math

Joey Yu Hsu

2024-04-30

Contents

index

math on bookdown started on 2024/01/28

Part I

by descipline

Chapter 1

mathematics

1.1 tool

- formula typesetting
  - TeX
    - LaTeX
      - pdfLaTeX
      - XeLaTeX
      - editor/tool:
        - LyX
        - OverLeaf
        - MathPix Snip
        - Micro$oft Office Word
          - WordTeX https://tomwildenhain.com/wordtex/
            - Pandoc dependent
          - https://superuser.com/questions/1114697/select-a-different-math-font-in-microsoft-word
          - https://www.youtube.com/watch?v=jlX_pThh7z8
        - Micro$oft Office PowerPoint
          - IguanaTeX https://www.jonathanleroux.org/software/iguanatex/
  - MathML
  - MathJax: JavaScript
- plot[3]
- symbolic computing
  - Maple: by MapleSoft
  - Mathematica: by Wolfram
- numeric computing
  - MatLab: by MathWorks

equivalence relation[11]

equivalence class[10]

partition[9]

1.2 discipline

Chapter 2

physics

2.1 discipline

- relativity
  - special relativity
    * Lorentz transformation[18]
  - general relativity
- analytic mechanics
  - Lagrangian mechanics
  - Hamiltonian mechanics
- electromagnetism
- quantum mechanics
- field theory

Chapter 3

plot

- LaTeX
  - TikZ[13]
    * https://tikz.dev/
    * TikZ-3Dplot
    * PGFplots[13.4]
      · https://tikz.dev/pgfplots/
      · https://pgfplots.sourceforge.net/gallery.html
      · https://pgfplots.net/
    * editor / export
      · https://zhuanlan.zhihu.com/p/660371706
      · offline
        · TikzEdt: WYSIWYG and live preview
        · TikZiT
      · online
        · OverLeaf
        · MathCha
        · GeoGebra Classic
      · Python
        · TikZplotLib / tikzplotlib[13.5]
          · matplotlib export to TikZ .tex
          · PyPI
          · GitHub
      · R
        · TikZDevice / tikzDevice
          · r chunk `engine='tikz'` knitr `out.width=if (knitr:::is_html_output()) '100%'`
          · CRAN
            · reference manual
            · vignette: TikZDevice - LaTeX Graphics for R
          · GitHub
    * TikZ library
  - xypic = xy-pic[14]
- OverLeaf
- MathCha
- GeoGebra
  - GeoGebra Classic: to export TikZ
  - GeoGebra Calculator Suite
- Python
  - MatPlotLib / matplotlib[27]
  - Seaborn / seaborn[27.1.3]
  - Plotly
  - Manim
- R
  - Modern Statistical Graphics
  - ggplot2[30]
    * Modern Statistical Graphics section 5.1
  - GraphViz `.gv`
  - Mermaid `.mmd`

    * about
    * JavaScript based diagramming and charting tool that renders Markdown-inspired text definitions to create and modify diagrams dynamically
  - Shiny
    * R Markdown Guide section 5.1
  - tool
    * Jamovi

neural network plot/draw https://github.com/ashishpatel26/Tools-to-Design-or-Visualize-Architecture-of-Neural-Network

Chapter 4

programming language

4.1 discipline

- Python[12]
- JavaScript
- SQL = structured query language
- R[19]
  - RMarkdown
    * Bookdown
  - knitr: engine
    * TikZ[13]
  - reticulate: Python
  - Jamovi
- C#
  - web
    * MVC
    * .NET
  - desktop
    * UWP = Universal Windows Platform
    * WPF = Windows Presentation Foundation
    * WinForms = Windows Forms
  - 3D/game
    * Unity

4.2 learning map

- W3School
- SoloLearn
- Codecademy

Chapter 5

machine learning

5.1 Shai Ben-David

https://www.youtube.com/playlist?list=PLPW2keNyw-usgvmR7FTQ3ZRjfLs5jT4BO

5.2 deep learning

我妻幸長

Esc = Einstein summation convention

$$W\boldsymbol{x} = \begin{pmatrix} w_{00} & w_{01} & \cdots & w_{0n} \\ w_{10} & w_{11} & \cdots & w_{1n} \\ \vdots & \vdots & \ddots & \vdots \\ w_{m0} & w_{m1} & \cdots & w_{mn} \end{pmatrix} \begin{pmatrix} x_0 \\ x_1 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} \sum_{\nu=0}^{n} w_{0\nu} x_\nu \\ \sum_{\nu=0}^{n} w_{1\nu} x_\nu \\ \vdots \\ \sum_{\nu=0}^{n} w_{m\nu} x_\nu \end{pmatrix} \overset{\text{Esc}}{=} \begin{pmatrix} w_{0\nu} x_\nu \\ w_{1\nu} x_\nu \\ \vdots \\ w_{m\nu} x_\nu \end{pmatrix} = w_{\mu\nu} x_\nu$$

$$\begin{aligned}
\boldsymbol{y} = \begin{pmatrix} y_0 \\ y_1 \\ \vdots \\ y_\mu \end{pmatrix} = y_\mu =& w_{\mu\nu} x_\nu = \begin{pmatrix} w_{0\nu} x_\nu \\ w_{1\nu} x_\nu \\ \vdots \\ w_{m\nu} x_\nu \end{pmatrix} = \begin{pmatrix} w_{00} & w_{01} & \cdots & w_{0n} \\ w_{10} & w_{11} & \cdots & w_{1n} \\ \vdots & \vdots & \ddots & \vdots \\ w_{m0} & w_{m1} & \cdots & w_{mn} \end{pmatrix} \begin{pmatrix} x_0 \\ x_1 \\ \vdots \\ x_n \end{pmatrix} = W\boldsymbol{x} \\
\boldsymbol{y}^\intercal = \begin{pmatrix} y_0 \\ y_1 \\ \vdots \\ y_m \end{pmatrix}^\intercal = y_\mu^\intercal = (w_{\mu\nu} x_\nu)^\intercal =& \begin{pmatrix} w_{0\nu} x_\nu \\ w_{1\nu} x_\nu \\ \vdots \\ w_{m\nu} x_\nu \end{pmatrix}^\intercal = \left[\begin{pmatrix} w_{00} & w_{01} & \cdots & w_{0n} \\ w_{10} & w_{11} & \cdots & w_{1n} \\ \vdots & \vdots & \ddots & \vdots \\ w_{m0} & w_{m1} & \cdots & w_{mn} \end{pmatrix} \begin{pmatrix} x_0 \\ x_1 \\ \vdots \\ x_n \end{pmatrix} \right]^\intercal = [W\boldsymbol{x}]^\intercal \\
=& x_\nu^\intercal w_{\mu\nu}^\intercal = \begin{pmatrix} w_{0\nu} x_\nu \\ w_{1\nu} x_\nu \\ \vdots \\ w_{m\nu} x_\nu \end{pmatrix}^\intercal = \begin{pmatrix} x_0 \\ x_1 \\ \vdots \\ x_n \end{pmatrix}^\intercal \begin{pmatrix} w_{00} & w_{01} & \cdots & w_{0n} \\ w_{10} & w_{11} & \cdots & w_{1n} \\ \vdots & \vdots & \ddots & \vdots \\ w_{m0} & w_{m1} & \cdots & w_{mn} \end{pmatrix}^\intercal = \boldsymbol{x}^\intercal W^\intercal
\end{aligned}$$

$$\boldsymbol{x}^\intercal W^\intercal = \begin{pmatrix} x_0 \\ x_1 \\ \vdots \\ x_n \end{pmatrix}^\intercal \begin{pmatrix} w_{00} & w_{01} & \cdots & w_{0n} \\ w_{10} & w_{11} & \cdots & w_{1n} \\ \vdots & \vdots & \ddots & \vdots \\ w_{m0} & w_{m1} & \cdots & w_{mn} \end{pmatrix}^\intercal = \begin{pmatrix} w_{0\nu} x_\nu \\ w_{1\nu} x_\nu \\ \vdots \\ w_{m\nu} x_\nu \end{pmatrix}^\intercal = x_\nu^\intercal w_{\mu\nu}^\intercal = y_\mu^\intercal = \begin{pmatrix} y_0 \\ y_1 \\ \vdots \\ y_m \end{pmatrix}^\intercal = \boldsymbol{y}^\intercal$$

$$W\boldsymbol{x} = \begin{pmatrix} w_{00} & w_{01} & \cdots & w_{0n} \\ w_{10} & w_{11} & \cdots & w_{1n} \\ \vdots & \vdots & \ddots & \vdots \\ w_{m0} & w_{m1} & \cdots & w_{mn} \end{pmatrix} \begin{pmatrix} x_0 \\ x_1 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} \sum\limits_{\nu=0}^{n} w_{0\nu} x_\nu \\ \sum\limits_{\nu=0}^{n} w_{1\nu} x_\nu \\ \vdots \\ \sum\limits_{\nu=0}^{n} w_{m\nu} x_\nu \end{pmatrix} \overset{\text{Esc}}{=} \begin{pmatrix} w_{0\nu} x_\nu \\ w_{1\nu} x_\nu \\ \vdots \\ w_{m\nu} x_\nu \end{pmatrix} = w_{\mu\nu} x_\nu, \begin{cases} x_0 = 1 \\ w_{\mu 0} = b_\mu \end{cases}$$

$$= W\boldsymbol{x} = \begin{pmatrix} b_0 & w_{01} & \cdots & w_{0n} \\ b_1 & w_{11} & \cdots & w_{1n} \\ \vdots & \vdots & \ddots & \vdots \\ b_m & w_{m1} & \cdots & w_{mn} \end{pmatrix} \begin{pmatrix} 1 \\ x_1 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} b_0 + \sum\limits_{j=1}^{n} w_{0j} x_j \\ b_1 + \sum\limits_{j=1}^{n} w_{1j} x_j \\ \vdots \\ b_m + \sum\limits_{j=1}^{n} w_{mj} x_j \end{pmatrix} \overset{\text{Esc}}{=} \begin{pmatrix} b_0 + w_{0j} x_j \\ b_1 + w_{1j} x_j \\ \vdots \\ b_m + w_{mj} x_j \end{pmatrix} = b_\mu + w_{\mu j} x_j$$

$$\boldsymbol{y} = \begin{pmatrix} y_0 \\ y_1 \\ \vdots \\ y_\mu \end{pmatrix} = y_\mu = w_{\mu\nu} x_\nu = \begin{pmatrix} w_{0\nu} x_\nu \\ w_{1\nu} x_\nu \\ \vdots \\ w_{m\nu} x_\nu \end{pmatrix} = \begin{pmatrix} w_{00} & w_{01} & \cdots & w_{0n} \\ w_{10} & w_{11} & \cdots & w_{1n} \\ \vdots & \vdots & \ddots & \vdots \\ w_{m0} & w_{m1} & \cdots & w_{mn} \end{pmatrix} \begin{pmatrix} x_0 \\ x_1 \\ \vdots \\ x_n \end{pmatrix} = W\boldsymbol{x}, \begin{cases} x_0 = 1 \\ w_{\mu 0} = b_\mu \end{cases}$$

$$= \boldsymbol{y} = \begin{pmatrix} y_0 \\ y_1 \\ \vdots \\ y_\mu \end{pmatrix} = y_\mu = w_{\mu j} x_j + b_\mu = \begin{pmatrix} b_0 + w_{0j} x_j \\ b_1 + w_{1j} x_j \\ \vdots \\ b_m + w_{mj} x_j \end{pmatrix} = \begin{pmatrix} b_0 & w_{01} & \cdots & w_{0n} \\ b_1 & w_{11} & \cdots & w_{1n} \\ \vdots & \vdots & \ddots & \vdots \\ b_m & w_{m1} & \cdots & w_{mn} \end{pmatrix} \begin{pmatrix} 1 \\ x_1 \\ \vdots \\ x_n \end{pmatrix} = W\boldsymbol{x}, \begin{cases} 1 = x_0 \\ b_\mu = w_{\mu 0} \end{cases}$$

$$\boldsymbol{y}^\mathsf{T} = \begin{pmatrix} y_0 \\ y_1 \\ \vdots \\ y_m \end{pmatrix}^\mathsf{T} = y_\mu^\mathsf{T} = (w_{\mu\nu} x_\nu)^\mathsf{T} = \begin{pmatrix} w_{0\nu} x_\nu \\ w_{1\nu} x_\nu \\ \vdots \\ w_{m\nu} x_\nu \end{pmatrix}^\mathsf{T} = \left[\begin{pmatrix} w_{00} & w_{01} & \cdots & w_{0n} \\ w_{10} & w_{11} & \cdots & w_{1n} \\ \vdots & \vdots & \ddots & \vdots \\ w_{m0} & w_{m1} & \cdots & w_{mn} \end{pmatrix} \begin{pmatrix} x_0 \\ x_1 \\ \vdots \\ x_n \end{pmatrix} \right]^\mathsf{T} = [W\boldsymbol{x}]^\mathsf{T}$$

$$= x_\nu^\mathsf{T} w_{\mu\nu}^\mathsf{T} = \begin{pmatrix} w_{0\nu} x_\nu \\ w_{1\nu} x_\nu \\ \vdots \\ w_{m\nu} x_\nu \end{pmatrix}^\mathsf{T} = \begin{pmatrix} x_0 \\ x_1 \\ \vdots \\ x_n \end{pmatrix}^\mathsf{T} \begin{pmatrix} w_{00} & w_{01} & \cdots & w_{0n} \\ w_{10} & w_{11} & \cdots & w_{1n} \\ \vdots & \vdots & \ddots & \vdots \\ w_{m0} & w_{m1} & \cdots & w_{mn} \end{pmatrix}^\mathsf{T} = \boldsymbol{x}^\mathsf{T} W^\mathsf{T}, \begin{cases} x_0 = 1 \\ w_{\mu 0} = b_\mu \end{cases}$$

$$= b_\mu^\mathsf{T} + x_j^\mathsf{T} w_{\mu j}^\mathsf{T} = \begin{pmatrix} b_0 + w_{0j} x_j \\ b_1 + w_{1j} x_j \\ \vdots \\ b_m + w_{mj} x_j \end{pmatrix}^\mathsf{T} = \begin{pmatrix} 1 \\ x_1 \\ \vdots \\ x_n \end{pmatrix}^\mathsf{T} \begin{pmatrix} b_0 & w_{01} & \cdots & w_{0n} \\ b_1 & w_{11} & \cdots & w_{1n} \\ \vdots & \vdots & \ddots & \vdots \\ b_m & w_{m1} & \cdots & w_{mn} \end{pmatrix}^\mathsf{T} = \boldsymbol{x}^\mathsf{T} W^\mathsf{T}$$

$$\boldsymbol{x}^\mathsf{T} W^\mathsf{T} = \begin{pmatrix} x_0 \\ x_1 \\ \vdots \\ x_n \end{pmatrix}^\mathsf{T} \begin{pmatrix} w_{00} & w_{01} & \cdots & w_{0n} \\ w_{10} & w_{11} & \cdots & w_{1n} \\ \vdots & \vdots & \ddots & \vdots \\ w_{m0} & w_{m1} & \cdots & w_{mn} \end{pmatrix}^\mathsf{T} = \begin{pmatrix} w_{0\nu} x_\nu \\ w_{1\nu} x_\nu \\ \vdots \\ w_{m\nu} x_\nu \end{pmatrix}^\mathsf{T} = x_\nu^\mathsf{T} w_{\mu\nu}^\mathsf{T} = y_\mu^\mathsf{T} = \begin{pmatrix} y_0 \\ y_1 \\ \vdots \\ y_m \end{pmatrix}^\mathsf{T} = \boldsymbol{y}^\mathsf{T}$$

$$\boldsymbol{x}^\mathsf{T} W^\mathsf{T} = \begin{pmatrix} 1 \\ x_1 \\ \vdots \\ x_n \end{pmatrix}^\mathsf{T} \begin{pmatrix} b_0 & w_{01} & \cdots & w_{0n} \\ b_1 & w_{11} & \cdots & w_{1n} \\ \vdots & \vdots & \ddots & \vdots \\ b_m & w_{m1} & \cdots & w_{mn} \end{pmatrix}^\mathsf{T} = \begin{pmatrix} b_0 + w_{0j} x_j \\ b_1 + w_{1j} x_j \\ \vdots \\ b_m + w_{mj} x_j \end{pmatrix}^\mathsf{T} = b_\mu^\mathsf{T} + x_j^\mathsf{T} w_{\mu j}^\mathsf{T} = y_\mu^\mathsf{T} = \begin{pmatrix} y_0 \\ y_1 \\ \vdots \\ y_m \end{pmatrix}^\mathsf{T} = \boldsymbol{y}^\mathsf{T}$$

$$W\boldsymbol{x} = \begin{pmatrix} w_{00} & w_{01} & \cdots & w_{0n} \\ w_{10} & w_{11} & \cdots & w_{1n} \\ \vdots & \vdots & \ddots & \vdots \\ w_{m0} & w_{m1} & \cdots & w_{mn} \end{pmatrix} \begin{pmatrix} x_0 \\ x_1 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} \sum\limits_{\nu=0}^{n} w_{0\nu} x_\nu \\ \sum\limits_{\nu=0}^{n} w_{1\nu} x_\nu \\ \vdots \\ \sum\limits_{\nu=0}^{n} w_{m\nu} x_\nu \end{pmatrix} \overset{\text{Esc}}{=} \begin{pmatrix} w_{0\nu} x_\nu \\ w_{1\nu} x_\nu \\ \vdots \\ w_{m\nu} x_\nu \end{pmatrix} = w_{\mu\nu} x_\nu, \begin{cases} x_0 = 1 \\ w_{\mu 0} = b_\mu \end{cases}$$

$$= W\boldsymbol{x} = \begin{pmatrix} b_0 & w_{01} & \cdots & w_{0n} \\ b_1 & w_{11} & \cdots & w_{1n} \\ \vdots & \vdots & \ddots & \vdots \\ b_m & w_{m1} & \cdots & w_{mn} \end{pmatrix} \begin{pmatrix} 1 \\ x_1 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} b_0 + \sum\limits_{j=1}^{n} w_{0j} x_j \\ b_1 + \sum\limits_{j=1}^{n} w_{1j} x_j \\ \vdots \\ b_m + \sum\limits_{j=1}^{n} w_{mj} x_j \end{pmatrix} \overset{\text{Esc}}{=} \begin{pmatrix} b_0 + w_{0j} x_j \\ b_1 + w_{1j} x_j \\ \vdots \\ b_m + w_{mj} x_j \end{pmatrix} = b_\mu + w_{\mu j} x_j$$

$$\boldsymbol{y} = \begin{pmatrix} y_0 \\ y_1 \\ \vdots \\ y_\mu \end{pmatrix} = y_\mu = w_{\mu\nu} x_\nu = \begin{pmatrix} w_{0\nu} x_\nu \\ w_{1\nu} x_\nu \\ \vdots \\ w_{m\nu} x_\nu \end{pmatrix} = \begin{pmatrix} w_{00} & w_{01} & \cdots & w_{0n} \\ w_{10} & w_{11} & \cdots & w_{1n} \\ \vdots & \vdots & \ddots & \vdots \\ w_{m0} & w_{m1} & \cdots & w_{mn} \end{pmatrix} \begin{pmatrix} x_0 \\ x_1 \\ \vdots \\ x_n \end{pmatrix} = W\boldsymbol{x}, \begin{cases} x_0 = 1 \\ w_{\mu 0} = b_\mu \end{cases}$$

$$= \boldsymbol{y} = \begin{pmatrix} y_0 \\ y_1 \\ \vdots \\ y_\mu \end{pmatrix} = y_\mu = w_{\mu j} x_j + b_\mu = \begin{pmatrix} b_0 + w_{0j} x_j \\ b_1 + w_{1j} x_j \\ \vdots \\ b_m + w_{mj} x_j \end{pmatrix} = \begin{pmatrix} b_0 & w_{01} & \cdots & w_{0n} \\ b_1 & w_{11} & \cdots & w_{1n} \\ \vdots & \vdots & \ddots & \vdots \\ b_m & w_{m1} & \cdots & w_{mn} \end{pmatrix} \begin{pmatrix} 1 \\ x_1 \\ \vdots \\ x_n \end{pmatrix} = W\boldsymbol{x}, \begin{cases} 1 = x_0 \\ b_\mu = w_{\mu 0} \end{cases}$$

$$\boldsymbol{y}^\intercal = \begin{pmatrix} y_0 \\ y_1 \\ \vdots \\ y_m \end{pmatrix}^\intercal = y_\mu^\intercal = (w_{\mu\nu} x_\nu)^\intercal = \begin{pmatrix} w_{0\nu} x_\nu \\ w_{1\nu} x_\nu \\ \vdots \\ w_{m\nu} x_\nu \end{pmatrix}^\intercal = \left[\begin{pmatrix} w_{00} & w_{01} & \cdots & w_{0n} \\ w_{10} & w_{11} & \cdots & w_{1n} \\ \vdots & \vdots & \ddots & \vdots \\ w_{m0} & w_{m1} & \cdots & w_{mn} \end{pmatrix} \begin{pmatrix} x_0 \\ x_1 \\ \vdots \\ x_n \end{pmatrix} \right]^\intercal = [W\boldsymbol{x}]^\intercal$$

$$= x_\nu^\intercal w_{\mu\nu}^\intercal = x_\nu^\intercal w_{\nu\mu} = \begin{pmatrix} w_{0\nu} x_\nu \\ w_{1\nu} x_\nu \\ \vdots \\ w_{m\nu} x_\nu \end{pmatrix}^\intercal = \begin{pmatrix} x_0 \\ x_1 \\ \vdots \\ x_n \end{pmatrix}^\intercal \begin{pmatrix} w_{00} & w_{01} & \cdots & w_{0n} \\ w_{10} & w_{11} & \cdots & w_{1n} \\ \vdots & \vdots & \ddots & \vdots \\ w_{m0} & w_{m1} & \cdots & w_{mn} \end{pmatrix}^\intercal = \boldsymbol{x}^\intercal W^\intercal$$

$$= b_\mu^\intercal + x_j^\intercal w_{\mu j}^\intercal = b_\mu^\intercal + x_j^\intercal w_{j\mu} = \begin{pmatrix} b_0 + w_{0j} x_j \\ b_1 + w_{1j} x_j \\ \vdots \\ b_m + w_{mj} x_j \end{pmatrix}^\intercal = \begin{pmatrix} 1 \\ x_1 \\ \vdots \\ x_n \end{pmatrix}^\intercal \begin{pmatrix} b_0 & w_{01} & \cdots & w_{0n} \\ b_1 & w_{11} & \cdots & w_{1n} \\ \vdots & \vdots & \ddots & \vdots \\ b_m & w_{m1} & \cdots & w_{mn} \end{pmatrix}^\intercal$$

$$\boldsymbol{x}^\intercal W^\intercal = \begin{pmatrix} x_0 \\ x_1 \\ \vdots \\ x_n \end{pmatrix}^\intercal \begin{pmatrix} w_{00} & w_{01} & \cdots & w_{0n} \\ w_{10} & w_{11} & \cdots & w_{1n} \\ \vdots & \vdots & \ddots & \vdots \\ w_{m0} & w_{m1} & \cdots & w_{mn} \end{pmatrix}^\intercal = \begin{pmatrix} w_{0\nu} x_\nu \\ w_{1\nu} x_\nu \\ \vdots \\ w_{m\nu} x_\nu \end{pmatrix}^\intercal = x_\nu^\intercal w_{\nu\mu} = y_\mu^\intercal = \begin{pmatrix} y_0 \\ y_1 \\ \vdots \\ y_m \end{pmatrix}^\intercal = \boldsymbol{y}^\intercal$$

$$\boldsymbol{x}^\intercal W^\intercal = \begin{pmatrix} 1 \\ x_1 \\ \vdots \\ x_n \end{pmatrix}^\intercal \begin{pmatrix} b_0 & w_{01} & \cdots & w_{0n} \\ b_1 & w_{11} & \cdots & w_{1n} \\ \vdots & \vdots & \ddots & \vdots \\ b_m & w_{m1} & \cdots & w_{mn} \end{pmatrix}^\intercal = \begin{pmatrix} b_0 + w_{0j} x_j \\ b_1 + w_{1j} x_j \\ \vdots \\ b_m + w_{mj} x_j \end{pmatrix}^\intercal = b_\mu^\intercal + x_j^\intercal w_{j\mu} = y_\mu^\intercal = \begin{pmatrix} y_0 \\ y_1 \\ \vdots \\ y_m \end{pmatrix}^\intercal$$

$$\boldsymbol{y} = \begin{pmatrix} y_0 \\ y_1 \\ \vdots \\ y_\mu \end{pmatrix} = y_\mu = w_{\mu\nu} x_\nu = \begin{pmatrix} w_{0\nu} x_\nu \\ w_{1\nu} x_\nu \\ \vdots \\ w_{m\nu} x_\nu \end{pmatrix} = \begin{pmatrix} w_{00} & w_{01} & \cdots & w_{0n} \\ w_{10} & w_{11} & \cdots & w_{1n} \\ \vdots & \vdots & \ddots & \vdots \\ w_{m0} & w_{m1} & \cdots & w_{mn} \end{pmatrix} \begin{pmatrix} x_0 \\ x_1 \\ \vdots \\ x_n \end{pmatrix} = W\boldsymbol{x}, \begin{cases} x_0 = 1 \\ w_{\mu 0} = b_\mu \end{cases}$$

$$= \boldsymbol{y} = \begin{pmatrix} y_0 \\ y_1 \\ \vdots \\ y_\mu \end{pmatrix} = y_\mu = w_{\mu j} x_j + b_\mu = \begin{pmatrix} b_0 + w_{0j} x_j \\ b_1 + w_{1j} x_j \\ \vdots \\ b_m + w_{mj} x_j \end{pmatrix} = \begin{pmatrix} b_0 & w_{01} & \cdots & w_{0n} \\ b_1 & w_{11} & \cdots & w_{1n} \\ \vdots & \vdots & \ddots & \vdots \\ b_m & w_{m1} & \cdots & w_{mn} \end{pmatrix} \begin{pmatrix} 1 \\ x_1 \\ \vdots \\ x_n \end{pmatrix} = W\boldsymbol{x}, \begin{cases} 1 = x_0 \\ b_\mu = w_{\mu 0} \end{cases}$$

$$\boldsymbol{x}^\intercal W^\intercal = \begin{pmatrix} x_0 \\ x_1 \\ \vdots \\ x_n \end{pmatrix}^\intercal \begin{pmatrix} w_{00} & w_{01} & \cdots & w_{0n} \\ w_{10} & w_{11} & \cdots & w_{1n} \\ \vdots & \vdots & \ddots & \vdots \\ w_{m0} & w_{m1} & \cdots & w_{mn} \end{pmatrix}^\intercal = \begin{pmatrix} w_{0\nu} x_\nu \\ w_{1\nu} x_\nu \\ \vdots \\ w_{m\nu} x_\nu \end{pmatrix}^\intercal = x_\nu^\intercal w_{\mu\nu}^\intercal = x_\nu^\intercal w_{\nu\mu} = y_\mu^\intercal = \begin{pmatrix} y_0 \\ y_1 \\ \vdots \\ y_m \end{pmatrix}^\intercal = \boldsymbol{y}^\intercal$$

$$= \boldsymbol{x}^\intercal W^\intercal = \begin{pmatrix} 1 \\ x_1 \\ \vdots \\ x_n \end{pmatrix}^\intercal \begin{pmatrix} b_0 & w_{01} & \cdots & w_{0n} \\ b_1 & w_{11} & \cdots & w_{1n} \\ \vdots & \vdots & \ddots & \vdots \\ b_m & w_{m1} & \cdots & w_{mn} \end{pmatrix}^\intercal = \begin{pmatrix} b_0 + w_{0j} x_j \\ b_1 + w_{1j} x_j \\ \vdots \\ b_m + w_{mj} x_j \end{pmatrix}^\intercal = b_\mu^\intercal + x_j^\intercal w_{j\mu} = y_\mu^\intercal = \begin{pmatrix} y_0 \\ y_1 \\ \vdots \\ y_m \end{pmatrix}^\intercal = \boldsymbol{y}^\intercal$$

$$\sigma\left(\boldsymbol{y}\right) = \sigma \begin{pmatrix} y_0 \\ y_1 \\ \vdots \\ y_\mu \end{pmatrix} = \sigma\left(y_\mu\right) = \sigma\left(w_{\mu\nu} x_\nu\right) = \sigma \begin{pmatrix} w_{0\nu} x_\nu \\ w_{1\nu} x_\nu \\ \vdots \\ w_{m\nu} x_\nu \end{pmatrix} = \sigma \left(\begin{pmatrix} w_{00} & w_{01} & \cdots & w_{0n} \\ w_{10} & w_{11} & \cdots & w_{1n} \\ \vdots & \vdots & \ddots & \vdots \\ w_{m0} & w_{m1} & \cdots & w_{mn} \end{pmatrix} \begin{pmatrix} x_0 \\ x_1 \\ \vdots \\ x_n \end{pmatrix} \right) = \sigma\left(W\boldsymbol{x}\right), \begin{cases} x_0 = 1 \\ w_{\mu 0} = b_\mu \end{cases}$$

$$= \sigma\left(\boldsymbol{y}\right) = \sigma \begin{pmatrix} y_0 \\ y_1 \\ \vdots \\ y_\mu \end{pmatrix} = \sigma\left(y_\mu\right) = \sigma\left(w_{\mu j} x_j + b_\mu\right) = \sigma \begin{pmatrix} b_0 + w_{0j} x_j \\ b_1 + w_{1j} x_j \\ \vdots \\ b_m + w_{mj} x_j \end{pmatrix} = \sigma \left(\begin{pmatrix} b_0 & w_{01} & \cdots & w_{0n} \\ b_1 & w_{11} & \cdots & w_{1n} \\ \vdots & \vdots & \ddots & \vdots \\ b_m & w_{m1} & \cdots & w_{mn} \end{pmatrix} \begin{pmatrix} 1 \\ x_1 \\ \vdots \\ x_n \end{pmatrix} \right) = \sigma\left(W\boldsymbol{x}\right), \begin{cases} 1 = x_0 \\ b_\mu = w_{\mu 0} \end{cases}$$

$$\left(\boldsymbol{x}^\intercal W^\intercal\right)\varsigma = \begin{pmatrix} x_0 \\ x_1 \\ \vdots \\ x_n \end{pmatrix}^\intercal \begin{pmatrix} w_{00} & w_{01} & \cdots & w_{0n} \\ w_{10} & w_{11} & \cdots & w_{1n} \\ \vdots & \vdots & \ddots & \vdots \\ w_{m0} & w_{m1} & \cdots & w_{mn} \end{pmatrix}^\intercal \varsigma = \begin{pmatrix} w_{0\nu} x_\nu \\ w_{1\nu} x_\nu \\ \vdots \\ w_{m\nu} x_\nu \end{pmatrix}^\intercal \varsigma = \left(x_\nu^\intercal w_{\mu\nu}^\intercal\right)\varsigma = \left(x_\nu^\intercal w_{\nu\mu}\right)\varsigma = \left(y_\mu^\intercal\right)\varsigma = \begin{pmatrix} y_0 \\ y_1 \\ \vdots \\ y_m \end{pmatrix}^\intercal \varsigma = \left(\boldsymbol{y}^\intercal\right)\varsigma$$

$$= \left(\boldsymbol{x}^\intercal W^\intercal\right)\varsigma = \begin{pmatrix} 1 \\ x_1 \\ \vdots \\ x_n \end{pmatrix}^\intercal \begin{pmatrix} b_0 & w_{01} & \cdots & w_{0n} \\ b_1 & w_{11} & \cdots & w_{1n} \\ \vdots & \vdots & \ddots & \vdots \\ b_m & w_{m1} & \cdots & w_{mn} \end{pmatrix}^\intercal \varsigma = \begin{pmatrix} b_0 + w_{0j} x_j \\ b_1 + w_{1j} x_j \\ \vdots \\ b_m + w_{mj} x_j \end{pmatrix}^\intercal \varsigma = \left(b_\mu^\intercal + x_j^\intercal w_{j\mu}\right)\varsigma = \left(y_\mu^\intercal\right)\varsigma = \begin{pmatrix} y_0 \\ y_1 \\ \vdots \\ y_m \end{pmatrix}^\intercal \varsigma = \left(\boldsymbol{y}^\intercal\right)\varsigma$$

$$\sigma\left(\boldsymbol{y}\right)=\sigma\begin{pmatrix}y_0\\y_1\\\vdots\\y_\mu\end{pmatrix}=\sigma\left(y_\mu\right)=\sigma\left(w_{\mu\nu}x_\nu\right)=\sigma\begin{pmatrix}w_{0\nu}x_\nu\\w_{1\nu}x_\nu\\\vdots\\w_{m\nu}x_\nu\end{pmatrix}=\sigma\left(\begin{pmatrix}w_{00}&w_{01}&\cdots&w_{0n}\\w_{10}&w_{11}&\cdots&w_{1n}\\\vdots&\vdots&\ddots&\vdots\\w_{m0}&w_{m1}&\cdots&w_{mn}\end{pmatrix}\begin{pmatrix}x_0\\x_1\\\vdots\\x_n\end{pmatrix}\right)=\sigma\left(W\boldsymbol{x}\right),\begin{cases}x_0=1\\w_{\mu0}=b_\mu\end{cases}$$

$$=\sigma\boldsymbol{y}=\sigma_\mu\begin{pmatrix}y_0\\y_1\\\vdots\\y_\mu\end{pmatrix}=\sigma_\mu y_\mu=\sigma_\mu w_{\mu\nu}x_\nu=\sigma_\mu\begin{pmatrix}w_{0\nu}x_\nu\\w_{1\nu}x_\nu\\\vdots\\w_{m\nu}x_\nu\end{pmatrix}=\sigma\begin{pmatrix}w_{00}&w_{01}&\cdots&w_{0n}\\w_{10}&w_{11}&\cdots&w_{1n}\\\vdots&\vdots&\ddots&\vdots\\w_{m0}&w_{m1}&\cdots&w_{mn}\end{pmatrix}\begin{pmatrix}x_0\\x_1\\\vdots\\x_n\end{pmatrix}=\sigma W\boldsymbol{x},\begin{cases}x_0=1\\w_{\mu0}=b_\mu\end{cases}$$

$$=\sigma\left(\boldsymbol{y}\right)=\sigma\begin{pmatrix}y_0\\y_1\\\vdots\\y_\mu\end{pmatrix}=\sigma\left(y_\mu\right)=\sigma\left(w_{\mu j}x_j+b_\mu\right)=\sigma\begin{pmatrix}b_0+w_{0j}x_j\\b_1+w_{1j}x_j\\\vdots\\b_m+w_{mj}x_j\end{pmatrix}=\sigma\left(\begin{pmatrix}b_0&w_{01}&\cdots&w_{0n}\\b_1&w_{11}&\cdots&w_{1n}\\\vdots&\vdots&\ddots&\vdots\\b_m&w_{m1}&\cdots&w_{mn}\end{pmatrix}\begin{pmatrix}1\\x_1\\\vdots\\x_n\end{pmatrix}\right)=\sigma\left(W\boldsymbol{x}\right),\begin{cases}1=x_0\\b_\mu=w_{\mu0}\end{cases}$$

$$=\sigma\boldsymbol{y}=\sigma_\mu\begin{pmatrix}y_0\\y_1\\\vdots\\y_\mu\end{pmatrix}=\sigma_\mu y_\mu=\sigma_\mu\left(w_{\mu j}x_j+b_\mu\right)=\sigma_\mu\begin{pmatrix}b_0+w_{0j}x_j\\b_1+w_{1j}x_j\\\vdots\\b_m+w_{mj}x_j\end{pmatrix}=\sigma\begin{pmatrix}b_0&w_{01}&\cdots&w_{0n}\\b_1&w_{11}&\cdots&w_{1n}\\\vdots&\vdots&\ddots&\vdots\\b_m&w_{m1}&\cdots&w_{mn}\end{pmatrix}\begin{pmatrix}1\\x_1\\\vdots\\x_n\end{pmatrix}=\sigma W\boldsymbol{x},\begin{cases}1=x_0\\b_\mu=w_{\mu0}\end{cases}$$

$$\left(\boldsymbol{x}^\mathsf{T}W^\mathsf{T}\right)\varsigma=\begin{pmatrix}x_0\\x_1\\\vdots\\x_n\end{pmatrix}^\mathsf{T}\begin{pmatrix}w_{00}&w_{01}&\cdots&w_{0n}\\w_{10}&w_{11}&\cdots&w_{1n}\\\vdots&\vdots&\ddots&\vdots\\w_{m0}&w_{m1}&\cdots&w_{mn}\end{pmatrix}^\mathsf{T}\varsigma=\begin{pmatrix}w_{0\nu}x_\nu\\w_{1\nu}x_\nu\\\vdots\\w_{m\nu}x_\nu\end{pmatrix}^\mathsf{T}\varsigma=\left(x_\nu^\mathsf{T}w_{\mu\nu}^\mathsf{T}\right)\varsigma=\left(x_\nu^\mathsf{T}w_{\nu\mu}\right)\varsigma=\left(y_\mu^\mathsf{T}\right)\varsigma=\begin{pmatrix}y_0\\y_1\\\vdots\\y_m\end{pmatrix}^\mathsf{T}\varsigma=\left(\boldsymbol{y}^\mathsf{T}\right)\varsigma$$

$$=\boldsymbol{x}^\mathsf{T}W^\mathsf{T}\varsigma=\begin{pmatrix}x_0\\x_1\\\vdots\\x_n\end{pmatrix}^\mathsf{T}\begin{pmatrix}w_{00}&w_{01}&\cdots&w_{0n}\\w_{10}&w_{11}&\cdots&w_{1n}\\\vdots&\vdots&\ddots&\vdots\\w_{m0}&w_{m1}&\cdots&w_{mn}\end{pmatrix}^\mathsf{T}\varsigma=\begin{pmatrix}w_{0\nu}x_\nu\\w_{1\nu}x_\nu\\\vdots\\w_{m\nu}x_\nu\end{pmatrix}^\mathsf{T}\varsigma=x_\nu^\mathsf{T}w_{\mu\nu}^\mathsf{T}\varsigma_\mu=x_\nu^\mathsf{T}w_{\nu\mu}\varsigma_\mu=y_\mu^\mathsf{T}\varsigma_\mu=\begin{pmatrix}y_0\\y_1\\\vdots\\y_m\end{pmatrix}^\mathsf{T}\varsigma_\mu=\boldsymbol{y}^\mathsf{T}\varsigma$$

$$=\left(\boldsymbol{x}^\mathsf{T}W^\mathsf{T}\right)\varsigma=\begin{pmatrix}1\\x_1\\\vdots\\x_n\end{pmatrix}^\mathsf{T}\begin{pmatrix}b_0&w_{01}&\cdots&w_{0n}\\b_1&w_{11}&\cdots&w_{1n}\\\vdots&\vdots&\ddots&\vdots\\b_m&w_{m1}&\cdots&w_{mn}\end{pmatrix}^\mathsf{T}\varsigma=\begin{pmatrix}b_0+w_{0j}x_j\\b_1+w_{1j}x_j\\\vdots\\b_m+w_{mj}x_j\end{pmatrix}^\mathsf{T}\varsigma=\left(b_\mu^\mathsf{T}+x_j^\mathsf{T}w_{j\mu}\right)\varsigma=\left(y_\mu^\mathsf{T}\right)\varsigma=\begin{pmatrix}y_0\\y_1\\\vdots\\y_m\end{pmatrix}^\mathsf{T}\varsigma=\left(\boldsymbol{y}^\mathsf{T}\right)\varsigma$$

$$=\boldsymbol{x}^\mathsf{T}W^\mathsf{T}\varsigma=\begin{pmatrix}1\\x_1\\\vdots\\x_n\end{pmatrix}^\mathsf{T}\begin{pmatrix}b_0&w_{01}&\cdots&w_{0n}\\b_1&w_{11}&\cdots&w_{1n}\\\vdots&\vdots&\ddots&\vdots\\b_m&w_{m1}&\cdots&w_{mn}\end{pmatrix}^\mathsf{T}\varsigma=\begin{pmatrix}b_0+w_{0j}x_j\\b_1+w_{1j}x_j\\\vdots\\b_m+w_{mj}x_j\end{pmatrix}^\mathsf{T}\varsigma=\left(b_\mu^\mathsf{T}+x_j^\mathsf{T}w_{j\mu}\right)\varsigma_\mu=y_\mu^\mathsf{T}\varsigma_\mu=\begin{pmatrix}y_0\\y_1\\\vdots\\y_m\end{pmatrix}^\mathsf{T}\varsigma_\mu=\boldsymbol{y}^\mathsf{T}\varsigma$$

$$\sigma\boldsymbol{y}=\sigma_\mu\begin{pmatrix}y_0\\y_1\\\vdots\\y_\mu\end{pmatrix}=\sigma_\mu y_\mu=\sigma_\mu w_{\mu\nu}x_\nu=\sigma_\mu\begin{pmatrix}w_{0\nu}x_\nu\\w_{1\nu}x_\nu\\\vdots\\w_{m\nu}x_\nu\end{pmatrix}=\sigma\begin{pmatrix}w_{00}&w_{01}&\cdots&w_{0n}\\w_{10}&w_{11}&\cdots&w_{1n}\\\vdots&\vdots&\ddots&\vdots\\w_{m0}&w_{m1}&\cdots&w_{mn}\end{pmatrix}\begin{pmatrix}x_0\\x_1\\\vdots\\x_n\end{pmatrix}=\sigma W\boldsymbol{x},\begin{cases}x_0=1\\w_{\mu0}=b_\mu\end{cases}$$

$$=\sigma\boldsymbol{y}=\sigma_\mu\begin{pmatrix}y_0\\y_1\\\vdots\\y_\mu\end{pmatrix}=\sigma_\mu y_\mu=\sigma_\mu\left(w_{\mu j}x_j+b_\mu\right)=\sigma_\mu\begin{pmatrix}b_0+w_{0j}x_j\\b_1+w_{1j}x_j\\\vdots\\b_m+w_{mj}x_j\end{pmatrix}=\sigma\begin{pmatrix}b_0&w_{01}&\cdots&w_{0n}\\b_1&w_{11}&\cdots&w_{1n}\\\vdots&\vdots&\ddots&\vdots\\b_m&w_{m1}&\cdots&w_{mn}\end{pmatrix}\begin{pmatrix}1\\x_1\\\vdots\\x_n\end{pmatrix}=\sigma W\boldsymbol{x},\begin{cases}1=x_0\\b_\mu=w_{\mu0}\end{cases}$$

$$\boldsymbol{x}^\mathsf{T}W^\mathsf{T}\varsigma=\begin{pmatrix}x_0\\x_1\\\vdots\\x_n\end{pmatrix}^\mathsf{T}\begin{pmatrix}w_{00}&w_{01}&\cdots&w_{0n}\\w_{10}&w_{11}&\cdots&w_{1n}\\\vdots&\vdots&\ddots&\vdots\\w_{m0}&w_{m1}&\cdots&w_{mn}\end{pmatrix}^\mathsf{T}\varsigma=\begin{pmatrix}w_{0\nu}x_\nu\\w_{1\nu}x_\nu\\\vdots\\w_{m\nu}x_\nu\end{pmatrix}^\mathsf{T}\varsigma=x_\nu^\mathsf{T}w_{\mu\nu}^\mathsf{T}\varsigma_\mu=x_\nu^\mathsf{T}w_{\nu\mu}\varsigma_\mu=y_\mu^\mathsf{T}\varsigma_\mu=\begin{pmatrix}y_0\\y_1\\\vdots\\y_m\end{pmatrix}^\mathsf{T}\varsigma_\mu=\boldsymbol{y}^\mathsf{T}\varsigma$$

$$=\boldsymbol{x}^\mathsf{T}W^\mathsf{T}\varsigma=\begin{pmatrix}1\\x_1\\\vdots\\x_n\end{pmatrix}^\mathsf{T}\begin{pmatrix}b_0&w_{01}&\cdots&w_{0n}\\b_1&w_{11}&\cdots&w_{1n}\\\vdots&\vdots&\ddots&\vdots\\b_m&w_{m1}&\cdots&w_{mn}\end{pmatrix}^\mathsf{T}\varsigma=\begin{pmatrix}b_0+w_{0j}x_j\\b_1+w_{1j}x_j\\\vdots\\b_m+w_{mj}x_j\end{pmatrix}^\mathsf{T}\varsigma=\left(b_\mu^\mathsf{T}+x_j^\mathsf{T}w_{j\mu}\right)\varsigma_\mu=y_\mu^\mathsf{T}\varsigma_\mu=\begin{pmatrix}y_0\\y_1\\\vdots\\y_m\end{pmatrix}^\mathsf{T}\varsigma_\mu=\boldsymbol{y}^\mathsf{T}\varsigma$$

$$\boldsymbol{z} = \sigma \boldsymbol{y} = \sigma_\mu \begin{pmatrix} y_0 \\ y_1 \\ \vdots \\ y_\mu \end{pmatrix} = \sigma_\mu y_\mu = \sigma_\mu w_{\mu\nu} x_\nu = \sigma_\mu \begin{pmatrix} w_{0\nu} x_\nu \\ w_{1\nu} x_\nu \\ \vdots \\ w_{m\nu} x_\nu \end{pmatrix} = \sigma \begin{pmatrix} w_{00} & w_{01} & \cdots & w_{0n} \\ w_{10} & w_{11} & \cdots & w_{1n} \\ \vdots & \vdots & \ddots & \vdots \\ w_{m0} & w_{m1} & \cdots & w_{mn} \end{pmatrix} \begin{pmatrix} x_0 \\ x_1 \\ \vdots \\ x_n \end{pmatrix} = \sigma W \boldsymbol{x}, \begin{cases} x_0 = 1 \\ w_{\mu 0} = b_\mu \end{cases}$$

$$= \boldsymbol{z} = z_\mu = \sigma_\mu \begin{pmatrix} y_0 \\ y_1 \\ \vdots \\ y_\mu \end{pmatrix} = \sigma_\mu y_\mu = \sigma_\mu \left(w_{\mu j} x_j + b_\mu \right) = \sigma_\mu \begin{pmatrix} b_0 + w_{0j} x_j \\ b_1 + w_{1j} x_j \\ \vdots \\ b_m + w_{mj} x_j \end{pmatrix} = \sigma \begin{pmatrix} b_0 & w_{01} & \cdots & w_{0n} \\ b_1 & w_{11} & \cdots & w_{1n} \\ \vdots & \vdots & \ddots & \vdots \\ b_m & w_{m1} & \cdots & w_{mn} \end{pmatrix} \begin{pmatrix} 1 \\ x_1 \\ \vdots \\ x_n \end{pmatrix} = \sigma W \boldsymbol{x}, \begin{cases} 1 = x_0 \\ b_\mu = w_{\mu 0} \end{cases}$$

$$\boldsymbol{x}^\intercal W^\intercal \varsigma = \begin{pmatrix} x_0 \\ x_1 \\ \vdots \\ x_n \end{pmatrix}^\intercal \begin{pmatrix} w_{00} & w_{01} & \cdots & w_{0n} \\ w_{10} & w_{11} & \cdots & w_{1n} \\ \vdots & \vdots & \ddots & \vdots \\ w_{m0} & w_{m1} & \cdots & w_{mn} \end{pmatrix}^\intercal \varsigma = \begin{pmatrix} w_{0\nu} x_\nu \\ w_{1\nu} x_\nu \\ \vdots \\ w_{m\nu} x_\nu \end{pmatrix}^\intercal \varsigma = x_\nu^\intercal w_{\mu\nu}^\intercal \varsigma_\mu = x_\nu^\intercal w_{\nu\mu} \varsigma_\mu = y_\mu^\intercal \varsigma_\mu = \begin{pmatrix} y_0 \\ y_1 \\ \vdots \\ y_m \end{pmatrix}^\intercal \varsigma_\mu = \boldsymbol{y}^\intercal \varsigma = \boldsymbol{z}^\intercal$$

$$= \boldsymbol{x}^\intercal W^\intercal \varsigma = \begin{pmatrix} 1 \\ x_1 \\ \vdots \\ x_n \end{pmatrix}^\intercal \begin{pmatrix} b_0 & w_{01} & \cdots & w_{0n} \\ b_1 & w_{11} & \cdots & w_{1n} \\ \vdots & \vdots & \ddots & \vdots \\ b_m & w_{m1} & \cdots & w_{mn} \end{pmatrix}^\intercal \varsigma = \begin{pmatrix} b_0 + w_{0j} x_j \\ b_1 + w_{1j} x_j \\ \vdots \\ b_m + w_{mj} x_j \end{pmatrix}^\intercal \varsigma = \left(b_\mu^\intercal + x_j^\intercal w_{j\mu} \right) \varsigma_\mu = y_\mu^\intercal \varsigma_\mu = \begin{pmatrix} y_0 \\ y_1 \\ \vdots \\ y_m \end{pmatrix}^\intercal \varsigma_\mu = z_\mu^\intercal = \boldsymbol{z}^\intercal$$

$$\begin{aligned} \boldsymbol{z} =& \sigma \boldsymbol{y} = \sigma_\mu w_{\mu\nu} x_\nu = \sigma W \boldsymbol{x}, \begin{cases} x_0 = 1 \\ w_{\mu 0} = b_\mu \end{cases} \\ = z_\mu =& \sigma_\mu y_\mu = \sigma_\mu \left(w_{\mu j} x_j + b_\mu \right), \begin{cases} 1 = x_0 \\ b_\mu = w_{\mu 0} \end{cases} \\ \boldsymbol{x}^\intercal W^\intercal \varsigma =& x_\nu^\intercal w_{\mu\nu}^\intercal \varsigma_\mu = x_\nu^\intercal w_{\nu\mu} \varsigma_\mu = \boldsymbol{y}^\intercal \varsigma = \boldsymbol{z}^\intercal \\ =& \left(b_\mu^\intercal + x_j^\intercal w_{j\mu} \right) \varsigma_\mu = y_\mu^\intercal \varsigma_\mu = z_\mu^\intercal \end{aligned}$$

