, \ldots, A^{n-1}\mathbf{b}\right\}$ where $\mathbf{b}$ is a real vector and A is a real matrix.
Used to solve $A\mathbf{x} = \mathbf{y}$. Assume $\mathbf{b} \approx \mathbf{x}$, make the ansatz

$$x \approx \sum_{m=0}^{n-1} c_m A^m \mathbf{b} = P(A)\mathbf{b} \tag{12}$$

Minimize the residual

$$||\mathbf{r}||^2 = ||\mathbf{y} - \sum_m c_m A^m \mathbf{b}||^2 \tag{13}$$

$$= ||\mathbf{y}||^2 - 2\sum_m c_m \langle \mathbf{y}|A^m\mathbf{b}\rangle + \sum_{mk} c_m c_k \langle A^m\mathbf{b}|A^k\mathbf{b}\rangle \tag{14}$$

$$\frac{d||\mathbf{r}||^2}{dc_m} = 0 \rightarrow \sum_k \langle A^m\mathbf{b}|A^k\mathbf{b}\rangle c_k = \langle \mathbf{y}|A^m\mathbf{b}\rangle \tag{15}$$

Orthonormalizing the vectors in $K_n(A,b)$ gives, in general, better numerical stability.

Lanczos' method

Transform Hermitian $n \times n$ matrix A to tridiagonal $m \times m$ matrix T.

1. Start with a vector $||\mathbf{x}_1|| = 1$
2. Initialize
   2.1 $\mathbf{s}_1 = A\mathbf{x}_1$
   2.2 $\alpha_1 = \mathbf{s}_1^\dagger \mathbf{x}_1$
   2.3 $\mathbf{w}_1 = \mathbf{s}_1 - \alpha_1 \mathbf{x}_1$
3. For $i = 2, \ldots, m$
   3.1 $\beta_i = ||\mathbf{w}_{i-1}||$
   3.2 If $\beta_i \neq 0$ Then $\mathbf{x}_i = \mathbf{w}_{i-1}/\beta_i$
   Else panic/stop/pick random orthogonal vector
   3.3 $\mathbf{s}_i = A\mathbf{x}_i$
   3.4 $\alpha_i = \mathbf{s}_i^\dagger \mathbf{x}_i$
   3.5 $\mathbf{w}_i = \mathbf{s}_i - \alpha_i \mathbf{x}_i - \beta_i \mathbf{x}_{i-1}$

We do not need A explicitly!

Lanczos' method

We get the unitary matrix V and the tridiagonal matrix T such that $T = V^\dagger A V$.

$$V = \begin{pmatrix} \uparrow & \uparrow & & \uparrow \\ \mathbf{x}_1 & \mathbf{x}_2 & \dots & \mathbf{x}_m \\ \downarrow & \downarrow & & \downarrow \end{pmatrix} \tag{16}$$

$$T = \begin{pmatrix} \alpha_1 & \beta_2 & 0 & \dots \\ \beta_2 & \alpha_2 & \beta_3 & \dots \\ 0 & \beta_3 & \alpha_4 & \dots \\ \vdots & \vdots & \vdots & \ddots \end{pmatrix} \tag{17}$$

Davidson type methods

Davidson matrix diagonalization for one root.

1. Guess a start vector such that $||\mathbf{c}_1|| = 1$
2. For $i = 1, \dots$
   - 2.1 Orthonormalize $\mathbf{c}_i$ against $\mathbf{c}_j$, $j = 1, \dots, i-1$.
   - 2.2 Compute $\mathbf{s}_i = H\mathbf{c}_i$ (σ vector)
   - 2.3 Form $\tilde{H}_{ij} = \langle \mathbf{s}_i | \mathbf{c}_j \rangle$ (only last row)
   - 2.4 Diagonalize $\tilde{H}$; $\quad \tilde{H}\mathbf{v}^{(k)} = \lambda_k \mathbf{v}^{(k)}$.
   - 2.5 Select root: $\mathbf{v}^{(m)}$ and λ_m
   - 2.6 Form residual $\mathbf{r}_i = \sum_j v_j^{(m)} (\mathbf{s}_j - \lambda_m \mathbf{c}_j)$
   - 2.7 $\mathbf{c}_{i+1} = (H^{(0)} - \lambda)^{-1} \mathbf{r}_i$

Improvements to Davidson's method

- Use several roots simultaneously. Davidson-Liu method.
- Use a better preconditioner, do full diagonalization of a submatrix of H.
- Skip normalization.
- Replace preconditioner $(H^{(0)} - \lambda)^{-1}\mathbf{r}$ with $c_1(H^{(0)} - \lambda)^{-1}\mathbf{r} + c_2(H^{(0)} - \lambda)^{-1}\mathbf{c}$
- Inverse iteration