---

matrix calculus[54]

4-15

wrong or incompatible transpose

$$\begin{aligned} \boldsymbol{x}^\intercal W &= \begin{pmatrix} x_0 & x_1 & \cdots & x_m \end{pmatrix} \begin{pmatrix} w_{00} & w_{01} & \cdots & w_{0n} \\ w_{10} & w_{11} & \cdots & w_{1n} \\ \vdots & \vdots & \ddots & \vdots \\ w_{m0} & w_{m1} & \cdots & w_{mn} \end{pmatrix} \\ &= \begin{pmatrix} x_0 \\ x_1 \\ \vdots \\ x_m \end{pmatrix}^\intercal \begin{pmatrix} w_{00} & w_{01} & \cdots & w_{0n} \\ w_{10} & w_{11} & \cdots & w_{1n} \\ \vdots & \vdots & \ddots & \vdots \\ w_{m0} & w_{m1} & \cdots & w_{mn} \end{pmatrix} = \begin{pmatrix} \sum_{\mu=1}^{m} x_\mu w_{\mu 0} \\ \sum_{\mu=1}^{m} x_\mu w_{\mu 1} \\ \vdots \\ \sum_{\mu=1}^{m} x_\mu w_{\mu n} \end{pmatrix}^\intercal \\ &\overset{\text{Einstein summation convention}}{=} \begin{pmatrix} x_0 \\ x_1 \\ \vdots \\ x_m \end{pmatrix}^\intercal \begin{pmatrix} w_{00} & w_{01} & \cdots & w_{0n} \\ w_{10} & w_{11} & \cdots & w_{1n} \\ \vdots & \vdots & \ddots & \vdots \\ w_{m0} & w_{m1} & \cdots & w_{mn} \end{pmatrix} = \begin{pmatrix} x_\mu w_{\mu 0} \\ x_\mu w_{\mu 1} \\ \vdots \\ x_\mu w_{\mu n} \end{pmatrix}^\intercal \\ &= x_\mu^\intercal w_{\mu\nu} = (x_\mu w_{\mu\nu})^\intercal ? \end{aligned}$$

4-18

wrong or incompatible transpose

$$
\begin{aligned}
\boldsymbol{x}^{\mathsf{T}}W &= \begin{pmatrix} x_0 \\ x_1 \\ \vdots \\ x_m \end{pmatrix}^{\mathsf{T}} \begin{pmatrix} w_{00} & w_{01} & \cdots & w_{0n} \\ w_{10} & w_{11} & \cdots & w_{1n} \\ \vdots & \vdots & \ddots & \vdots \\ w_{m0} & w_{m1} & \cdots & w_{mn} \end{pmatrix} = \begin{pmatrix} x_\mu w_{\mu 0} \\ x_\mu w_{\mu 1} \\ \vdots \\ x_\mu w_{\mu n} \end{pmatrix}^{\mathsf{T}}, \begin{cases} x_0 = 1 \\ w_{0\nu} = b_\nu \end{cases} \\
&= \begin{pmatrix} 1 \\ x_1 \\ \vdots \\ x_m \end{pmatrix}^{\mathsf{T}} \begin{pmatrix} b_0 & b_1 & \cdots & b_n \\ w_{10} & w_{11} & \cdots & w_{1n} \\ \vdots & \vdots & \ddots & \vdots \\ w_{m0} & w_{m1} & \cdots & w_{mn} \end{pmatrix} = \begin{pmatrix} x_i w_{i0} + b_0 \\ x_i w_{i1} + b_1 \\ \vdots \\ x_i w_{in} + b_n \end{pmatrix}^{\mathsf{T}}, \begin{cases} 1 = x_0 \\ b_\nu = w_{0\nu} \end{cases}
\end{aligned}
$$

wrong or incompatible transpose

$$
\begin{aligned}
\sigma\left(\boldsymbol{x}^{\mathsf{T}}W\right) =&\sigma\left(\begin{pmatrix} x_0 \\ x_1 \\ \vdots \\ x_m \end{pmatrix}^{\mathsf{T}} \begin{pmatrix} w_{00} & w_{01} & \cdots & w_{0n} \\ w_{10} & w_{11} & \cdots & w_{1n} \\ \vdots & \vdots & \ddots & \vdots \\ w_{m0} & w_{m1} & \cdots & w_{mn} \end{pmatrix}\right) = \sigma\left(\begin{pmatrix} x_\mu w_{\mu 0} \\ x_\mu w_{\mu 1} \\ \vdots \\ x_\mu w_{\mu n} \end{pmatrix}^{\mathsf{T}}\right) = \begin{pmatrix} \sigma_0\left(x_\mu w_{\mu 0}\right) \\ \sigma_1\left(x_\mu w_{\mu 1}\right) \\ \vdots \\ \sigma_n\left(x_\mu w_{\mu n}\right) \end{pmatrix}^{\mathsf{T}}, \begin{cases} x_0 = 1 \\ w_{0\nu} = b_\nu \end{cases} \\
=&\sigma\left(\begin{pmatrix} 1 \\ x_1 \\ \vdots \\ x_m \end{pmatrix}^{\mathsf{T}} \begin{pmatrix} b_0 & b_1 & \cdots & b_n \\ w_{10} & w_{11} & \cdots & w_{1n} \\ \vdots & \vdots & \ddots & \vdots \\ w_{m0} & w_{m1} & \cdots & w_{mn} \end{pmatrix}\right) = \sigma\left(\begin{pmatrix} x_i w_{i0} + b_0 \\ x_i w_{i1} + b_1 \\ \vdots \\ x_i w_{in} + b_n \end{pmatrix}^{\mathsf{T}}\right) = \sigma_\nu\left(x_\mu w_{\mu\nu}\right), \begin{cases} 1 = x_0 \\ b_\nu = w_{0\nu} \end{cases}
\end{aligned}
$$

Part II

by date

Chapter 6

A Minimal Book Example

6.1 About

This is a *sample* book written in **Markdown**. You can use anything that Pandoc's Markdown supports; for example, a math equation $a^2 + b^2 = c^2$.

6.1.1 Usage

Each **bookdown** chapter is an .Rmd file, and each .Rmd file can contain one (and only one) chapter. A chapter *must* start with a first-level heading: `# A good chapter`, and can contain one (and only one) first-level heading.

Use second-level and higher headings within chapters like: `## A short section` or `### An even shorter section`.

The `index.Rmd` file is required, and is also your first book chapter. It will be the homepage when you render the book.

6.1.2 Render book

You can render the HTML version of this example book without changing anything:

1. Find the **Build** pane in the RStudio IDE, and

2. Click on **Build Book**, then select your output format, or select "All formats" if you'd like to use multiple formats from the same book source files.

Or build the book from the R console:

```
bookdown::render_book()
```

To render this example to PDF as a `bookdown::pdf_book`, you'll need to install XeLaTeX. You are recommended to install TinyTeX (which includes XeLaTeX): https://yihui.org/tinytex/.

6.1.3 Preview book

As you work, you may start a local server to live preview this HTML book. This preview will update as you edit the book when you save individual .Rmd files. You can start the server in a work session by using the RStudio add-in "Preview book", or from the R console:

```
bookdown::serve_book()
```

6.2 Hello bookdown

All chapters start with a first-level heading followed by your chapter title, like the line above. There should be only one first-level heading (`#`) per .Rmd file.

6.2.1 A section

All chapter sections start with a second-level (`##`) or higher heading followed by your section title, like the sections above and below here. You can have as many as you want within a chapter.

An unnumbered section

Chapters and sections are numbered by default. To un-number a heading, add a `{.unnumbered}` or the shorter `{-}` at the end of the heading, like in this section.

6.3 Cross-references

Cross-references make it easier for your readers to find and link to elements in your book.

6.3.1 Chapters and sub-chapters

There are two steps to cross-reference any heading:

1. Label the heading: `# Hello world {#nice-label}`.
   - Leave the label off if you like the automated heading generated based on your heading title: for example, `# Hello world` = `# Hello world {#hello-world}`.
   - To label an un-numbered heading, use: `# Hello world {-#nice-label}` or `{# Hello world .unnumbered}`.
2. Next, reference the labeled heading anywhere in the text using `\@ref(nice-label)`; for example, please see Chapter 6.3.
   - If you prefer text as the link instead of a numbered reference use: any text you want can go here.

6.3.2 Captioned figures and tables

Figures and tables *with captions* can also be cross-referenced from elsewhere in your book using `\@ref(fig:chunk-label)` and `\@ref(tab:chunk-label)`, respectively.

See Figure 6.1.

```
par(mar = c(4, 4, .1, .1))
plot(pressure, type = 'b', pch = 19)
```

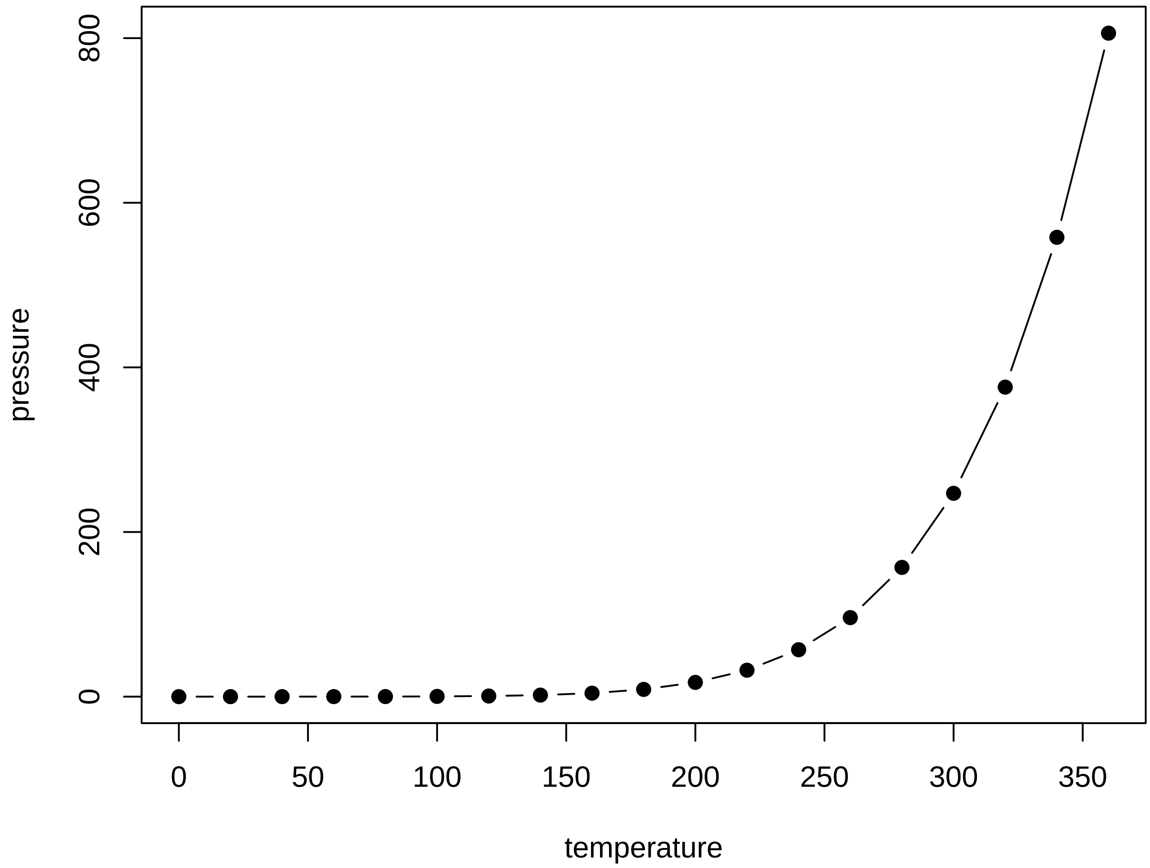


Figure 6.1: Here is a nice figure!

Don't miss Table 6.1.

```
knitr::kable(
  head(pressure, 10), caption = 'Here is a nice table!',
```

Table 6.1: Here is a nice table!

temperature	pressure
0	0.0002
20	0.0012
40	0.0060
60	0.0300
80	0.0900
100	0.2700
120	0.7500
140	1.8500
160	4.2000
180	8.8000

```
  booktabs = TRUE
)
```

6.4 Parts

You can add parts to organize one or more book chapters together. Parts can be inserted at the top of an .Rmd file, before the first-level chapter heading in that same file.

Add a numbered part: `# (PART) Act one {-}` (followed by `# A chapter`)

Add an unnumbered part: `# (PART\*) Act one {-}` (followed by `# A chapter`)

Add an appendix as a special kind of un-numbered part: `# (APPENDIX) Other stuff {-}` (followed by `# A chapter`). Chapters in an appendix are prepended with letters instead of numbers.

6.5 Footnotes and citations

6.5.1 Footnotes

Footnotes are put inside the square brackets after a caret `^[]`. Like this one [1].

6.5.2 Citations

Reference items in your bibliography file(s) using `@key`.

For example, we are using the **bookdown** package[1] (check out the last code chunk in index.Rmd to see how this citation key was added) in this sample book, which was built on top of R Markdown and **knitr**[2] (this citation was added manually in an external file book.bib). Note that the `.bib` files need to be listed in the index.Rmd with the YAML `bibliography` key.

The RStudio Visual Markdown Editor can also make it easier to insert citations: https://rstudio.github.io/visual-markdown-editing/#/citations

6.6 Blocks

6.6.1 Equations

Here is an equation.

$$f\left(k\right) = \binom{n}{k} p^k \left(1-p\right)^{n-k} \tag{6.1}$$

You may refer to using `\@ref(eq:binom)`, like see Equation (6.1).

[1] This is a footnote.

6.6.2 Theorems and proofs

Labeled theorems can be referenced in text using `\@ref(thm:tri)`, for example, check out this smart theorem 6.1.

Theorem 6.1. *For a right triangle, if c denotes the* length *of the hypotenuse and a and b denote the lengths of the **other** two sides, we have*

$$a^2 + b^2 = c^2$$

Read more here https://bookdown.org/yihui/bookdown/markdown-extensions-by-bookdown.html.

6.6.3 Callout blocks

The R Markdown Cookbook provides more help on how to use custom blocks to design your own callouts: https://bookdown.org/yihui/rmarkdown-cookbook/custom-blocks.html

6.7 Sharing your book

6.7.1 Publishing

HTML books can be published online, see: https://bookdown.org/yihui/bookdown/publishing.html

6.7.2 404 pages

By default, users will be directed to a 404 page if they try to access a webpage that cannot be found. If you'd like to customize your 404 page instead of using the default, you may add either a `_404.Rmd` or `_404.md` file to your project root and use code and/or Markdown syntax.

6.7.3 Metadata for sharing

Bookdown HTML books will provide HTML metadata for social sharing on platforms like Twitter, Facebook, and LinkedIn, using information you provide in the `index.Rmd` YAML. To setup, set the `url` for your book and the path to your `cover-image` file. Your book's `title` and `description` are also used.

This `gitbook` uses the same social sharing data across all chapters in your book- all links shared will look the same.

Specify your book's source repository on GitHub using the `edit` key under the configuration options in the `_output.yml` file, which allows users to suggest an edit by linking to a chapter's source file.

Read more about the features of this output format here:

https://pkgs.rstudio.com/bookdown/reference/gitbook.html

Or use:

```
?bookdown::gitbook
```

Chapter 7

test

https://bookdown.org/yihui/rmarkdown-cookbook/verbatim-code-chunks.html

7.1 RStudio

7.1.1 writer options

https://rstudio.github.io/visual-markdown-editing/markdown.html#writer-options

7.1.1.1 line wrapping

https://rstudio.github.io/visual-markdown-editing/markdown.html#line-wrapping

7.1.1.2 ensuring the same markdown between source / visual mode

https://stackoverflow.com/questions/71775027/rstudio-switch-markdown-editing-mode-between-source-and-visual-changes-special

canonical mode

https://rstudio.github.io/visual-markdown-editing/markdown.html#canonical-mode

```
---
title: "My Document"
editor_options:
  markdown:
    wrap: 72
    references:
      location: block
    canonical: true
---
```

7.1.2 Rtools

Rtools43 for Windows https://cran.r-project.org/bin/windows/Rtools/rtools43/rtools.html

7.1.3 addins

https://github.com/rstudio/addinexamples

```
if (!requireNamespace("devtools", quietly = TRUE))
  install.packages("devtools")

devtools::install_github("rstudio/htmltools")
devtools::install_github("rstudio/shiny")
devtools::install_github("rstudio/miniUI")
```

7.1.4 Git

commit: filename or extension is too long

https://stackoverflow.com/questions/22575662/filename-too-long-in-git-for-windows

https://stackoverflow.com/questions/55327408/how-to-fix-git-for-windows-error-could-not-lock-config-file-c-file-path-to-g

7.2 RMarkdown

R Markdown 指南 https://cosname.github.io/rmarkdown-guide/index.html

https://www.rstudio.com/wp-content/uploads/2015/02/rmarkdown-cheatsheet.pdf

https://slides.yihui.org/2020-taipei-satrday-rmarkdown.html#1

7.2.1 Pandoc link

https://pandoc.org/chunkedhtml-demo/8.16-links-1.html

https://stackoverflow.com/questions/39281266/use-internal-links-in-rmarkdown-html-output

https://community.rstudio.com/t/how-to-hyperlink-between-different-rmd-files-in-rmarkdown/62289

7.2.2 URL

https://stackoverflow.com/questions/29787850/how-do-i-add-a-url-to-r-markdown

```
[I'm an inline-style link](https://www.google.com)

[I'm an inline-style link with title](https://www.google.com "Google's Homepage")

[I'm a reference-style link][Arbitrary case-insensitive reference text]

[I'm a relative reference to a repository file](../blob/master/LICENSE)

[You can use numbers for reference-style link definitions][1]

Or leave it empty and use the [link text itself]

Some text to show that the reference links can follow later.

[arbitrary case-insensitive reference text]: https://www.mozilla.org
[1]: http://slashdot.org
[link text itself]: http://www.reddit.com
```

7.2.3 arrow

https://reimbar.org/dev/arrows/

Up arrow: `↑`

Down arrow: `↓`

Left arrow: `←`

Right arrow: `→`

Double headed arrow: `↔`

7.2.4 superscript and subscript

$\text{script}^{\text{superscript}}{}_{\text{subscript}}$

```
script^superscript^
```

$\text{script}^{\text{superscript}}$

```
~subscript~
```

$\text{script}_{\text{subscript}}$

7.2.4.1 LaTeX

https://tex.stackexchange.com/questions/580824/subscript-not-distinguished-enough

https://tex.stackexchange.com/questions/262295/make-subscript-size-smaller-always

7.2.5 equation

https://stackoverflow.com/questions/26049762/erroneous-nesting-of-equation-structures-in-using-beginalign-in-a-multi-l

7.2.5.1 proof QED

https://math.meta.stackexchange.com/questions/3582/qed-for-mathjax-here-on-stackexchange

`\tag*{$\Box$}`

$$a^2 + b^2 = c^2 \tag*{$\Box$}$$

`\tag*{$\blacksquare$}`

$$a^2 + b^2 = c^2 \tag*{$\blacksquare$}$$

7.2.6 image

https://stackoverflow.com/questions/25166624/insert-picture-table-in-r-markdown

7.2.6.1 DiagrammeR / mermaid flowchart

```
Error: Functions that produce HTML output found in document targeting latex output.
Please change the output type of this document to HTML.
If your aiming to have some HTML widgets shown in non-HTML format as a screenshot,
please install webshot or webshot2 R package for knitr to do the screenshot.
Alternatively, you can allow HTML output in non-HTML formats
by adding this option to the YAML front-matter of
your rmarkdown file:

  always_allow_html: true

Note however that the HTML output will not be visible in non-HTML formats.
```

https://bookdown.org/yihui/rmarkdown-cookbook/diagrams.html#diagrams

https://stackoverflow.com/questions/40803017/how-to-include-diagrammer-mermaid-flowchart-in-a-rmarkdown-file

```
{r}
library(DiagrammeR)
mermaid("
graph LR
    A-->B
",
width = 100
)
```

https://github.com/rich-iannone/DiagrammeR/issues/364

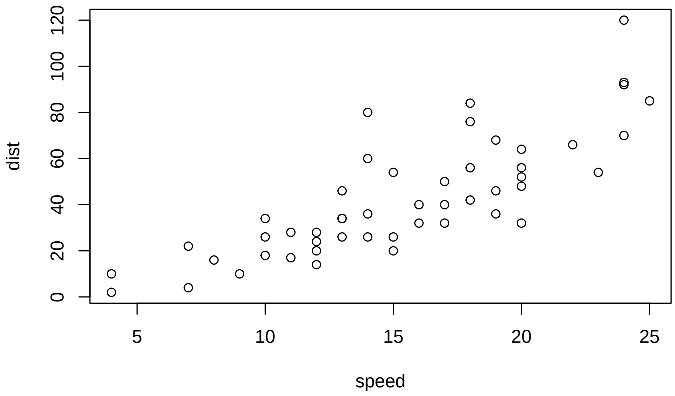
https://stackoverflow.com/questions/55994210/how-to-solve-diagrammer-waste-of-space-issue-in-rmarkdown

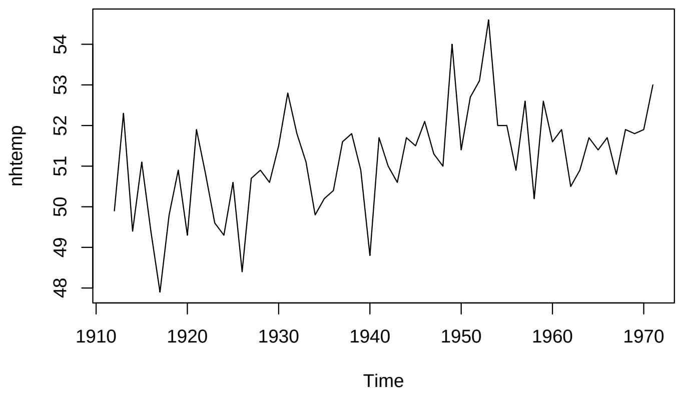
7.2.6.2 multiple images / figures in the same line

https://cosname.github.io/rmarkdown-guide/rmarkdown-base.html#element-figure

```
{r, fig.show = "hold", out.width = "50%"}
plot(cars)
plot(nhtemp)
```

```r
plot(cars)
plot(nhtemp)
```

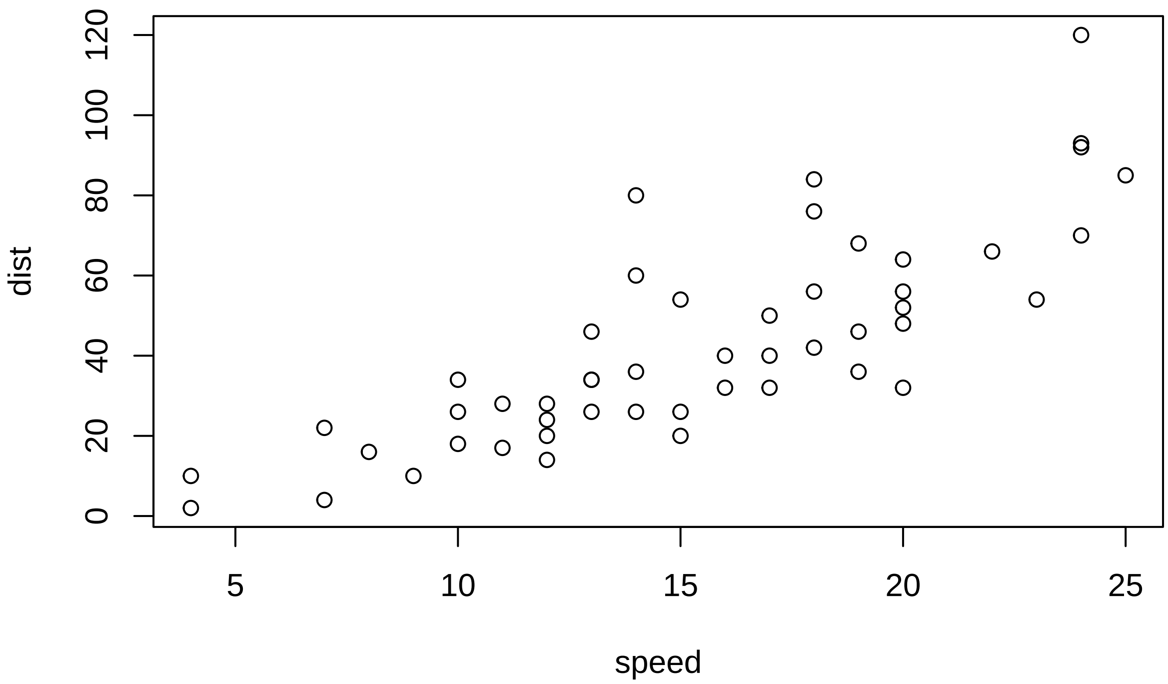




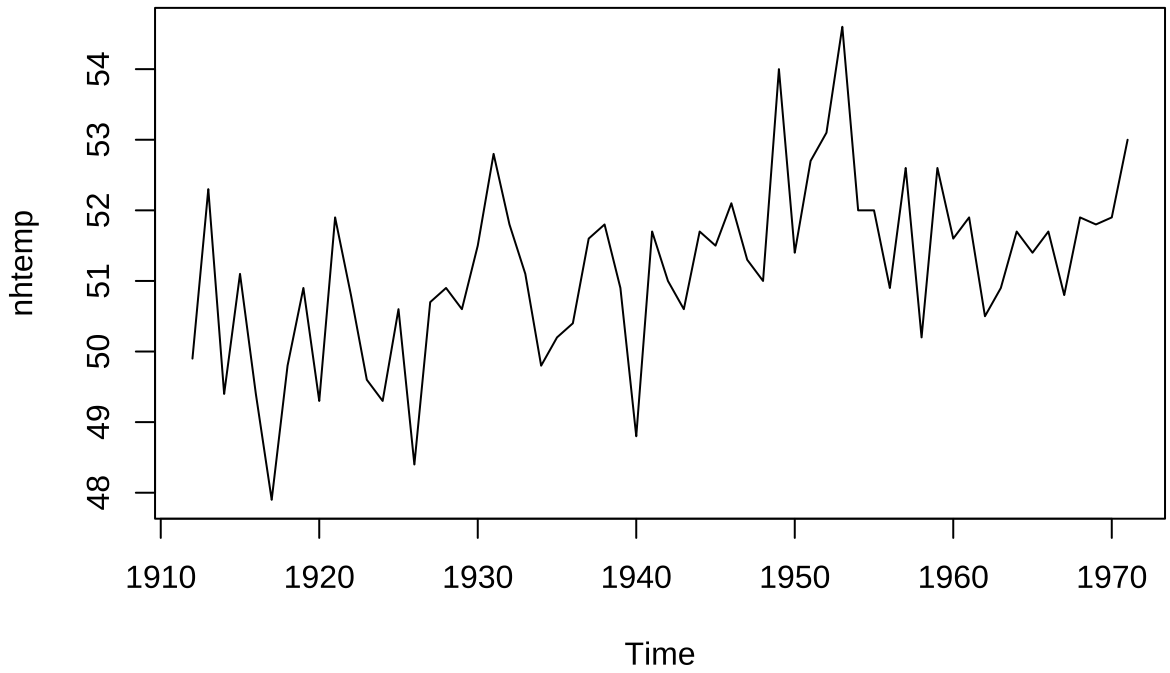
cf.

{r}
plot(cars)
plot(nhtemp)

```r
plot(cars)
```




```r
plot(nhtemp)
```



7.2.6.3 figure size

https://sebastiansauer.github.io/figure_sizing_knitr/

YAML in index.Rmd

```
---
title: "My Document"
output: html_document:
fig_width: 6
fig_height: 4
---
```

first R-chunk in your RMD document

```
knitr::opts_chunk$set(fig.width=12, fig.height=8)
```

width, height and options

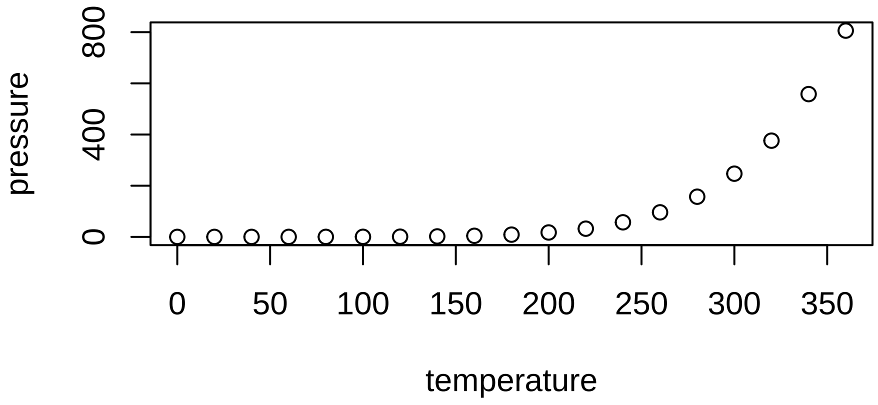
```
```{r fig.height = 3, fig.width = 5

plot(pressure)

```
```

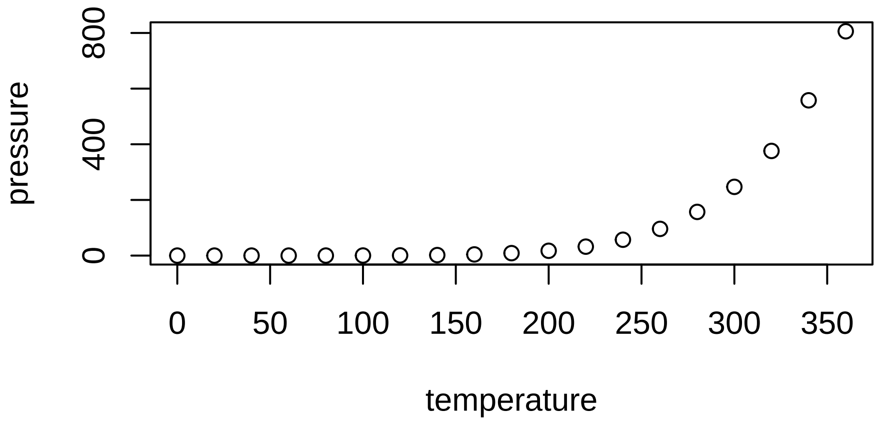
```
{r fig.height = 3, fig.width = 5
```

```
plot(pressure)
```



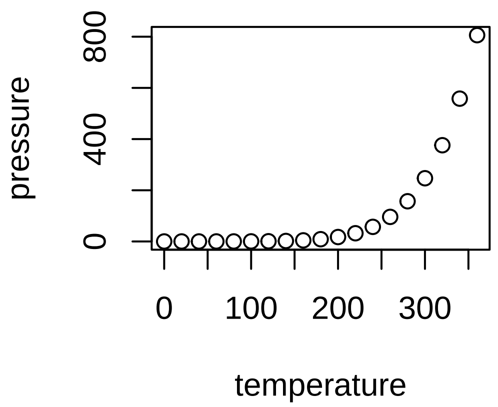

```
{r fig.height = 3, fig.width = 3, fig.align = "center"
```

```r
plot(pressure)
```




```
{r fig.width = 5, fig.asp = .62
```

```r
plot(pressure)
```

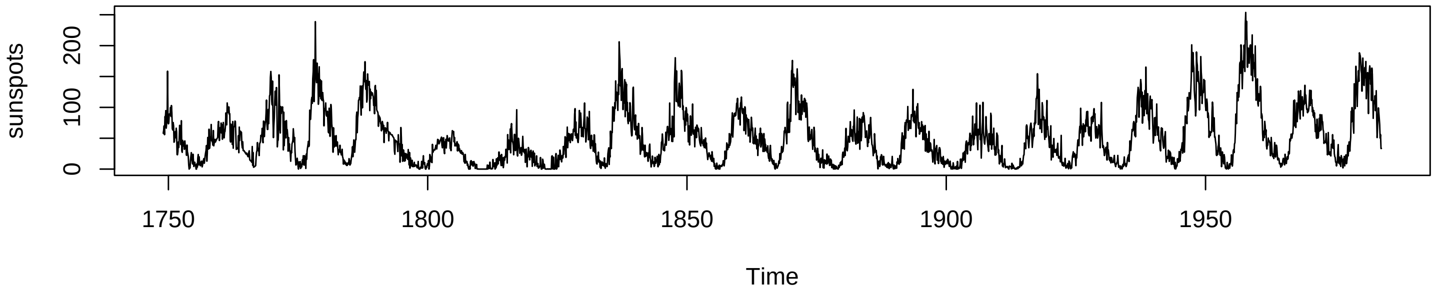


```
<center>
![](https://bookdown.org/yihui/rmarkdown-cookbook/images/cover.png){width=20%}
</center>
```

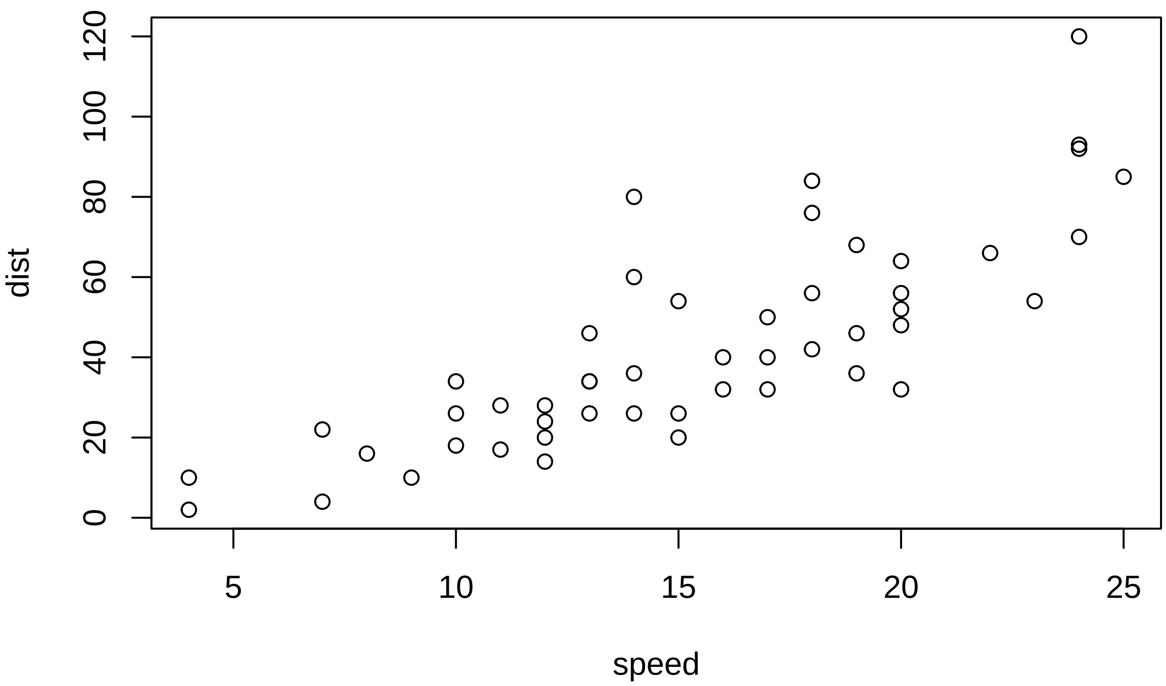
7.2.6.3.1 `knitr` https://yihui.org/knitr/options/

https://bookdown.org/yihui/rmarkdown/tufte-figures.html

```
par(mar = c(4, 4, .1, .2)); plot(sunspots)
```




```
plot(cars)
```

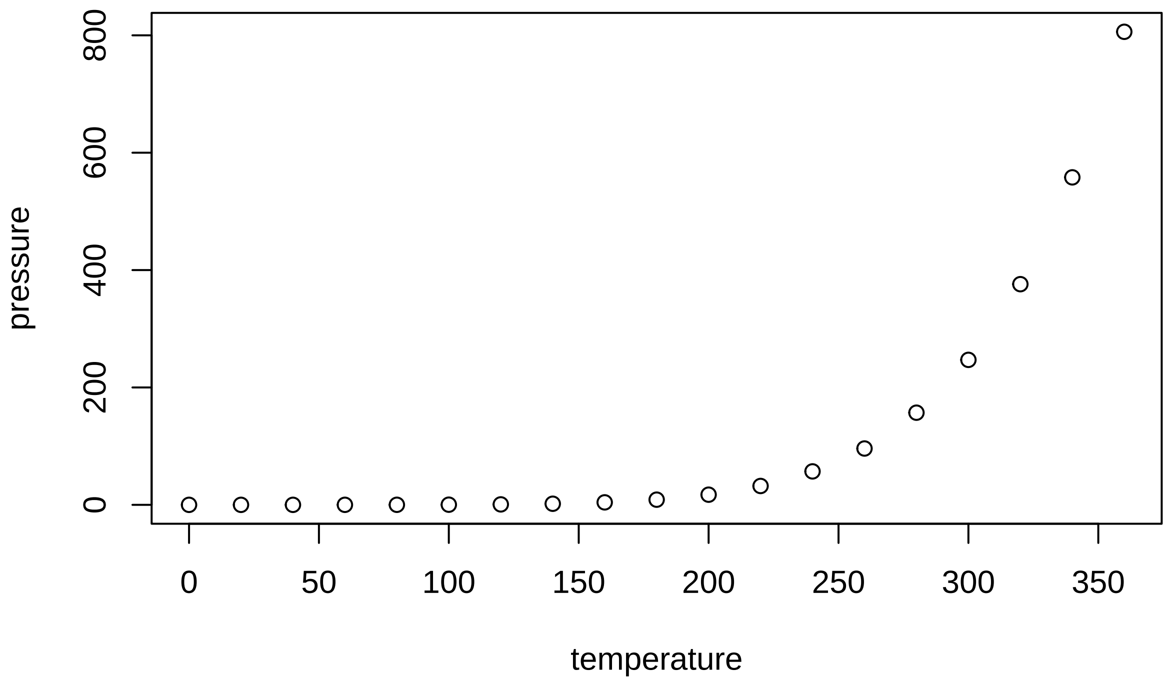



```
We know from _the first fundamental theorem of calculus_ that
for $x$ in $[a, b]$:
$$\frac{d}{dx}\left( \int_{a}^{x} f(u)\,du\right)=f(x).$$
```

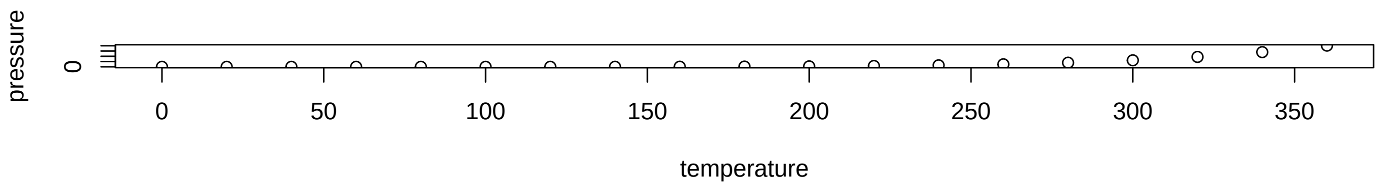
7.2.6.3.2 `out.width` vs. `fig.width` https://stackoverflow.com/questions/29657777/how-to-make-fig-width-and-out-width-consistent-with-knitr

when chunk option `cache=FALSE` is set, then `out.width` has no effect because no PDF output is created. Hence one has to specify exact measures in inches for `fig.width` and `fig.height` for each chunk

https://stackoverflow.com/questions/59567235/a-ggmap-too-small-when-rendered-within-a-rmd-file



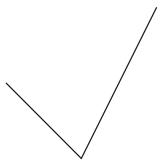

```
plot(pressure)
```



problem: `out.width='100%'` causing `LaTeX Error: Not in outer par mode.`

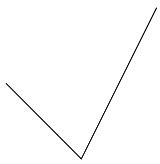
solution: `out.width=if (knitr:::is_html_output()) '100%'`

```
\begin{tikzpicture}
  \draw (-1,1)--(0,0)--(1,2);
\end{tikzpicture}
```



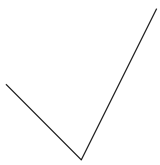
`fig.width=10, fig.height=2`

```
\begin{tikzpicture}
  \draw (-1,1)--(0,0)--(1,2);
\end{tikzpicture}
```



`out.width=if (knitr:::is_html_output()) '100%'`

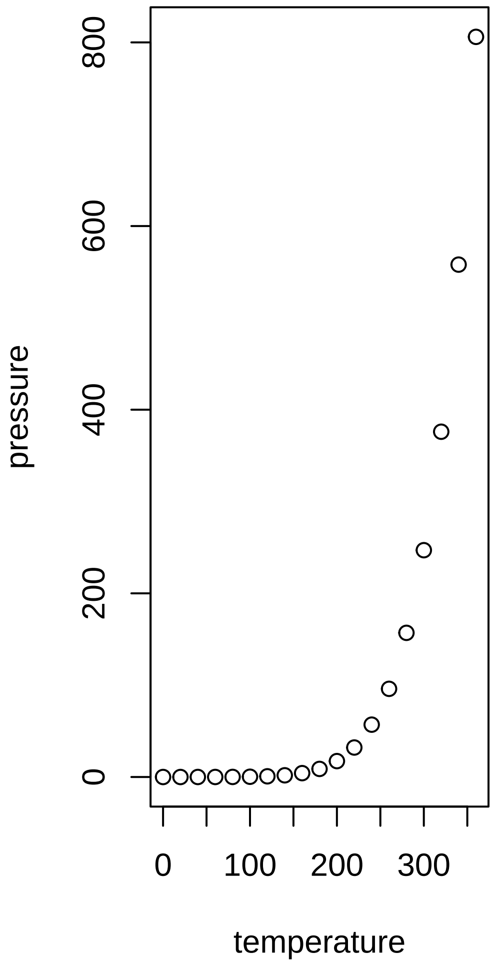
```
\begin{tikzpicture}
  \draw (-1,1)--(0,0)--(1,2);
\end{tikzpicture}
```



7.2.6.4 dynamic knitr plot width and height

https://stackoverflow.com/questions/15365829/dynamic-height-and-width-for-knitr-plots

```
plot(pressure)
```



7.2.6.5 web image in PDF

https://stackoverflow.com/questions/46331896/how-can-i-insert-an-image-from-internet-to-the-pdf-file-produced-by-r-bookdown-i

```
cover_url = 'https://bookdown.org/yihui/bookdown/images/cover.jpg'
if (!file.exists(cover_file <- xfun::url_filename(cover_url)))
  xfun::download_file(cover_url)
knitr::include_graphics(if (knitr::pandoc_to('html')) cover_url else cover_file)
```

7.2.6.6 SVG

https://stackoverflow.com/questions/50165404/how-to-make-a-pdf-using-bookdown-including-svg-images

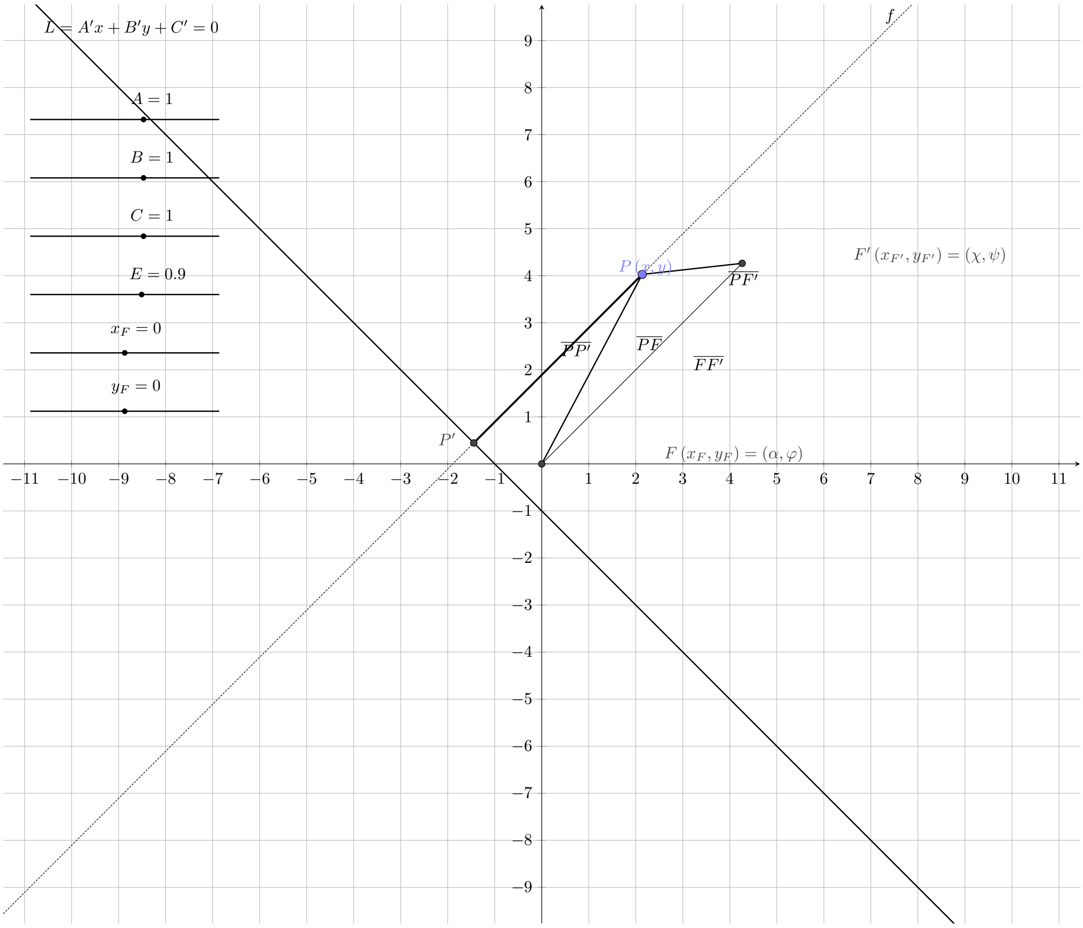
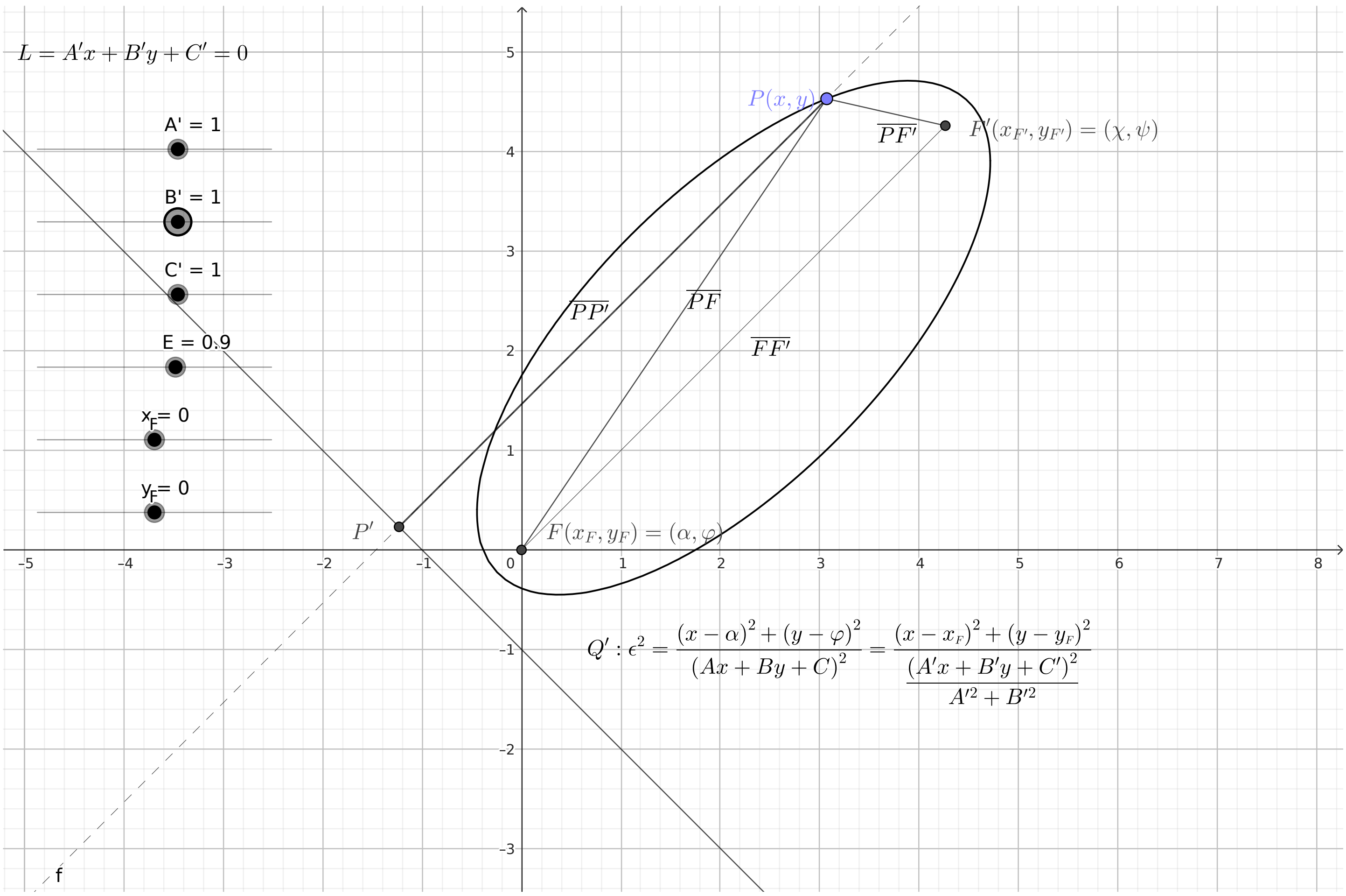


Figure 7.1: conic sections



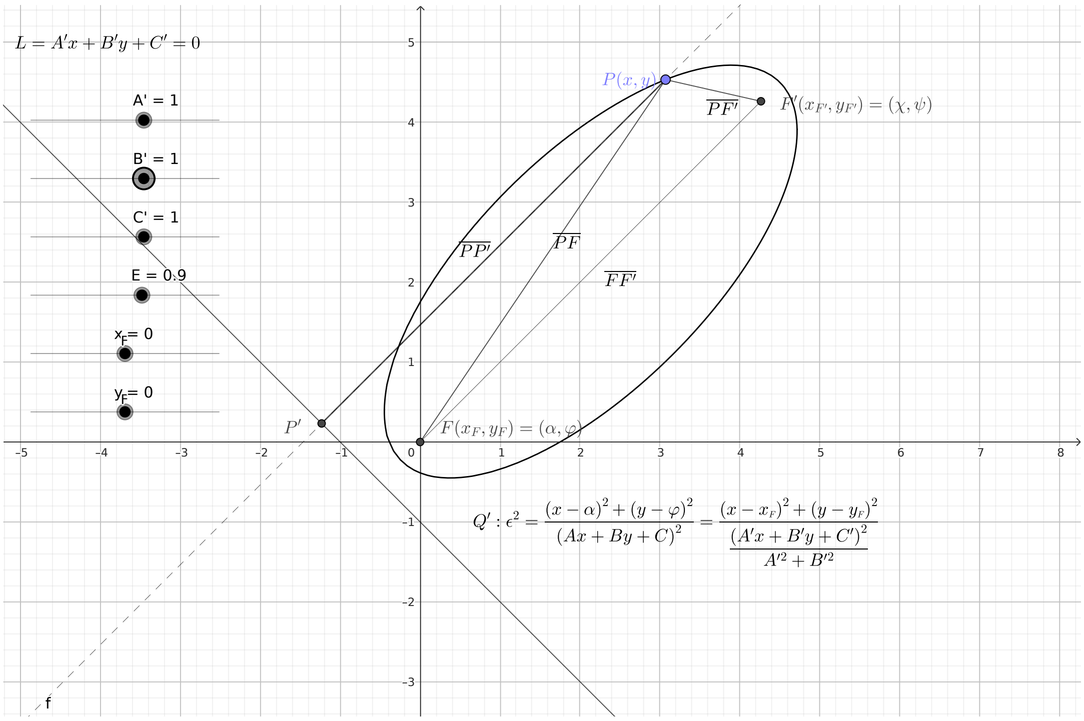


Figure 7.2: conic sections

https://stackoverflow.com/questions/34064292/is-it-possible-to-include-svg-image-in-pdf-document-rendered-by-rmarkdown

7.2.7 horizontal rule

```
***
```

horizontal rule (or slide break)

```r
dim(iris)
```

```
## [1] 150   5
```

7.2.8 footnote

7.2.9 hyperlink

PDF pandoc internal link will lose focus

equivalence relation [11] equivalence relation[1] equivalence relation[11]

equivalence class [10] equivalence class[2] equivalence class[10]

partition [9] partition[3] partition[9]

- LaTeX
  - TikZ[13]
    - TikZ-3Dplot
    - PGFplots
  - xypic = xy-pic[4]
- OverLeaf
- MathCha
- GeoGebra
- Python

[1] {11} equivalence relation
[2] {10} equivalence class
[3] {9} partition
[4] {14} xy-pic

  - MatPlotLib
  - Seaborn
  - Plotly

7.2.10 code chunk

7.2.10.1 code folding

https://cosname.github.io/rmarkdown-guide/rmarkdown-document.html#html-code-folding

7.2.11 xaringan

slide realtime preview with RStudio addin Infinite Moon Reader in RStudio viewer

https://github.com/yihui/xaringan

https://www.youtube.com/watch?v=3n9nASHg9gc

7.3 Bookdown

7.3.1 system locale

https://bookdown.org/tpemartin/ntpu-programming-for-data-science/appendix-d-.html

`Sys.getlocale()`

Windows

`Sys.setlocale(category = "LC_ALL", locale = "UTF-8")`

MacOS

`Sys.setlocale(category = "LC_ALL", locale = "en_US.UTF-8")`

https://bookdown.org/yihui/rmarkdown-cookbook/multi-column.html

7.3.2 `render_book()`

https://bookdown.org/yihui/bookdown/build-the-book.html

```
render_book(input = ".", output_format = NULL, ..., clean = TRUE,
  envir = parent.frame(), clean_envir = !interactive(),
  output_dir = NULL, new_session = NA, preview = FALSE,
  config_file = "_bookdown.yml")
```

7.3.3 `serve_book()`

https://bookdown.org/yihui/bookdown/serve-the-book.html

```
serve_book(dir = ".", output_dir = "_book", preview = TRUE,
  in_session = TRUE, quiet = FALSE, ...)
```

7.3.4 LaTeX

7.3.4.1 hyperlink, URL, href

https://www.baeldung.com/cs/latex-hyperref-url-hyperlinks

https://www.omdte.com/小技巧讓-facebook和-line顯示中文網址，網址不再變亂碼/

7.3.4.2 ugly mathptmx $\sum$

PDF LaTeX `\usepackage{fdsymbol}` to have `\overrightharpoon` vector; however, there are too many side effects, including ugly mathptmx $\sum$, ...

```
\usepackage{fdsymbol} % vector over accent, but will use mathptmx
% replace the rather ugly mathptmx \sum operator with the equivalent Computer Modern one
\let\sum\relax
\DeclareSymbolFont{CMlargesymbols}{OMX}{cmex}{m}{n}
\DeclareMathSymbol{\sum}{\mathop}{CMlargesymbols}{"50}
```

https://tex.stackexchange.com/questions/315102/different-sum-signs

https://tex.stackexchange.com/questions/275038/how-to-replace-mathptmx-sum-with-cm-sum

https://tex.stackexchange.com/questions/391410/calligraphic-symbols-are-too-fancy-with-mathptmx-package

https://blog.csdn.net/kongtaoxing/article/details/131005044

In `preamble.tex`, add

```
% replace the rather ugly mathptmx \sum operator with the equivalent Computer Modern one
\let\sum\relax
\DeclareSymbolFont{CMlargesymbols}{OMX}{cmex}{m}{n}
\DeclareMathSymbol{\sum}{\mathop}{CMlargesymbols}{"50}

\DeclareMathAlphabet{\mathcal}{OMS}{cmsy}{m}{n}
\DeclareSymbolFont{largesymbols}{OMX}{cmex}{m}{n}
```

7.3.4.3 LaTeX package in HTML document

https://github.com/rstudio/rmarkdown/issues/1829

```
---
title: "assignment"
author: "author"
output: html_document
---

$$
  \require{cancel}
  \cancel{x}
$$
```

$$\cancel{x}$$

https://stackoverflow.com/questions/18189175/how-to-use-textup-with-mathjax

> `\textup` is not available in MathJax. You can replace it with `\mathrm`, **but `\mathrm` does not interpret spaces**.

7.3.5 depth of table of contents `toc_depth`

https://stackoverflow.com/questions/49009212/how-to-change-toc-depth-in-r-bookdown-gitbook

```
bookdown::gitbook:
    toc_depth: 2
```

https://stackoverflow.com/questions/68537309/how-can-i-specific-the-initial-level-to-have-my-table-of-contents-be-expanded-to

```
toc:
  collapse: section
```

7.3.6 multi-column layout / two columns

https://bookdown.org/yihui/rmarkdown-cookbook/multi-column.html

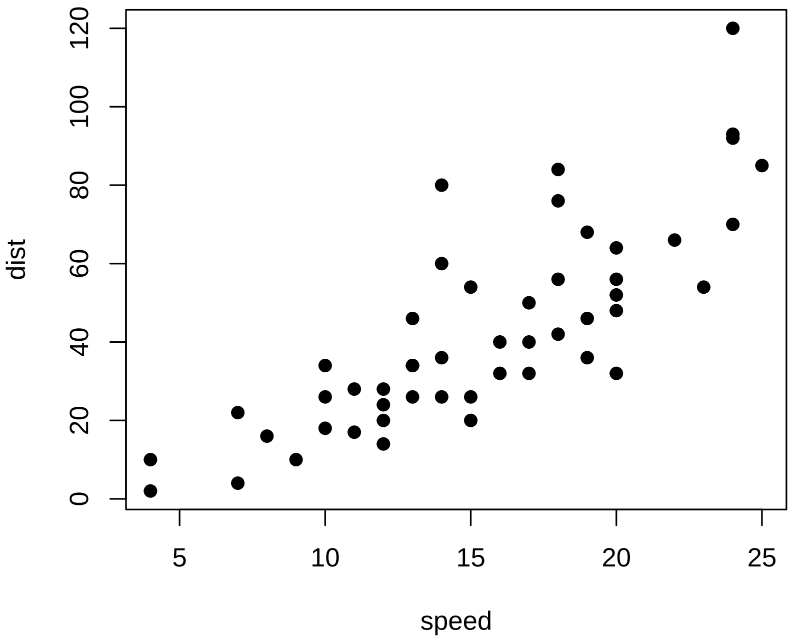
7.3.6.1 for both HTML and PDF

figure size[7.2.6.3]

> Below is a Div containing three child Divs side by side. The Div in the middle is empty, just to add more space between the left and right Divs.

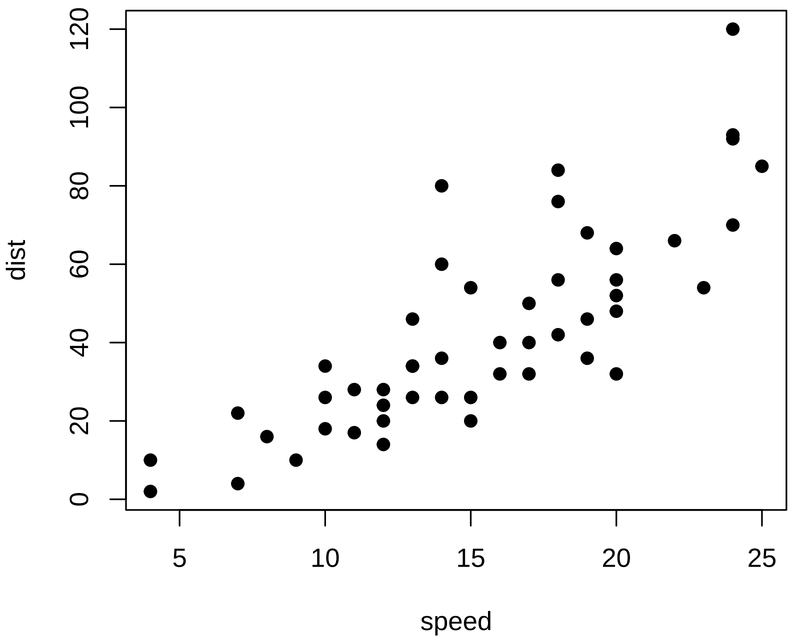
```
:::::: {.cols data-latex=""}

::: {.col data-latex="{0.55\textwidth}"}
![](202401280001-test_files/figure-latex/unnamed-chunk-32-1.pdf)<!-- -->
```

```
:::

::: {.col data-latex="{0.05\textwidth}"}
\
<!-- an empty Div (with a white space), serving as
a column separator -->
:::

::: {.col data-latex="{0.4\textwidth}"}
The figure on the left-hand side shows the `cars` data.

Lorem ipsum dolor sit amet, consectetur adipiscing elit, sed do
eiusmod tempor incididunt ut labore et dolore magna aliqua. Ut
enim ad minim veniam, quis nostrud exercitation ullamco laboris
nisi ut aliquip ex ea commodo consequat. Duis aute irure dolor
in reprehenderit in voluptate velit esse cillum dolore eu fugiat
nulla pariatur.
:::
::::::

{r, echo=FALSE, fig.width=5, fig.height=4}
```



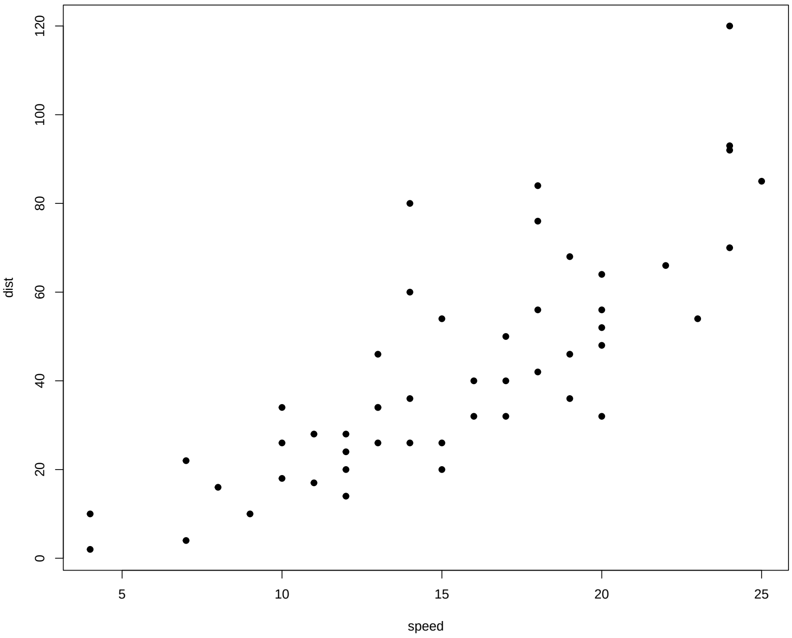
The figure on the left-hand side shows the `cars` data.
Lorem ipsum dolor sit amet, consectetur adipiscing elit, sed do eiusmod tempor incididunt ut labore et dolore magna aliqua. Ut enim ad minim veniam, quis nostrud exercitation ullamco laboris nisi ut aliquip ex ea commodo consequat. Duis aute irure dolor in reprehenderit in voluptate velit esse cillum dolore eu fugiat nulla pariatur.

```
{r, echo=FALSE, fig.width=10, fig.height=8, out.width = "100%"}
```



The figure on the left-hand side shows the `cars` data.
Lorem ipsum dolor sit amet, consectetur adipiscing elit, sed do eiusmod tempor incididunt ut labore et dolore magna aliqua. Ut enim ad minim veniam, quis nostrud exercitation ullamco laboris nisi ut aliquip ex ea commodo consequat. Duis aute irure dolor in reprehenderit in voluptate velit esse cillum dolore eu fugiat nulla pariatur.

`{r, echo=FALSE, fig.width=10, fig.height=8}`



The figure on the left-hand side shows the `cars` data.
Lorem ipsum dolor sit amet, consectetur adipiscing elit, sed do eiusmod tempor incididunt ut labore et dolore magna aliqua. Ut enim ad minim veniam, quis nostrud exercitation ullamco laboris nisi ut aliquip ex ea commodo consequat. Duis aute irure dolor in reprehenderit in voluptate velit esse cillum dolore eu fugiat nulla pariatur.

7.3.6.2 multi-column `fig.cap` must use `fig.pos="H"`

https://community.rstudio.com/t/adding-fig-cap-caption-text-to-code-chunk-causes-figure-to-print-at-top-of-page-instead-of-where-it-should-be/30297

https://bookdown.org/yihui/rmarkdown-cookbook/figure-placement.html

to avoid `LaTeX Error: Not in outer par mode` for caption in multi-column LaTeX PDF

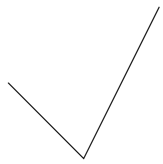
in `output.yml` add `extra_dependencies: ["float"]` under `bookdown::pdf_book:`

include first chunk `knitr::opts_chunk$set(fig.pos = "H", out.extra = "")` in .Rmd

add `out.width=if (knitr:::is_html_output()) '50%'` for TikZ chunk

thus complete chunk beginning with `{r, echo=FALSE, cache=TRUE, engine='tikz', fig.ext=if (knitr:::is_latex_output( 'pdf' else 'png', fig.width=10, fig.height=2, out.width=if (knitr:::is_html_output()) '100%', fig.cap=''}`

```
\begin{tikzpicture}
  \draw (-1,1)--(0,0)--(1,2);
\end{tikzpicture}
```



LaTeX package `caption`

https://tex.stackexchange.com/questions/128485/how-to-make-a-caption-via-captionof-and-extra-margins-adhere-to-minipage-marg

What is the different between using `\captionof{table}{ABC}` and `\caption{ABC}`?

https://tex.stackexchange.com/questions/514286/what-is-the-different-between-using-captionoftableabc-and-captionabc

side-by-side table

https://stackoverflow.com/questions/73745714/how-to-print-gt-tbl-tables-side-by-side-with-knitr-kable

R ternary operator

https://stackoverflow.com/questions/8790143/does-the-ternary-operator-exist-in-r

7.3.6.3 caption above figure

https://stackoverflow.com/questions/56979022/caption-above-figure-in-html-rmarkdown

`fig.topcaption=TRUE`

7.3.6.4 for only HTML

7.3.6.4.1 CSS flex

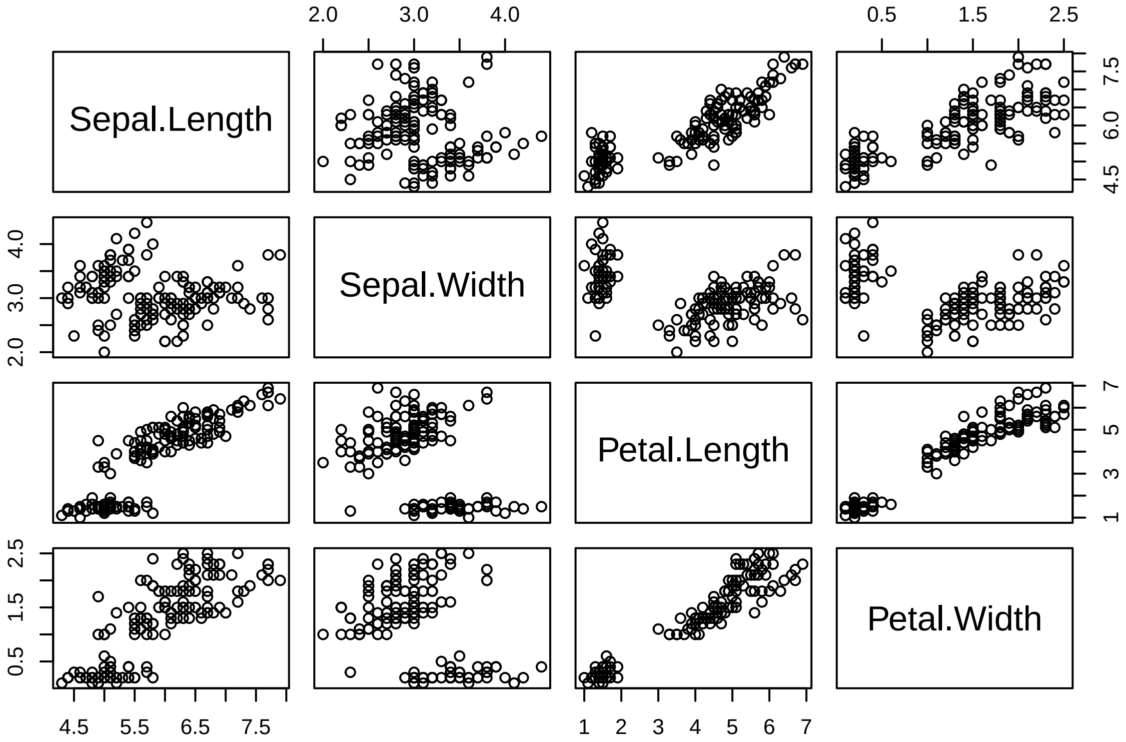
Here is the **first** Div.

```
str(iris)
```

```
## 'data.frame':	150 obs. of  5 variables:
##  $ Sepal.Length: num  5.1 4.9 4.7 4.6 5 5.4 4.6 5 4.4 4.9 ...
##  $ Sepal.Width : num  3.5 3 3.2 3.1 3.6 3.9 3.4 3.4 2.9 3.1 ...
##  $ Petal.Length: num  1.4 1.4 1.3 1.5 1.4 1.7 1.4 1.5 1.4 1.5 ...
##  $ Petal.Width : num  0.2 0.2 0.2 0.2 0.2 0.4 0.3 0.2 0.2 0.1 ...
##  $ Species     : Factor w/ 3 levels "setosa","versicolor",..: 1 1 1 1 1 1 1 1 1 1 ...
```

And this block will be put on the right:

```
plot(iris[, -5])
```



7.3.6.4.2 CSS grid

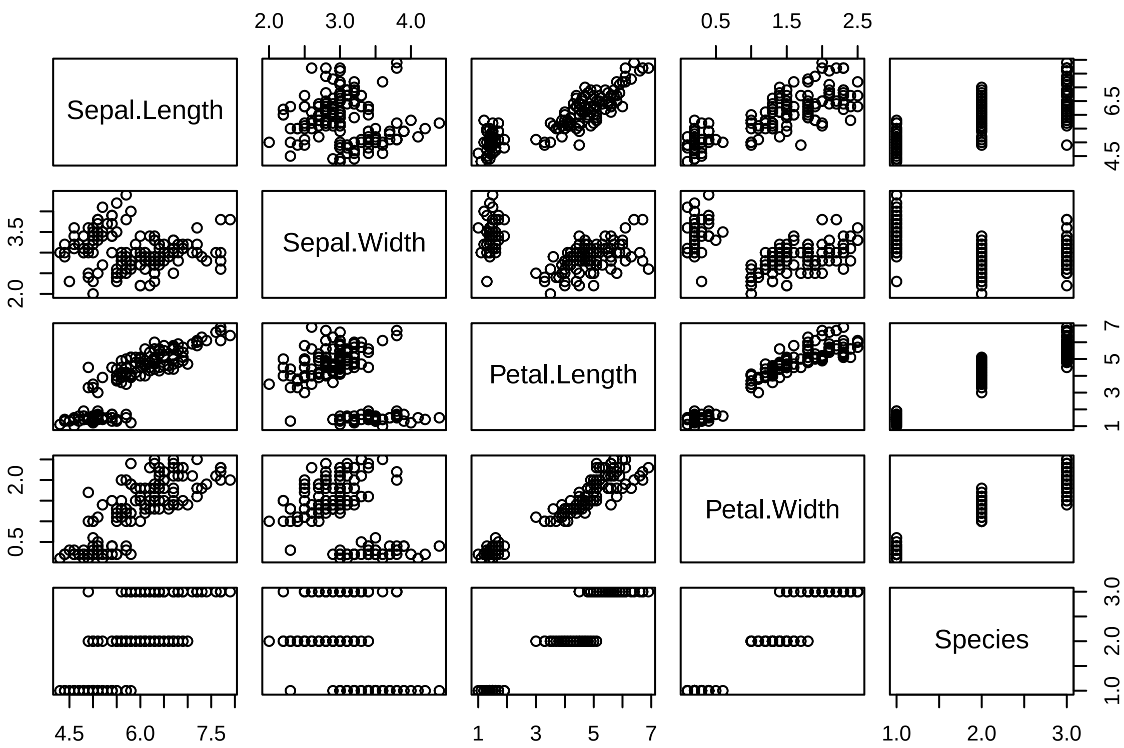
https://github.com/yihui/knitr/issues/1743

https://medium.com/enjoy-life-enjoy-coding/css-所以我說那個版能不能好切一點-grid-基本用法-cd763091cf70

```
head(iris)
```

```
##   Sepal.Length Sepal.Width Petal.Length Petal.Width Species
## 1          5.1         3.5          1.4         0.2  setosa
## 2          4.9         3.0          1.4         0.2  setosa
## 3          4.7         3.2          1.3         0.2  setosa
## 4          4.6         3.1          1.5         0.2  setosa
## 5          5.0         3.6          1.4         0.2  setosa
## 6          5.4         3.9          1.7         0.4  setosa
```

```r
plot(iris)
```



7.4 conditional block/chunk for either HTML or PDF, and Chinese issue

https://stackoverflow.com/questions/76240244/bookdown-conditional-display-of-text-and-code-blocks-latex-pdf-vs-html

等價關係 equivalence relation

$$
\begin{aligned}
& R \text{ is an equivalence relation over } A \times B \\
\Leftrightarrow & \begin{cases} R = \sim = \{\langle x,y\rangle | x \sim y\} \subseteq A \times B & \text{(e) equivalence 等價} \\ \vdots & \vdots \end{cases} \\
\Leftrightarrow & \begin{cases} R = \{\langle x,y\rangle | xRy\} \subseteq A \times B & (R)\,\text{relation} \\ \forall \langle x,y\rangle \in R\,(xRx) & (r)\,\text{reflexive} \\ \forall \langle x,y\rangle \in R\,(xRy \Rightarrow yRx) & (s)\,\text{symmetric} \\ \forall \langle x,y\rangle, \langle y,z\rangle \in R \left(\begin{cases} xRy \\ yRz \end{cases} \Rightarrow xRz \right) & (t)\,\text{transitive} \end{cases}
\Leftrightarrow \begin{cases} R = \{\langle x,y\rangle | xRy\} \subseteq A \times B & \text{關係} \\ \forall \langle x,y\rangle \in R\,(\langle x,x\rangle \in R) & \text{自反} \\ \forall \langle x,y\rangle \in R\,(\langle y,x\rangle \in R) & \text{對稱} \\ \forall \langle x,y\rangle, \langle y,z\rangle \in R\,(\langle x,z\rangle \in R) & \text{遞移} \end{cases}
\end{aligned}
$$

7.5 video embedding

https://stackoverflow.com/questions/42543206/r-markdown-compile-error

always_allow_html: true

```r
install.packages("webshot")
webshot::install_phantomjs()
```

however webshot not work

Error: cannot find bilibili.com

https://cran.r-project.org/web/packages/vembedr/vignettes/embed.html

```
## embed_youtube("0LFg5dvP0oc")
```

7.5.1 timestamp

- YouTube: https://www.youtube.com/embed/%7BvideoID%7D?start=%7Bsecond%7D
- BiliBili: https://player.bilibili.com/player.html?bvid=%7BvideoID%7D&autoplay=0&t=%7Bsecond%7D

7.6 equation term coloring

LaTeX annotation by TikZ[62]

7.6.1 font color

RegEx replacement in RStudio for `{\\color{(\w+)}` in LyX to be replaced with `\\color{$1}{` in HTML document, and remain the same for PDF document

In HTML document, if no `{}` for text range, only the first following term will take effect

`\color{orange}x=y`

$$\color{orange}x=y$$

`\color{orange}` and `\color{cyan}` are better color for HTML GitBook `White` and `Night` themes and PDF

`\color{cyan}{x=y}`

$$\color{orange}x=y$$

`\color{cyan}{x=y}`

$$\color{cyan}{x=y}$$

```
::: {show-in="html"}

$$
\dfrac{\colorbox{#FFD1DC}{$\epsilon^{2}\left(y_{{\scriptscriptstyle F}}-y_{{\scriptscriptstyle
↪  L}}\right)^{2}$}}{1-\epsilon^{2}}
$$

:::

::: {show-in="pdf"}

$$
\dfrac{\colorbox{red!50}{\text{\ensuremath{\epsilon^{2}\left(y_{{\scriptscriptstyle
↪  F}}-y_{{\scriptscriptstyle L}}\right)^{2}}}}}{1-\epsilon^{2}}
$$

:::
```

7.6.2 background color

https://bookdown.org/yihui/rmarkdown-cookbook/font-color.html

LaTex color

https://latexcolor.com/

https://www.overleaf.com/learn/latex/Using_colors_in_LaTeX

https://latex-tutorial.com/color-latex/#:~:text=To%20summarize%2C%20pyellow!50efined%20colors%20in,when%20loading%20the

LaTex color methods

color frame

https://tex.stackexchange.com/questions/582748/highlight-equation-with-boxes-and-arrows color box

https://tex.stackexchange.com/questions/567739/how-to-move-and-size-colorbox

color box with round corners

https://tex.stackexchange.com/questions/568880/color-box-with-rounded-corners

highlighting

https://tex.stackexchange.com/questions/318991/highlighting-math

https://forum.remnote.io/t/highlighting-latex-formulas/149

LyX

https://tex.stackexchange.com/questions/250069/create-a-color-box https://latexlyx.blogspot.com/2013/12/lyx.html

https://tex.stackexchange.com/questions/635486/prevent-lyx-from-escaping-math-in-color-box-title

Bookdown - conditional display of text and code blocks (LaTeX/PDF vs. HTML) https://stackoverflow.com/questions/76240244/bookdown-conditional-display-of-text-and-code-blocks-latex-pdf-vs-html

$$F = ma$$

https://community.rstudio.com/t/highlighting-text-inline-in-rmarkdown-or-bookdown-pdf/35118/4

$$F = ma$$

$$F = F$$

$$F = ma \tag{7.1}$$

$$F = ma$$

$$Y = \beta_0 + \beta_1 X_1 + \ldots + \beta_n X_n$$

7.7 link and reference

https://stackoverflow.com/questions/57469501/cross-referencing-bookdownhtml-document2-not-working

$$E = mc^2 \tag{7.2}$$

`\@ref(nice-label)` 7.8

`[link to partition][partition]` link to partition

`[partition] \@ref(partition)`

partition [#partition] (9) @ref(#partition)

`[equivalence class] \@ref(equivalence-class)`

equivalence class (10)

equivalence class [#equivalence class] (@ref(equivalence class)) @ref(#equivalence class)

[equivalence-class] [#equivalence-class] (10) @ref(#equivalence-class)

X [equivalence-class.html] [equivalence-class.html#equivalence-class] (@ref(equivalence-class.html)) @ref(equivalence-class.html#equivalence-class)

equivalence relation [#equivalence relation] (@ref(equivalence relation)) @ref(#equivalence relation)

[equivalence-relation] [#equivalence-relation] (11) @ref(#equivalence-relation)

X [equivalence-relation.html] [equivalence-relation.html#equivalence-relation] (@ref(equivalence-relation.html)) @ref(equivalence-relation.html#equivalence-relation)

7.8 number and reference equations

https://stackoverflow.com/questions/71595882/rstudio-error-in-windows-running-pdflatex-exe-on-file-name-tex-exit-code-10

https://bookdown.org/yihui/rmarkdown/bookdown-markdown.html#equations

`\#eq:emc \@ref(eq:emc)`

https://stackoverflow.com/questions/55923290/consistent-math-equation-numbering-in-bookdown-across-pdf-docx-html-output

$$
\begin{aligned}
&C \text{ is an equivalence class of } a \text{ on } A\\
&\Leftrightarrow [a]_\sim = C = \left\{ x \middle| \begin{cases} a \in A \\ x \in A \\ x \sim a \\ \sim \text{ is an equivalence relation over } A \times A = A^2 \end{cases} \right\} \subseteq A \neq \emptyset \\
&\Leftrightarrow [a] = [a]_\sim = \left\{ x \middle| \begin{cases} a \in A \\ x \in A \\ x \sim a \\ \sim \text{ is an equivalence relation on } A \end{cases} \right\} \subseteq A \neq \emptyset \\
&\Rightarrow [a]_\sim = \{x | x \sim a\} \subseteq A \neq \emptyset
\end{aligned} \tag{7.3}
$$

https://bookdown.org/yihui/rmarkdown/bookdown-markdown.html#cross-referencing

This cross reference is the Fig. 7.4

https://stackoverflow.com/questions/51595939/bookdown-cross-reference-figure-in-another-file

I ran into the same issue and came up with this solution if you aim at compiling 2 different pdfs. It relies on LaTeX's xr package for cross references: https://stackoverflow.com/a/52532269/576684

7.9 footnote

```
noun^[This is a footnote]

noun[^202401260000-test-cross-link-1]

[^202401260000-test-cross-link-1]: This is a footnote.
```

noun[5]

7.10 citation

https://stackoverflow.com/questions/48965247/use-csl-file-for-pdf-output-in-bookdown/49145699#49145699

citation 1[3] citation 2[3]

citation 3[4] citation 4[4]

[5] This is a footnote.

7.10.1 citation in `fig.cap`

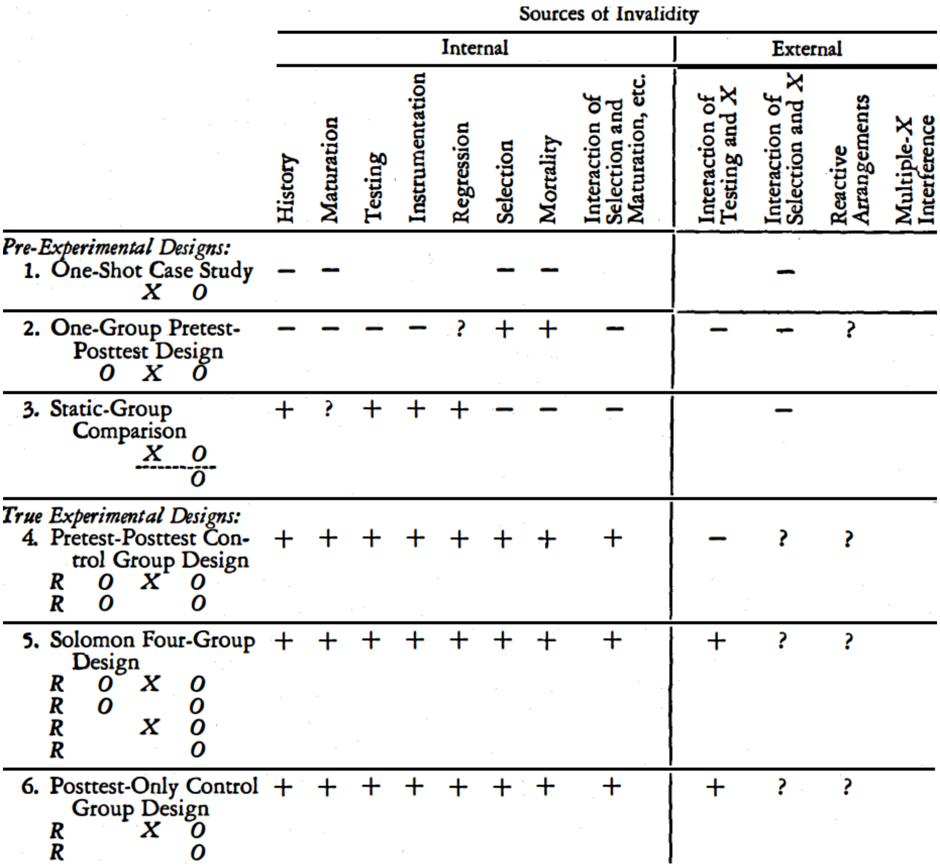
https://tex.stackexchange.com/questions/591882/citation-within-a-latex-figure-caption-in-rmarkdown

`(ref:rudolph) *nice* cite: [@Lam94].`

`(ref:campbell1963) *nice* cite: [@campbell1963].`

`(ref:campbell1963) ([@campbell1963]`

`(ref:campbell1963) \ [@campbell1963]`



	Sources of Invalidity											
	Internal								External			
	History	Maturation	Testing	Instrumentation	Regression	Selection	Mortality	Interaction of Selection and Maturation, etc.	Interaction of Testing and X	Interaction of Selection and X	Reactive Arrangements	Multiple-X Interference
Pre-Experimental Designs: 1. One-Shot Case Study X O	−	−				−	−			−		
2. One-Group Pretest-Posttest Design O X O	−	−	−	−	?	+	+	−	−	−	?	
3. Static-Group Comparison X O / O	+	?	+	+	+	−	−	−		−		
True Experimental Designs: 4. Pretest-Posttest Control Group Design R O X O / R O O	+	+	+	+	+	+	+	+	−	?	?	
5. Solomon Four-Group Design R O X O / R O O / R X O / R O	+	+	+	+	+	+	+	+	+	?	?	
6. Posttest-Only Control Group Design R X O / R O	+	+	+	+	+	+	+	+	+	?	?	

Figure 7.3: pre- and true experimental designs ([5] p.8)

7.10.2 backreference

https://community.rstudio.com/t/how-to-create-a-backreference-to-place-of-citation-in-rmarkdown/84866

https://blog.csdn.net/RobertChenGuangzhi/article/details/50455429

https://latex.org/forum/viewtopic.php?t=3722

7.11 environment for definition, theorem, proof

https://bookdown.org/yihui/rmarkdown/bookdown-markdown.html

https://github.com/rstudio/rstudio/issues/5264

`@howthebodyworks` Ideally, previews of such equations should also work inside a theorem, although I could survive without that.

https://github.com/rstudio/rstudio/issues/8773

Theorem 7.1 (Theorem Name)**.** *Here is my theorem.*

Proof Name. Here is my proof. □

Theorem 7.2 (Pythagorean theorem)**.** *For a right triangle, if c denotes the length of the hypotenuse and a and b denote the lengths of the other two sides, we have*

$$a^2 + b^2 \overset{7.2}{=} c^2$$

Definition 7.1 (Definition Name)**.** Here is my definition.

number and reference equations

(7.3)

(7.2)

7.2

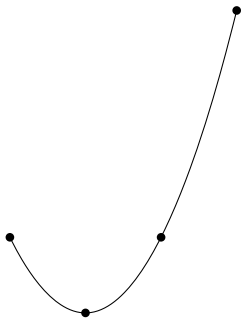


Figure 7.4: parabola arc with points

7.12 slide or presentation

7.12.1 Xaringan and Infinite Moon Reader

https://rpubs.com/RW1304/xarigan-zh

https://slides.yihui.org/xaringan/#1

https://slides.yihui.org/xaringan/zh-CN.html#1

https://github.com/yihui/xaringan/tree/master

https://bookdown.org/yihui/rmarkdown/xaringan.html

7.12.2 ioslides

https://www.youtube.com/watch?v=gkyjTcpCITM

https://bookdown.org/yihui/rmarkdown/ioslides-presentation.html

https://stackoverflow.com/questions/63749683/how-to-set-up-theorem-environment-in-the-rmarkdown-presentation

```
---
title: "Theorem demo"
output:
  ioslides_presentation:
    css: style.css
---
```

```

/* theorem environment _ plain */

/*
.theorem {
  display: block;
  font-style: italic;
  font-size: 24px;
  font-family: "Times New Roman";
  color: black;
}
.theorem::before {
```

```
  content: "Theorem. ";
  font-weight: bold;
  font-style: normal;
}
.theorem[text]::before {
  content: "Theorem (" attr(text) ") ";
}
.theorem p {
  display: inline;
}
*/

/* theorem environment _ Copenhagen style */

/*
.theorem {
  display: block;
  font-style: italic;
  font-size: 24px;
  font-family: "Times New Roman";
  color: black;
  border-radius: 10px;
  background-color: rgb(222,222,231);
  box-shadow: 5px 10px 8px #888888;
}
.theorem::before {
  content: "Theorem. ";
  font-weight: bold;
  font-style: normal;
  display: inline-block;
  width: -webkit-fill-available;
  color: white;
  border-radius: 10px 10px 0 0;
  padding: 10px 5px 5px 15px;
  background-color: rgb(38, 38, 134);
}
.theorem p {
  padding: 15px 15px 15px 15px;
}
*/
```

7.12.3 PowerPoint

https://bookdown.org/yihui/rmarkdown/powerpoint-presentation.html

Chapter 8

test2

8.1 verbatim

https://community.rstudio.com/t/continued-issues-with-new-verbatim-in-rstudio/139737

https://bookdown.org/yihui/rmarkdown-cookbook/verbatim-code-chunks.html

````
```r
1 + 1
```

```
## [1] 2
```
````

````
We can output arbitrary content **verbatim**.

```r
1 + 1
```

```
## [1] 2
```

The content can contain inline code like
78.5398163, too.
````

Chapter 9

partition

$$
\begin{aligned}
&\{A_i\}_{i\in I} = \{A_i | i \in I\} \text{ is a partition of a set } A\\
\Leftrightarrow &\begin{cases}\forall i \in I\,(A_i \neq \emptyset)\\ A = \bigcup\limits_{i\in I} A_i\\ \forall i, j \in I\,(i \neq j \Rightarrow A_i \cap A_j = \emptyset)\end{cases}
\end{aligned}
$$

https://proofwiki.org/wiki/Definition:Set_Partition

Chapter 10

equivalence class

$$
\begin{aligned}
&C \text{ is an equivalence class of } a \text{ on } A \\
&\Leftrightarrow [a]_\sim = C = \left\{ x \middle| \begin{cases} a \in A \\ x \in A \\ x \sim a \\ \sim \text{ is an equivalence relation over } A \times A = A^2 \end{cases} \right\} \subseteq A \neq \emptyset \\
&\Leftrightarrow [a] = [a]_\sim = \left\{ x \middle| \begin{cases} a \in A \\ x \in A \\ x \sim a \\ \sim \text{ is an equivalence relation on } A \end{cases} \right\} \subseteq A \neq \emptyset \\
&\Rightarrow [a]_\sim = \{x | x \sim a\} \subseteq A \neq \emptyset
\end{aligned}
$$

where the definition of equivalence relation can be found in 11.

Chapter 11

equivalence relation

等價關係 equivalence relation

$$
\begin{aligned}
&R \text{ is an equivalence relation over } A \times B \\
\Leftrightarrow &\begin{cases}
R = \sim = \{\langle x, y\rangle | x \sim y\} \subseteq A \times B & \text{(e) equivalence 等價} \\
\vdots & \vdots
\end{cases} \\
\Leftrightarrow &\begin{cases}
R = \{\langle x, y\rangle | xRy\} \subseteq A \times B & (R)\,\text{relation} \\
\forall \langle x, y\rangle \in R\,(xRx) & (r)\,\text{reflexive} \\
\forall \langle x, y\rangle \in R\,(xRy \Rightarrow yRx) & (s)\,\text{symmetric} \\
\forall \langle x, y\rangle, \langle y, z\rangle \in R \left(\begin{cases} xRy \\ yRz \end{cases} \Rightarrow xRz\right) & (t)\,\text{transitive}
\end{cases}
\Leftrightarrow \begin{cases}
R = \{\langle x, y\rangle | xRy\} \subseteq A \times B & \text{關係} \\
\forall \langle x, y\rangle \in R\,(\langle x, x\rangle \in R) & \text{自反} \\
\forall \langle x, y\rangle \in R\,(\langle y, x\rangle \in R) & \text{對稱} \\
\forall \langle x, y\rangle, \langle y, z\rangle \in R\,(\langle x, z\rangle \in R) & \text{遞移}
\end{cases}
\end{aligned}
$$

Chapter 12

Python

12.1 using Python in R / RMarkdown

https://bookdown.org/yihui/rmarkdown/language-engines.html

```
names(knitr::knit_engines$get())
```

```
##  [1] "awk"         "bash"        "coffee"      "gawk"        "groovy"
##  [6] "haskell"     "lein"        "mysql"       "node"        "octave"
## [11] "perl"        "php"         "psql"        "Rscript"     "ruby"
## [16] "sas"         "scala"       "sed"         "sh"          "stata"
## [21] "zsh"         "asis"        "asy"         "block"       "block2"
## [26] "bslib"       "c"           "cat"         "cc"          "comment"
## [31] "css"         "ditaa"       "dot"         "embed"       "eviews"
## [36] "exec"        "fortran"     "fortran95"   "go"          "highlight"
## [41] "js"          "julia"       "python"      "R"           "Rcpp"
## [46] "sass"        "scss"        "sql"         "stan"        "targets"
## [51] "tikz"        "verbatim"    "theorem"     "lemma"       "corollary"
## [56] "proposition" "conjecture"  "definition"  "example"     "exercise"
## [61] "hypothesis"  "proof"       "remark"      "solution"    "glue"
## [66] "glue_sql"    "gluesql"
```

https://rstudio.github.io/reticulate/articles/python_packages.html

```
x = 'hello, python world!'
print(x.split(' '))
```

```
## ['hello,', 'python', 'world!']
```

```
library(reticulate)
virtualenv_python()
```

```
library(reticulate)
# conda_list()
```

```
library(reticulate)
virtualenv_list()
```

https://rstudio.github.io/reticulate/reference/install_python.html

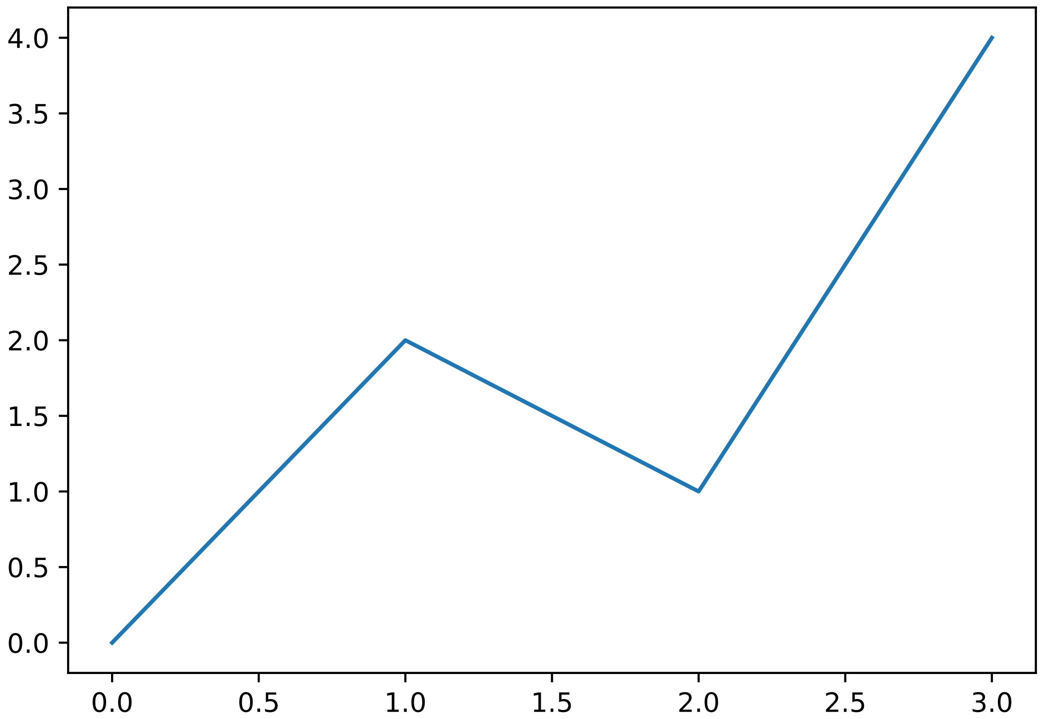
```
library(reticulate)
version <- "3.9.12"
# install_python(version)

## create a new environment
# virtualenv_create("r-reticulate", version = version)

# use_virtualenv("r-reticulate")

## install MatPlotLib
```

```
# virtualenv_install("r-reticulate", "matplotlib")

## import MatPlotLib (it will be automatically discovered in "r-reticulate")
matplotlib <- import("matplotlib")
```

copy `C:\Users\RW\AppData\Local\r-reticulate\r-reticulate\pyenv\pyenv-win\versions\3.9.12\tcl\tcl8.6` and `C:\Users\RW\AppData\Local\r-reticulate\r-reticulate\pyenv\pyenv-win\versions\3.9.12\tcl\tk8.6` two folders to the folder `C:\Users\RW\AppData\Local\r-reticulate\r-reticulate\pyenv\pyenv-win\versions\3.9.12\Lib`

```
# library(reticulate)
# use_virtualenv("r-reticulate")
# # matplotlib <- import("matplotlib")
# matplotlib$use("Agg", force = TRUE)
```

```
import matplotlib.pyplot as plt
plt.plot([0, 2, 1, 4])
plt.show()
```



12.2 SoloLearn

https://www.sololearn.com/

https://www.sololearn.com/en/learn/courses/python-intermediate

12.3 list comprehension

https://www.sololearn.com/en/learn/courses/python-intermediate/lesson/1188906590?p=1

```
cubes = [i**3 for i in range(5)]

print(cubes)
```

```
## [0, 1, 8, 27, 64]
```

12.4 functional programming

- pure function
- lambda
- map
- filter
- generator
- decorator
- recursion
- *args
- **kwargs

12.5 object-oriented programming = OOP

- class
- inheritance
- magic method
- operator overloading
- data hiding
- static method
- property

Chapter 13

TikZ

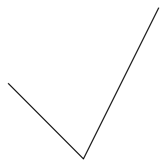
- Ti*k*Z
  - PGFplots[13.4]
  - tikzplotlib[13.5]: Python[12] matplotlib[27] export to TikZ[13] .tex

multi-column 7.3.6

```
knitr::opts_chunk$set(fig.pos = "H", out.extra = "")
```

https://bookdown.org/yihui/rmarkdown-cookbook/html-scroll.html

```
\begin{tikzpicture}
  \draw (-1,1)--(0,0)--(1,2);
\end{tikzpicture}
```



How to speed up bookdown generation?

https://stackoverflow.com/questions/56541371/how-to-speed-up-bookdown-generation

TikZ and PGFplots

What's the relation between packages PGFplots and TikZ?

https://tex.stackexchange.com/questions/285925/whats-the-relation-between-packages-pgfplots-and-tikz

https://www.youtube.com/watch?v=bQugbYq0BVA

https://www.youtube.com/watch?v=ft4Kg9emK1k&list=PLg5nrpKdkk2DWcg3scb75AknF7DJXs8lk&index=18

```
\begin{tikzpicture}
  \def\a{1.5} % amplitude
  \def\b{2}   % frequency
  \draw[->] (-0.2,0)--(4.2,0) node[right,
   ↪  font=\small] {$x$};
  \draw[->] (0,-4)--(0,0.5) node[above] {$y$};

   ↪  \draw[domain=0:4,smooth,variable=\t,blue,thick]
   ↪
    plot ({\a * (\b*\t -
↪  sin(deg(\b*\t)))},{-\a * (1 -
↪  cos(deg(\b*\t)))});
  % \node[above] at (2, 0.5) {Brachistochrone
   ↪  Curve};
  \node[above, font=\footnotesize] at (2, 1)
   ↪  {Brachistochrone Curve};
  \node[above, font=\footnotesize] at (2, 0)
   ↪  {$\begin{aligned}
& x=r(t-\sin t) \\
& y=r(1-\cos t)
\end{aligned}$};
\end{tikzpicture}
```

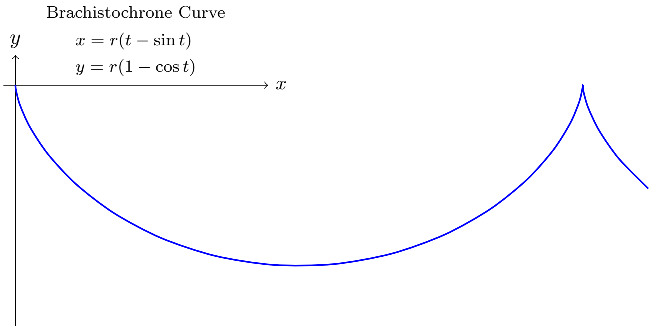


Figure 13.1: Brachistochrone Curve

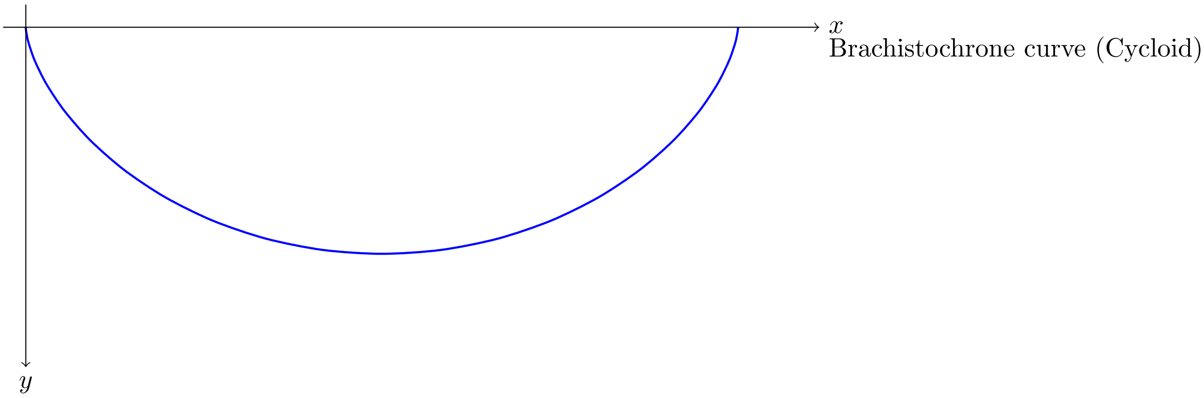


Figure 13.2: Brachistochrone Curve

13.1 2D

https://zhuanlan.zhihu.com/p/127155579?utm_psn=1741479950987960320

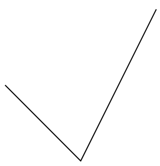
```
\begin{tikzpicture}
  \draw (-1,1)--(0,0)--(1,2);
\end{tikzpicture}
```

```
\begin{tikzpicture}
  \draw (-1,1)--(0,0)--(1,2);
\end{tikzpicture}
```

---

out.width=if (knitr:::is_html_output()) '20%'

```
\begin{tikzpicture}
  \draw (-1,1)--(0,0)--(1,2);
\end{tikzpicture}
```



```
\begin{tikzpicture}
  \draw[rounded corners]
  ↪  (-1,1)--(0,0)--(1,2)--(-1,1);
\end{tikzpicture}
```

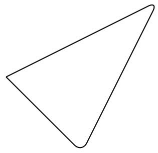


Figure 13.3: rounded corner pseudo-closed triangle

```
\begin{tikzpicture}
  \draw[rounded corners]
  ↪  (-1,1)--(0,0)--(1,2)--cycle;
\end{tikzpicture}
```

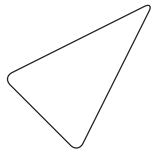


Figure 13.4: rounded corner triangle

```
\begin{tikzpicture}
  \draw[rounded corners]
  ↪  (-1,1)--(0,0)--(1,2)--cycle;
  \draw[rounded corners]
  ↪  (-1,1)--(0,0)--(1,2)--(-1,1);
\end{tikzpicture}
```

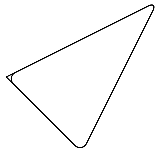


Figure 13.5: triangle vs. pseudo-closed triangle

```
\begin{tikzpicture}
  \draw (0,0) rectangle (4,2);
\end{tikzpicture}
```

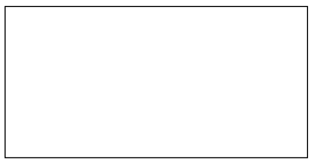


Figure 13.6: rectangle

```
\begin{tikzpicture}
  \draw (0,0) rectangle (2,2);
\end{tikzpicture}
```

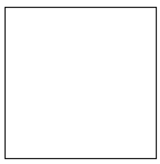


Figure 13.7: square

```
\begin{tikzpicture}
  \draw (0,0) circle (1);
\end{tikzpicture}
```

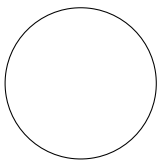


Figure 13.8: circle

```
\begin{tikzpicture}
  \draw (0,0) circle (1);
  \draw (0,0) rectangle (2,2);
\end{tikzpicture}
```

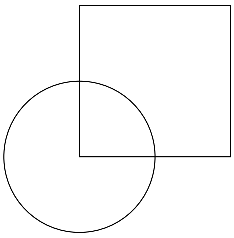


Figure 13.9: circle and square

```
\begin{tikzpicture}
  \draw (1,1) ellipse (2 and 1);
\end{tikzpicture}
```

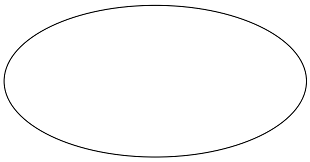


Figure 13.10: ellipse

```
\begin{tikzpicture}
  \draw (1 ,1) arc (0:270:1);
  \draw (6 ,1) arc (0:270:2 and 1);
\end{tikzpicture}
```

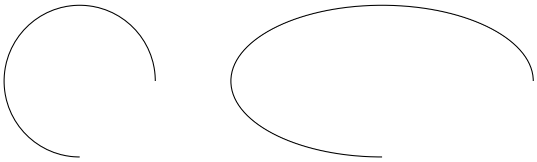


Figure 13.11: circle and ellipse arcs

```
\begin{tikzpicture}
  \draw (-1,1) parabola bend (0,0) (2,4);
\end{tikzpicture}
```

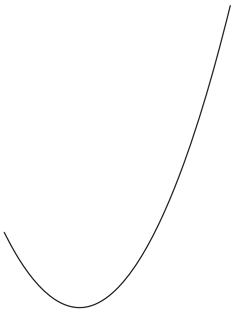


Figure 13.12: parabola arc

```
\begin{tikzpicture}
  \draw (-1,1) parabola bend (0,0) (2,4);
  \filldraw
    (-1,1) circle (.05)
    ( 0,0) circle (.05)
    ( 1,1) circle (.05)
    ( 2,4) circle (.05);
\end{tikzpicture}
```

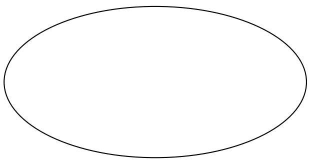


Figure 13.13: parabola arc with points

```
\begin{tikzpicture}
  \draw[step=20pt] (0,0) grid (3,2);
  \draw[help lines ,step=20pt] (4,0) grid
  ↪  (7,2);
\end{tikzpicture}
```

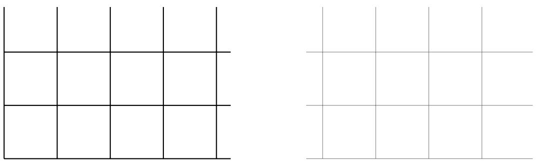


Figure 13.14: grid and help lines

```
\begin{tikzpicture}[scale=0.25]
  \draw[->] (0,0)--(9,0);
  \draw[<-] (0,1)--(9,1);
  \draw[<->] (0,2)--(9,2);
  \draw[>->>] (0,3)--(9,3);
  \draw[|<->|] (0,4)--(9,4);
\end{tikzpicture}
```

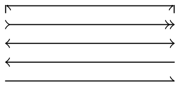


Figure 13.15: arrows

```
\begin{tikzpicture}
  \draw[line width =2pt] (0,6)--(9,6);
  \draw[dotted]          (0,5)--(9,5);
  \draw[densely dotted]  (0,4)--(9,4);
  \draw[loosely dotted]  (0,3)--(9,3);
  \draw[dashed]          (0,2)--(9,2);
  \draw[densely dashed]  (0,1)--(9,1);
  \draw[loosely dashed]  (0,0)--(9,0);
\end{tikzpicture}
```

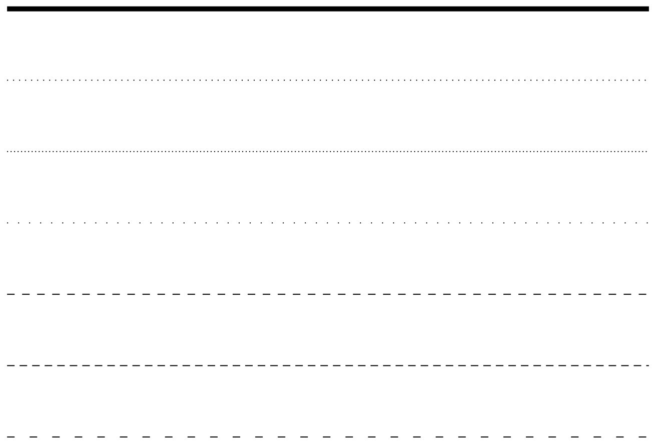


Figure 13.16: lines

```
\begin{tikzpicture}[dline/.style={color= blue, line width=2pt}]
  \draw[dline] (0,0)--(9,0);
\end{tikzpicture}
```


Figure 13.17: head styling

```
\begin{tikzpicture}
  \draw (0,0) rectangle (2,2);
  \draw[shift={( 3, 0)}] (0,0) rectangle (2,2);
  \draw[shift={( 0, 3)}] (0,0) rectangle (2,2);
  \draw[shift={( 0,-3)}] (0,0) rectangle (2,2);
  \draw[shift={(-3, 0)}] (0,0) rectangle (2,2);
  \draw[shift={( 3, 3)}] (0,0) rectangle (2,2);
  \draw[shift={(-3, 3)}] (0,0) rectangle (2,2);
  \draw[shift={( 3,-3)}] (0,0) rectangle (2,2);
  \draw[shift={(-3,-3)}] (0,0) rectangle (2,2);
\end{tikzpicture}
```

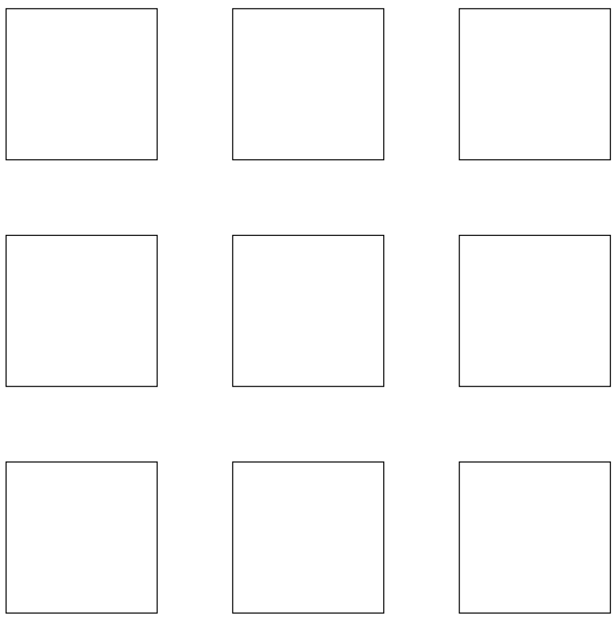


Figure 13.18: transform: shift

```
\begin{tikzpicture}
  \draw (0,0) rectangle (2,2);
  \draw[xshift= 100pt] (0,0) rectangle (2,2);
  \draw[xshift=-100pt] (0,0) rectangle (2,2);
  \draw[yshift= 100pt] (0,0) rectangle (2,2);
  \draw[yshift=-100pt] (0,0) rectangle (2,2);
\end{tikzpicture}
```

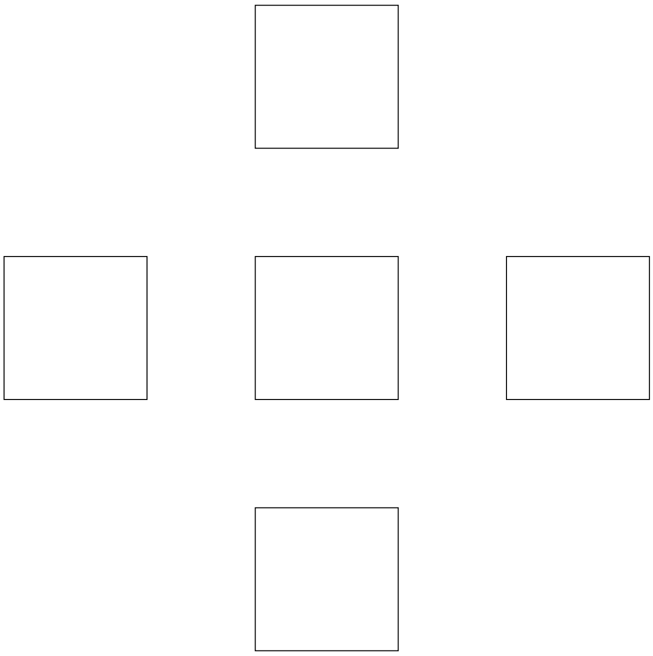


Figure 13.19: transform: shift x, y

```
\begin{tikzpicture}
  \draw (0,0) rectangle (2,2);
  \draw[xshift= 100pt, xscale=1.5] (0,0)
  ↪  rectangle (2,2);
  \draw[yshift= 100pt, xscale=0.5] (0,0)
  ↪  rectangle (2,2);
  \draw[xshift=-100pt, yscale=1.5] (0,0)
  ↪  rectangle (2,2);
  \draw[yshift=-100pt, yscale=0.5] (0,0)
  ↪  rectangle (2,2);
\end{tikzpicture}
```

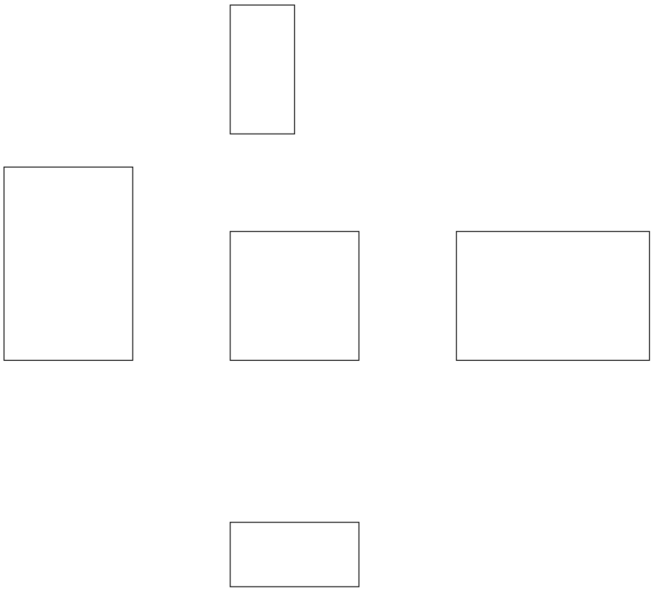


Figure 13.20: transform: scale x, y

```
\begin{tikzpicture}
  \draw (0,0) rectangle (2,2);
  \draw[xshift= 100pt, xscale=1.5] (0,0)
  ↪  rectangle (2,2);
  \draw[yshift= 100pt, yscale=1.5] (0,0)
  ↪  rectangle (2,2);
  \draw[xshift=-100pt, xscale=0.5] (0,0)
  ↪  rectangle (2,2);
  \draw[yshift=-100pt, yscale=0.5] (0,0)
  ↪  rectangle (2,2);
\end{tikzpicture}
```

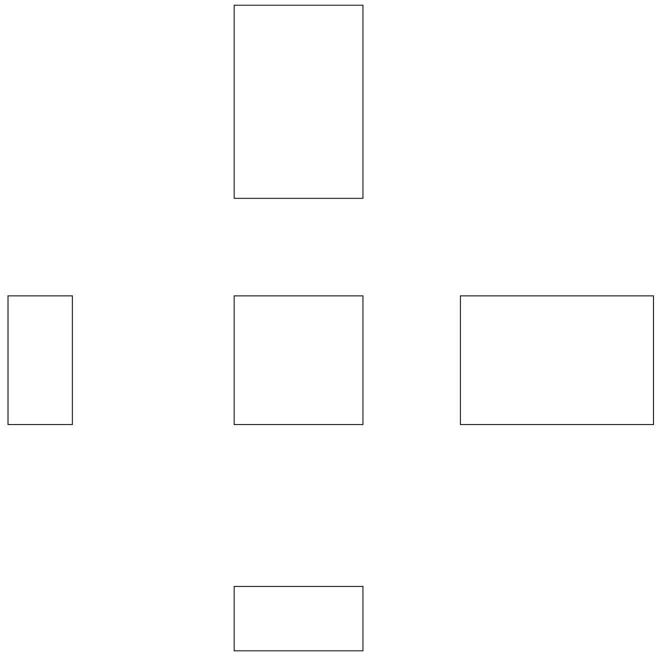


Figure 13.21: transform: scale

```
\begin{tikzpicture}
  \draw (0,0) rectangle (2,2);
  \draw[xshift=125pt,rotate=45] (0,0)
  ↪  rectangle (2,2);
  \draw[xshift=175pt,rotate around={45:(2
  ↪  ,2)}] (0,0) rectangle (2,2);
\end{tikzpicture}
```

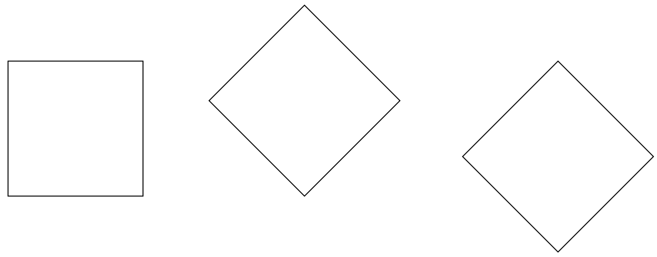


Figure 13.22: transform: rotate

```
\begin{tikzpicture}
  \draw (0,0) rectangle (2,2);
  \draw[xshift=70pt,xslant=1] (0,0) rectangle
  ↪  (2,2);
  \draw[yshift=70pt,yslant=1] (0,0) rectangle
  ↪  (2,2);
\end{tikzpicture}
```

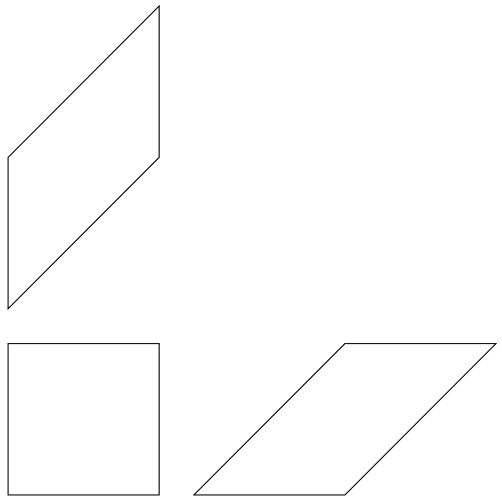


Figure 13.23: transform: slant

```
\tikzset{
  box/.style={
    draw=blue,
    rectangle,
    rounded corners=5pt,
    minimum width=50pt,
    minimum height=20pt,
    inner sep=5pt
  }
}
\begin{tikzpicture}
  \node[box] (1) at (0,0) {1};
  \node[box] (2) at (4,0) {2};
  \node[box] (3) at (8,0) {3};
  \draw[->] (1)--(2);
  \draw[->] (2)--(3);
  \node at (2,1) {a};
  \node at (6,1) {b};
\end{tikzpicture}
```

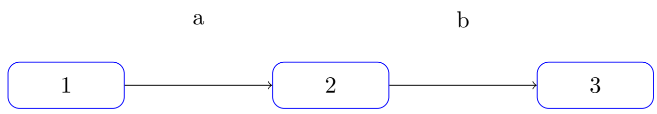


Figure 13.24: flowchart

```
\tikzset{
  box/.style={
    draw=blue,
    fill=blue!20,
    rectangle,
    rounded corners=5pt,
    minimum height=20pt,
    inner sep=5pt
  }
}
\begin{tikzpicture}
  \node[box] {1}
      child {node[box] {2}}
      child {node[box] {3}
          child {node[box] {4}}
          child {node[box] {5}}
          child {node[box] {6}}
      };
\end{tikzpicture}
```

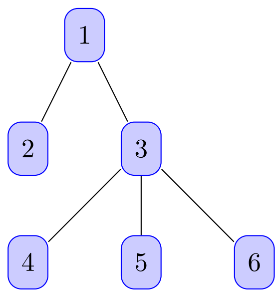


Figure 13.25: tree

```
\begin{tikzpicture}
  \draw[->] (-0.2,0)--(6,0) node[right] {$x$};
  \draw[->] (0,-0.2)--(0,6) node[above]
  ↪  {$f(x)$};
  \draw[domain=0:4] plot (\x ,{0.1* exp(\x)})
  ↪  node[right] {$f(x)=\frac{1}{10}e^x$};
\end{tikzpicture}
```

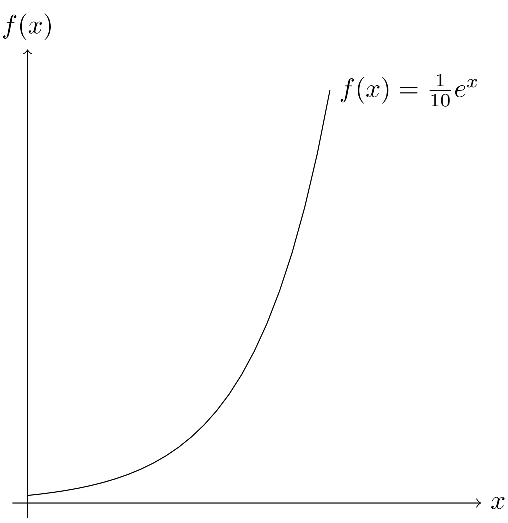


Figure 13.26: function plot

https://stackoverflow.com/questions/64897575/tikz-libraries-in-bookdown

It turns out that you can simply put the `\usetikzlibrary{...}` command directly before the `\begin{tikzpicture}` and everything works fine :)

https://stackoverflow.com/questions/56211210/r-markdown-document-with-html-docx-output-using-latex-package-bbm

https://tex.stackexchange.com/questions/171711/how-to-include-latex-package-in-r-markdown

13.2 3D

https://zhuanlan.zhihu.com/p/431732330?utm_psn=1741857547550638080

https://github.com/RRWWW/Stereometry

```
\begin{tikzpicture}
  \coordinate (A) at ( 1, 1, 1);
  \coordinate (B) at ( 1, 1,-1);
  \coordinate (C) at ( 1,-1,-1);
  \coordinate (D) at ( 1,-1, 1);
  \coordinate (E) at (-1,-1, 1);
  \coordinate (F) at (-1,-1,-1);
  \coordinate (G) at (-1, 1,-1);
  \coordinate (H) at (-1, 1, 1);
  \draw (A) node[right=1pt] {$A$}--
        (B) node[right=1pt] {$B$}--
        (C) node[right=1pt] {$C$}--
        (D) node[right=1pt] {$D$}--
        (E) node[left= 1pt] {$E$}--
        (F) node[right=1pt] {$F$}--
        (G) node[right=1pt] {$G$}--
        (H) node[left= 1pt] {$H$}--
        (A) node[right=1pt] {$A$};
\end{tikzpicture}
```

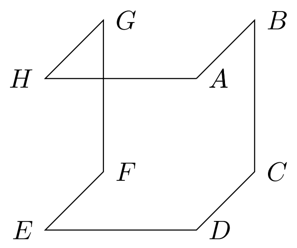


Figure 13.27: cube

```
\usetikzlibrary{patterns}
\usetikzlibrary{3d,calc}
\tdplotsetmaincoords{45}{45}
\begin{tikzpicture}[tdplot_main_coords]
  \coordinate (A) at ( 1, 1, 1);
  \coordinate (B) at ( 1, 1,-1);
  \coordinate (C) at ( 1,-1,-1);
  \coordinate (D) at ( 1,-1, 1);
  \coordinate (E) at (-1,-1, 1);
  \coordinate (F) at (-1,-1,-1);
  \coordinate (G) at (-1, 1,-1);
  \coordinate (H) at (-1, 1, 1);
  \draw (A) node[right=1pt] {$A$}--
        (B) node[right=1pt] {$B$}--
        (C) node[right=1pt] {$C$}--
        (D) node[right=1pt] {$D$}--
        (E) node[left= 1pt] {$E$}--
        (F) node[right=1pt] {$F$}--
        (G) node[right=1pt] {$G$}--
        (H) node[left= 1pt] {$H$}--
        (A) node[right=1pt] {$A$};
\end{tikzpicture}
```

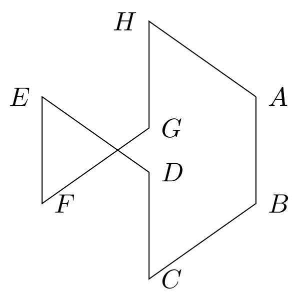


Figure 13.28: cube rotate

```
\usetikzlibrary{patterns}
\usetikzlibrary{3d,calc}
%\tdplotsetmaincoords{70}{110}
\begin{tikzpicture}[rotate around y=-15,
↪  rotate around z=7]
  \coordinate (A) at ( 1, 1, 1);
  \coordinate (B) at ( 1, 1,-1);
  \coordinate (C) at ( 1,-1,-1);
  \coordinate (D) at ( 1,-1, 1);
  \coordinate (E) at (-1,-1, 1);
  \coordinate (F) at (-1,-1,-1);
  \coordinate (G) at (-1, 1,-1);
  \coordinate (H) at (-1, 1, 1);
  \draw (A) node[right=1pt] {$A$}--
        (B) node[right=1pt] {$B$}--
        (C) node[right=1pt] {$C$}--
        (D) node[right=1pt] {$D$}--
        (E) node[left= 1pt] {$E$}--
        (F) node[right=1pt] {$F$}--
        (G) node[right=1pt] {$G$}--
        (H) node[left= 1pt] {$H$}--
        (A) node[right=1pt] {$A$};
\end{tikzpicture}
```

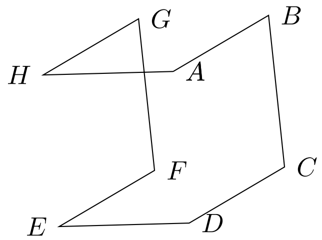


Figure 13.29: cube rotate around

https://tex.stackexchange.com/questions/388621/optimizing-perspective-tikz-graphic

```
\usetikzlibrary{patterns}
\usetikzlibrary{3d,calc}
\begin{tikzpicture}[y={(.5cm,.7cm)}]
  \coordinate (A) at ( 1, 1, 1);
  \coordinate (B) at ( 1, 1,-1);
  \coordinate (C) at ( 1,-1,-1);
  \coordinate (D) at ( 1,-1, 1);
  \coordinate (E) at (-1,-1, 1);
  \coordinate (F) at (-1,-1,-1);
  \coordinate (G) at (-1, 1,-1);
  \coordinate (H) at (-1, 1, 1);
  \draw (A) node[right=1pt] {$A$}--
        (B) node[right=1pt] {$B$}--
        (C) node[right=1pt] {$C$}--
        (D) node[right=1pt] {$D$}--
        (E) node[left= 1pt] {$E$}--
        (F) node[right=1pt] {$F$}--
        (G) node[right=1pt] {$G$}--
        (H) node[left= 1pt] {$H$}--
        (A) node[right=1pt] {$A$};
\end{tikzpicture}
```

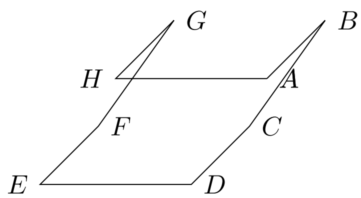


Figure 13.30: cube perspective slant

https://github.com/XiangyunHuang/bookdown-broken/blob/master/index.Rmd

```
\smartdiagramset{planet color=gray!40!white,
uniform color list=gray!40!white for 10 items}
\smartdiagram[bubble diagram]{Basic skills,
  Edit~/\\ (RStudio),
  Organize~/\\ (bookdown),
  Cooperate~/\\ (Git),
  Typeset~/\\ (LaTeX/Pandoc),
  Compile~/\\ (GitHub Action)}
```

Figure 13.31: modern statistics plot skills

13.3 plots of functions

https://tikz.dev/tikz-plots

> A warning before we get started: If you are looking for an easy way to create a normal plot of a function with scientific axes, ignore this section and instead look at the `pgfplots` package or at the `datavisualization` command from Part VI.

https://tikz.dev/tikz-plots#sec-22.5

```
\begin{tikzpicture}[domain=0:4]
  \draw[very thin,color=gray] (-0.1,-1.1) grid
  ↪  (3.9,3.9);

  \draw[->] (-0.2,0) -- (4.2,0) node[right]
  ↪  {$x$};
  \draw[->] (0,-1.2) -- (0,4.2) node[above]
  ↪  {$f(x)$};

  \draw[color=red]    plot (\x,\x)
  ↪  node[right] {$f(x) =x$};
  % \x r means to convert '\x' from degrees to
  ↪  _r_adians:
  \draw[color=blue]   plot (\x,{sin(\x r)})
  ↪  node[right] {$f(x) = \sin x$};
  \draw[color=orange] plot (\x,{0.05*exp(\x)})
  ↪  node[right] {$f(x) = \frac{1}{20}
  ↪  \mathrm e^x$};
\end{tikzpicture}
```

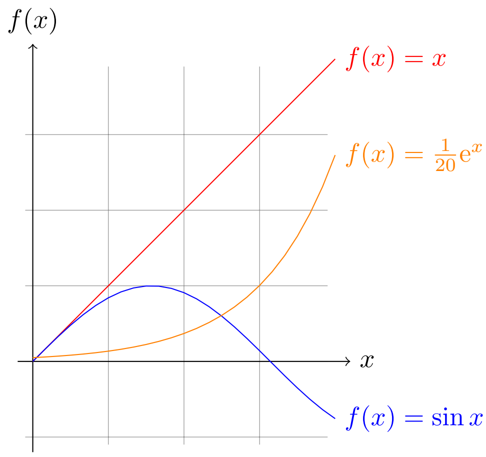


Figure 13.32: plots of functions

```
\tikz
↪  \draw[scale=0.5,domain=-3.141:3.141,smooth,variable=\t]
  plot ({\t*sin(\t r)},{\t*cos(\t r)});
```

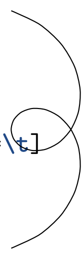


Figure 13.33: 2D parametric function

```
\tikz \draw[domain=0:360,
            smooth,
            variable=\t]
  plot ({sin(\t)},\t/360,{cos(\t)});
```

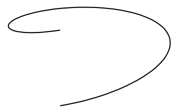


Figure 13.34: 3D parametric function

13.4 PGFplots

axis similar to matplotlib figure anatomy, see Fig: 27.1

https://tikz.dev/pgfplots/

https://tikz.dev/pgfplots/tutorial1

> Not so common is `\pgfplotsset{compat=1.5}` . A statement like this should always be used in order to (a) benefit from a more or less recent feature set and (b) avoid changes to your picture if you recompile it with a later version of pgfplots.

```
\pgfplotsset{width=7cm,compat=1.18}
\begin{tikzpicture}
\begin{axis}[
]
    % density of Normal distribution:
    \addplot [
        red,
        domain=-3e-3:3e-3,
        samples=201,
    ]
        {exp(-x^2 / (2e-3^2)) / (1e-3 *
↪  sqrt(2*pi))};
\end{axis}
\end{tikzpicture}
```

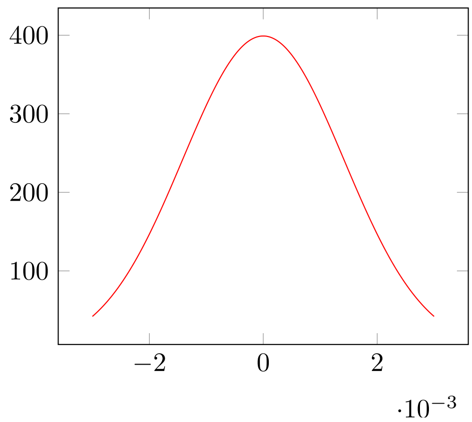


Figure 13.35: PGFplots: normal distribution

13.4.1 the axis environments

https://tikz.dev/pgfplots/reference-axis

`\pgfplotsset{every linear axis/.append style={...}}`

```
\begin{tikzpicture}
    \begin{axis}[
        no markers,
        axis x line = center,
        axis y line = center,
        xlabel = {$ x $}, xlabel style =
↪   {right},
        ylabel = {$ y $}, ylabel style =
↪   {above},
        xmin = -8, xmax = 8,
        ymin = 0, ymax = 0.45,
        hide obscured x ticks=false, % for
↪   origin x tick label i.e. xtick = {0}
        xtick={-4, 0,  4},
        xticklabels={$
↪   \mu_{\scriptscriptstyle{1}} $,
                        $
                        ↪   \mu_{\scriptscriptstyle{0}}
                        ↪   $,
                        $
                        ↪   \mu_{\scriptscriptstyle{1}}
                        ↪   $},
        %extra x ticks={0},
        ytick = \empty,
        x = 1cm, y = 5cm, % x y scaling
        %grid = major,
        domain = -10:10,
        samples = 1000
    ]
  \end{axis}
\end{tikzpicture}
```

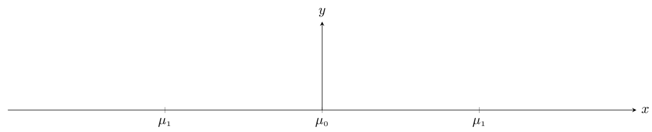


Figure 13.36: beginaxis

https://tex.stackexchange.com/questions/134959/axis-lines-middle-and-axis-lines-center

No, there is no difference.

13.4.2 axis descriptions

https://tikz.dev/pgfplots/reference-axisdescription

13.4.2.1 placement of axis descriptions

13.4.2.1.1 coordinate system `axis description cs` https://tikz.dev/pgfplots/reference-axisdescription#pgfp.axis_description_cs

`\addplot {x};` can change to `\addplot {x^2};` still with auto blue dots

`small dot style,pin=angle:LaTeX label   at PGFplots axis coordinate system   ;`

```
\begin{tikzpicture}
    \tikzset{
        every
↪   pin/.style={fill=yellow!50!white,rectangle,rounded
↪   corners=3pt,font=\tiny},
        small
↪   dot/.style={fill=black,circle,scale=0.3},
    }
    \begin{axis}[
        clip=false,
        title=How \texttt{axis description cs}
↪   works,
    ]
        \addplot {x};
        %small dot style,pin=angle:LaTeX label
        ↪   at PGFplots axis coordinate system
        ↪   ;
        \node [small dot,pin=120:{$(0,0)$}]
        ↪   at (axis description cs:0,0)
        ↪   {};
        \node [small dot,pin=-30:{$(1,1)$}]
        ↪   at (axis description cs:1,1)
        ↪   {};
        \node [small
        ↪   dot,pin=-90:{$(1.03,0.5)$}] at
        ↪   (axis description cs:1.03,0.5) {};
        \node [small
        ↪   dot,pin=125:{$(0.5,0.5)$}]  at
        ↪   (axis description cs:0.5,0.5)  {};
    \end{axis}
\end{tikzpicture}
```

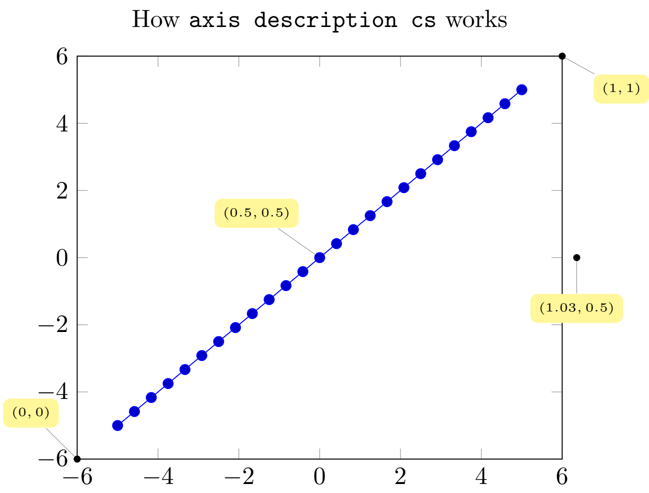


Figure 13.37: PGFplots: x

```
\begin{tikzpicture}
    \tikzset{
        every
↪   pin/.style={fill=yellow!50!white,rectangle,rounded
↪   corners=3pt,font=\tiny},
        small
↪   dot/.style={fill=black,circle,scale=0.3},
    }
    \begin{axis}[
        clip=false,
        title=How \texttt{axis description cs}
↪   works,
    ]
        \addplot {x^2};
        %small dot style,pin=angle:LaTeX label
        ↪   at PGFplots axis coordinate system
        ↪   ;
        \node [small dot,pin=120:{$(0,0)$}]
        ↪   at (axis description cs:0,0)
        ↪   {};
        \node [small dot,pin=-30:{$(1,1)$}]
        ↪   at (axis description cs:1,1)
        ↪   {};
        \node [small
        ↪   dot,pin=-90:{$(1.03,0.5)$}] at
        ↪   (axis description cs:1.03,0.5) {};
        \node [small
        ↪   dot,pin=125:{$(0.5,0.5)$}]  at
        ↪   (axis description cs:0.5,0.5)  {};
    \end{axis}
\end{tikzpicture}
```

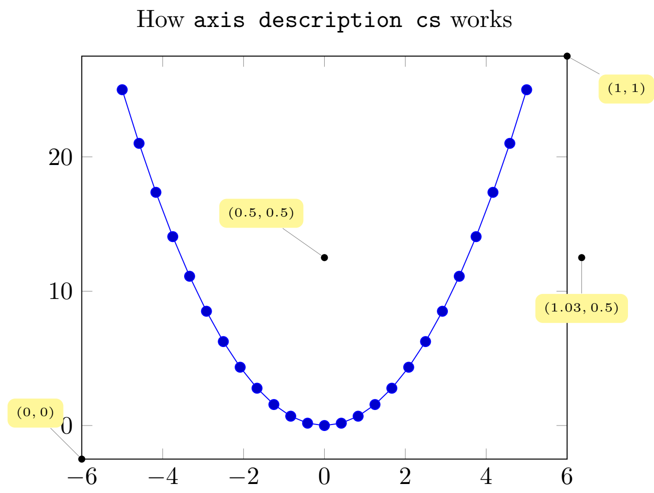


Figure 13.38: PGFplots: x^2

13.4.2.1.2 legend https://tikz.dev/pgfplots/reference-axisdescription#sec-4.9.4

13.4.2.1.3 tick option https://tikz.dev/pgfplots/reference-tickoptions

13.4.3 declare function

https://tikz.dev/pgfplots/utility-commands#pgf/declare_function

```
\begin{tikzpicture}
\begin{axis}[
    declare function={
        C=4;
        square(\t)=(\t)^2 + C;
        },
    ]
    \addplot+ [samples=2] {C*x};
    \addplot {square(x)};
\end{axis}
\end{tikzpicture}
```

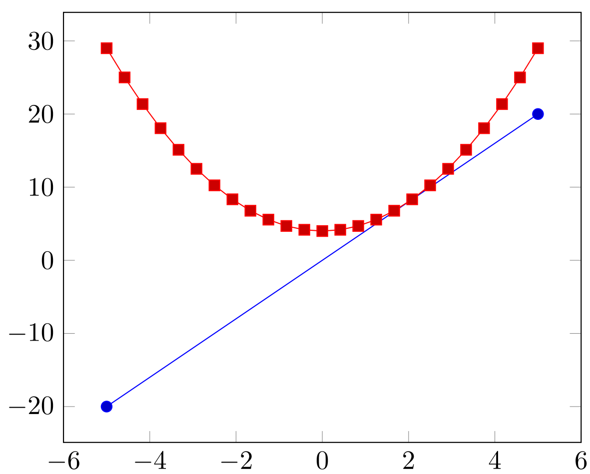


Figure 13.39: declare function

13.4.3.1 pgfmathparse

https://tikz.dev/math-parsing

https://tikz.dev/math-parsing#sec-94.1

This macro parses and returns the result without units in the macro 0.017. Example: `\pgfmathparse{2pt+3.5pt}` will set `\pgfmathresult` to the text `5.5`.

`\pgfmathsqrt{x}` = `\pgfmathparse{sqrt(x)}`

`\pgfmathln{x}` = `\pgfmathparse{ln(x)}`

...

13.4.3.2 pgfmathdeclarefunction

like `pgfplotsinvokeforeach`

replaces any occurrence of `#1` inside of (math image)command(math image) once for every element in (math image)list(math image). Thus, it actually assumes that (math image)command(math image) is like a `\newcommand` body.

```
% pgfmathdeclarefunction{name}{num_var}{%
%% #1 = \mu, #2 = \sigma
\pgfmathdeclarefunction{gauss}{2}{%
    \pgfmathparse{1/(#2*sqrt(2*pi))*exp(-((x-#1)^2)/(2*#2^2))}%
}
```

```
% pgfmathdeclarefunction{name}{num_var}{%
%% #1 = \mu, #2 = \sigma
\pgfmathdeclarefunction{gauss}{2}{%
    \pgfmathparse{1/(#2*sqrt(2*pi))*exp(-((x-#1)^2)/(2*#2^2))}%
}

\begin{tikzpicture}
    \begin{axis}[
        no markers,
        axis x line = center,
        axis y line = center,
        xlabel = {$ x $}, xlabel style = {right},
        ylabel = {$ y $}, ylabel style = {above},
        xmin = -8, xmax = 8,
        ymin = 0, ymax = 0.45,
        hide obscured x ticks=false, % for origin x tick label i.e. xtick = {0}
        xtick={-4, 0,  4},
        xticklabels={$ \mu_{\scriptscriptstyle{1}} $,
                     $ \mu_{\scriptscriptstyle{0}} $,
                     $ \mu_{\scriptscriptstyle{1}} $},
        %extra x ticks={0},
        ytick = \empty,
        x = 1cm, y = 5cm, % x y scaling
        %grid = major,
        domain = -10:10,
        samples = 1000
    ]
        \addplot [fill=cyan!20, draw=none, domain=-10:-2] {gauss(-4, 1)} \closedcycle;
        \addplot [fill=cyan!20, draw=none, domain=   2:10] {gauss( 4, 1)} \closedcycle;
        \addplot [very thick, cyan!50!black] {gauss(-4, 1)};
        \addplot [very thick, cyan!50!black] {gauss( 0, 1)};
        \addplot [very thick, cyan!50!black] {gauss( 4, 1)};
        %\node [anchor=north] at (axis cs: 0, -0.01) {$ \mu $};
        %\node at (axis cs: -4, -0.02) {$ \mu $};
        \draw [dashed, thin] (axis cs: -4, 0) -- (axis cs: -4, 1);
        \draw [dashed, thin] (axis cs:  4, 0) -- (axis cs: 4, 1);
    \end{axis}
\end{tikzpicture}
```

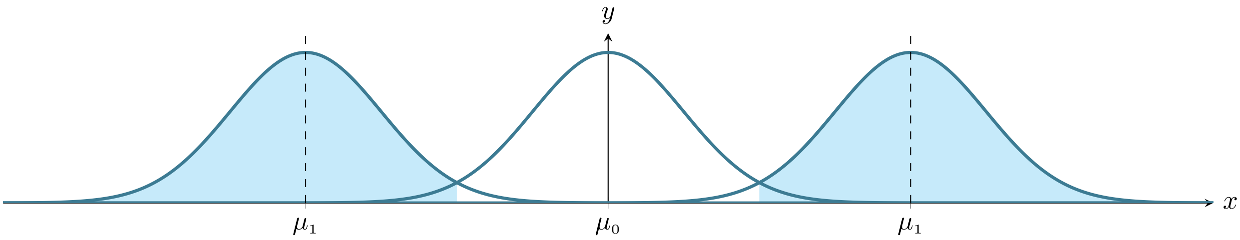


Figure 13.40: PGFmathDeclareFunction: normal distributions hypothesis testing

https://tex.stackexchange.com/questions/43610/plotting-bell-shaped-curve-in-tikz-pgf

```
% pgfmathdeclarefunction{name}{num_var}{%
%% #1 = \mu, #2 = \sigma
\pgfmathdeclarefunction{gauss}{2}{%
  \pgfmathparse{1/(#2*sqrt(2*pi))*exp(-((x-#1)^2)/(2*#2^2))}%
}

\begin{tikzpicture}
\begin{axis}[
  no markers, domain=0:10, samples=100,
  axis lines*=left, xlabel=$x$, ylabel=$y$,
  every axis y label/.style={at=(current axis.above origin),anchor=south},
  every axis x label/.style={at=(current axis.right of origin),anchor=west},
  height=5cm, width=12cm,
  xtick={4,6.5}, ytick=\empty,
  enlargelimits=false, clip=false, axis on top,
  grid = major
  ]
  \addplot [fill=cyan!20, draw=none, domain=0:5.96] {gauss(6.5,1)} \closedcycle;
  \addplot [very thick,cyan!50!black] {gauss(4,1)};
  \addplot [very thick,cyan!50!black] {gauss(6.5,1)};
  \draw [yshift=-0.6cm, latex-latex](axis cs:4,0) -- node [fill=white] {$1.96\sigma$} (axis
  ↪  cs:5.96,0);
\end{axis}

\end{tikzpicture}
```

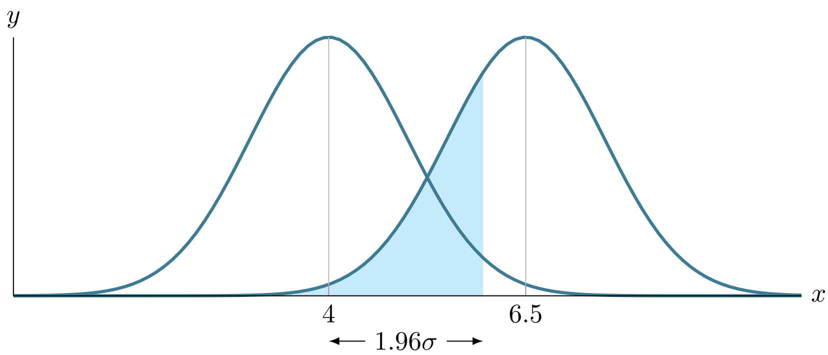


Figure 13.41: PGFmathDeclareFunction: normal distributions

13.4.4 `|-` and `-|` in TikZ

https://tex.stackexchange.com/questions/401425/tikz-what-exactly-does-the-the-notation-for-arrows-do

> `(a -| b)` where `a` and `b` are named nodes or coordinates. This means the coordinate that is at the y-coordinate of `a`, and x-coordinate of `b`. Similarly, `(a |- b)` has the x-coordinate of `a` and y-coordinate of `b`.

13.4.5 `pgfplotsinvokeforeach`

https://tikz.dev/pgfplots/pgfplotstable-miscellaneous#/pgfplotsinvokeforeach like `\foreach` in TikZ

> A variant of `\pgfplotsforeachungrouped` (and such also of `\foreach`) which replaces any occurrence of `#1` inside of (math image)command(math image) once for every element in (math image)list(math image). Thus, it actually assumes that (math image)command(math image) is like a `\newcommand` body.

13.4.6 interpolation dashed lines

https://tex.stackexchange.com/questions/193259/what-is-the-easiest-way-to-accomplish-textual-tick-labels-in-tikz

`interpa = (10,10), interpb = (30,30), interp = interpolation`

`({axis cs:0,0}|-interp#1) = (x of (0,0), y of (interpa)) = (0, 10), ...`

```
\begin{tikzpicture}
\begin{axis}[
    axis lines=left,
    xmin = 0, xmax = 40,
    ymin = 0, ymax = 40,
    xtick={10,30},
    xticklabels={$V_i=10$,$V_f=30$},
    ytick={10,30},
    yticklabels={$P_i=10$,$P_f=30$},
    xlabel={Volume},
    ylabel={Pressure}
  ]
  \addplot[very thick,-latex ]
 ↪  coordinates{(10,10) (30,30)}
    % interpa = (10,10), interpb = (30,30),
    ↪  interp = interpolation
    coordinate[at start](interpa)coordinate[at
↪   end](interpb);
  \pgfplotsinvokeforeach {a,b} {
    \draw[ultra thin, dashed]
      % ({axis cs:0,0}|-interp#1) = (x of
      ↪  (0,0), y of (interpa)) = (0, 10),
      ↪  ...
      ({axis
↪   cs:0,0}|-interp#1)--(interp#1)--(interp#1|-{axis
↪   cs:0,0});
    }
\end{axis}
\end{tikzpicture}
```

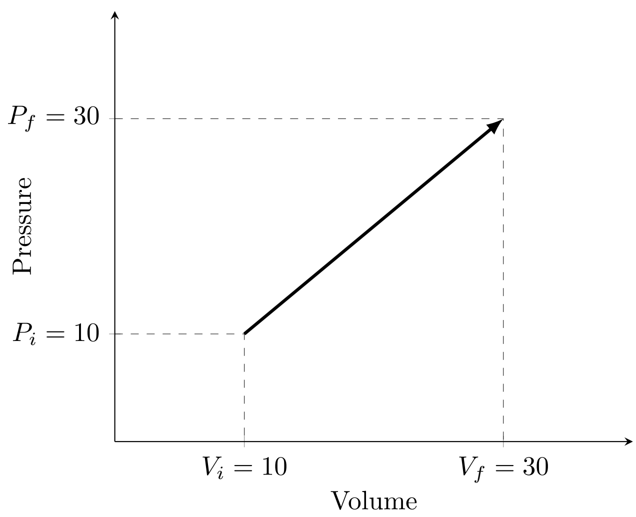


Figure 13.42: tick texts and interpolation dashed lines

13.4.7 Zewbie

https://zhuanlan.zhihu.com/p/551874337

axis similar to matplotlib figure anatomy, see Fig: 27.1

13.4.7.1 coordinate axis/axes fine-tuing

```
\begin{tikzpicture}
    \begin{axis}
        % empty
    \end{axis}
\end{tikzpicture}
```

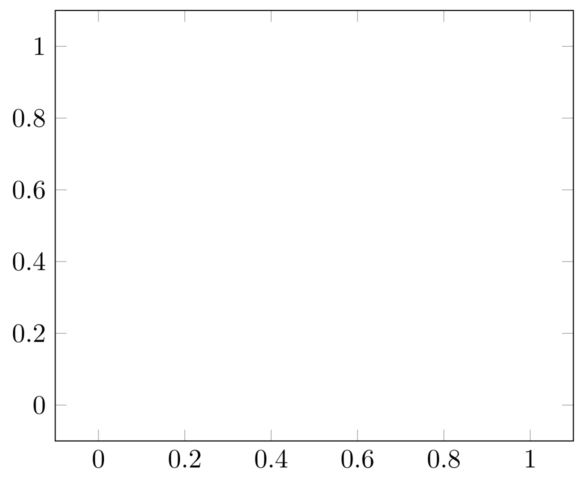


Figure 13.43: PGFplots: 2D axis/axes

13.4.7.1.1 range

```
\begin{tikzpicture}
    \begin{axis}[
        xmin = -1, xmax = 11,
        ymin = -3, ymax = 1,
    ]
        % empty
    \end{axis}
\end{tikzpicture}
```

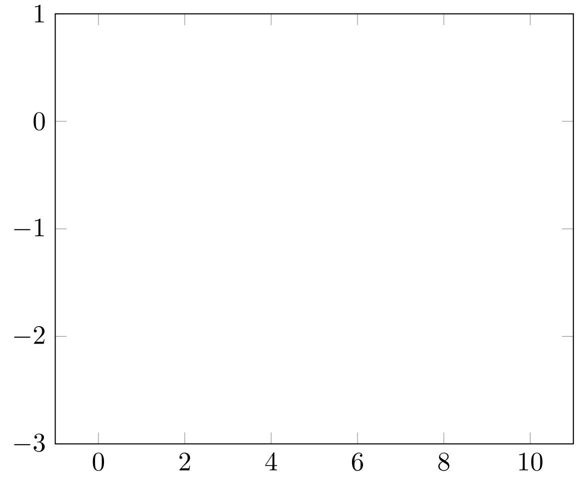


Figure 13.44: PGFplots: axis/axes range

13.4.7.1.2 scaling `axis equal image` equivalent to `unit vector ratio = {1 1 1}`

```
\begin{tikzpicture}
    \begin{axis}[
        xmin = -1, xmax = 11,
        ymin = -3, ymax = 1,
        axis equal image, % unit vector ratio
↪   = {1 1 1},
    ]
        % empty
    \end{axis}
\end{tikzpicture}
```

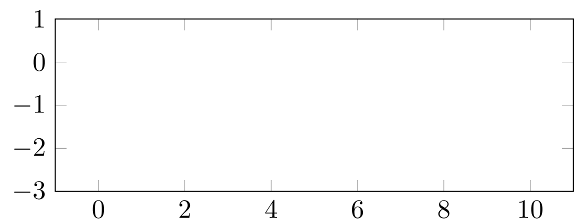


Figure 13.45: PGFplots: axis/axes range

`scale only axis`

`width` x axis length, `height` y axis length

```
\begin{tikzpicture}
    \begin{axis}[
        xmin = -1, xmax = 11,
        ymin = -3, ymax = 1,
        scale only axis,
        width = 5cm, height = 7cm,
    ]
        % empty
    \end{axis}
\end{tikzpicture}
```

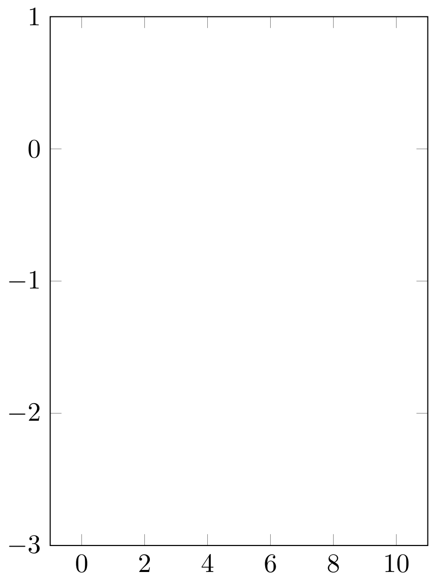


Figure 13.46: PGFplots: axis/axes range

---

x x unit vector length, y y unit vector length

```
\begin{tikzpicture}
    \begin{axis}[
        xmin = -1, xmax = 11,
        ymin = -3, ymax = 1,
        x = 1cm, y = 1cm,
    ]
        % empty
    \end{axis}
\end{tikzpicture}
```

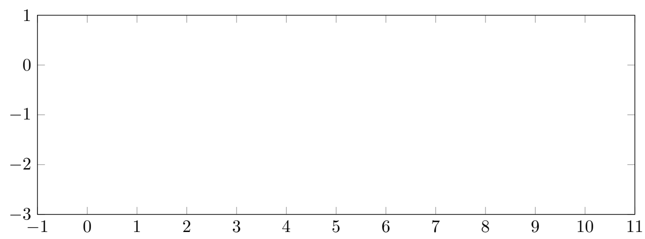


Figure 13.47: PGFplots: axis/axes range

13.4.7.1.3 direction vector

```
\begin{tikzpicture}
    \begin{axis}[
        xmin = -1, xmax = 11,
        ymin = -3, ymax = 1,
        x={(.2cm,-.1cm)}, y={(.5cm,.5cm)},
    ]
        % empty
    \end{axis}
\end{tikzpicture}
```

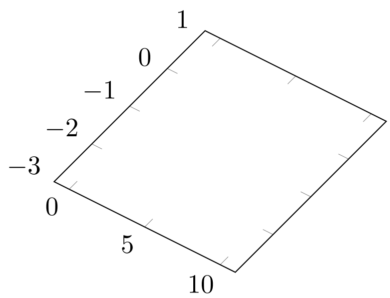


Figure 13.48: PGFplots: axis/axes range

---

unit vector ratio = {1 1}

```
\begin{tikzpicture}
    \begin{axis}[
        xmin = -1, xmax = 11,
        ymin = -3, ymax = 1,
        unit vector ratio = {1 1},
    ]
        % empty
    \end{axis}
\end{tikzpicture}
```

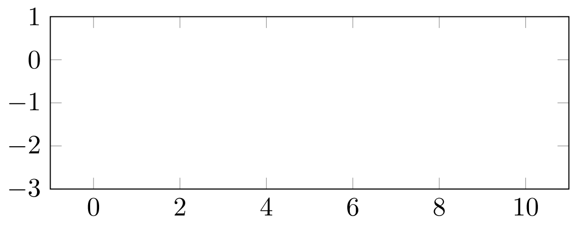


Figure 13.49: PGFplots: axis/axes range

13.4.7.1.4 axis style `axis lines` to assign all axes, either `axis lines = box`(default), `axis lines = center`(axis lines with arrows, `center`, x: `bottom`, `top`, y:), or `axis lines = none`(not shown), or even axis lines wihtout arrows `axis lines *= center`

`axis x line`, `axis y line` to assign the respect axis, e.g. `axis x line = center`

`axis lines = center`:

```
\begin{tikzpicture}
    \begin{axis}[
        axis lines = center,
        xmin = -1, xmax = 11,
        ymin = -3, ymax = 1,
    ]
        % empty
    \end{axis}
\end{tikzpicture}
```

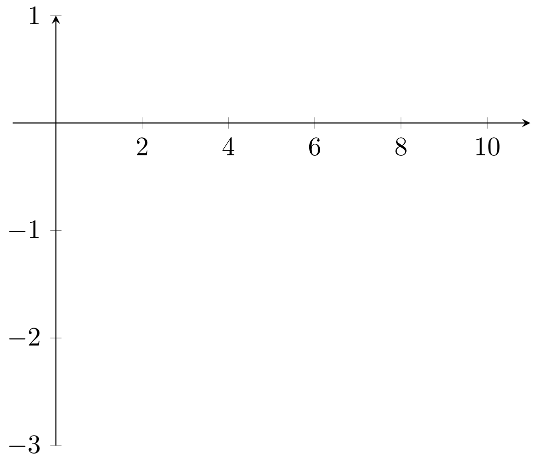


Figure 13.50: PGFplots: axis/axes range

`axis lines *= center`:

x axis line without arrow, y axis box

```
\begin{tikzpicture}
    \begin{axis}[
        axis x line* = center, % x axis line without arrow, y axis box
        xmin = -1, xmax = 11,
        ymin = -3, ymax = 1,
    ]
        % empty
    \end{axis}
\end{tikzpicture}
```

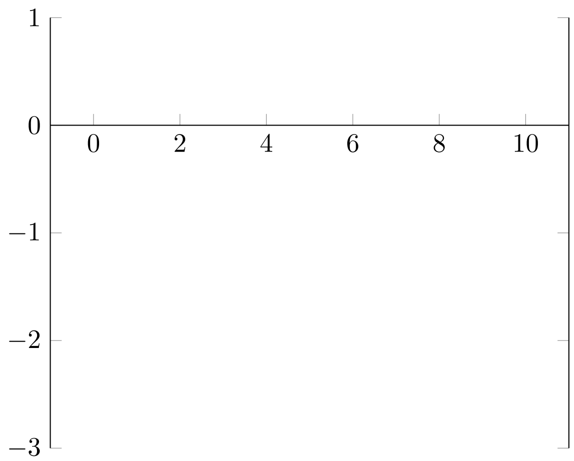


Figure 13.51: PGFplots: axis/axes range

---

x axis line with arrow, y axis line without arrow

```
\begin{tikzpicture}
    \begin{axis}[
        axis x line = center, % x axis line
↪   with arrow
        axis y line* = center, % y axis line
↪   without arrow
        xmin = -1, xmax = 11,
        ymin = -3, ymax = 1,
    ]
        % empty
    \end{axis}
\end{tikzpicture}
```

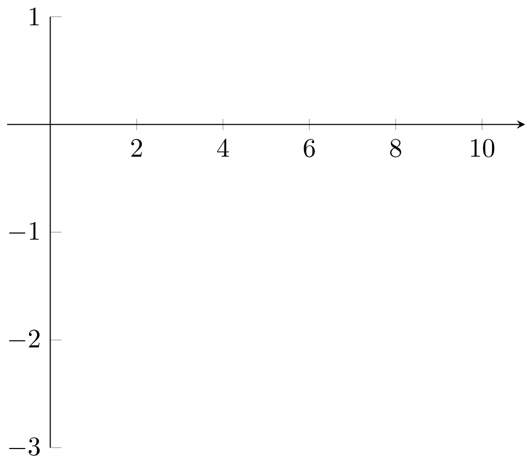


Figure 13.52: PGFplots: axis/axes range

13.4.7.1.5 axis discontinuity `crunch`, `parallel`, `none`

`crunch`

```
\begin{tikzpicture}
    \begin{axis}[
        axis x line = bottom,
        axis y line = center,
        xmin = -2, xmax = 10,
        ymin =  0, ymax = 12,
        axis y discontinuity = crunch,
    ]
        % empty
    \end{axis}
\end{tikzpicture}
```

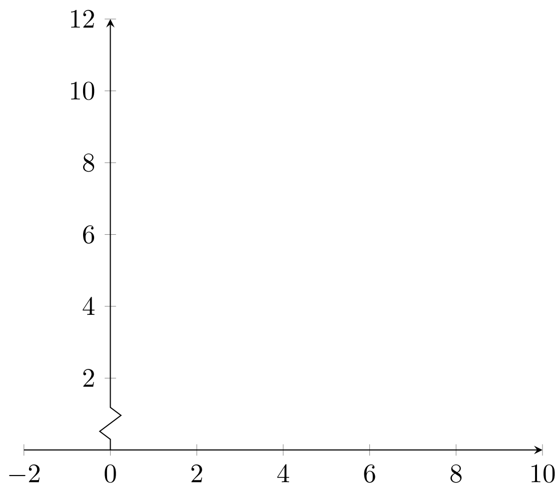


Figure 13.53: PGFplots: axis/axes range

`parallel`

```
\begin{tikzpicture}
    \begin{axis}[
        axis x line = bottom,
        axis y line = center,
        xmin = -2, xmax = 10,
        ymin =  0, ymax = 12,
        axis y discontinuity = parallel,
    ]
        % empty
    \end{axis}
\end{tikzpicture}
```

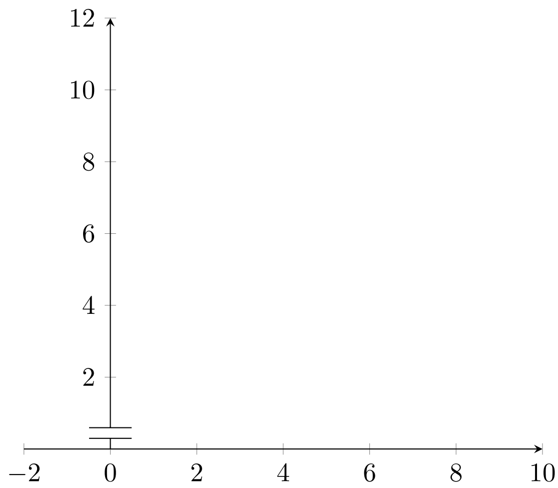


Figure 13.54: PGFplots: axis/axes range

13.4.7.1.6 tick `tick pos` `ticklabel pos`

```
\begin{tikzpicture}
    \begin{axis}[
        xtick pos = upper,
        yticklabel pos = upper,
    ]
        % empty
    \end{axis}
\end{tikzpicture}
```

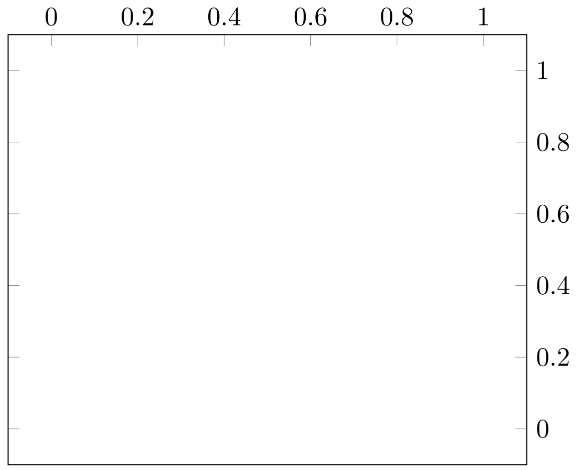


Figure 13.55: PGFplots: axis/axes range

`tick distance`

`tick align`: `inside`, `center`, `outside`

```
\begin{tikzpicture}
    \begin{axis}[
        axis lines=center,
        xmin=-1,xmax=3,
        ymin=-3,ymax=3,
        xtick distance=.8,
        ytick distance=1.1,
        tick align=inside,
    ]
        % empty
    \end{axis}
\end{tikzpicture}
```

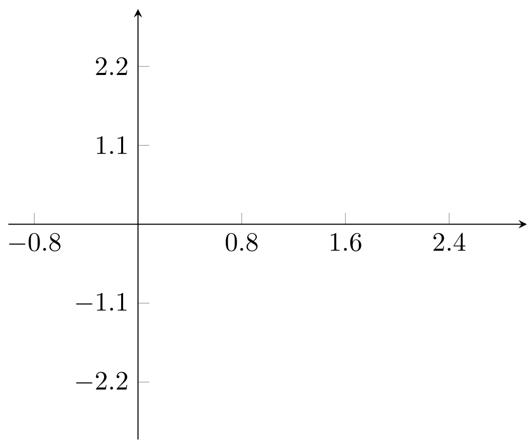


Figure 13.56: PGFplots: axis/axes range

`minor tick num`

```
\begin{tikzpicture}
    \begin{axis}[
        axis y line=none,axis x line=center,
        ymin=0,ymax=0,xmin=-3,xmax=3,
        xtick distance=2,tick align=inside,
        minor tick num=2,
    ]
        % empty
    \end{axis}
\end{tikzpicture}
```

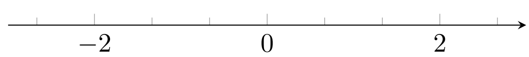


Figure 13.57: PGFplots: axis/axes range

`xtick=\empty|data|{<coordinates>}`

```
\begin{tikzpicture}
    \begin{axis}[
        axis y line=none,axis x line=center,
        ymin=0,ymax=0,xmin=-3,xmax=3,
        xtick={-2.5,0,1},minor
↪   xtick={1/3,2/3},
        tick align=inside,
        ]
        % empty
    \end{axis}
\end{tikzpicture}
```

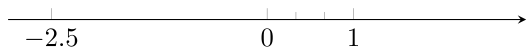


Figure 13.58: PGFplots: axis/axes range

`extra x ticks={<coordinates>}`

```
\begin{tikzpicture}
    \begin{axis}[
        axis y line=none,axis x line=center,
        ymin=0,ymax=0,xmin=-2.3,xmax=4.9,
        xtick distance=2,minor tick num=1,
        extra x ticks={e,pi},
        extra x tick labels={$e$,$\pi$},
        tick align=inside,
    ]
        % empty
    \end{axis}
\end{tikzpicture}
```

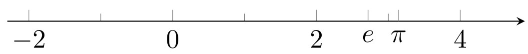


Figure 13.59: PGFplots: axis/axes range

`ticklabels={<labels>}` `extra x tick labels={<labels>}`

`hide obscured x ticks = false` for origin x tick label

```
\begin{tikzpicture}
    \begin{axis}[
        axis y line=none,axis x line=center,

↪   ymin=0,ymax=0,xmin=-2.3*pi,xmax=2.3*pi,
        xtick distance=pi,

↪   xticklabels={$-2\pi$,$-\pi$,$0$,$\pi$,$2\pi$},
    ]
        % empty
    \end{axis}
\end{tikzpicture}
```

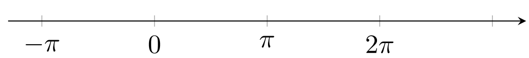


Figure 13.60: PGFplots: axis/axes range

13.4.7.2 `addplot+`

What does addplot+ do exactly?

https://tex.stackexchange.com/questions/275959/what-does-addplot-do-exactly

https://tikz.dev/pgfplots/reference-addplot#/addplot+

Every `\addplot` directive receives a pre-defined style (line color, marker style etc) through a pre-defined cycle list that is automatically chosen depending on the index of the current `\addplot` instruction. If you want to add some of your styles manually (like I want red colour instead of blue, say), you can add them through options to `\addplot` like `\addplot[<your options>]`. Now the question is whether you want your own style (your options) to be appended to or replace one of these cycle lists assigned. This is decided by the + sign. If you use `\addplot+ [<your options>]`, your style is appended to and by `\addplot[<your options>]` , your options will replace the assigned cycle list.

13.4.7.3 point

`only marks` only points without lines

zero y axis range `ymin=0, ymax=0` and `axis y line=none`, making 1D x axis

```
\begin{tikzpicture}
    \begin{axis}[
        xlabel=$x$,
        axis y line=none,
        axis x line=center,
        tick align=inside,
        xmin=-1.5, xmax=4.9, ymin=0, ymax=0,
        xtick distance=1,
        x=1cm
        ]
        \addplot+ coordinates {(e,0)}
            node [pin=90:{$e$}] {};
        \addplot+ coordinates {(pi,0)}
            node [pin=90:{$\pi$}] {};
        \addplot+ coordinates {(-1,0)};
    \end{axis}
\end{tikzpicture}
```

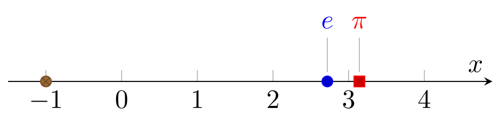


Figure 13.61: PGFplots: 1D points with pins

```
\begin{tikzpicture}
    \begin{axis}[
        xlabel=$x$, ylabel=$y$,
        axis lines=center,
        tick align=inside,
        xmin=-1.5, xmax=4.9, ymin=-3.3,
↪   ymax=3.9,
        xtick distance=1, ytick distance=1,
        axis equal image
        ]
        \addplot+ [only marks] coordinates {
            (-1,-2) (pi,pi/4) (3,2) (0,0)};
        \addplot+ coordinates {(2,1)};
        \addplot+ coordinates {(3,-2)};
    \end{axis}
\end{tikzpicture}
```

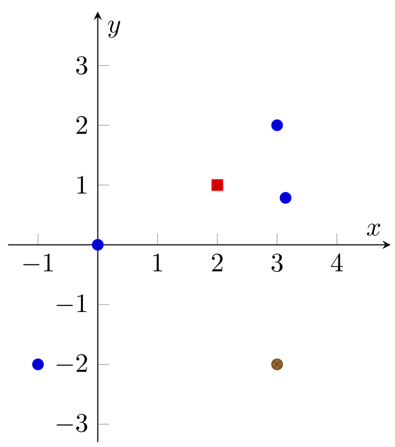


Figure 13.62: PGFplots: 2D points

```
\begin{tikzpicture}
    \begin{axis}[axis equal image]
        \addplot+ [only marks]
            table [x=xdata,y=ydata]
↪   {data/func.dat};
    \end{axis}
\end{tikzpicture}
```

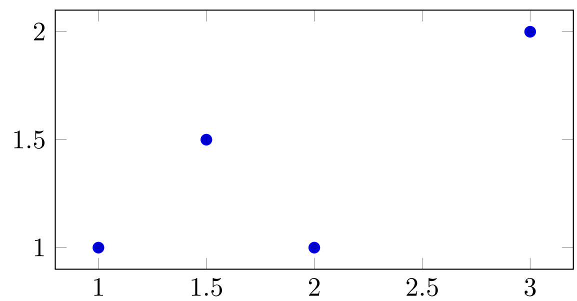


Figure 13.63: PGFplots: data points

`axis equal image` to fix aspect ratio 1:1

```
\begin{tikzpicture}
    \begin{axis}[title=5 sampling points,
        xlabel=$x$,ylabel=$y$,
        axis equal image]
        \addplot+ [only
        ↪  marks,domain=-2:2,samples=5]
            {x^2};
    \end{axis}
\end{tikzpicture}
```

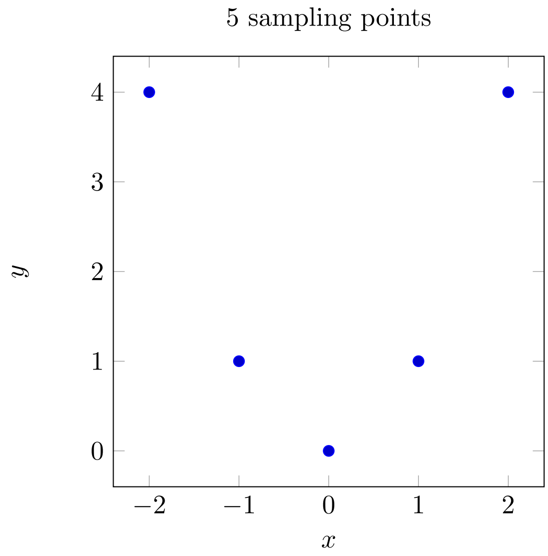


Figure 13.64: PGFplots: function sampling 5 points

```
\begin{tikzpicture}
    \begin{axis}[title=55 sampling points,
        xlabel=$x$,ylabel=$y$,
        axis equal image]
        \addplot+ [only
        ↪  marks,domain=-2:2,samples=55]
            {x^2};
    \end{axis}
\end{tikzpicture}
```

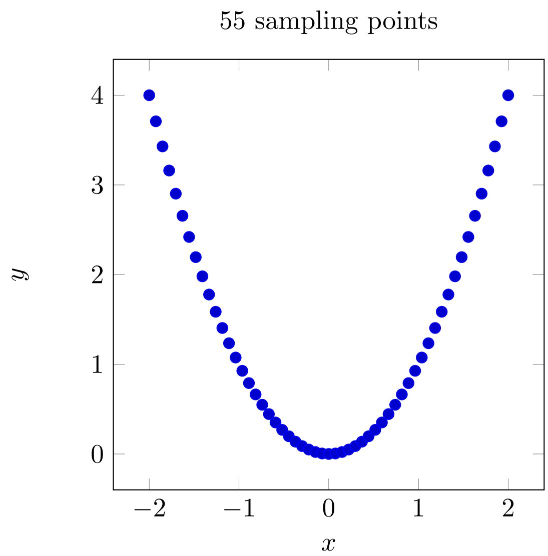


Figure 13.65: PGFplots: function sampling 55 points

```
\begin{tikzpicture}
    \begin{axis}[
        trig format plots=rad,  % angle in
↪   radian
        axis equal image]
        \addplot+ [only
        ↪  marks,variable=t,domain=0:pi*3/2,
            samples=20] ({cos(t)},{sin(t)});
    \end{axis}
\end{tikzpicture}
```

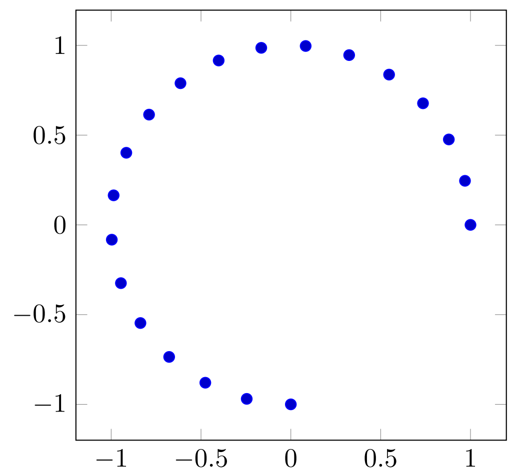


Figure 13.66: PGFplots: parametric function sampling

```
\begin{tikzpicture}
    \begin{axis}[
        xlabel=$x$,ylabel=$y$,zlabel=$z$,
        axis lines=center,
        tick align=inside,
        xmin=-1.5,xmax=3.9,
        ymin=-1.5,ymax=3.9,
        zmin=-0.5,zmax=3.9,
        xtick distance=1,
        ytick distance=1,
        ztick distance=1,
        % width=10cm,
        % scale only axis,
        view={120}{30},  % perspective angle
        axis equal image,]
        \addplot3+ coordinates{(1,0,0)};
        \addplot3+ coordinates{(0,1,0)};
        \addplot3+ coordinates{(0,0,1)};
    \end{axis}
\end{tikzpicture}
```

Figure 13.67: PGFplots: 3D points

```
\begin{tikzpicture}
    \begin{axis}[
        xlabel=$x$,ylabel=$y$,zlabel=$z$,
        grid=major
        ]
        \addplot3+ [only marks] {x^2+y^2};
    \end{axis}
\end{tikzpicture}
```

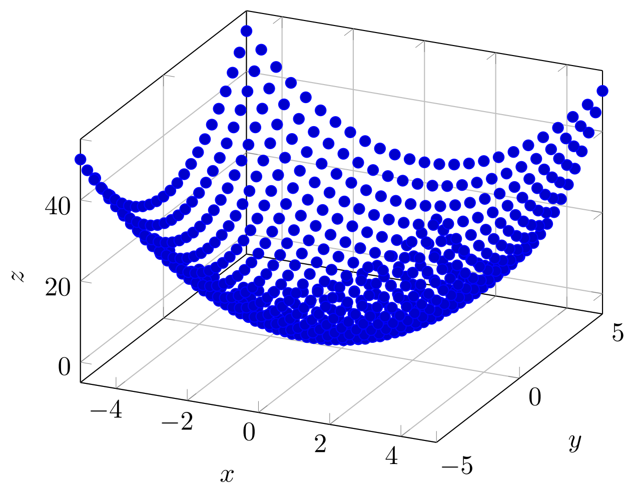


Figure 13.68: PGFplots: 3D function sampling points

13.4.7.4 line

```
\begin{tikzpicture}
    \begin{axis}[ytick=data]
        \addplot coordinates {
            (1,0.15) (2,0.21) (3,0.33) (4,0.4)

            (2.5,.1) (3.5,.1)
        };
    \end{axis}
\end{tikzpicture}
```

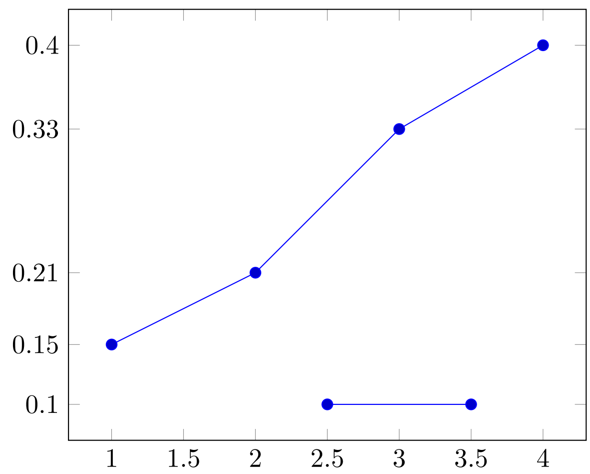


Figure 13.69: lines connecting adjacent points

`smooth` to make smooth curves or lines

```
\begin{tikzpicture}
    \begin{axis}[ytick=data]
        \addplot+ [smooth] coordinates {
            (1,0.15) (2,0.21) (3,0.33)
↪   (4,0.4)};
    \end{axis}
\end{tikzpicture}
```

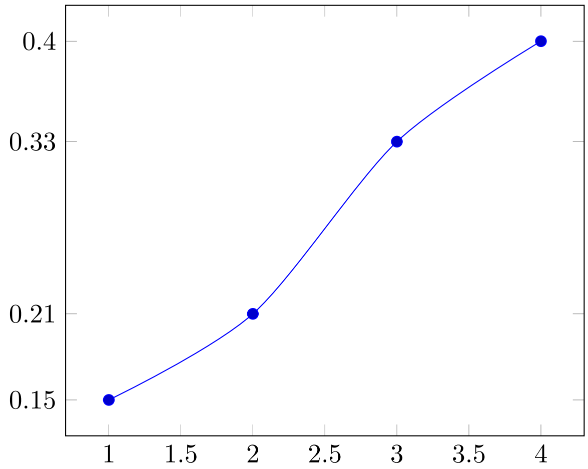


Figure 13.70: smooth lines connecting adjacent points

`no markers` to make no markers or points on the curves or lines

`\addlegendentry` to add legends

```
\begin{tikzpicture}
    \begin{axis}[
        trig format plots=rad,
        xtick distance=pi/2, ytick distance=1,

↪   xticklabels={$-\pi$,$-\pi/2$,0,$\pi/2$,$\pi$},
        width=10cm, scale only axis,
        axis equal image]
        \addplot+ [domain=-pi:pi] {sin(x)};
        \addlegendentry{$\sin(x)$}
        \addplot+ [no
        ↪   markers,domain=-pi:pi,samples=100]
            {cos(x)};
        \addlegendentry{$\cos(x)$}
    \end{axis}
\end{tikzpicture}
```

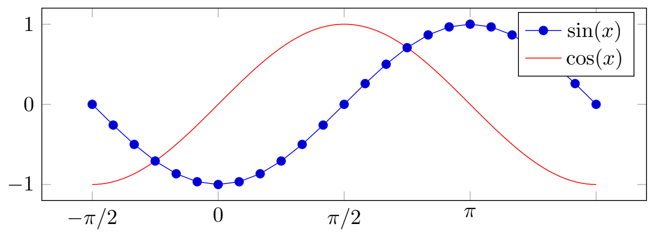


Figure 13.71: $\sin(x)$ and $\cos(x)$

```
\begin{tikzpicture}
    \begin{axis}[
        trig format plots=rad,
        ymax=2.3,
        axis equal image]
        \addplot+ [no markers,
                   variable=t,
                   domain=0:2*pi,
                   samples=100]
↪   ({2*cos(t)},{sin(t)});
        \node
        ↪   [pin=90:$\dfrac{x^2}{4}+y^2{=}1$]
        at ({2*cos(45)},{sin(45)}) {};
    \end{axis}
\end{tikzpicture}
```

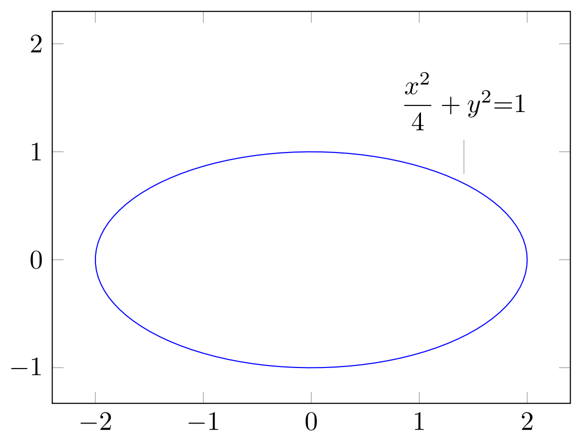


Figure 13.72: $\dfrac{x^2}{4}+y^2{=}1$

```
\begin{tikzpicture}
    \begin{axis}[trig format plots=rad]
        \addplot3+ [variable=t,
                    domain=0:3*pi,
                    samples=100,
                    samples y=0]
            ({cos(t)},{sin(t)},{t});
    \end{axis}
\end{tikzpicture}
```

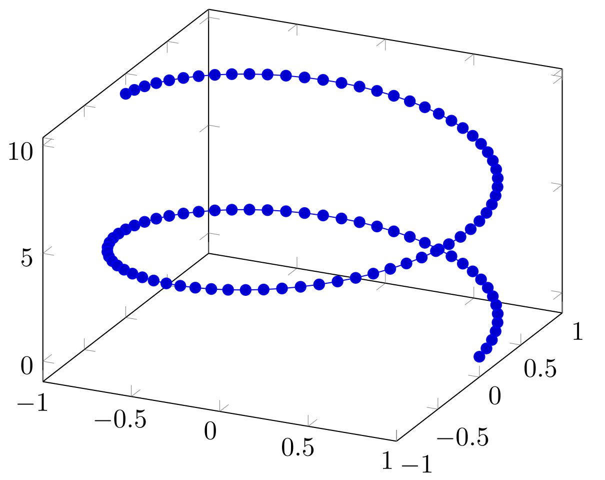


Figure 13.73: $(\cos(t), \sin(t), t)$

13.4.7.5 plane

`[mesh]`

```
\begin{tikzpicture}
    \begin{axis}
        \addplot3+ [no markers, mesh] {x*y};
    \end{axis}
\end{tikzpicture}
```

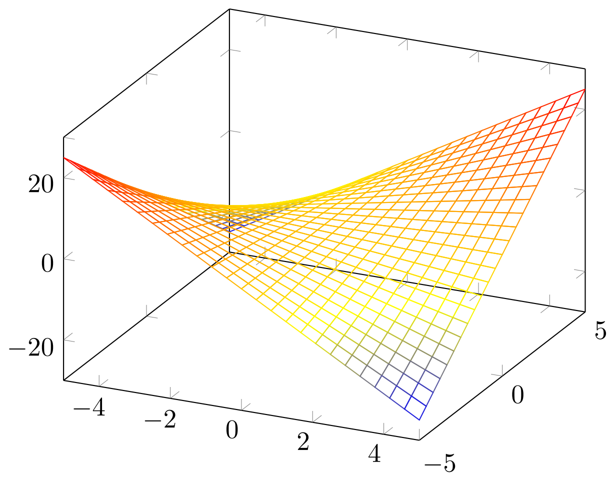


Figure 13.74: $f(x, y) = xy$

`[surf]` surface

```
\begin{tikzpicture}
    \begin{axis}[
        view={120}{30},
        trig format plots=rad,
        width=10cm,
        scale only axis,
        axis equal image
        ]
        \addplot3+ [no markers,
                    surf,
                    domain=0:2*pi,
                    domain y=0:pi]

↪   ({sin(y)*cos(x)},{sin(y)*sin(x)},{cos(y)});
    \end{axis}
\end{tikzpicture}
```

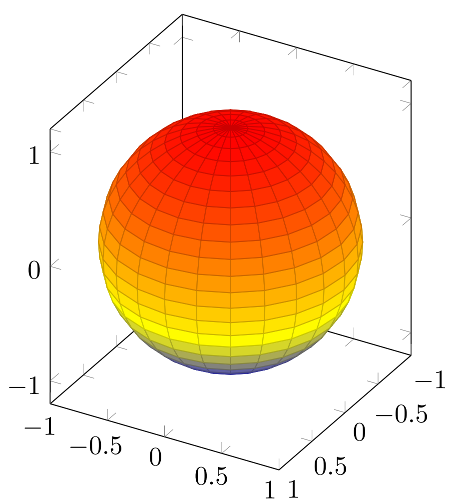


Figure 13.75: $x^2 + y^2 + z^2 = 1$

13.4.7.6 polar coordinate

`data cs=polar|polarrad`

```
\begin{tikzpicture}
    \begin{axis}[
      title={$\rho=\cos 2\theta$},
      axis equal image
      ]
        \addplot+ [no markers,
                   data cs=polar,
                   domain=0:360,
                   samples=360
                   ] (\x,{cos(2*\x)});
    \end{axis}
\end{tikzpicture}
```

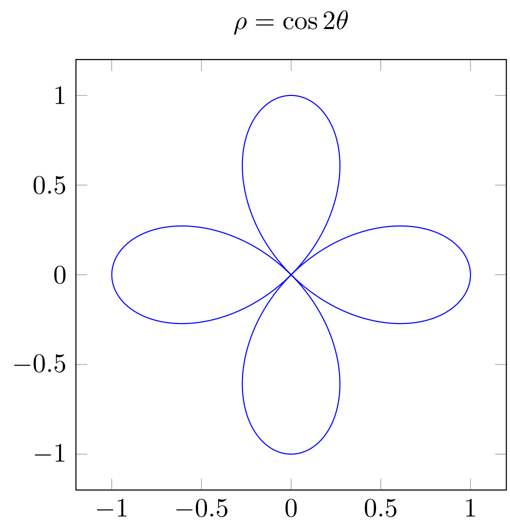


Figure 13.76: polar coordinate $\rho = \cos 2\theta$

`\usepgfplotslibrary{polar}` to use `\begin{polaraxis}`

```
\usepgfplotslibrary{polar}
\begin{tikzpicture}
    \begin{polaraxis}
        \addplot+ coordinates
            {(0,0) (60,1) (90,{sqrt(3)/2})} --
↪   cycle;
    \end{polaraxis}
\end{tikzpicture}
```

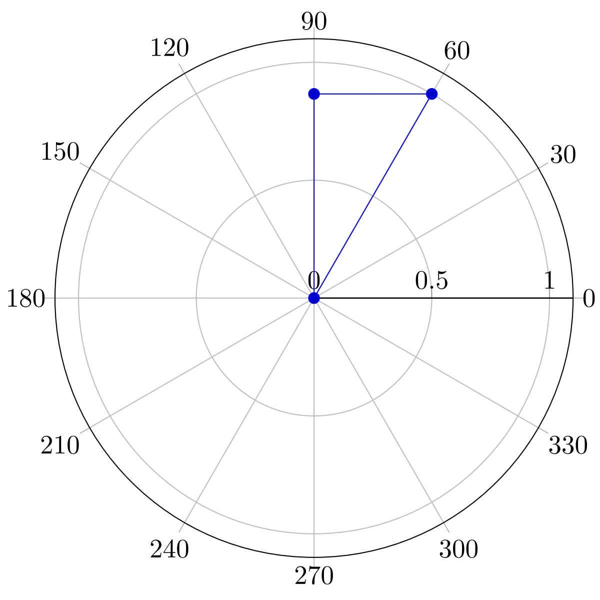


Figure 13.77: polar coordinate axes

https://zhuanlan.zhihu.com/p/128341873

13.4.8 Arnav Bandekar: Using pgfplots to make economic graphs in LaTeX

https://towardsdatascience.com/using-pgfplots-to-make-economic-graphs-in-latex-bcdc8e27c0eb

13.4.9 PGFplots gallery

https://pgfplots.sourceforge.net/gallery.html

```
\begin{tikzpicture}
\begin{axis}[
    xmin=-3,   xmax=3,
    ymin=-3,   ymax=3,
    extra x ticks={-1,1},
    extra y ticks={-2,2},
    extra tick style={grid=major},
]
    \draw[red] \pgfextra{

      ↪  \pgfpathellipse{\pgfplotspointaxisxy{0}{0}}
        {\pgfplotspointaxisdirectionxy{1}{0}}
        {\pgfplotspointaxisdirectionxy{0}{2}}
      % see also the documentation of
      % 'axis direction cs' which
      % allows a simpler way to draw this
      ↪  ellipse
    };
    \draw[blue] \pgfextra{

      ↪  \pgfpathellipse{\pgfplotspointaxisxy{0}{0}}
        {\pgfplotspointaxisdirectionxy{1}{1}}
        {\pgfplotspointaxisdirectionxy{0}{2}}
    };
    \addplot [only marks,mark=*] coordinates {
    ↪  (0,0) };
\end{axis}
\end{tikzpicture}
```

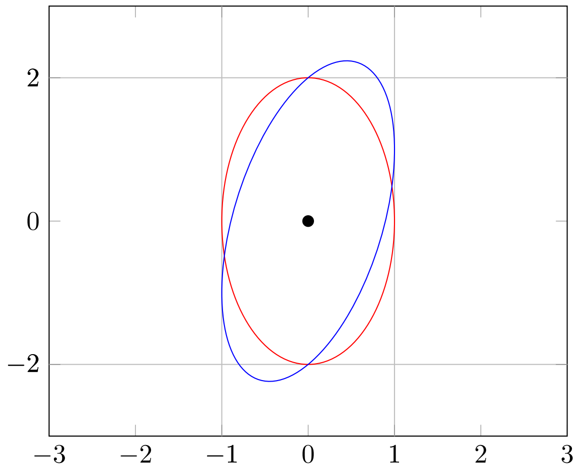


Figure 13.78: declare function

13.5 TikZplotLib / tikzplotlib

Python[12]

```
library(reticulate)
```

```
## Warning: package 'reticulate' was built under R version 4.2.3
```

```
# virtualenv_list()
# virtualenv_python()
# use_virtualenv("r-reticulate")

# conda_list()
use_condaenv(condaenv = 'sandbox-3.9')

## install TikZplotLib
# virtualenv_install("r-reticulate", "tikzplotlib")

## import TikZplotLib (it will be automatically discovered in "r-reticulate")
tikzplotlib <- import("tikzplotlib")
```

```
Error: ImportError: cannot import name 'common_texification' from 'matplotlib.backends.backend_pgf' (D:\Users
```

https://github.com/NixOS/nixpkgs/issues/289305

> The “solution” is to use **matplotlib 3.6**, but I guess in nixpkgs a single version is used at a time. The last working upgrade is from 0911608 I guess (I tried using virtualenv + pip + nix-ld + export LD_LIBRARY_PATH=“$LD_LIBRARY_PATH$:NIX_LD_LIBRARY_PATH”)

https://stackoverflow.com/questions/60882638/install-a-particular-version-of-python-package-in-a-virtualenv-created-with-reti

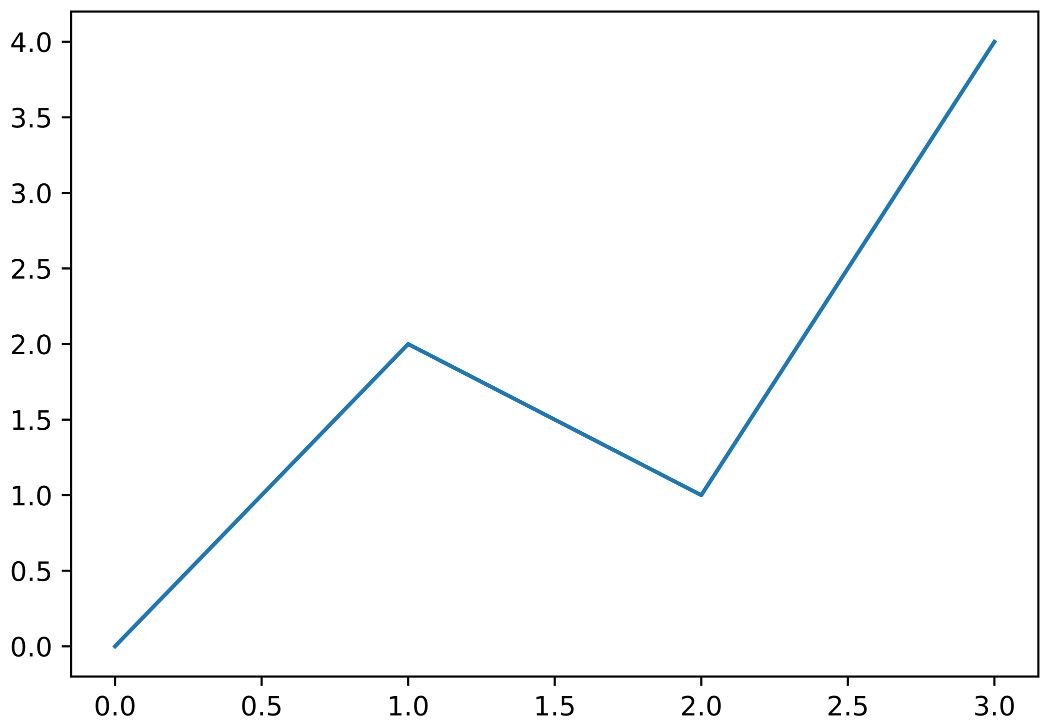
```
#reticulate::virtualenv_install(packages = c("matplotlib==3.6.0"))
```

```
reticulate::conda_list()
```

```
##                  name                                          python
## 1                base                         D:\\Anaconda3/python.exe
## 2            fiftyone          D:\\Anaconda3\\envs\\fiftyone/python.exe
## 3               keras             D:\\Anaconda3\\envs\\keras/python.exe
## 4             labelme           D:\\Anaconda3\\envs\\labelme/python.exe
## 5               manim             D:\\Anaconda3\\envs\\manim/python.exe
## 6              mmyolo            D:\\Anaconda3\\envs\\mmyolo/python.exe
## 7        r-reticulate      D:\\Anaconda3\\envs\\r-reticulate/python.exe
## 8  rsconnect-jupyter D:\\Anaconda3\\envs\\rsconnect-jupyter/python.exe
## 9             sandbox           D:\\Anaconda3\\envs\\sandbox/python.exe
## 10        sandbox-3.9       D:\\Anaconda3\\envs\\sandbox-3.9/python.exe
```

```
reticulate::use_condaenv(condaenv = 'sandbox-3.9')
```

```
import matplotlib.pyplot as plt

plt.plot([0, 2, 1, 4])
plt.show()
```



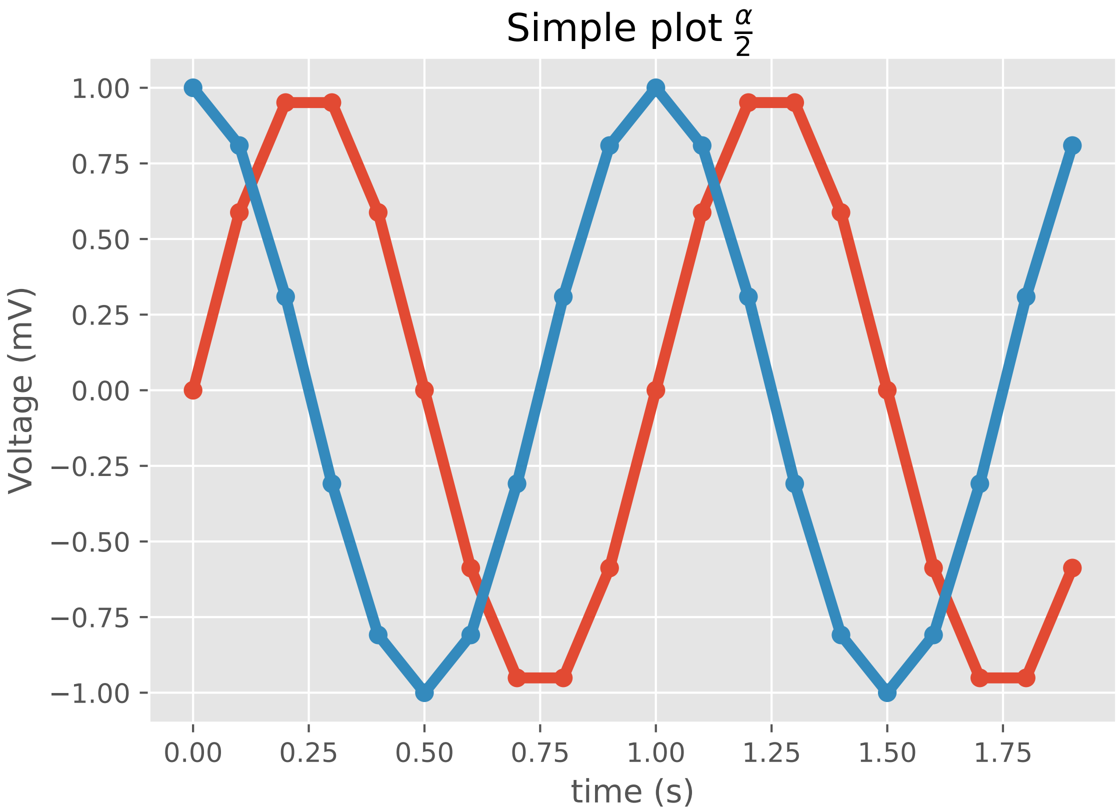

```
import tikzplotlib

# tikzplotlib.save("test.tex")
tikzplotlib.get_tikz_code()
```

```
## '% This file was created with tikzplotlib v0.10.1.\n\\begin{tikzpicture}\n\n\\end{tikzpicture}\n'
```

```
import matplotlib.pyplot as plt

plt.close()
```

```
import matplotlib.pyplot as plt
import numpy as np

plt.style.use("ggplot")
```

```
t = np.arange(0.0, 2.0, 0.1)
s = np.sin(2 * np.pi * t)
s2 = np.cos(2 * np.pi * t)
plt.plot(t, s, "o-", lw=4.1)
plt.plot(t, s2, "o-", lw=4.1)
plt.xlabel("time (s)")
plt.ylabel("Voltage (mV)")
plt.title("Simple plot $\\frac{\\alpha}{2}$")
plt.grid(True)
plt.show()
```




```
import tikzplotlib

# tikzplotlib.save("test.tex")
tikzplotlib.get_tikz_code()
```

```
## '% This file was created with tikzplotlib v0.10.1.\n\\begin{tikzpicture}\n\n\\end{tikzpicture}\n'
```

13.6 animation

https://zhuanlan.zhihu.com/p/338402487

Chapter 14

xy-pic

https://bookdown.org/yihui/rmarkdown-cookbook/install-latex-pkgs.html

`tinytex::install_tinytex()`

the following xymatrix from LaTeX package xy for xy-pic is not shown or rendered in HTML:

`$\LaTeX$` can only be used in HTML, not PDF

$$
\begin{tikzcd}
U \arrow[drr, "\varphi"] \arrow[dr, "(x,y)"] \arrow[ddr, "\psi"'] & & \\
& X \times_Z Y \arrow[r, "p"'] \arrow[d, "q"] & X \arrow[d, "f"] \\
& Y \arrow[r, "g"] & Z
\end{tikzcd}
$$

Chapter 15

statistics

15.1 Hung Hung

population

普查 vs. 統計

random variable

$$X$$

sample has randomness

probability function

$$\mathrm{P}_X(E) \in [0, 1]$$

event = subset of sample space

$$E$$

input event with output probability in 0 to 1

$$\mathrm{P}_X : \{E_i\}_{i \in I} \to [0, 1]$$

target of interest

- probability function

but events are hard to be listed or enumerated

$$X : \{\omega_i\}_{i \in I} \to \mathbb{R}$$

CDF = cumulative distribution function

$$F_X(x) = \mathrm{P}_X((-\infty, x]) = \mathrm{P}_X(X \le x)$$

real function is much easier to be operable, there is differentiation or difference operation

target of interest

- CDF = cumulative distribution function
- probability function

target of interest

- $\mathrm{P}_X(\cdot)$ PF = probability function
  - $f_X(x) = \mathrm{P}_X(X = x)$ PMF = probability mass function
  - $f_X(x) = \frac{\mathrm{d}}{\mathrm{d}x}\mathrm{P}_X(X \le x)$ PDF = probability density function

- $F_X(x)$ CDF = cumulative distribution function[15.1.2.1]
- $M_X(\xi)$ MGF = moment generating function[15.1.2.7.1]
- $\varphi_X(\xi)$ CF = characteristic function[15.1.2.7.2]

$$\begin{array}{ccccccc} X & \sim & F_X(x) & \xleftrightarrow{\text{FToC}} & f_X(x) & \leftrightarrow & \mathrm{P}_X \\ & \text{inversion formula} \ : & \updownarrow & \forall \xi \approx 0 \left[M_X(\xi) \in \mathbb{R} \right] \ \wedge & \uparrow\downarrow & \searrow\nwarrow & \hookleftarrow \wedge \ \mathrm{supp}\left(f_X\right) \ \text{ is bounded} \\ & & \varphi_X(\xi) & \leftrightarrows & M_X(\xi) & \rightarrow & \{\mu_n | n \in \mathbb{N}\} \end{array}$$

In population,

$$X \sim F_X(x)$$

by sampling,

$$X_1, \cdots, X_i, \cdots, X_n = X_i \sim F_X(x)$$

or

$$X_1, \cdots, X_i, \cdots, X_n = X_i \overset{\text{i.i.d.}}{\sim} F_X(x)$$

i.i.d. = independently identically distributed

and inference back

parametrically

$$\widehat{X} \sim \widehat{F}_X(x) = \widehat{F}_X(x|\theta)$$

or nonparametrically

$$\hat{X} \sim \hat{F}_X(x)$$

inference is function of samples, or called random function, to estimate unknown parameters

$$\widehat{\Theta} \leftarrow (X_1, \cdots, X_i, \cdots, X_n) = T(X_1, \cdots, X_n) = T(\cdots, X_i, \cdots) = T(X_i)$$

correspondng CDF for inference or estimation function of sampling random variables

$$T(X_1, \cdots, X_n) = T \sim F_T(t)$$

wish to be unbiased and consistent

$$\begin{cases} \mathrm{E}\left(\widehat{\Theta}\right) = \theta \Leftrightarrow \mathrm{E}\left(\widehat{\Theta}\right) - \theta = 0 & \text{unbiasedness} \\ \mathrm{V}\left(\widehat{\Theta}\right) = 0 & \text{consistency} \end{cases}$$

unbiasedness usually harder than consistency, thus usually first considered consistency.

modeling or parameterizing with unknown parameter θ

$$F_X(x) \overset{M}{=} F_X(x|\theta) = F_X(x;\theta)$$

parameterization is to reduce unknown parameters from infinite ones to finite ones

e.g. for normally distributed data

$$f_X(x) \overset{M}{=} f_X(x|\theta) = \frac{\mathrm{e}^{\frac{-1}{2}\left(\frac{x-\mu}{\sigma}\right)^2}}{\sqrt{2\pi\sigma^2}} = f_X\left(x|\mu,\sigma^2\right) = f_X(x|\mu,\sigma)$$

the price of parameterization is guess wrong model.

For some non-negative data, instead of normal distribution, consider distributions skewed to the right

- gamma
- exponential
- Weibull
- log-normal

topics

1. P_X probability theory
2. $f_X(x) \overset{M}{=} f_X(x|\theta) = f_X(x;\theta)$ various univariable distribution
3. $f_{\boldsymbol{X}}(\boldsymbol{x}) \overset{M}{=} f_{\boldsymbol{X}}(\boldsymbol{x}|\boldsymbol{\theta}) = f_{\boldsymbol{X}}(\boldsymbol{x};\boldsymbol{\theta})$ multivariable distribution
4. $T(X_1, \cdots, X_n)$ inference
   - point estimation
   $$\widehat{\mu} = \begin{cases} \overline{X} & \to \mu \\ \text{median}(X_1, \cdots, X_n) & \to \mu \\ \vdots \end{cases}$$
   - interval estimation = hypothesis testing
   $$\begin{cases} H_0 : \theta = \theta_0 \\ H_1 : \theta \neq \theta_0 \end{cases} \leftarrow T \in \{0, 1\}$$
5. how to find T
6. behavior of random function $T(X_1, \cdots, X_n) = T \sim F_T(t)$
   - statistical properties of T
   - asymptotic properties
   $$n \to \infty \begin{cases} \text{CLT} = \text{central limit theorem} \\ \text{LLN} = \text{law of large number} \end{cases}$$

15.1.1 probability theory

Definition 15.1. sample space: The set S of all possible outcomes of an experiment is called the sample space

$$S = \{\omega_i\}_{i \in I}$$

Definition 15.2. event: An event E is any collection of possible outcomes of an experiment, i.e. any subset of S

$$E \subseteq S$$

set operation

commutativity, associativity, distributivity

De Morgan law

pairwise disjoint = mutually exclusive

partition

15.1.1.1 probability function

probability function axioms = probability function definition

Kolmogorov axioms of probability[6] p.72

Definition 15.3. probability function: Given a sample space S and its event E, a probability function is a function P satisfying

$$\begin{cases} \mathrm{P}(S) = 1 \\ \forall E \subseteq S \, (\mathrm{P}(E) \geq 0) \\ E_1, \cdots, E_i, \cdots \text{ are pairwise disjoint} \Rightarrow \mathrm{P}\left(\bigcup_{i \in I} E_i\right) = \sum_{i \in I} E_i \end{cases}$$

tossing a dice

theorems

$$\mathrm{P}(\emptyset) = 0$$

$$\mathrm{P}(E) \leq 1$$

$$\mathrm{P}\left(E^{\mathrm{C}}\right) = \mathrm{P}\left(\overline{E}\right) = 1 - \mathrm{P}(E)$$

$$\mathrm{P}\left(E_2 \cap \overline{E}_1\right) = \mathrm{P}(E_2) - \mathrm{P}(E_2 \cap E_1)$$

$$E_1 \subseteq E_2 \Rightarrow \mathrm{P}(E_1) \leq \mathrm{P}(E_2)$$

addition rule[6] p.75 and extended addition rule[6] p.76

inclusion-exclusion principle = sieve principle

$$\mathrm{P}(E_1 \cup E_2) = \mathrm{P}(E_1) + \mathrm{P}(E_2) - \mathrm{P}(E_1 \cap E_2)$$

$$\mathrm{P}(E_1 \cup E_2 \cup E_3) = \mathrm{P}(E_1) + \mathrm{P}(E_2) + \mathrm{P}(E_3) - \mathrm{P}(E_1 \cap E_2) - \mathrm{P}(E_2 \cap E_3) - \mathrm{P}(E_3 \cap E_1) + \mathrm{P}(E_1 \cap E_2 \cap E_3)$$

$$\mathrm{P}\left(\bigcup_{i=1}^{n} E_i\right) = \sum_{k=1}^{n}\left((-1)^{k-1} \sum_{1 \leq i_1 < \cdots < i_k \leq n} \mathrm{P}\left(\bigcap_{i \in \{i_1, \ldots, i_k\}} E_i\right)\right)$$

symmetric difference[6] p.75

union probability upper-bounded by sum of indivisual probability

$$\mathrm{P}(E_1 \cup E_2) = \mathrm{P}(E_1) + \mathrm{P}(E_2) - \mathrm{P}(E_1 \cap E_2) \leq \mathrm{P}(E_1) + \mathrm{P}(E_2)$$

$$E_1 \cap E_2 = \emptyset \Leftrightarrow \mathrm{P}(E_1 \cup E_2) = \mathrm{P}(E_1) + \mathrm{P}(E_2)$$

Boole inequality

$$\mathrm{P}\left(\bigcup_{i \in I} E_i\right) \leq \sum_{i \in I} E_i$$

$$\mathrm{P}\left(\left.\widehat{H_0}\right| H_0\right) = \mathrm{P}(\text{reject } H_0 | H_0 \text{ is true}) = \alpha = \text{type 1 error}$$

multiple hypothesis testing

How to control the family-wise error rate?

Ideally,

FWER = family-wise error rate

$$
\begin{aligned}
\alpha =& \mathrm{P}\left(\overbrace{H_0^1} \cup \cdots \cup \overbrace{H_0^m} \middle| H_0^1 \cap \cdots \cap H_0^m\right) = \mathrm{P}\left(\text{reject any } H_0^i \middle| \text{any } H_0^j \text{ is true}\right)\\
=& \mathrm{P}\left(\bigcup_{i=1}^m \overbrace{H_0^i} \middle| \bigcap_{j=1}^m H_0^j\right) = 1 - \mathrm{P}\left(\overbrace{\bigcup_{i=1}^m \overbrace{H_0^i}} \middle| \bigcap_{j=1}^m H_0^j\right)\\
=& 1 - \mathrm{P}\left(\bigcap_{i=1}^m \underbrace{H_0^i} \middle| \bigcap_{j=1}^m H_0^j\right) = 1 - \mathrm{P}\left(\text{not to reject any } H_0^i \middle| \bigcap_{j=1}^m H_0^j\right)\\
&\overset{H_0^j \text{ pairwise independent}}{=} 1 - \prod_{i=1}^m \mathrm{P}\left(\underbrace{H_0^i} \middle| \bigcap_{j=1}^m H_0^j\right) = 1 - \prod_{i=1}^m \left(1 - \mathrm{P}\left(\overbrace{H_0^i} \middle| \bigcap_{j=1}^m H_0^j\right)\right)\\
&\overset{\forall i,j\left[\mathrm{P}\left(\overbrace{H_0^i} \middle| \bigcap_{j=1}^m H_0^j\right) = \alpha_0\right]}{=} 1 - \prod_{i=1}^m (1 - \alpha_0) = 1 - (1 - \alpha_0)^m\\
\alpha =& 1 - (1 - \alpha_0)^m\\
\alpha_0 =& 1 - (1 - \alpha)^{\frac{1}{m}} = 1 - \sqrt[m]{1 - \alpha}\\
&\Downarrow\\
&\text{set } \mathrm{P}\left(\overbrace{H_0^i} \middle| \bigcap_{j=1}^m H_0^j\right) = \alpha_0 = 1 - \sqrt[m]{1 - \alpha}
\end{aligned}
$$

But condition H_0^j pairwise independent is too strong.

Practically,

$$
\begin{aligned}
\alpha =& \mathrm{P}\left(\overbrace{H_0^1} \cup \cdots \cup \overbrace{H_0^m} \middle| H_0^1 \cap \cdots \cap H_0^m\right) = \mathrm{P}\left(\text{reject any} H_0^i \middle| \text{any } H_0^j \text{ is true}\right)\\
=& \mathrm{P}\left(\bigcup_{i=1}^m \overbrace{H_0^i} \middle| \bigcap_{j=1}^m H_0^j\right) \overset{\mathrm{P}\left(\bigcup_{i \in I} E_i\right) \le \sum_{i \in I} E_i}{\underset{\text{Boole inequality}}{\le}} \sum_{i=1}^m \mathrm{P}\left(\overbrace{H_0^i} \middle| \bigcap_{j=1}^m H_0^j\right) \underset{\Uparrow}{=} \sum_{i=1}^m \alpha_0 = m\alpha_0 \underset{\Downarrow}{=} \alpha\\
\text{let } \forall i,j & \left[\mathrm{P}\left(\overbrace{H_0^i} \middle| \bigcap_{j=1}^m H_0^j\right) = \alpha_0\right] \Rightarrow \sum_{i=1}^m \mathrm{P}\left(\overbrace{H_0^i} \middle| \bigcap_{j=1}^m H_0^j\right) = \sum_{i=1}^m \alpha_0 = m\alpha_0 \Rightarrow \alpha_0 = \frac{\alpha}{m}\\
&\Downarrow\\
&\text{set } \mathrm{P}\left(\overbrace{H_0^i} \middle| \bigcap_{j=1}^m H_0^j\right) = \alpha_0 = \frac{\alpha}{m}
\end{aligned}
$$

Bonferroni correction

$$
\mathrm{P}\left(\overbrace{H_0^i} \middle| \bigcap_{j=1}^m H_0^j\right) = \frac{\alpha}{m} \Rightarrow \mathrm{P}\left(\bigcup_{i=1}^m \overbrace{H_0^i} \middle| \bigcap_{j=1}^m H_0^j\right) \le \alpha
$$

Bonferroni inequality[6] p.77

Bonferroni inequality and Boole inequality are equivalent inequalities

birthday problem[6] p.78

15.1.1.2 conditional probability

15.1.2 univariable distribution

$$
\begin{cases}
\mathrm{P}_X\left(X \in E\right) & \forall E \subseteq S\\
\mathrm{P}_X\left(X \le x\right) & \forall x \in \mathbb{R}
\end{cases}
$$

$$\begin{cases} \mathrm{P}_X\left(X \in E\right) & \forall E \subseteq S \\ \mathrm{P}_X\left(X \leq x\right) = \mathrm{P}_X\left(\left(-\infty, x\right]\right) = F_X\left(x\right) & \forall x \in \mathbb{R} \end{cases}$$

$$\begin{aligned} &\mathrm{P}_X\left(X \leq x\right) = \mathrm{P}_X\left(\left(-\infty, x\right]\right) \\ =&\mathrm{P}_X\left(\bigcup_{\epsilon>0}\left(-\infty, x-\epsilon\right]\right) = \lim_{\epsilon\to 0}\mathrm{P}_X\left(\left(-\infty, x-\epsilon\right]\right) \\ \leftrightarrow&\mathrm{P}_X\left(\left(-\infty, x\right)\right) = \mathrm{P}_X\left(E\right), E = \left(-\infty, x\right) \end{aligned}$$

15.1.2.1 cumulative distribution function

CDF = cumulative distribution function

$$F_X\left(x\right) = \mathrm{P}_X\left(\left(-\infty, x\right]\right) = \mathrm{P}_X\left(X \leq x\right)$$
$$X \sim \mathrm{P}_X \leftrightarrow F_X\left(x\right)$$

$$X \sim F_X\left(x\right) \leftrightarrow \mathrm{P}_X$$

Definition 15.4. CDF = cumulative distribution function: A cumulative distribution function is a function $F : \mathbb{R} \to [0, 1]$ satisfying

$$F_X\left(x\right) = \mathrm{P}_X\left(\left(-\infty, x\right]\right) = \mathrm{P}_X\left(X \leq x\right)$$

Theorem 15.1. *CDF = cumulative distribution function: $F\left(x\right)$ is a cumulative distribution function iff*

$$\begin{cases} \begin{cases} \lim_{x\to-\infty} F\left(x\right) = 0 & \lim_{x\to+\infty} F\left(x\right) = 1 \end{cases} & (01)\ [0,1] \\ \forall x_1 < x_2 \left[F\left(x_1\right) \leq F\left(x_2\right)\right] & (nd)\text{ non-decreasing} \\ \lim_{x\to x_0^+} F\left(x\right) = F\left(x_0\right) & (rc)\text{ right-continuous} \end{cases}$$

Definition 15.5. RV = r.v. = random variable

$$\begin{cases} X \text{ is a continuous RV} & \lim_{x\to x_0} F_X\left(x\right) = F_X\left(x_0\right) \\ X \text{ is a discrete RV} & F_X \text{ is a step function of } x \end{cases}$$

[6] p.103

Definition 15.6. RV = r.v. = random variable

[6] p.104

Definition 15.7. range of r.v. = range of RV = the range of a random variable

$$\begin{aligned} \mathcal{R}_X =& \left\{ x \middle| \begin{cases} \omega \in S \\ x \in X\left(\omega\right) \end{cases} \right\} \\ =& \left\{x \middle| \forall\omega \in S\left[x \in X\left(\omega\right)\right]\right\} \\ =& \left\{x \middle| x \in X\left(\Omega\right)\right\} = X\left(\Omega\right) \end{aligned}$$

15.1.2.2 probability density function

$$\begin{cases} \mathrm{P}_X\left(X \leq x\right) = \mathrm{P}_X\left(\left(-\infty, x\right]\right) = F_X\left(x\right) \\ \mathrm{P}_X\left(X = x\right) = \mathrm{P}_X\left(x\right) = ? \end{cases}$$

Definition 15.8. PDF = probability density function

PMF = probability mass function

$$\begin{cases} f_X(x) = \dfrac{\mathrm{d}}{\mathrm{d}x} F_X(x) & X \text{ continuous RV} \\ f_X(x) = F_X(x) - F_X(x^-) & X \text{ discrete RV} \end{cases}$$

$$\begin{cases} f_X(x) = \text{derivative of } F_X(x) & X \text{ continuous} \\ f_X(x) = \text{difference of } F_X(x) & X \text{ discrete} \end{cases}$$

$$\begin{cases} f_X(x) = \dfrac{\mathrm{d}}{\mathrm{d}x} F_X(x) & \Leftrightarrow F_X(x) = \int\limits_{-\infty}^{x} f_X(t)\,\mathrm{d}t \\ f_X(x) = F_X(x) - F_X(x^-) & \Leftrightarrow F_X(x) = \sum\limits_{t \le x} f_X(t) \end{cases}$$

$$\begin{cases} X \sim \mathrm{P}_X \leftrightarrow F_X(x) \leftrightarrow f_X(x) & \text{e.g. probability theory} \\ X \sim F_X(x) \leftrightarrow \mathrm{P}_X & \Rightarrow F_X(x) \overset{M}{=} F_X(x|\theta) \text{ e.g. survial analysis} \\ X \sim f_X(x) \leftrightarrow F_X(x) \leftrightarrow \mathrm{P}_X & \Rightarrow f_X(x) \overset{M}{=} f_X(x|\theta) \text{ e.g. general statistics} \end{cases}$$

Theorem 15.2. *PDF = probability density function or PMF = probability mass function: $f(x)$ is a probability density function or probability mass function iff*

$$\begin{cases} \forall x \in \mathbb{R}\,[f(x) \ge 0] \\ \begin{cases} \int\limits_{-\infty}^{+\infty} f(x)\,\mathrm{d}x = 1 & t \text{ continuous} \\ \sum\limits_{x \in X(\Omega)} f(x) = 1 & t \text{ discrete} \end{cases} \end{cases}$$

$$\forall E \subseteq S \left[\mathrm{P}_X(X \in E) = \begin{cases} \int_{x \in E} f_X(x)\,\mathrm{d}x & X \text{ continuous} \\ \sum\limits_{x \in E} f_X(x) & X \text{ discrete} \end{cases} \right]$$

$$\begin{aligned} \mathrm{P}_X(X = x) &= \lim_{\epsilon \to 0} \mathrm{P}_X([x-\epsilon, x+\epsilon]) \\ &= \lim_{\epsilon \to 0} \mathrm{P}_X(x - \epsilon \le X \le x + \epsilon) \\ &= \lim_{\epsilon \to 0} [F_X(x+\epsilon) - F_X(x-\epsilon)] \\ &= \begin{cases} F_X(x) - F_X(x) = 0 & X \text{ continuous} \\ F_X(x) - F_X(x^-) = f_X(x) & X \text{ discrete} \end{cases} \end{aligned}$$

$$X \sim F_X(x) \leftrightarrow \mathrm{P}_X$$

$$Y = g(X)$$

$$\begin{cases} Y \sim F_Y(y) \leftrightarrow f_Y(y) & \Rightarrow F_Y(y) \overset{M}{=} F_Y(y|\theta) \\ Y \sim f_Y(y) \leftrightarrow F_Y(y) \leftrightarrow \mathrm{P}_Y & \Rightarrow f_Y(y) \overset{M}{=} f_Y(y|\theta) \end{cases}$$

15.1.2.3 range vs. support

Definition 15.9. range of r.v. = range of RV = the range of a random variable

$$\begin{aligned} \mathcal{R}_X &= \left\{ x \middle| \begin{cases} \omega \in S \\ x \in X(\omega) \end{cases} \right\} \\ &= \{x | \forall \omega \in S\,[x \in X(\omega)]\} \\ &= \{x | x \in X(\Omega)\} = X(\Omega) \end{aligned}$$

Definition 15.10. support

$$\operatorname{supp}(f)=\left\{x\,\middle|\,\begin{cases}f:D\to\mathcal{R}\\x\in D\\f_X(x)\neq 0\end{cases}\right\}$$

Definition 15.11. support of r.v. = support of RV = the support of a random variable

$$\operatorname{supp}(f_X)=\left\{x\,\middle|\,\begin{cases}x\in X(\Omega)\\f_X(x)\neq 0\end{cases}\right\}\overset{f_X(x)\geq 0}{=}\left\{x\,\middle|\,\begin{cases}x\in X(\Omega)\\f_X(x)>0\end{cases}\right\}$$

15.1.2.4 continuous monotone transformation

Theorem 15.3. *Random variable Y is monotone transformation of random variable X, i.e.* $\begin{cases}X\sim F_X(x)\leftrightarrow f_X(x)\\Y=g(X)\begin{cases}\forall x_1<x_2\left[g(x_1)<g(x_2)\right]\\\forall x_1<x_2\left[g(x_1)>g(x_2)\right]\end{cases}\end{cases}$ *then*

$$f_Y(y)=f_X\left(g^{-1}(y)\right)\left|\frac{\mathrm{d}g^{-1}(y)}{\mathrm{d}y}\right|$$

Proof:

$$\begin{aligned}F_Y(y)=&\mathrm{P}_Y(Y\leq y)\\=&\mathrm{P}(g(X)\leq y)\begin{cases}\forall x_1<x_2\left[g(x_1)<g(x_2)\right]&\Leftrightarrow\forall g(x_1)<g(x_2)\left[x_1<x_2\right]\\\forall x_1<x_2\left[g(x_1)>g(x_2)\right]&\Leftrightarrow\forall g(x_1)>g(x_2)\left[x_1<x_2\right]\end{cases}\\=&\begin{cases}\mathrm{P}_X\left(X\leq g^{-1}(y)=x\right)&\forall y_1<y_2\left[g^{-1}(y_1)<g^{-1}(y_2)\right]\\\mathrm{P}_X\left(X\geq g^{-1}(y)=x\right)&\forall y_1>y_2\left[g^{-1}(y_1)<g^{-1}(y_2)\right]\end{cases}\\=&\begin{cases}\mathrm{P}_X\left(X\leq g^{-1}(y)=x\right)&\forall y_1<y_2\left[g^{-1}(y_1)<g^{-1}(y_2)\right]\\\mathrm{P}_X\left(X\geq g^{-1}(y)=x\right)&\forall y_1<y_2\left[g^{-1}(y_1)>g^{-1}(y_2)\right]\end{cases}\\=&\begin{cases}F_X\left(g^{-1}(y)\right)&\forall y_1<y_2\left[g^{-1}(y_1)<g^{-1}(y_2)\right]\\1-F_X\left(g^{-1}(y)\right)&\forall y_1<y_2\left[g^{-1}(y_1)>g^{-1}(y_2)\right]\end{cases}\\F_Y(y)=&\begin{cases}F_X\left(g^{-1}(y)\right)&\forall y_1<y_2\left[g^{-1}(y_1)<g^{-1}(y_2)\right]\\1-F_X\left(g^{-1}(y)\right)&\forall y_1<y_2\left[g^{-1}(y_1)>g^{-1}(y_2)\right]\end{cases}\end{aligned}$$

$$\begin{aligned}
f_Y(y) = \frac{\mathrm{d}}{\mathrm{d}y} F_Y(y) &= \begin{cases} \dfrac{\mathrm{d}}{\mathrm{d}y} F_X\left(g^{-1}(y)\right) & \forall y_1 < y_2 \left[g^{-1}(y_1) < g^{-1}(y_2)\right] \\ \dfrac{\mathrm{d}}{\mathrm{d}y}\left[1 - F_X\left(g^{-1}(y)\right)\right] & \forall y_1 < y_2 \left[g^{-1}(y_1) > g^{-1}(y_2)\right] \end{cases} \\
&= \begin{cases} \dfrac{\mathrm{d}F_X\left(g^{-1}(y)\right)}{\mathrm{d}g^{-1}(y)} \dfrac{\mathrm{d}g^{-1}(y)}{\mathrm{d}y} & \forall y_1 < y_2 \left[g^{-1}(y_1) < g^{-1}(y_2)\right] \\ \dfrac{-\mathrm{d}F_X\left(g^{-1}(y)\right)}{\mathrm{d}g^{-1}(y)} \dfrac{\mathrm{d}g^{-1}(y)}{\mathrm{d}y} & \forall y_1 < y_2 \left[g^{-1}(y_1) > g^{-1}(y_2)\right] \end{cases} \\
&= \begin{cases} \dfrac{\mathrm{d}F_X\left(g^{-1}(y)\right)}{\mathrm{d}g^{-1}(y)} \dfrac{\mathrm{d}g^{-1}(y)}{\mathrm{d}y} & \forall y_1 < y_2 \left[g^{-1}(y_1) < g^{-1}(y_2)\right] \\ \dfrac{\mathrm{d}F_X\left(g^{-1}(y)\right)}{\mathrm{d}g^{-1}(y)} \dfrac{-\mathrm{d}g^{-1}(y)}{\mathrm{d}y} & \forall y_1 < y_2 \left[g^{-1}(y_1) > g^{-1}(y_2)\right] \end{cases} \\
&= \begin{cases} f_X\left(g^{-1}(y)\right) \dfrac{\mathrm{d}g^{-1}(y)}{\mathrm{d}y} & \begin{cases} \dfrac{\mathrm{d}g^{-1}(y)}{\mathrm{d}y} \geq 0 \\ f_X\left(g^{-1}(y)\right) \geq 0 \end{cases} \Rightarrow f_Y(y) \geq 0 \\ f_X\left(g^{-1}(y)\right) \dfrac{-\mathrm{d}g^{-1}(y)}{\mathrm{d}y} & \begin{cases} \dfrac{-\mathrm{d}g^{-1}(y)}{\mathrm{d}y} \geq 0 \\ f_X\left(g^{-1}(y)\right) \geq 0 \end{cases} \Rightarrow f_Y(y) \geq 0 \end{cases} \\
f_Y(y) &= \begin{cases} f_X\left(g^{-1}(y)\right) \dfrac{\mathrm{d}g^{-1}(y)}{\mathrm{d}y} & \dfrac{\mathrm{d}g^{-1}(y)}{\mathrm{d}y} \geq 0 \\ f_X\left(g^{-1}(y)\right) \dfrac{-\mathrm{d}g^{-1}(y)}{\mathrm{d}y} & \dfrac{-\mathrm{d}g^{-1}(y)}{\mathrm{d}y} \geq 0 \end{cases} \\
f_Y(y) &= f_X\left(g^{-1}(y)\right) \left| \frac{\mathrm{d}g^{-1}(y)}{\mathrm{d}y} \right|
\end{aligned}$$

□

segment $g(X)$ into monotone functions

For example, $\begin{cases} g(x) = x^2 \\ Y = g(X) \end{cases} \Rightarrow Y = g(X) = X^2,$

$$\begin{cases} Y = g(X) = X^2 \\ X \in (-\infty, +\infty) \end{cases}$$

$$\begin{aligned}
Y &= g(X) = X^2 \\
\Rightarrow Y &= \begin{cases} X^2 = g(X) & X \geq 0 \Leftrightarrow X \in [0, +\infty) \Rightarrow \forall X_1 < X_2 \left[X_1^2 < X_2^2\right] \\ X^2 = g(X) & X < 0 \Leftrightarrow X \in (-\infty, 0) \Rightarrow \forall X_1 < X_2 \left[X_1^2 > X_2^2\right] \end{cases} \\
\Rightarrow X &= \begin{cases} \sqrt{Y} = g^{-1}(Y) & X \geq 0 \Rightarrow \forall X_1^2 < X_2^2 \left[X_1 < X_2\right] \Rightarrow \forall Y_1 < Y_2 \left[X_1 < X_2\right] \\ -\sqrt{Y} = g^{-1}(Y) & X < 0 \Rightarrow \forall X_1^2 < X_2^2 \left[X_1 > X_2\right] \Rightarrow \forall Y_1 < Y_2 \left[X_1 > X_2\right] \end{cases} \\
\Rightarrow X &= \begin{cases} \sqrt{Y} = g^{-1}(Y) & Y \in [0, \infty) \Rightarrow X \geq 0 \Rightarrow \forall Y_1 < Y_2 \left[g^{-1}(Y_1) < g^{-1}(Y_2)\right] \\ -\sqrt{Y} = g^{-1}(Y) & Y \in [0, \infty) \Rightarrow X < 0 \Rightarrow \forall Y_1 < Y_2 \left[g^{-1}(Y_1) > g^{-1}(Y_2)\right] \end{cases}
\end{aligned}$$

$$\begin{aligned}
F_Y(y) =& \mathrm{P}_Y(Y \le y) = \mathrm{P}\left(X^2 \le y\right) \\
=& \mathrm{P}\left(\left\{X^2 \le y\right\} \cap (\{X < 0\} \cup \{X \ge 0\})\right) \\
=& \mathrm{P}\left(\left(\left\{X^2 \le y\right\} \cap \{X < 0\}\right) \cup \left(\left\{X^2 \le y\right\} \cap \{X \ge 0\}\right)\right) \\
=& \mathrm{P}\left(\left\{X^2 \le y\right\} \cap \{X < 0\}\right) + \mathrm{P}\left(\left\{X^2 \le y\right\} \cap \{X \ge 0\}\right) \\
& - \mathrm{P}\left(\left(\left\{X^2 \le y\right\} \cap \{X < 0\}\right) \cap \left(\left\{X^2 \le y\right\} \cap \{X \ge 0\}\right)\right) \\
=& \mathrm{P}\left(\left\{X^2 \le y\right\} \cap \{X < 0\}\right) + \mathrm{P}\left(\left\{X^2 \le y\right\} \cap \{X \ge 0\}\right) - \mathrm{P}(\emptyset) \\
=& \mathrm{P}\left(\left\{X^2 \le y\right\} \cap \{X < 0\}\right) + \mathrm{P}\left(\left\{X^2 \le y\right\} \cap \{X \ge 0\}\right) - 0 \\
=& \mathrm{P}\left(\left\{X^2 \le y\right\} \cap \{X < 0\}\right) + \mathrm{P}\left(\left\{X^2 \le y\right\} \cap \{X \ge 0\}\right) \\
=& \mathrm{P}\left(\{-X \le \sqrt{y}\} \cap \{X < 0\}\right) + \mathrm{P}\left(\{X \le \sqrt{y}\} \cap \{X \ge 0\}\right) \\
=& \mathrm{P}\left(\{X \ge -\sqrt{y}\} \cap \{X < 0\}\right) + \mathrm{P}\left(\{X \le \sqrt{y}\} \cap \{X \ge 0\}\right) \\
=& \mathrm{P}_X(-\sqrt{y} \le X < 0) + \mathrm{P}_X(0 \le X \le \sqrt{y}) \\
=& \left[F_X(0) - F_X(-\sqrt{y})\right] + \left[F_X(\sqrt{y}) - F_X(0)\right] \\
=& F_X(\sqrt{y}) - F_X(-\sqrt{y})
\end{aligned}$$

$\square$

Another example, $\begin{cases} Y = g(X) = X^2 \\ X \in [-1, \infty) \end{cases}$,

$$\begin{aligned}
& Y = g(X) = X^2 \\
\Rightarrow & Y = \begin{cases} X^2 = g(X) & X \ge 0 \Leftrightarrow X \in [0, +\infty) \Rightarrow \forall X_1 < X_2 \left[X_1^2 < X_2^2\right] \\ X^2 = g(X) & -1 \le X < 0 \Leftrightarrow X \in [-1, 0) \Rightarrow \forall X_1 < X_2 \left[X_1^2 > X_2^2\right] \end{cases} \\
\Rightarrow & X = \begin{cases} \sqrt{Y} = g^{-1}(Y) & X \in [0, \infty) \Rightarrow \forall X_1^2 < X_2^2 \left[X_1 < X_2\right] \Rightarrow \forall Y_1 < Y_2 \left[X_1 < X_2\right] \\ -\sqrt{Y} = g^{-1}(Y) & X \in [-1, 0) \Rightarrow \forall X_1^2 < X_2^2 \left[X_1 > X_2\right] \Rightarrow \forall Y_1 < Y_2 \left[X_1 > X_2\right] \end{cases} \\
\Rightarrow & X = \begin{cases} \sqrt{Y} = g^{-1}(Y) & Y \in [0, \infty) \Rightarrow X \in [0, \infty) \Rightarrow \forall Y_1 < Y_2 \left[g^{-1}(Y_1) < g^{-1}(Y_2)\right] \\ -\sqrt{Y} = g^{-1}(Y) & Y \in (0, 1] \Rightarrow X \in [-1, 0) \Rightarrow \forall Y_1 < Y_2 \left[g^{-1}(Y_1) > g^{-1}(Y_2)\right] \end{cases} \\
\Rightarrow & X = \begin{cases} \sqrt{Y} = g^{-1}(Y) & Y \in (1, \infty) \Rightarrow X \in (1, \infty) \Rightarrow \forall Y_1 < Y_2 \left[g^{-1}(Y_1) < g^{-1}(Y_2)\right] \\ \sqrt{Y} = g^{-1}(Y) & Y \in [0, 1] \Rightarrow X \in [0, 1] \Rightarrow \forall Y_1 < Y_2 \left[g^{-1}(Y_1) < g^{-1}(Y_2)\right] \\ -\sqrt{Y} = g^{-1}(Y) & Y \in (0, 1] \Rightarrow X \in [-1, 0) \Rightarrow \forall Y_1 < Y_2 \left[g^{-1}(Y_1) > g^{-1}(Y_2)\right] \end{cases}
\end{aligned}$$

$$
\begin{aligned}
F_Y(y) = \mathrm{P}_Y(Y \le y) = \mathrm{P}\left(X^2 \le y\right) & \begin{cases} Y = g(X) = X^2 \\ X \in [-1, \infty) \end{cases} \\
=& \mathrm{P}\left(\left\{X^2 \le y\right\} \cap (\{X < 0\} \cup \{X \ge 0\})\right) = \cdots \text{ as } \begin{cases} Y = g(X) = X^2 \\ X \in (-\infty, +\infty) \end{cases} \\
=& \mathrm{P}\left(\left\{X^2 \le y\right\} \cap \{X < 0\}\right) + \mathrm{P}\left(\left\{X^2 \le y\right\} \cap \{X \ge 0\}\right) \\
=& \mathrm{P}\left(\{-X \le \sqrt{y}\} \cap \{X < 0\}\right) + \mathrm{P}\left(\{X \le \sqrt{y}\} \cap \{X \ge 0\}\right) \\
=& \begin{cases} \mathrm{P}\left(\{-X \le \sqrt{y}\} \cap \{X < 0\} \cap \{X > 1\}\right) + \mathrm{P}\left(\{X \le \sqrt{y}\} \cap \{X \ge 0\} \cap \{X > 1\}\right) & Y \in (1, \infty) \Rightarrow X \in (1, \infty) \\ \mathrm{P}\left(\{-X \le \sqrt{y}\} \cap \{X < 0\} \cap \{X \ge -1\}\right) + \mathrm{P}\left(\{X \le \sqrt{y}\} \cap \{X \ge 0\} \cap \{X \le 1\}\right) & Y \in [0, 1] \Rightarrow \begin{cases} X \in [0, 1] \\ X \in [-1, 0) \end{cases} \end{cases} \\
=& \begin{cases} \mathrm{P}(\emptyset) + \mathrm{P}_X\left(1 < X \le \sqrt{y}\right) & Y \in (1, \infty) \Rightarrow X \in (1, \infty) \\ \mathrm{P}\left(\{X \ge -\sqrt{y}\} \cap \{X < 0\} \cap \{X \ge -1\}\right) + \mathrm{P}_X\left(0 \le X \le \min\left\{\sqrt{y}, 1\right\}\right) & Y \in [0, 1] \Rightarrow \begin{cases} X \in [0, 1] \\ X \in [-1, 0) \end{cases} \end{cases} \\
=& \begin{cases} 0 + \left[F_X\left(\sqrt{y}\right) - F_X(1)\right] & Y \in (1, \infty) \Rightarrow X \in (1, \infty) \\ \mathrm{P}_X\left(\max\left\{-1, -\sqrt{y}\right\} \le X < 0\right) + \mathrm{P}_X\left(0 \le X \le \sqrt{y}\right) & Y \in [0, 1] \Rightarrow \begin{cases} X \in [0, 1] \\ X \in [-1, 0) \end{cases} \end{cases} \\
=& \begin{cases} F_X\left(\sqrt{y}\right) - F_X(1) & Y \in (1, \infty) \Rightarrow X \in (1, \infty) \\ \mathrm{P}_X\left(-\sqrt{y} \le X < 0\right) + \left[F_X\left(\sqrt{y}\right) - F_X(0)\right] & Y \in [0, 1] \Rightarrow \begin{cases} X \in [0, 1] \\ X \in [-1, 0) \end{cases} \end{cases} \\
=& \begin{cases} F_X\left(\sqrt{y}\right) - F_X(1) & y > 1 \\ F_X\left(\sqrt{y}\right) - F_X\left(-\sqrt{y}\right) & -1 \le y \le 1 \end{cases}
\end{aligned}
$$

□

15.1.2.5 discrete monotone transformation

$$
\begin{cases} Y = g(X) = X^2 \\ X \text{ discrete} \Rightarrow \qquad Y \text{ discrete} \end{cases}
$$

$$
\begin{aligned}
f_Y(y) =& \mathrm{P}_Y(Y = y) \\
=& \mathrm{P}\left(X^2 = y\right) \\
=& \mathrm{P}\left(\{X = \sqrt{y}\} \cup \{X = -\sqrt{y}\}\right) \\
=& \mathrm{P}\left(\{X = \sqrt{y}\}\right) + \mathrm{P}\left(\{X = -\sqrt{y}\}\right) - \mathrm{P}\left(\{X = \sqrt{y}\} \cap \{X = -\sqrt{y}\}\right) \\
=& \mathrm{P}_X\left(X = \sqrt{y}\right) + \mathrm{P}_X\left(X = -\sqrt{y}\right) - \mathrm{P}(\emptyset) \\
=& \mathrm{P}_X\left(X = \sqrt{y}\right) + \mathrm{P}_X\left(X = -\sqrt{y}\right) - 0 \\
=& \mathrm{P}_X\left(X = \sqrt{y}\right) + \mathrm{P}_X\left(X = -\sqrt{y}\right) \\
=& f_X\left(\sqrt{y}\right) + f_X\left(-\sqrt{y}\right)
\end{aligned}
$$

□

Theorem 15.4. *discrete monotone transformation*

$$
\begin{gathered}
\begin{cases} Y = g(X) \\ X \text{ discrete} \Rightarrow \quad Y \text{ discrete} \end{cases} \\
\Downarrow \\
f_Y(y) = \sum_{\{x \mid g(x) = y\}} f_X(x) = \sum_{\{x \mid x = g^{-1}(y)\}} f_X(x)
\end{gathered}
$$

Proof:

$$\begin{aligned} f_Y(y) =& \mathrm{P}_Y(Y=y) \\ =& \mathrm{P}(g(X)=y) = \sum_{t\in\{x|g(x)=y\}} f_X(t) = \sum_{x\in\{x|g(x)=y\}} f_X(x) = \sum_{\{x|g(x)=y\}} f_X(x) \\ =& \mathrm{P}_X\left(X=g^{-1}(y)\right) = \sum_{t\in\{x|x=g^{-1}(y)\}} f_X(t) = \sum_{x\in\{x|x=g^{-1}(y)\}} f_X(x) = \sum_{\{x|x=g^{-1}(y)\}} f_X(x) \end{aligned}$$

$$\begin{array}{ccccccccc} f_Y(y) & = & \mathrm{P}_Y(Y=y) & & & & & & \\ & = & \mathrm{P}(g(X)=y) & = & \sum\limits_{t\in\{x|g(x)=y\}} f_X(t) & = & \sum\limits_{x\in\{x|g(x)=y\}} f_X(x) & = & \sum\limits_{\{x|g(x)=y\}} f_X(x) \\ & = & \mathrm{P}_X\left(X=g^{-1}(y)\right) & = & \sum\limits_{t\in\{x|x=g^{-1}(y)\}} f_X(t) & = & \sum\limits_{x\in\{x|x=g^{-1}(y)\}} f_X(x) & = & \sum\limits_{\{x|x=g^{-1}(y)\}} f_X(x) \end{array}$$

$\square$

Theorem 15.5. *probability integral transformation*

$$\begin{gathered} \begin{cases} \begin{cases} X \text{ continuous} & (c) \\ X \sim F_X(x) & (d) \end{cases} \\ Y = F_X(X) & (t) \end{cases} \\ \Downarrow \\ F_Y(y) = y, \forall y \in [0,1] \\ \Updownarrow \text{def.} \\ Y \sim U = U(y) \Leftrightarrow Y \sim U(y) \Leftrightarrow Y \text{ is uniformly distributed on } [0,1] \end{gathered}$$

Proof:

$$\begin{aligned} F_Y(y) =& \mathrm{P}_Y(Y\le y) \overset{(t)}{=} \mathrm{P}(F_X(X)\le y), \forall x_1<x_2\left[F_X(x_1)<F_X(x_2)\right] \Rightarrow \begin{cases} \exists F_X^{-1}: Y\to X \\ \forall y_1<y_2\left[F_X^{-1}(y_1)<F_X^{-1}(y_2)\right] \end{cases} \\ =& \mathrm{P}_X\left(X\le F_X^{-1}(y)=x\right) = \mathrm{P}_X(X\le x), x=F_X^{-1}(y) \\ =& \mathrm{P}_X(X\le x) = F_X(x) \overset{x=F_X^{-1}(y)}{=} F_X\left(F_X^{-1}(y)\right) = y \\ F_Y(y) =& y \end{aligned}$$

$\square$

Note:

According to Theorem 15.5,

$$\begin{aligned} & U = F_X(X) \overset{15.5}{\sim} U(u) \text{ on } [0,1] \\ \Rightarrow & X = F_X^{-1}(U) \wedge X \sim F_X(x) \Rightarrow F_X^{-1}(U) = X \sim F_X(x) \Rightarrow F_X^{-1}(U) \sim F_X(x) \\ \Rightarrow & X = F_X^{-1}(U) \sim F_X(x) \end{aligned}$$

i.e. uniform random variables subsituted into the inverse of F_X,
we can get random variables follow $F_X(x)$

15.1.2.6 expected value

$$\mathrm{E}(g(X)) = \mathrm{E}[g(X)] = \mathrm{E}g(X) = \mathbb{E}[g(X)] = \mathbb{E}g(X)$$

Definition 15.12. expected value: The expected value of a random variable $g(X)$ is

$$\mathrm{E}(g(X)) = \mathrm{E}[g(X)] = \begin{cases} \int\limits_{-\infty}^{+\infty} g(x) f_X(x)\,\mathrm{d}x & X \text{ continuous} \\ \sum\limits_{x\in X(\Omega)} g(x) f_X(x) & X \text{ discrete} \end{cases}$$

[6] p.126

Definition 15.13. expected value or expectation function: The expected value of a random variable X is

$$\mathrm{E}(X) = \mathrm{E}[X] = \begin{cases} \int\limits_{-\infty}^{+\infty} x f_X(x)\,\mathrm{d}x = 1 & X \text{ continuous} \\ \sum\limits_{x \in X(\Omega)} x f_X(x) = 1 & X \text{ discrete} \end{cases}$$

Theorem 15.6. *the rule of the lazy statistician*

the law of the unconscious statistician = the LOTUS

$$\mathrm{E}(g(X)) = \mathrm{E}[g(X)] = \begin{cases} \int\limits_{-\infty}^{+\infty} g(x) f_X(x)\,\mathrm{d}x & X \text{ continuous} \\ \sum\limits_{x \in X(\Omega)} g(x) f_X(x) & X \text{ discrete} \end{cases}$$

Proof:[7] p.162 for p.119

Discrete case:

to be proved

Continuous case:

to be proved

□

By linearity of $\int$ and $\sum$, expected values have the following properties or theorems,

- $\mathrm{E}[a_1 g_1(X_1) + a_2 g_2(X_2) + c] = a_1 \mathrm{E}[g_1(X_1)] + a_2 \mathrm{E}[g_2(X_2)] + c$
- $\forall x \in \mathbb{R}\,[g(x) \geq 0] \Rightarrow \mathrm{E}[g(X)] \geq 0$
- $\forall x \in \mathbb{R}\,[g_1(x) \geq g_2(x)] \Rightarrow \mathrm{E}[g_1(X)] \geq \mathrm{E}[g_2(X)]$
- $\forall x \in \mathbb{R}\,[a \leq g(x) \leq b] \Rightarrow a \leq \mathrm{E}[g(X)] \leq b$

Theorem 15.7. $\mathrm{E}[X]$ *minimizes Euclidean distance* $\mathrm{E}\left[(X-b)^2\right]$ *over* b*, i.e.*

$$\mathrm{E}[X] = \arg\min_b \mathrm{E}\left[(X-b)^2\right]$$

Proof:

$$
\begin{aligned}
\mathrm{E}\left[(X-b)^2\right] =& \mathrm{E}\left[(X-\mathrm{E}[X]+\mathrm{E}[X]-b)^2\right]\\
=& \mathrm{E}\left[\{(X-\mathrm{E}[X])+(\mathrm{E}[X]-b)\}^2\right]\\
=& \mathrm{E}\left[(X-\mathrm{E}[X])^2+2(X-\mathrm{E}[X])(\mathrm{E}[X]-b)+(\mathrm{E}[X]-b)^2\right]\\
=& \mathrm{E}\left[(X-\mathrm{E}[X])^2\right]+2(\mathrm{E}[X]-b)\,\mathrm{E}[(X-\mathrm{E}[X])]+\mathrm{E}\left[(\mathrm{E}[X]-b)^2\right]\\
=& \mathrm{E}\left[(X-\mathrm{E}[X])^2\right]+2(\mathrm{E}[X]-b)\,\mathrm{E}[X-\mathrm{E}[X]]+(\mathrm{E}[X]-b)^2\\
=& \mathrm{E}\left[(X-\mathrm{E}[X])^2\right]+2(\mathrm{E}[X]-b)(\mathrm{E}[X]-\mathrm{E}[X])+(\mathrm{E}[X]-b)^2\\
=& \mathrm{E}\left[(X-\mathrm{E}[X])^2\right]+2(\mathrm{E}[X]-b)\,0+(\mathrm{E}[X]-b)^2\\
=& \mathrm{E}\left[(X-\mathrm{E}[X])^2\right]+0+(\mathrm{E}[X]-b)^2\\
=& \mathrm{E}\left[(X-\mathrm{E}[X])^2\right]+(\mathrm{E}[X]-b)^2 \overset{(\mathrm{E}[X]-b)^2\geq 0}{\geq} \mathrm{E}\left[(X-\mathrm{E}[X])^2\right]\\
\mathrm{E}\left[(X-b)^2\right] \geq& \mathrm{E}\left[(X-\mathrm{E}[X])^2\right]\\
&\Downarrow\\
\mathrm{E}\left[(X-b)^2\right] =& \mathrm{E}\left[(X-\mathrm{E}[X])^2\right] \text{ holds if } (\mathrm{E}[X]-b)^2=0 \Rightarrow b=\mathrm{E}[X] \Rightarrow \mathrm{E}[X]=\arg\min_b \mathrm{E}\left[(X-b)^2\right]
\end{aligned}
$$

□

Note:

When $b=\mathrm{E}[X]$, $\mathrm{E}\left[(X-b)^2\right]$ has minimum loss $\mathrm{E}\left[(X-\mathrm{E}[X])^2\right]=\mathrm{V}[X]=\mathrm{V}(X)$, i.e. defintion of variance appears.

Theorem 15.8. median $[X]$ *minimizes* $\mathrm{E}[|X-b|]$ *over* b*, i.e.*

$$\text{median}[X]=\arg\min_b \mathrm{E}[|X-b|]$$

Proof:

to be proved

□

Note:

When $b=\text{median}[X]$, $\mathrm{E}\left[(X-b)^2\right]$ has minimum loss $\mathrm{E}[|X-\text{median}[X]|]$, i.e. defintion of MAD(mean absolute deviation) in robust statistics appears.

Definition 15.14. indicator function

$$
\begin{aligned}
1(E)=1(x\in E)=1(\{x\in E\})=1(\{x|x\in E\})=&\begin{cases}1 & E\\ 0 & \overline{E}\end{cases}=\begin{cases}1 & \text{if } E\\ 0 & \text{if } \overline{E}=E^{\mathrm{C}}\end{cases}\\
=&\begin{cases}1 & \text{if event } E \text{ occurs}\\ 0 & \text{if event } E \text{ does not occur}\end{cases}
\end{aligned}
$$

Note:

Theorem 15.9. *probability as expected value*

$$\mathrm{P}_X(E) = \mathrm{P}(x \in E) = \mathrm{E}[1(X \in E)]$$

Proof:

$$\begin{aligned}
\mathrm{P}_X(E) = \mathrm{P}(x \in E) &= \int_{x \in E} f_X(x)\,\mathrm{d}x = \int_E f_X(x)\,\mathrm{d}x \\
&= \int 1(x \in E) f_X(x)\,\mathrm{d}x \\
&= \int g(x) f_X(x)\,\mathrm{d}x, g(x) = 1(x \in E) \\
&= \mathrm{E}[g(X)], g(X) = 1(X \in E) \\
&= \mathrm{E}[1(X \in E)] \\
\mathrm{P}_X(E) = \mathrm{P}(x \in E) &= \mathrm{E}[1(X \in E)]
\end{aligned}$$

$\square$

Iverson bracket https://en.wikipedia.org/wiki/Iverson_bracket

$$\begin{cases} v(p(x)) = \mathrm{T} & \Leftrightarrow [p(x)] = 1 \\ v(p(x)) = \mathrm{F} & \Leftrightarrow [p(x)] = 0 \end{cases}$$

$$[p(x)] = \begin{cases} 1 & v(p(x)) = \mathrm{T} \\ 0 & v(\neg p(x)) = \mathrm{T} \end{cases} = \begin{cases} 1 & p(x) \\ 0 & \neg p(x) \end{cases}$$

negation = NOT

$$[\neg p] = 1 - [p]$$

in set theory or domain of events,

$$1\left(\overline{E}\right) = 1 - 1(E)$$

conjunction = AND

$$[p \wedge q] = [p][q]$$

in set theory or domain of events,

$$1(E_1 \cap E_2) = 1(E_1)\,1(E_2)$$

disjunction = OR

$$[p \vee q] = [p] + [q] - [p][q] = [p] + [q] - [p \wedge q]$$

Proof:

in set theory or domain of events,

$$\begin{aligned}
1(E_1 \cup E_2) &\overset{\text{de Moivre}}{=} 1\left(\overline{\overline{E}_1 \cap \overline{E}_2}\right) \\
&= 1 - 1\left(\overline{E}_1 \cap \overline{E}_2\right) = 1 - 1\left(\overline{E}_1\right) 1\left(\overline{E}_2\right) \\
&= 1 - [1 - 1(E_1)][1 - 1(E_2)] \\
&= 1 - [1 - 1(E_1)][1 - 1(E_2)] \\
&= 1 - [1 - 1(E_1) - 1(E_2) + 1(E_1)\,1(E_2)] \\
&= 1(E_1) + 1(E_2) - 1(E_1)\,1(E_2) \\
&= 1(E_1) + 1(E_2) - 1(E_1 \cap E_2)
\end{aligned}$$

$$1\left(E_1 \cup E_2\right) = 1\left(E_1\right) + 1\left(E_2\right) - 1\left(E_1\right)1\left(E_2\right) = 1\left(E_1\right) + 1\left(E_2\right) - 1\left(E_1 \cap E_2\right)$$

$\square$

implication = conditional

$$\begin{aligned}[p \to q] &= [\neg p \vee q] \\ &= [\neg p] + [q] - [\neg p]\,[q] \\ &= 1 - [p] + [q] - (1 - [p])\,[q] \\ &= 1 - [p] + [p]\,[q]\end{aligned}$$

exclusive disjunction = XOR

$$\begin{aligned}[p \veebar q] = [p \oplus q] &= |[p] - [q]| = ([p] - [q])^2 \\ &= [p]\,(1 - [q]) + (1 - [p])\,[q]\end{aligned}$$

biconditional = XNOR

$$[p \leftrightarrow q] = [p \odot q] = [\neg\,(p \oplus q)] = [\neg\,(p \veebar q)] = ([p] + (1 - [q]))\,((1 - [p]) + [q])$$

Kronecker delta

$$\delta_{ij} = [i = j]$$

single-argument notation

$$\delta_i = \delta_{i0} = \begin{cases} 1 & i = j = 0 \\ 0 & i \neq j = 0 \end{cases}$$

sign function

$$\operatorname{sgn}(x) = \begin{cases} 1 & x > 0 \\ 0 & x = 0 = [x > 0] - [x < 0] \\ -1 & x < 0 \end{cases}$$

absolute function

$$\begin{aligned}|x| &= \begin{cases} x & x \geq 0 \\ -x & x < 0 \end{cases} = \begin{cases} x & x > 0 \\ -x & x \leq 0 \end{cases} = \begin{cases} x & x > 0 \\ 0 & x = 0 \\ -x & x < 0 \end{cases} \\ &= \begin{cases} x \cdot 1 & x > 0 \\ x \cdot 0 & x = 0 \\ x \cdot (-1) & x < 0 \end{cases} = \begin{cases} x \cdot \operatorname{sgn}(x) & x > 0 \\ x \cdot \operatorname{sgn}(x) & x = 0 \\ x \cdot \operatorname{sgn}(x) & x < 0 \end{cases} \\ &= x \cdot \operatorname{sgn}(x) = x\,([x > 0] - [x < 0]) = x\,[x > 0] - x\,[x < 0]\end{aligned}$$

binary min and max function

$$\max(x, y) = x\,[x > y] + y\,[x \leq y]$$

$$\min(x, y) = x\,[x \leq y] + y\,[x > y]$$

binary max function

$$\max\left(x,y\right)=\frac{x+y+\left|x-y\right|}{2}$$

floor and ceiling functions

floor function

$$\begin{aligned}\left\lfloor x\right\rfloor =&n,\, n\leq x<n+1\\=&\sum_{n\in\mathbb{N}}n\left[n\leq x<n+1\right]\end{aligned}$$

ceiling function

$$\begin{aligned}\left\lceil x\right\rceil =&n,\, n-1<x\leq n\\=&\sum_{n\in\mathbb{N}}n\left[n-1<x\leq n\right]\end{aligned}$$

Heaviside step function

$$\mathrm{H}\left(x\right)=\begin{cases}1 & x>0\\0 & x\leq0\end{cases}=\left[x>0\right]=1_{\left(0,\infty\right)}\left(x\right)$$

or conveniently define “unit step function”

$$u\left(x\right)=\begin{cases}1 & x\geq0\\0 & x<0\end{cases}=\left[x\geq0\right]=1_{\left[0,\infty\right)}\left(x\right)$$

ramp function = rectified linear unit activation function = ReLU

$$\mathrm{ReLU}\left(x\right)=\begin{cases}x & x\geq0\\0 & x<0\end{cases}=x\left[x\geq0\right]$$

indicator function

$$A\subseteq X\Rightarrow\begin{cases}1_{A}:X\rightarrow\left\{0,1\right\} & \Leftrightarrow x\in X\overset{1_{A}}{\rightarrow}\left\{0,1\right\}\\1_{A}\left(x\right)=\begin{cases}1 & x\in A\\0 & x\notin A\end{cases} & =\left[x\in A\right]=\begin{cases}1 & v\left(x\in A\right)=\mathrm{T}\\0 & v\left(\neg\left(x\in A\right)\right)=\mathrm{T}\end{cases}\end{cases}$$

$A,B\subseteq\Omega$,

$$A=B\Leftrightarrow1_{A}=1_{B}$$

$$A=\Omega\Leftrightarrow1_{A}\left(x\right)=1$$

$$A=\emptyset\Leftrightarrow1_{A}\left(x\right)=0$$

Theorem 15.10. *subset indicator order*

$$A\subset B\Rightarrow1_{A}\left(x\right)\leq1_{B}\left(x\right)$$

Proof:

$$\begin{aligned}
&\forall x\left(1_A(x)=1\Rightarrow 1_B(x)=1\right)\\
\Leftrightarrow&\forall x\left(\neg 1_A(x)=1\vee 1_B(x)=1\right)\\
\Leftrightarrow&\forall x\neg\left(1_A(x)=1\wedge\neg 1_B(x)=1\right)\\
\Rightarrow&\neg\exists x\left(1_A(x)=1\wedge 1_B(x)=0\right)\\
\Rightarrow&\neg\exists x\left(1_B(x)=0<1=1_A(x)\right)\\
\Rightarrow&\neg\exists x\left(1_B(x)<1_A(x)\right)\\
\Rightarrow&\forall x\left(1_B(x)\geq 1_A(x)\right)
\end{aligned}$$

□

---

in set theory or domain of events,

$$1\left(E_1\cap E_2\right)=1\left(E_1\right)1\left(E_2\right)$$

$$1\left(\overline{E}\right)=1-1\left(E\right)$$

$$1\left(E_1\cup E_2\right)=1\left(E_1\right)+1\left(E_2\right)-1\left(E_1\right)1\left(E_2\right)=1\left(E_1\right)+1\left(E_2\right)-1\left(E_1\cap E_2\right)$$

---

expectation in many perspectives

$$Y=g\left(X\right)$$

$$\int_{-\infty}^{+\infty}y\,f_Y\left(y\right)\mathrm{d}y=\mathrm{E}\left[Y\right]=\mathrm{E}\left[g\left(X\right)\right]=\int_{-\infty}^{+\infty}g\left(x\right)f_X\left(x\right)\mathrm{d}x$$

$$\mathrm{E}_Y\left[Y\right]=\int_{-\infty}^{+\infty}y\,f_Y\left(y\right)\mathrm{d}y=\mathrm{E}\left[Y\right]=\mathrm{E}\left[g\left(X\right)\right]=\int_{-\infty}^{+\infty}g\left(x\right)f_X\left(x\right)\mathrm{d}x=\mathrm{E}_X\left[g\left(X\right)\right]$$

$$\mathrm{E}_Y\left[Y\right]=\mathrm{E}_X\left[g\left(X\right)\right]$$

$$\begin{aligned}
&\mathrm{E}\left[a_1g_1\left(X_1\right)+a_2g_2\left(X_2\right)+c\right],\begin{cases}Y_1=g_1\left(X_1\right)\\Y_2=g_2\left(X_2\right)\end{cases}\\
=&\mathrm{E}\left[a_1Y_1+a_2Y_2+c\right]
\end{aligned}$$

$$\begin{aligned}
&\mathrm{E}\left[a_1g_1\left(X_1\right)+a_2g_2\left(X_2\right)+c\right]=a_1\mathrm{E}\left[g_1\left(X_1\right)\right]+a_2\mathrm{E}\left[g_2\left(X_2\right)\right]+c\\
=&\mathrm{E}\left[a_1Y_1+a_2Y_2+c\right]=a_1\mathrm{E}\left[Y_1\right]+a_2\mathrm{E}\left[Y_2\right]+c
\end{aligned}$$

$$a_1\mathrm{E}\left[g_1\left(X_1\right)\right]+a_2\mathrm{E}\left[g_2\left(X_2\right)\right]+c=a_1\mathrm{E}_{X_1}\left[g_1\left(X_1\right)\right]+a_2\mathrm{E}_{X_2}\left[g_2\left(X_2\right)\right]+c$$

$$a_1\mathrm{E}\left[Y_1\right]+a_2\mathrm{E}\left[Y_2\right]+c=a_1\mathrm{E}_{Y_1}\left[Y_1\right]+a_2\mathrm{E}_{Y_1}\left[Y_2\right]+c$$

$$a_1\mathrm{E}_{X_1}\left[g_1\left(X_1\right)\right]+a_2\mathrm{E}_{X_2}\left[g_2\left(X_2\right)\right]+c=a_1\mathrm{E}_{Y_1}\left[Y_1\right]+a_2\mathrm{E}_{Y_1}\left[Y_2\right]+c$$

15.1.2.7 moment

Definition 15.15. n^{th} moment: For each integer n, the n^{th} moment of X is $\mathrm{E}[X^n]$.

The n^{th} central moment of X is $\mu_n = \mathrm{E}[(X - \mathrm{E}[X])^n]$.

$$\mathrm{E}[X^n] = \begin{cases} \int\limits_{-\infty}^{+\infty} x^n f_X(x)\,\mathrm{d}x & X \text{ continuous} \\ \sum\limits_{x \in X(\Omega)} x^n f_X(x) & X \text{ discrete} \end{cases}$$

$$\mu = \mathrm{E}\left[X^1\right] = \mathrm{E}[X] = \begin{cases} \int\limits_{-\infty}^{+\infty} x^1 f_X(x)\,\mathrm{d}x & X \text{ continuous} \\ \sum\limits_{x \in X(\Omega)} x^1 f_X(x) & X \text{ discrete} \end{cases} = \begin{cases} \int\limits_{-\infty}^{+\infty} x\, f_X(x)\,\mathrm{d}x & X \text{ continuous} \\ \sum\limits_{x \in X(\Omega)} x\, f_X(x) & X \text{ discrete} \end{cases}$$

$$\mu_n = \mathrm{E}[(X - \mathrm{E}[X])^n] = \begin{cases} \int\limits_{-\infty}^{+\infty} (x-\mu)^n f_X(x)\,\mathrm{d}x & X \text{ continuous} \\ \sum\limits_{x \in X(\Omega)} (x-\mu)^n f_X(x) & X \text{ discrete} \end{cases}$$

1^{st} moment of X = mean

$$\mu = \mathrm{E}\left[X^1\right] = \mathrm{E}[X] = \begin{cases} \int\limits_{-\infty}^{+\infty} x^1 f_X(x)\,\mathrm{d}x & X \text{ continuous} \\ \sum\limits_{x \in X(\Omega)} x^1 f_X(x) & X \text{ discrete} \end{cases} = \begin{cases} \int\limits_{-\infty}^{+\infty} x\, f_X(x)\,\mathrm{d}x & X \text{ continuous} \\ \sum\limits_{x \in X(\Omega)} x\, f_X(x) & X \text{ discrete} \end{cases}$$

1^{st} central moment of $X = 0$

$$\begin{aligned}
\mu_1 = \mathrm{E}\left[(X - \mathrm{E}[X])^1\right] &= \begin{cases} \int\limits_{-\infty}^{+\infty} (x-\mu)^1 f_X(x)\,\mathrm{d}x & X \text{ continuous} \\ \sum\limits_{x \in X(\Omega)} (x-\mu)^1 f_X(x) & X \text{ discrete} \end{cases} = \begin{cases} \int\limits_{-\infty}^{+\infty} (x-\mu) f_X(x)\,\mathrm{d}x & X \text{ continuous} \\ \sum\limits_{x \in X(\Omega)} (x-\mu) f_X(x) & X \text{ discrete} \end{cases} \\
&= \begin{cases} \int\limits_{-\infty}^{+\infty} x\, f_X(x)\,\mathrm{d}x - \int\limits_{-\infty}^{+\infty} \mu\, f_X(x)\,\mathrm{d}x & X \text{ continuous} \\ \sum\limits_{x \in X(\Omega)} x\, f_X(x) - \sum\limits_{x \in X(\Omega)} \mu\, f_X(x) & X \text{ discrete} \end{cases} \\
&= \begin{cases} \mathrm{E}[X] - \mu \int\limits_{-\infty}^{+\infty} f_X(x)\,\mathrm{d}x & X \text{ continuous} \\ \mathrm{E}[X] - \mu \sum\limits_{x \in X(\Omega)} f_X(x) & X \text{ discrete} \end{cases} = \begin{cases} \mathrm{E}[X] - \mu \cdot 1 & X \text{ continuous} \\ \mathrm{E}[X] - \mu \cdot 1 & X \text{ discrete} \end{cases} \\
&= \begin{cases} \mathrm{E}[X] - \mu & X \text{ continuous} \\ \mathrm{E}[X] - \mu & X \text{ discrete} \end{cases} = \begin{cases} \mathrm{E}[X] - \mathrm{E}[X] & X \text{ continuous} \\ \mathrm{E}[X] - \mathrm{E}[X] & X \text{ discrete} \end{cases} \\
&= \begin{cases} 0 & X \text{ continuous} \\ 0 & X \text{ discrete} \end{cases} = 0 \\
\mathrm{E}[(X - \mathrm{E}[X])] &= 0 \\
\mathrm{E}[X - \mathrm{E}[X]] &= 0
\end{aligned}$$

$$\forall X\, (\mathrm{E}[X - \mathrm{E}[X]] = 0)$$

For normal distribution, actually for any distribution,

$$\begin{gathered} X \sim \mathrm{n}(0,1) = \mathcal{N}\left(0, 1^2\right) \\ \Downarrow \\ \mathrm{E}[X - \mathrm{E}[X]] = 0 \end{gathered}$$

2^{nd} central moment of X = variance

$$\mu_2 = \mathrm{E}\left[\left(X - \mathrm{E}\left[X\right]\right)^2\right] = \begin{cases} \int\limits_{-\infty}^{+\infty} \left(x-\mu\right)^2 f_X\left(x\right) \mathrm{d}x & X \text{ continuous} \\ \sum\limits_{x \in X(\Omega)} \left(x-\mu\right)^2 f_X\left(x\right) & X \text{ discrete} \end{cases}$$
$$= \mathrm{E}\left[\left(X - \mathrm{E}\left[X\right]\right)^2\right] = \mathrm{V}\left[X\right] = \mathrm{V}\left(X\right)$$

For normal distribution,

$$X \sim \mathrm{n}\left(0,1\right) = \mathcal{N}\left(0,1^2\right) = \mathcal{N}\left(\mu=0, \mathrm{V}^2\left[X\right]=1^2\right)$$
$$\Downarrow$$
$$\mathrm{V}\left[X\right] = \mathrm{V}\left(X\right) = 1$$

variance properties

$$\mathrm{V}\left[aX+b\right] = a^2 \mathrm{V}\left[X\right]$$

Proof:

to be proved

□

3^{rd} central moment of X

$$\mu_3 = \mathrm{E}\left[\left(X - \mathrm{E}\left[X\right]\right)^3\right] = \begin{cases} \int\limits_{-\infty}^{+\infty} \left(x-\mu\right)^3 f_X\left(x\right) \mathrm{d}x & X \text{ continuous} \\ \sum\limits_{x \in X(\Omega)} \left(x-\mu\right)^3 f_X\left(x\right) & X \text{ discrete} \end{cases}$$

skewness

偏度

$$\text{skewness}\left[X\right] = \frac{\mu_3}{\mu_2^{\frac{3}{2}}} = \frac{\mathrm{E}\left[\left(X - \mathrm{E}\left[X\right]\right)^3\right]}{\left(\mathrm{V}\left[X\right]\right)^{\frac{3}{2}}} = \frac{\mathrm{E}\left[\left(X - \mathrm{E}\left[X\right]\right)^3\right]}{\left(\mathrm{E}\left[\left(X - \mathrm{E}\left[X\right]\right)^2\right]\right)^{\frac{3}{2}}}$$
$$= \begin{cases} \dfrac{\int\limits_{-\infty}^{+\infty} \left(x-\mu\right)^3 f_X\left(x\right) \mathrm{d}x}{\left(\int\limits_{-\infty}^{+\infty} \left(x-\mu\right)^2 f_X\left(x\right) \mathrm{d}x\right)^{\frac{3}{2}}} & X \text{ continuous} \\ \dfrac{\sum\limits_{x \in X(\Omega)} \left(x-\mu\right)^3 f_X\left(x\right)}{\left(\sum\limits_{x \in X(\Omega)} \left(x-\mu\right)^2 f_X\left(x\right)\right)^{\frac{3}{2}}} & X \text{ discrete} \end{cases}$$

For normal distribution,

$$X \sim \mathrm{n}\left(0,1\right) = \mathcal{N}\left(0,1^2\right) = \mathcal{N}\left(\mu=0, \mathrm{V}^2\left[X\right]=1^2\right)$$
$$\Downarrow$$
$$\text{skewness}\left[X\right] = \frac{\mathrm{E}\left[\left(X - \mathrm{E}\left[X\right]\right)^3\right]}{\left(\mathrm{V}\left[X\right]\right)^{\frac{3}{2}}} = \frac{\mathrm{E}\left[\left(X - \mathrm{E}\left[X\right]\right)^3\right]}{1^{\frac{3}{2}}} = 0$$

Proof:

to be proved

□

4^{th} central moment of X

$$\mu_4 = \mathrm{E}\left[(X - \mathrm{E}[X])^4\right] = \begin{cases} \int\limits_{-\infty}^{+\infty} (x-\mu)^4 f_X(x)\,\mathrm{d}x & X \text{ continuous} \\ \sum\limits_{x \in X(\Omega)} (x-\mu)^4 f_X(x) & X \text{ discrete} \end{cases}$$

kurtosis

峰度

$$\begin{aligned} \mathrm{kurtosis}[X] =& \frac{\mu_4}{\mu_2^2} = \frac{\mathrm{E}\left[(X-\mathrm{E}[X])^4\right]}{(\mathrm{V}[X])^2} = \frac{\mathrm{E}\left[(X-\mathrm{E}[X])^4\right]}{\left(\mathrm{E}\left[(X-\mathrm{E}[X])^2\right]\right)^2} \\ =& \begin{cases} \dfrac{\int\limits_{-\infty}^{+\infty} (x-\mu)^4 f_X(x)\,\mathrm{d}x}{\left(\int\limits_{-\infty}^{+\infty} (x-\mu)^2 f_X(x)\,\mathrm{d}x\right)^2} & X \text{ continuous} \\ \dfrac{\sum\limits_{x \in X(\Omega)} (x-\mu)^4 f_X(x)}{\left(\sum\limits_{x \in X(\Omega)} (x-\mu)^2 f_X(x)\right)^2} & X \text{ discrete} \end{cases} \end{aligned}$$

For normal distribution,

$$\begin{gathered} X \sim \mathrm{n}(0,1) = \mathcal{N}\left(0,1^2\right) = \mathcal{N}\left(\mu = 0, \mathrm{V}^2[X] = 1^2\right) \\ \Downarrow \\ \mathrm{kurtosis}[X] = \frac{\mathrm{E}\left[(X-\mathrm{E}[X])^4\right]}{(\mathrm{V}[X])^2} = \frac{\left[(X-\mathrm{E}[X])^4\right]}{1^2} = 3 \end{gathered}$$

Proof:

to be proved

□

For normal distribution,

$$\begin{gathered} X \sim \mathrm{n}(0,1) = \mathcal{N}\left(0,1^2\right) = \mathcal{N}\left(\mu = 0, \mathrm{V}^2[X] = 1^2\right) \\ \Downarrow \\ \begin{cases} \mu = \mathrm{E}[X] & = 0 \\ \mu_1 = \mathrm{E}[X - \mathrm{E}[X]] & = 0 \\ \mathrm{variance}[X] = \mathrm{V}[X] = \mathrm{E}\left[(X-\mathrm{E}[X])^2\right] & = 1 \\ \mathrm{skewness}[X] = \dfrac{\mathrm{E}\left[(X-\mathrm{E}[X])^3\right]}{(\mathrm{V}[X])^{\frac{3}{2}}} & = 0 \\ \mathrm{kurtosis}[X] = \dfrac{\mathrm{E}\left[(X-\mathrm{E}[X])^4\right]}{(\mathrm{V}[X])^2} & = 3 \end{cases} \end{gathered}$$

$$\begin{aligned}
\mu =& \mathrm{E}[X] = 0\\
\mu_1 =& \mathrm{E}[X - \mathrm{E}[X]] = 0\\
\text{variance}[X] =& \mathrm{V}[X] = \mathrm{E}\left[(X - \mathrm{E}[X])^2\right] = 1\\
\text{skewness}[X] =& \frac{\mathrm{E}\left[(X - \mathrm{E}[X])^3\right]}{(\mathrm{V}[X])^{\frac{3}{2}}} = 0\\
\text{kurtosis}[X] =& \frac{\mathrm{E}\left[(X - \mathrm{E}[X])^4\right]}{(\mathrm{V}[X])^2} = 3
\end{aligned}$$

$$X \sim F_X(x) \leftrightarrow f_X(x) \to \{\mu_n | n \in \mathbb{N}\} = \left\{\mu_n \middle| \begin{cases} n \in \mathbb{N} \\ \mu_n = \mathrm{E}[(X - \mathrm{E}[X])^n] \end{cases} \right\}$$

15.1.2.7.1 moment generating function

Definition 15.16. MGF = moment generating function: The moment generating function of X is $M(\xi) = M_X(\xi) = \mathrm{E}\left[\mathrm{e}^{\xi X}\right]$, provided that the expression exists for $t \approx 0$.

$$M(t) = M_X(t) = \mathrm{E}\left[\mathrm{e}^{tX}\right]$$

$$M(\xi) = M_X(\xi) = \mathrm{E}\left[\mathrm{e}^{\xi X}\right]$$

Theorem 15.11. *moment generating function(MGF) generating moment*

$$M_X^{(n)}(\xi) = \mathrm{E}[X^n]$$

where

$$M_X^{(n)}(\xi) = \frac{\mathrm{d}^n}{\mathrm{d}\xi^n} M_X(\xi)$$

Proof:

to be proved

□

$$X \sim F_X(x) \leftrightarrow f_X(x) \to \{\mu_n | n \in \mathbb{N}\} = \left\{\mu_n \middle| \begin{cases} n \in \mathbb{N} \\ \mu_n = \mathrm{E}[(X - \mathrm{E}[X])^n] \end{cases} \right\}$$

$$\begin{array}{ccccccc}
X & \sim & F_X(x) & \leftrightarrow & f_X(x) & \to & \{\mu_n | n \in \mathbb{N}\} \\
 & & & & \downarrow & \nearrow & \\
 & & & & M_X(\xi) & &
\end{array}$$

$$\begin{array}{ccccccc}
X & \sim & F_X(x) & \leftrightarrow & f_X(x) & & \\
 & & & & \downarrow & \searrow & \\
 & & & & M_X(\xi) & \to & \{\mu_n | n \in \mathbb{N}\}
\end{array}$$

Theorem 15.12. *If X and Y have bounded support, then $\forall u\,[F_X(u) = F_Y(u)]$ iff $\forall n \in \mathbb{N}\,(\mathrm{E}[X^n] = \mathrm{E}[Y^n])$.*

$$\forall u\left[F_X(u)=F_Y(u)\right]\Rightarrow\forall n\in\mathbb{N}\left(\mathrm{E}\left[X^n\right]=\mathrm{E}\left[Y^n\right]\right)$$

$$\begin{cases}\forall n\in\mathbb{N}\left(\mathrm{E}\left[X^n\right]=\mathrm{E}\left[Y^n\right]\right)\\ \begin{cases}\operatorname{supp}\left(f_X\right)\text{ is bounded}\\ \operatorname{supp}\left(f_Y\right)\text{ is bounded}\end{cases}\end{cases}\Rightarrow\forall u\left[F_X(u)=F_Y(u)\right]$$

Proof:

to be proved

□

Theorem 15.13. *If $M_X(t)$ and $M_Y(t)$ exist, then $\forall u\left[F_X(u)=F_Y(u)\right]$ iff $\forall t\approx 0\left[M_X(t)=M_Y(t)\right]$.*

$$\forall u\left[F_X(u)=F_Y(u)\right]\Rightarrow\forall t\approx 0\left[M_X(t)=M_Y(t)\right]$$

$$\begin{cases}\forall t\approx 0\left[M_X(t)=M_Y(t)\right]\\ \begin{cases}\exists M_X(t)\in\mathbb{R}\\ \exists M_Y(t)\in\mathbb{R}\end{cases}\end{cases}\Rightarrow\forall u\left[F_X(u)=F_Y(u)\right]$$

Proof:

to be proved

□

$$\begin{array}{ccccccc} X & \sim & F_X(x) & \leftrightarrow & f_X(x) & & \\ & & & \looparrowright & \uparrow\downarrow & \searrow\nwarrow & \looparrowleft\wedge\operatorname{supp}\left(f_X\right)\text{ is bounded} \\ & & \forall\xi\approx 0\left[M_X(\xi)\in\mathbb{R}\right] & \wedge & M_X(\xi) & \rightarrow & \left\{\mu_n\middle|n\in\mathbb{N}\right\} \end{array}$$

15.1.2.7.2 characteristic function

Definition 15.17. CF = characteristic function: The characteristic function of X is $\varphi(\xi)=\varphi_X(\xi)=\mathrm{E}\left[\mathrm{e}^{\mathrm{i}\xi X}\right]$, provided that the expression always exists.

$$\varphi(t)=\varphi_X(t)=\mathrm{E}\left[\mathrm{e}^{\mathrm{i}tX}\right]$$

$$\varphi(\xi)=\varphi_X(\xi)=\mathrm{E}\left[\mathrm{e}^{\mathrm{i}\xi X}\right]$$

Note:

1. $\varphi(\xi)=\varphi_X(\xi)=\mathrm{E}\left[\mathrm{e}^{\mathrm{i}\xi X}\right]$ always exists.

$$\forall X\left(\varphi(\xi)=\varphi_X(\xi)=\mathrm{E}\left[\mathrm{e}^{\mathrm{i}\xi X}\right]\in\mathbb{R}\right)$$

2. moment generating function to characteristic function

$$M(\xi)=M_X(\xi)=\mathrm{E}\left[\mathrm{e}^{\xi X}\right]\in\mathbb{R}\Rightarrow M_X(\mathrm{i}\xi)=\varphi_X(\xi)$$

3. inversion theorem or inversion formula

For $a < b$,

$$\begin{aligned}&\lim_{T\to\infty}\frac{1}{2\pi}\int_{-T}^{+T}\frac{\mathrm{e}^{-ita}-\mathrm{e}^{-itb}}{\mathrm{i}t}\varphi_X(t)\,\mathrm{d}t\\ =&\mathrm{P}(a<X<b)+\frac{1}{2}\left[\mathrm{P}(X=a)+\mathrm{P}(X=b)\right]\end{aligned}$$

i.e. $\varphi_X(\xi)$ determines $F_X(x)$.

$$\begin{array}{cccccccc} X & \sim & F_X(x) & \leftrightarrow & & f_X(x) & & \\ & & \updownarrow & \forall\xi\approx 0\,[M_X(\xi)\in\mathbb{R}] \ \wedge & & \upharpoonleft\downharpoonright & \searrow\nwarrow & \hookleftarrow\!\!\!\!\!\!\not\ \ \wedge\ \mathrm{supp}\,(f_X)\ \text{ is bounded} \\ & & \varphi_X(\xi) & \leftrightarrows & & M_X(\xi) & \to & \{\mu_n | n\in\mathbb{N}\} \end{array}$$

$$\begin{array}{cccccccc} X & \sim & F_X(x) & \xleftrightarrow{\text{FToC}} & f_X(x) & \leftrightarrow & \mathrm{P}_X \\ & \text{inversion formula}\ : & \updownarrow & \forall\xi\approx 0\,[M_X(\xi)\in\mathbb{R}] \ \wedge & \upharpoonleft\downharpoonright & \searrow\nwarrow & \hookleftarrow\!\!\!\!\!\!\not\ \ \wedge\ \mathrm{supp}\,(f_X)\ \text{ is bounded} \\ & & \varphi_X(\xi) & \leftrightarrows & M_X(\xi) & \to & \{\mu_n | n\in\mathbb{N}\} \end{array}$$

MGF theorems

https://www.youtube.com/watch?v=fSbs6im6wqY&t=524s

Theorem 15.14. *If*

CLT

15.1.2.8 common families of distributions

mean and variance of discrete probability distributions[43.5]

15.1.2.8.1 discrete distribution

15.1.2.8.1.1 discrete uniform distribution

$$\begin{gathered} X\sim\mathcal{DU}(1,N)\\ \Updownarrow\\ \begin{cases} f_X(x|N)=\dfrac{1}{N}\\ x\in X(\Omega)=\{1,2,\cdots,N\}\end{cases}\\ \Downarrow\\ \begin{cases}\mathrm{E}[X]= & \dfrac{N+1}{2}\\ \mathrm{V}[X]= & \dfrac{(N+1)(N-1)}{12}\end{cases}\end{gathered}$$

$$\begin{aligned}\mathrm{E}[X]=&\sum_{x\in X(\Omega)}x\,f_X(x|N)=\sum_{x=1}^{N}x\frac{1}{N}\\ &=\frac{1}{N}\sum_{x=1}^{N}x=\frac{1}{N}\frac{N(N+1)}{2}=\frac{N+1}{2}\\ \mathrm{E}[X]=&\frac{N+1}{2}\end{aligned}$$

$$
\begin{aligned}
\mathrm{V}[X]=&\sum_{x\in X(\Omega)}(x-\mathrm{E}[X])^2 f_X(x|N)=\sum_{x=1}^{N}\left(x-\frac{N+1}{2}\right)^2\frac{1}{N}\\
=&\frac{1}{N}\sum_{x=1}^{N}\left[x^2-(N+1)x+\left(\frac{N+1}{2}\right)^2\right]\\
=&\frac{1}{N}\left[\sum_{x=1}^{N}x^2-(N+1)\sum_{x=1}^{N}x+N\left(\frac{N+1}{2}\right)^2\right]\\
=&\frac{1}{N}\left[\frac{N(N+1)(2N+1)}{6}-(N+1)\frac{N(N+1)}{2}+N\left(\frac{N+1}{2}\right)^2\right]\\
=&(N+1)\left[\frac{2N+1}{6}-\frac{N+1}{2}+\frac{N+1}{4}\right]=(N+1)\frac{4N+2-(3N+3)}{12}\\
=&\frac{(N+1)(N-1)}{12}\\
\mathrm{V}[X]=&\frac{(N+1)(N-1)}{12}
\end{aligned}
$$

$$
X\sim\mathcal{DU}(a,b)\Updownarrow\begin{cases}f_X(x|N)=\dfrac{1}{N}=\dfrac{1}{b-a+1} & N=b-a+1\\ x\in X(\Omega)=\{a,a+1,\cdots,b\}\end{cases}\Updownarrow F_X(x|a,b)=\frac{\lfloor x\rfloor-a+1}{b-a+1}
$$

15.1.2.8.1.2 hypergeometric distribution

$$
\begin{gathered}
X\sim\mathcal{HG}(N,M,K)\\
\Updownarrow\\
\begin{cases}f_X(x|N,M,K)=\dfrac{\dbinom{N-M}{K-x}\dbinom{M}{x}}{\dbinom{N}{K}}\\ x\in X(\Omega)=\{\max\{0,K-(N-M)\},\cdots,\min\{K,M\}\}\end{cases}\\
\Downarrow K\ll N,M\\
x\in X(\Omega)=\{0,1,\cdots,K\}\\
\Downarrow\\
\begin{cases}\mathrm{E}[X]= & \dfrac{KM}{N}\\ \mathrm{V}[X]= & \dfrac{KM}{N}\dfrac{(N-M)(N-K)}{N(N-1)}\end{cases}
\end{gathered}
$$

15.1.2.8.1.3 Bernoulli distribution

$$
\begin{gathered}
X\sim\mathcal{B}(p),p=\mathrm{P}(X=1)\\
\Updownarrow\\
\begin{cases}f_X(x|p)=(1-p)^{1-x}p^x\\ x\in X(\Omega)=\{0,1\}\end{cases}\\
\Downarrow\\
\begin{cases}\mathrm{E}[X]= & p\\ \mathrm{V}[X]= & p(1-p)\\ M_X(\xi)= & (1-p)+p\mathrm{e}^{\xi}\end{cases}
\end{gathered}
$$

15.1.2.8.1.4 binomial distribution

independent and identical Bernoulli trials

$$
\begin{gathered}
X \sim \mathrm{b}\left(n,p\right), p = \mathrm{P}\left(X=1\right) \\
\Updownarrow \\
\begin{cases}
f_X\left(x|n,p\right) = \binom{n}{x}\left(1-p\right)^{n-x}p^x \\
x \in X\left(\Omega\right) = \left\{0,1,\cdots,n\right\}
\end{cases} \\
\Downarrow \\
\begin{cases}
\mathrm{E}\left[X\right] = & np \\
\mathrm{V}\left[X\right] = & np\left(1-p\right) \\
M_X\left(\xi\right) = & \left[\left(1-p\right)+p\mathrm{e}^{\xi}\right]^n
\end{cases}
\end{gathered}
$$

reparameterization technique

$$
\begin{aligned}
M_X\left(\xi\right) =& \mathrm{E}\left[\mathrm{e}^{\xi X}\right] = \sum_{x\in X(\Omega)} \mathrm{e}^{\xi x} f_X\left(x|n,p\right) \\
=& \sum_{x=1}^{n} \mathrm{e}^{\xi x}\binom{n}{x}\left(1-p\right)^{n-x}p^x = \sum_{x=1}^{n}\binom{n}{x}\left(1-p\right)^{n-x}\left(p\mathrm{e}^{\xi}\right)^x \\
=& \sum_{x=1}^{n}\binom{n}{x}\left[\frac{1-p}{\left(1-p\right)+p\mathrm{e}^{\xi}}\right]^{n-x}\left[\frac{p\mathrm{e}^{\xi}}{\left(1-p\right)+p\mathrm{e}^{\xi}}\right]^{x}\left[\left(1-p\right)+p\mathrm{e}^{\xi}\right]^{n-x}\left[\left(1-p\right)+p\mathrm{e}^{\xi}\right]^{x} \\
=& \sum_{x=1}^{n}\binom{n}{x}\left[\frac{1-p}{\left(1-p\right)+p\mathrm{e}^{\xi}}\right]^{n-x}\left[\frac{p\mathrm{e}^{\xi}}{\left(1-p\right)+p\mathrm{e}^{\xi}}\right]^{x}\left[\left(1-p\right)+p\mathrm{e}^{\xi}\right]^{n} \\
=& \left[\left(1-p\right)+p\mathrm{e}^{\xi}\right]^{n}\sum_{x=1}^{n}\binom{n}{x}\left[\frac{1-p}{\left(1-p\right)+p\mathrm{e}^{\xi}}\right]^{n-x}\left[\frac{p\mathrm{e}^{\xi}}{\left(1-p\right)+p\mathrm{e}^{\xi}}\right]^{x} \\
=& \left[\left(1-p\right)+p\mathrm{e}^{\xi}\right]^{n}\sum_{x=1}^{n}\binom{n}{x}\left[p^*\right]^{n-x}\left[1-p^*\right]^{x}, p^* = \frac{1-p}{\left(1-p\right)+p\mathrm{e}^{\xi}} \\
=& \left[\left(1-p\right)+p\mathrm{e}^{\xi}\right]^{n}\sum_{x=1}^{n} f_X\left(x|n,p^*\right), X \sim \mathrm{b}\left(n,p^*\right) \\
=& \left[\left(1-p\right)+p\mathrm{e}^{\xi}\right]^{n}\cdot 1 = \left[\left(1-p\right)+p\mathrm{e}^{\xi}\right]^{n} \\
M_X\left(\xi\right) =& \left[\left(1-p\right)+p\mathrm{e}^{\xi}\right]^{n}
\end{aligned}
$$

15.1.2.8.1.5 Poisson distribution

count = number of events

an unbounded discrete distribution we first see or met

$$
\begin{gathered}
X \sim \mathcal{P}\left(\lambda\right) \\
\Updownarrow \\
\begin{cases}
f_X\left(x|\lambda\right) = \dfrac{\mathrm{e}^{-\lambda}\lambda^x}{x!} \\
x \in X\left(\Omega\right) = \left\{0,1,\cdots\right\}
\end{cases} \\
\Downarrow \\
\begin{cases}
\mathrm{E}\left[X\right] = & \lambda \\
\mathrm{V}\left[X\right] = & \lambda = \mathrm{E}\left[X\right] \\
M_X\left(\xi\right) = & \exp\left[\lambda\left(\mathrm{e}^{\xi}-1\right)\right] = \mathrm{e}^{\lambda\left(\mathrm{e}^{\xi}-1\right)}
\end{cases}
\end{gathered}
$$

the Poisson postulates

$$\begin{cases} N_t & \text{a r.v. denoting the of events in } [0,t] \\ N_0 = N_{t=0} = 0 & \text{reset the count at the initial point} \\ \forall s < t \left[N_s \perp N_t - N_s \right] & \text{disjoint intervals independent} \\ N_s = N_{t+s} - N_t & \text{depends on length instead of initial point} \\ \lim\limits_{t \to 0} \dfrac{\mathrm{P}(N_t = 1)}{t} = \lambda & \Rightarrow \forall t \approx 0 \left[\mathrm{P}(N_t = 1) \approx \lambda t \right] \\ \lim\limits_{t \to 0} \dfrac{\mathrm{P}(N_t > 1)}{t} = 0 & \text{no coincidence for small } t \end{cases}$$

$$\text{solve the differential equations} \Downarrow \text{with probability axioms}$$

$$f_X(x|\lambda t) = \mathrm{P}(N_t = x) = \frac{\mathrm{e}^{-\lambda t} (\lambda t)^x}{x!}$$

15.1.2.8.1.6 negative binomial distribution also an unbounded discrete distribution

Count the number of independent and identical Bernoulli trials until r a fixed number of success.

$$X \sim \mathcal{NB}(r,p)$$
$$\Updownarrow$$
$$\begin{cases} f_X(x|r,p) = \dbinom{x-1}{r-1} (1-p)^{x-r} p^r \\ x \in X(\Omega) = \{r, r+1, \cdots\} \end{cases}$$
$$\Downarrow$$
$$\begin{cases} \mathrm{E}[X] = \\ \mathrm{V}[X] = \\ M_X(\xi) = \end{cases}$$

$$Y = X - r, X \sim \mathcal{NB}(r,p) \Rightarrow X = Y + r$$
$$\Updownarrow$$
$$\begin{cases} f_Y(y|r,p) = \dbinom{r+y-1}{r-1} (1-p)^y p^r \\ y \in Y(\Omega) = \{0, 1, \cdots\} \end{cases}$$
$$\Downarrow$$
$$\begin{cases} \mathrm{E}[Y] = & r\dfrac{1-p}{p} = r\left(\dfrac{1}{p} - 1\right) \\ \mathrm{V}[Y] = & r\dfrac{1-p}{p^2} = r\left(\dfrac{1}{p^2} - \dfrac{1}{p}\right) \\ M_Y(\xi) = & \left[\dfrac{p}{1-(1-p)\mathrm{e}^{\xi}}\right]^r \end{cases}$$

Note:

$$\begin{aligned} Y = X - r \Rightarrow f_Y(y|r,p) =& \mathrm{P}(Y = y) \\ =& \mathrm{P}(X - r = y) \\ =& \mathrm{P}(X = r + y) \end{aligned}$$

reparameterization technique

$$\begin{aligned}
M_Y(\xi) =& \mathrm{E}\left[\mathrm{e}^{\xi Y}\right] = \sum_{y\in Y(\Omega)} \mathrm{e}^{\xi y} f_Y(y|r,p) \\
=& \sum_{y=0}^{\infty} \mathrm{e}^{\xi y} \binom{r+y-1}{r-1} (1-p)^y p^r \\
=& p^r \sum_{y=0}^{\infty} \binom{r+y-1}{r-1} \left[(1-p)\,\mathrm{e}^{\xi}\right]^y \\
=& p^r \sum_{y=0}^{\infty} \binom{r+y-1}{r-1} [1-p^*]^y [p^*]^r \frac{1}{[p^*]^r}, 1-p^* = (1-p)\,\mathrm{e}^{\xi} \\
=& \left[\frac{p}{p^*}\right]^r \sum_{y=0}^{\infty} f_Y(y|r,p^*), Y \sim \mathcal{NB}(r,p^*), p^* = 1-(1-p)\,\mathrm{e}^{\xi} \\
=& \left[\frac{p}{p^*}\right]^r \cdot 1 = \left[\frac{p}{p^*}\right]^r = \left[\frac{p}{1-(1-p)\,\mathrm{e}^{\xi}}\right]^r
\end{aligned}$$

Note:

For sufficiently small t such that

$$0 \le p^* = (1-p)\,\mathrm{e}^{\xi} \le 1$$

or else $t \gg 1$

$$p^* = (1-p)\,\mathrm{e}^{\xi} > 1$$

15.1.2.8.1.7 geometric distribution

$$\begin{gathered}
X \sim \mathcal{NB}(r,p) \\
\Updownarrow \\
\begin{cases} f_X(x|r,p) = \binom{x-1}{r-1} (1-p)^{x-r} p^r \\ x \in X(\Omega) = \{r, r+1, \cdots\} \end{cases} \\
\Downarrow \\
\begin{cases} \mathrm{E}[X] = \\ \mathrm{V}[X] = \\ M_X(\xi) = \end{cases}
\end{gathered}$$

$r = 1$,

$$\begin{gathered}
X \sim \mathcal{NB}(r=1,p) = \mathcal{G}(p) \\
\Updownarrow \\
\begin{cases} f_X(x|r=1,p) = \left[\binom{x-1}{r-1} (1-p)^{x-r} p^r\right]_{r=1} = \binom{x-1}{1-1} (1-p)^{x-1} p^1 = (1-p)^{x-1} p \\ x \in X(\Omega) = \{r, r+1, \cdots\}_{r=1} = \{1, 2, \cdots\} \end{cases} \\
\Downarrow \\
\begin{cases} \mathrm{E}[X] = \\ \mathrm{V}[X] = \\ M_X(\xi) = \end{cases}
\end{gathered}$$

the only “memoryless” discrete distribution, and there is also a “memoryless” continuous distribution.

memoryless property

$$\forall s > t \left[\mathrm{P}(X > s | X > t) = \mathrm{P}(X > s - t)\right]$$

Survival depends on length instead of initial point; it might be proper assumption for stuff survival, but might not be proper for human suvival.

15.1.2.8.2 continuous distribution

15.1.2.8.2.1 uniform distribution = continuous uniform distribution

$$X \sim \mathcal{U}(a,b)$$
$$\Updownarrow$$
$$\begin{cases} f_X(x|a,b) = \dfrac{1(a \le x \le b)}{b-a} = \dfrac{1(x \in [a,b])}{b-a} \\ x \in X(\Omega) = [a,b] \end{cases}$$
$$\Downarrow$$
$$\begin{cases} \mathrm{E}[X] = & \dfrac{a+b}{2} \\ \mathrm{V}[X] = & \dfrac{(b-a)^2}{12} \\ M_X(\xi) = & \dfrac{\mathrm{e}^{b\xi} - \mathrm{e}^{a\xi}}{(b-a)\xi} \end{cases}$$

$a = 0, b = 1$

$$X \sim \mathcal{U}(a=0,b=1) = \mathcal{U}(0,1)$$
$$\Updownarrow$$
$$\begin{cases} f_X(x|a=0,b=1) = \left[\dfrac{1(a \le x \le b)}{b-a}\right]_{a=0,b=1} = 1(0 \le x \le 1) = 1(x \in [0,1]) \\ x \in X(\Omega) = [a,b]\,|_{a=0,b=1} = [0,1] \end{cases}$$
$$\Downarrow$$
$$\begin{cases} \mathrm{E}[X] = & \dfrac{1}{2} \\ \mathrm{V}[X] = & \dfrac{1}{12} \\ M_X(\xi) = & \dfrac{\mathrm{e}^{\xi} - 1}{\xi} \end{cases}$$

15.1.2.8.2.2 gamma distribution

15.1.2.8.2.3 exponential distribution

15.1.2.8.2.4 Chi-square distribution

15.1.2.8.2.5 Weibull distribution

15.1.2.8.2.6 normal distribution

$$X \sim \mathcal{N}\left(\mu, \sigma^2\right)$$
$$\Updownarrow$$
$$\begin{cases} f_X\left(x|\mu,\sigma^2\right) = \dfrac{1}{\sigma\sqrt{2\pi}}\mathrm{e}^{\frac{-1}{2}\left(\frac{x-\mu}{\sigma}\right)^2} = \dfrac{\mathrm{e}^{\frac{-1}{2}\left(\frac{x-\mu}{\sigma}\right)^2}}{\sigma\sqrt{2\pi}} = \dfrac{\mathrm{e}^{-(x-\mu)^2/\left(2\sigma^2\right)}}{\sigma\sqrt{2\pi}} = \dfrac{\exp\left\{\frac{-1}{2}\left(\frac{x-\mu}{\sigma}\right)^2\right\}}{\sigma\sqrt{2\pi}} \\ x \in X(\Omega) = \mathbb{R} \end{cases}$$
$$\Downarrow$$
$$\begin{cases} \mathrm{E}[X] = & \mu \\ \mathrm{V}[X] = & \sigma^2 \\ M_X(\xi) = & \mathrm{e}^{\left(\mu\xi + \frac{\sigma^2}{2}\xi^2\right)} = \mathrm{e}^{\mu\xi + \frac{\sigma^2}{2}\xi^2} \end{cases}$$

15.1.2.8.2.7 beta distribution

random success probability

15.1.2.8.2.8 Cauchy distribution

https://www.youtube.com/watch?v=ojZb6nZWdvI&t=33m34s
https://en.wikipedia.org/wiki/Cauchy_distribution

$$X \sim \mathcal{C}(\theta, \sigma)$$
$$\Updownarrow$$
$$\begin{cases} f_X(x|\theta,\sigma) = \dfrac{1}{\pi\sigma}\left[1+\left(\dfrac{x-\theta}{\sigma}\right)^2\right]^{-1} = \dfrac{1}{\pi\sigma\left[1+\left(\dfrac{x-\theta}{\sigma}\right)^2\right]} = \dfrac{1}{\pi}\left[\dfrac{\sigma}{(x-\theta)^2+\sigma^2}\right] \\ x \in X(\Omega) = \mathbb{R} \end{cases}$$
$$\Downarrow$$
$$\begin{cases} \mathrm{E}[X] & \text{diverges} \\ \mathrm{V}[X] & \text{diverges} \\ M_X(\xi) & \text{diverges} \end{cases}$$

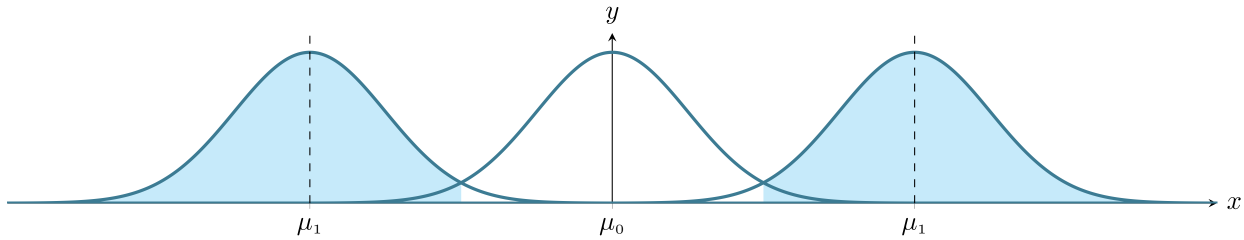


Figure 15.1: not completed: Cauchy distribution vs. normal distribution

heavy tail

t distribution is also heavy tail

box plot

15.1.2.8.2.9 log-normal distribution

15.1.2.9 exponential family

Definition 15.18. exponential family: A family of PDF/PMF is called exponential family if

$$f(x|\boldsymbol{\theta}) = h(x)\,c(\boldsymbol{\theta})\,\mathrm{e}^{\sum\limits_{j=1}^{k} w_j(\boldsymbol{\theta})t_j(x)} = h(x)\,c(\boldsymbol{\theta})\exp\left(\sum_{j=1}^{k} w_j(\boldsymbol{\theta})\,t_j(x)\right)$$

with $\boldsymbol{\theta} = \boldsymbol{\theta}(\theta_1,\cdots,\theta_k) = (\theta_1,\cdots,\theta_k)$ for some $h(x), c(\boldsymbol{\theta}), w_j(\boldsymbol{\theta}), t_j(x)$, where

$$h(x)\,c(\boldsymbol{\theta}) \geq 0 \Rightarrow f(x|\theta) \geq 0$$

and parameters $\boldsymbol{\theta}$ and statistic or real number x can be separated.

$$\mathcal{E}^f = \left\{ f \,\middle|\, f = f(x|\boldsymbol{\theta}) = h(x)\,c(\boldsymbol{\theta})\,\mathrm{e}^{\sum\limits_{j=1}^{k} w_j(\boldsymbol{\theta})t_j(x)} = h(x)\,c(\boldsymbol{\theta})\exp\left(\sum_{j=1}^{k} w_j(\boldsymbol{\theta})\,t_j(x)\right) \right\}$$

- exponential family
  - normal distribution[15.1.2.8.2.6]
  - gamma distribution[15.1.2.8.2.2]
  - log-normal distribution[15.1.2.8.2.9]
- non-exponential family
  - Cauchy distribution[15.1.2.8.2.8] https://www.youtube.com/watch?v=ojZb6nZWdvI&t=1h7m43s

15.1.2.9.1 binomial distribution with known n

$$f(x|p)=\binom{n}{x}(1-p)^{n-x}p^x$$

or

$$\underline{f(x|p)=\binom{n}{x}(1-p)^{n-x}p^x=f(x|n=n,p)}$$

not

$$\underline{f(x|n,p)=\binom{n}{x}(1-p)^{n-x}p^x}$$

$$\begin{aligned}f(x|p)=&\binom{n}{x}(1-p)^{n-x}p^x\\=&\binom{n}{x}(1-p)^n\left(\frac{p}{1-p}\right)^x\\=&\binom{n}{x}(1-p)^n\mathrm{e}^{\left(\ln\frac{p}{1-p}\right)x}\\=&h(x)\,c(\boldsymbol{\theta})\,\mathrm{e}^{\sum_{j=1}^{k}w_j(\boldsymbol{\theta})t_j(x)},\begin{cases}h(x)= & \binom{n}{x}\\c(\boldsymbol{\theta})=c(\theta_1)=c(p)= & (1-p)^n\\w_1(\boldsymbol{\theta})=w(\theta_1)=w(p)= & \ln\frac{p}{1-p}\\t_1(x)= & x\\k= & 1\end{cases}\end{aligned}$$

$$\Downarrow$$

$$f(x|p)=\binom{n}{x}(1-p)^{n-x}p^x\in\mathcal{E}^f=\left\{f\middle|f=f(x|\boldsymbol{\theta})=h(x)\,c(\boldsymbol{\theta})\exp\left(\sum_{j=1}^{k}w_j(\boldsymbol{\theta})\,t_j(x)\right)\right\}$$

why known n?

known n

$$\binom{n}{x}=h(x)$$

unknown n

$$\binom{n}{x}=h(x,n)\neq h_1(n)\,h_2(x)$$

15.1.2.9.2 continuous uniform distribution not in exponential family

$$X\sim\mathcal{U}(a,b)$$

$$\begin{aligned}f_X(x|a,b)=&\frac{1(x\in[a,b])}{b-a}\\=&\frac{1}{b-a},x\in[a,b]\end{aligned}$$

$$\frac{1}{b-a}=h(x)\,c(\boldsymbol{\theta})\,\mathrm{e}^{\sum_{j=1}^{k}w_j(\boldsymbol{\theta})t_j(x)},\begin{cases}h(x)= & 1\\c(\boldsymbol{\theta})=c(\theta_1,\theta_2)=c(a,b)= & \frac{1}{b-a}\\w_1(\theta_1)=w_2(\theta_2)= & 0\\t_1(x)=t_2(x)= & x\\k= & 2\end{cases}$$

however,

$$1\left(x\in[a,b]\right)\neq h\left(x\right)c\left(\boldsymbol{\theta}\right)=h\left(x\right)c\left(a,b\right)$$

thus

$$f_X\left(x|a,b\right)=\frac{1\left(x\in[a,b]\right)}{b-a}\notin\mathcal{E}^f=\left\{f\middle|f=f\left(x|\boldsymbol{\theta}\right)=h\left(x\right)c\left(\boldsymbol{\theta}\right)\exp\left(\sum_{j=1}^{k}w_j\left(\boldsymbol{\theta}\right)t_j\left(x\right)\right)\right\}$$

15.1.2.9.3 normal distribution is in exponential family

$$X\sim\mathcal{N}\left(\mu,\sigma^2\right)$$
$$\Updownarrow$$
$$\begin{cases}f_X\left(x|\mu,\sigma^2\right)=\dfrac{1}{\sigma\sqrt{2\pi}}\mathrm{e}^{\frac{-1}{2}\left(\frac{x-\mu}{\sigma}\right)^2}=\dfrac{\mathrm{e}^{\frac{-1}{2}\left(\frac{x-\mu}{\sigma}\right)^2}}{\sigma\sqrt{2\pi}}=\dfrac{\mathrm{e}^{-(x-\mu)^2/\left(2\sigma^2\right)}}{\sigma\sqrt{2\pi}}=\dfrac{\exp\left\{\frac{-1}{2}\left(\frac{x-\mu}{\sigma}\right)^2\right\}}{\sigma\sqrt{2\pi}}\\ x\in X\left(\Omega\right)=\mathbb{R}\end{cases}$$
$$\Downarrow$$
$$\begin{cases}\mathrm{E}\left[X\right]= & \mu\\ \mathrm{V}\left[X\right]= & \sigma^2\\ M_X\left(\xi\right)= & \mathrm{e}^{\left(\mu\xi+\frac{\sigma^2}{2}\xi^2\right)}=\mathrm{e}^{\mu\xi+\frac{\sigma^2}{2}\xi^2}\end{cases}$$

$$\begin{aligned}f_X\left(x|\mu,\sigma^2\right)=f_X\left(x|\mu,\sigma\right)=&\frac{1}{\sigma\sqrt{2\pi}}\mathrm{e}^{\frac{-1}{2}\left(\frac{x-\mu}{\sigma}\right)^2}\\ =&\frac{1}{\sigma\sqrt{2\pi}}\mathrm{e}^{\frac{-1}{2\sigma^2}\left(x^2-2x\mu+\mu^2\right)}=\frac{\mathrm{e}^{\frac{-1}{2}\left(\frac{\mu}{\sigma}\right)^2}}{\sigma\sqrt{2\pi}}\mathrm{e}^{\frac{-1}{2\sigma^2}x^2+\frac{\mu}{\sigma^2}x}=\frac{\mathrm{e}^{\frac{-1}{2}\left(\frac{\mu}{\sigma}\right)^2}}{\sigma\sqrt{2\pi}}\mathrm{e}^{\frac{\mu}{\sigma^2}x-\frac{1}{2\sigma^2}x^2}\\ =&h\left(x\right)c\left(\boldsymbol{\theta}\right)\mathrm{e}^{\sum_{j=1}^{k}w_j\left(\boldsymbol{\theta}\right)t_j\left(x\right)},\begin{cases}h\left(x\right)= & 1\\ c\left(\boldsymbol{\theta}\right)=c\left(\theta_1,\theta_2\right)=c\left(\mu,\sigma\right)= & \dfrac{\mathrm{e}^{\frac{-1}{2}\left(\frac{\mu}{\sigma}\right)^2}}{\sigma\sqrt{2\pi}}\\ w_1\left(\theta_1,\theta_2\right)=w_1\left(\mu,\sigma\right)= & \frac{\mu}{\sigma^2}\\ w_2\left(\theta_2\right)=w_2\left(\sigma\right)= & \frac{-1}{2\sigma^2}\\ t_1\left(x\right)= & x\\ t_2\left(x\right)= & x^2\\ k= & 2\end{cases}\end{aligned}$$
$$\Downarrow$$
$$f_X\left(x|\mu,\sigma\right)=\frac{1}{\sigma\sqrt{2\pi}}\mathrm{e}^{\frac{-1}{2}\left(\frac{x-\mu}{\sigma}\right)^2}\in\mathcal{E}^f=\left\{f\middle|f=f\left(x|\boldsymbol{\theta}\right)=h\left(x\right)c\left(\boldsymbol{\theta}\right)\exp\left(\sum_{j=1}^{k}w_j\left(\boldsymbol{\theta}\right)t_j\left(x\right)\right)\right\}$$

15.1.2.9.4 normal distribution with unknown equal mean and standard deviation in curved exponential family

$$X\sim\mathcal{N}\left(\mu,\sigma^2=\mu^2\right)=\mathcal{N}\left(\mu,\mu^2\right)$$
$$\Updownarrow$$
$$\begin{cases}f_X\left(x|\mu,\mu^2\right)=f_X\left(x|\mu\right)=\left[\dfrac{1}{\sigma\sqrt{2\pi}}\mathrm{e}^{\frac{-1}{2}\left(\frac{x-\mu}{\sigma}\right)^2}\right]_{\sigma=\mu}=\dfrac{1}{\mu\sqrt{2\pi}}\mathrm{e}^{\frac{-1}{2}\left(\frac{x-\mu}{\mu}\right)^2}\\ x\in X\left(\Omega\right)=\mathbb{R}\end{cases}$$
$$\Downarrow$$
$$\begin{cases}\mathrm{E}\left[X\right]= & \mu\\ \mathrm{V}\left[X\right]= & \mu^2\\ M_X\left(\xi\right)= & \mathrm{e}^{\left(\mu\xi+\frac{\mu^2}{2}\xi^2\right)}=\mathrm{e}^{\mu\xi+\frac{\mu^2}{2}\xi^2}\end{cases}$$

$$
\begin{aligned}
f_X\left(x|\mu,\mu^2\right) = f_X\left(x|\mu\right) =& \frac{1}{\mu\sqrt{2\pi}} \mathrm{e}^{\frac{-1}{2}\left(\frac{x-\mu}{\mu}\right)^2} \\
=& \frac{1}{\mu\sqrt{2\pi}} \mathrm{e}^{\frac{-1}{2\mu^2}\left(x^2-2x\mu+\mu^2\right)} = \frac{\mathrm{e}^{\frac{-1}{2}}}{\mu\sqrt{2\pi}} \mathrm{e}^{\frac{-1}{2\mu^2}x^2+\frac{\mu}{\mu^2}x} = \frac{\mathrm{e}^{\frac{-1}{2}}}{\mu\sqrt{2\pi}} \mathrm{e}^{\frac{1}{\mu}x-\frac{1}{2\mu^2}x^2} \\
=& h\left(x\right) c\left(\boldsymbol{\theta}\right) \mathrm{e}^{\sum_{j=1}^{k} w_j\left(\boldsymbol{\theta}\right) t_j\left(x\right)}, \begin{cases}
h\left(x\right) = & 1 \\
c\left(\boldsymbol{\theta}\right) = c\left(\theta_1\right) = c\left(\mu\right) = & \frac{\mathrm{e}^{\frac{-1}{2}}}{\mu\sqrt{2\pi}} \\
w_1\left(\theta_1\right) = w_1\left(\mu\right) = & \frac{1}{\mu} \\
w_2\left(\theta_1\right) = w_2\left(\mu\right) = & \frac{-1}{2\mu^2} \\
t_1\left(x\right) = & x \\
t_2\left(x\right) = & x^2 \\
k = & 2 > 1 = p
\end{cases} \\
& \Downarrow \begin{cases} p = \dim \boldsymbol{\theta} \\ k = \dim \boldsymbol{w} \end{cases}
\end{aligned}
$$

$$
\begin{aligned}
f_X\left(x|\mu,\sigma\right) = \frac{1}{\sigma\sqrt{2\pi}} \mathrm{e}^{\frac{-1}{2}\left(\frac{x-\mu}{\sigma}\right)^2} \in& \mathcal{C}^f = \left\{ f \middle| f = f\left(x|\boldsymbol{\theta}\right) = h\left(x\right) c\left(\boldsymbol{\theta}\right) \exp\left(\sum_{j=1}^{k>p} w_j\left(\boldsymbol{\theta}\right) t_j\left(x\right)\right)\right\} \\
\subset& \mathcal{E}^f = \left\{ f \middle| f = f\left(x|\boldsymbol{\theta}\right) = h\left(x\right) c\left(\boldsymbol{\theta}\right) \exp\left(\sum_{j=1}^{k} w_j\left(\boldsymbol{\theta}\right) t_j\left(x\right)\right)\right\}
\end{aligned}
$$

---

curved exponential family

$$
\mathcal{C}^f = \left\{ f \middle| f = f\left(x|\boldsymbol{\theta}\right) = h\left(x\right) c\left(\boldsymbol{\theta}\right) \exp\left(\sum_{j=1}^{k>p} w_j\left(\boldsymbol{\theta}\right) t_j\left(x\right)\right)\right\}
$$

---

exponential family

$$
\mathcal{E}^f = \left\{ f \middle| f = f\left(x|\boldsymbol{\theta}\right) = h\left(x\right) c\left(\boldsymbol{\theta}\right) \exp\left(\sum_{j=1}^{k} w_j\left(\boldsymbol{\theta}\right) t_j\left(x\right)\right)\right\}
$$

---

https://tex.stackexchange.com/questions/145969/filling-specified-area-by-random-dots-in-tikz

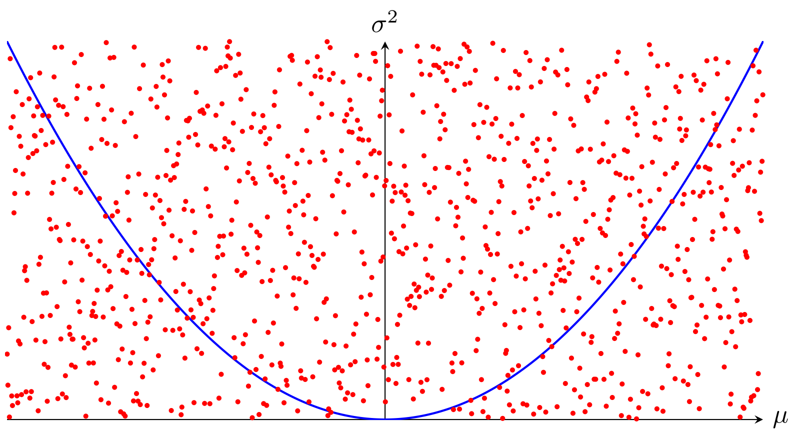


Figure 15.2: curved exponential family vs. exponential family

15.1.2.9.5 properties of exponential family

15.1.2.9.5.1 fundamentals of statistical inference

Lemma 15.1. *Leibnitz rule*

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}\theta}\int_{a(\theta)}^{b(\theta)} g(x,\theta)\,\mathrm{d}x =& \frac{\mathrm{d}\int_{a(\theta)}^{b(\theta)} g(x,\theta)\,\mathrm{d}x}{\mathrm{d}\theta}\\ =& g(x,b(\theta))\frac{\mathrm{d}b(\theta)}{\mathrm{d}\theta} - g(x,a(\theta))\frac{\mathrm{d}a(\theta)}{\mathrm{d}\theta} + \int_{a(\theta)}^{b(\theta)}\frac{\mathrm{d}}{\mathrm{d}\theta}g(x,\theta)\,\mathrm{d}x\\ =& g(x,b(\theta))\frac{\mathrm{d}b(\theta)}{\mathrm{d}\theta} - g(x,a(\theta))\frac{\mathrm{d}a(\theta)}{\mathrm{d}\theta} + \int_{a(\theta)}^{b(\theta)}\frac{\mathrm{d}g(x,\theta)}{\mathrm{d}\theta}\mathrm{d}x\end{aligned}$$

$$\frac{\mathrm{d}}{\mathrm{d}\theta}\int_{a(\theta)}^{b(\theta)} g(x,\theta)\,\mathrm{d}x = g(x,b(\theta))\frac{\mathrm{d}b(\theta)}{\mathrm{d}\theta} - g(x,a(\theta))\frac{\mathrm{d}a(\theta)}{\mathrm{d}\theta} + \int_{a(\theta)}^{b(\theta)}\frac{\mathrm{d}g(x,\theta)}{\mathrm{d}\theta}\mathrm{d}x$$

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}\theta}\int_{a(\theta)=a}^{b(\theta)=b} g(x,\theta)\,\mathrm{d}x =& g(x,b(\theta))\frac{\mathrm{d}b(\theta)}{\mathrm{d}\theta} - g(x,a(\theta))\frac{\mathrm{d}a(\theta)}{\mathrm{d}\theta} + \int_{a(\theta)}^{b(\theta)}\frac{\mathrm{d}g(x,\theta)}{\mathrm{d}\theta}\mathrm{d}x\\ =& g(x,b)\frac{\mathrm{d}b}{\mathrm{d}\theta} - g(x,a)\frac{\mathrm{d}a}{\mathrm{d}\theta} + \int_a^b \frac{\mathrm{d}g(x,\theta)}{\mathrm{d}\theta}\mathrm{d}x\\ =& g(x,b)\,0 - g(x,a)\,0 + \int_a^b \frac{\mathrm{d}g(x,\theta)}{\mathrm{d}\theta}\mathrm{d}x = 0+0+\int_a^b \frac{\mathrm{d}g(x,\theta)}{\mathrm{d}\theta}\mathrm{d}x\\ =& \int_a^b \frac{\mathrm{d}g(x,\theta)}{\mathrm{d}\theta}\mathrm{d}x\\ \frac{\mathrm{d}}{\mathrm{d}\theta}\int_{X(\Omega)\perp\theta} g(x,\theta)\,\mathrm{d}x =& \frac{\mathrm{d}}{\mathrm{d}\theta}\int_{a(\theta)=a}^{b(\theta)=b} g(x,\theta)\,\mathrm{d}x = \int_a^b \frac{\mathrm{d}g(x,\theta)}{\mathrm{d}\theta}\mathrm{d}x = \int_{X(\Omega)\perp\theta}\frac{\mathrm{d}g(x,\theta)}{\mathrm{d}\theta}\mathrm{d}x\\ \frac{\mathrm{d}}{\mathrm{d}\theta}\int_{X(\Omega)\perp\theta} g(x,\theta)\,\mathrm{d}x =& \int_{X(\Omega)\perp\theta}\frac{\mathrm{d}g(x,\theta)}{\mathrm{d}\theta}\mathrm{d}x\end{aligned}$$

$$\frac{\mathrm{d}}{\mathrm{d}\theta}\int_{X(\Omega)\perp\theta} g(x,\theta)\,\mathrm{d}x = \int_{X(\Omega)\perp\theta}\frac{\mathrm{d}g(x,\theta)}{\mathrm{d}\theta}\mathrm{d}x$$

---

point estimation

Lemma 15.2. *parameter-independent expectation:*

Assume the domain of X independent of θ, then

$$\mathrm{E}\left[\frac{\partial \ln f(X|\boldsymbol{\theta})}{\partial\theta_i}\right] = 0, \forall i = 1,\cdots,p$$

Proof:

$$\begin{aligned}\mathrm{E}\left[\frac{\partial \ln f(X|\boldsymbol{\theta})}{\partial\theta_i}\right] =& \int \frac{\partial \ln f(x|\boldsymbol{\theta})}{\partial\theta_i} f(x|\boldsymbol{\theta})\,\mathrm{d}x\\ =& \int \left[\frac{\partial \ln f(x|\boldsymbol{\theta})}{\partial\theta_i}\right] f(x|\boldsymbol{\theta})\,\mathrm{d}x\\ =& \int \left[\frac{1}{f(x|\boldsymbol{\theta})}\frac{\partial f(x|\boldsymbol{\theta})}{\partial\theta_i}\right] f(x|\boldsymbol{\theta})\,\mathrm{d}x = \int \frac{1}{f(x|\boldsymbol{\theta})}\frac{\partial f(x|\boldsymbol{\theta})}{\partial\theta_i} f(x|\boldsymbol{\theta})\,\mathrm{d}x\\ =& \int \frac{f(x|\boldsymbol{\theta})}{f(x|\boldsymbol{\theta})}\frac{\partial f(x|\boldsymbol{\theta})}{\partial\theta_i}\mathrm{d}x = \int 1\frac{\partial f(x|\boldsymbol{\theta})}{\partial\theta_i}\mathrm{d}x = \int \frac{\partial f(x|\boldsymbol{\theta})}{\partial\theta_i}\mathrm{d}x\\ \overset{\begin{cases}\text{Leibnitz rule}\\ x\perp\boldsymbol{\theta}\end{cases}}{=}& \frac{\partial}{\partial\theta_i}\int f(x|\boldsymbol{\theta})\,\mathrm{d}x\end{aligned}$$

□

Note:

$$
\begin{aligned}
\mathrm{E}\left[\frac{\partial \ln f(X|\boldsymbol{\theta})}{\partial \theta_i}\right] &= \int \frac{\partial \ln f(x|\boldsymbol{\theta})}{\partial \theta_i} f(x|\boldsymbol{\theta})\,\mathrm{d}x = \int \left(\frac{\frac{\partial f(x|\boldsymbol{\theta})}{\partial \theta_i}}{f(x|\boldsymbol{\theta})}\right) f(x|\boldsymbol{\theta})\,\mathrm{d}x \\
X \sim f_X(x|\boldsymbol{\theta}) \Rightarrow \mathrm{E}\left[\frac{\partial \ln f_X(X|\boldsymbol{\theta})}{\partial \theta_i}\right] &= \int \frac{\partial \ln f_X(x|\boldsymbol{\theta})}{\partial \theta_i} f_X(x|\boldsymbol{\theta})\,\mathrm{d}x = \int \left(\frac{\frac{\partial f_X(x|\boldsymbol{\theta})}{\partial \theta_i}}{f_X(x|\boldsymbol{\theta})}\right) f_X(x|\boldsymbol{\theta})\,\mathrm{d}x \\
&= \int \frac{\partial f(x|\boldsymbol{\theta})}{\partial \theta_i}\mathrm{d}x \\
X \sim f_X(x|\boldsymbol{\theta}) \Rightarrow \mathrm{E}\left[\frac{\partial \ln f_X(X|\boldsymbol{\theta}^*)}{\partial \theta_i}\right] &= \int \frac{\partial \ln f_X(x|\boldsymbol{\theta}^*)}{\partial \theta_i} f_X(x|\boldsymbol{\theta})\,\mathrm{d}x = \int \left(\frac{\frac{\partial f_X(x|\boldsymbol{\theta}^*)}{\partial \theta_i}}{f_X(x|\boldsymbol{\theta}^*)}\right) f(x|\boldsymbol{\theta})\,\mathrm{d}x \\
&= \int \frac{\partial f_X(x|\boldsymbol{\theta}^*)}{\partial \theta_i}\left(\frac{f_X(x|\boldsymbol{\theta})}{f_X(x|\boldsymbol{\theta}^*)}\right)\mathrm{d}x
\end{aligned}
$$

$$
\mathrm{E}\left[\frac{\partial \ln f_X(X|\boldsymbol{\theta}^*)}{\partial \theta_i}\right] = \mathrm{E}_{\boldsymbol{\theta}^*}\left[\frac{\partial \ln f_X(X|\boldsymbol{\theta}^*)}{\partial \theta_i}\right] = \int \frac{\partial \ln f_X(x|\boldsymbol{\theta}^*)}{\partial \theta_i} f_X(x|\boldsymbol{\theta}^*)\,\mathrm{d}x
$$

$$
\mathrm{E}\left[\frac{\partial \ln f_X(X|\boldsymbol{\theta}^*)}{\partial \theta_i}\right] = \mathrm{E}_{\boldsymbol{\theta}}\left[\frac{\partial \ln f_X(X|\boldsymbol{\theta}^*)}{\partial \theta_i}\right] = \int \frac{\partial \ln f_X(x|\boldsymbol{\theta}^*)}{\partial \theta_i} f_X(x|\boldsymbol{\theta})\,\mathrm{d}x
$$

$$
\mathrm{E}_{\boldsymbol{\theta}}\left[\frac{\partial \ln f_X(X|\boldsymbol{\theta}^*)}{\partial \theta_i}\right] = \int \frac{\partial \ln f_X(x|\boldsymbol{\theta}^*)}{\partial \theta_i} f_X(x|\boldsymbol{\theta})\,\mathrm{d}x
$$

$$
\mathrm{E}_{\boldsymbol{\theta}}\left[\frac{\partial \ln f_X(X|\boldsymbol{\theta}^*)}{\partial \theta_i}\right] \not\equiv 0
$$

$$
\mathrm{E}_{\boldsymbol{\theta}}\left[\frac{\partial \ln f_X(X|\boldsymbol{\theta}^*)}{\partial \theta_i}\right] \overset{\boldsymbol{\theta}^*=\boldsymbol{\theta}}{\equiv} 0
$$

as the fundamental to estimate parameters.

interval estimation

Lemma 15.3. *parameter-independent variance:*

Assume the domain of X independent of θ, then

$$
\mathrm{E}\left[\frac{\partial^2 \ln f(X|\boldsymbol{\theta})}{\partial \theta_i^2}\right] = -\mathrm{E}\left[\left(\frac{\partial \ln f(X|\boldsymbol{\theta})}{\partial \theta_i}\right)^2\right], \forall i = 1, \cdots, p
$$

Proof:

$$
\begin{aligned}
&\mathrm{E}\left[\frac{\partial^2 \ln f(X|\boldsymbol{\theta})}{\partial \theta_i^2}\right] = \int \frac{\partial^2 \ln f(X|\boldsymbol{\theta})}{\partial \theta_i^2} f(x|\boldsymbol{\theta})\,\mathrm{d}x = \int \frac{\partial}{\partial \theta_i}\frac{\partial \ln f(x|\boldsymbol{\theta})}{\partial \theta_i} f(x|\boldsymbol{\theta})\,\mathrm{d}x \\
&= \int \left\{\frac{\partial}{\partial \theta_i}\left(\frac{1}{f(x|\boldsymbol{\theta})}\frac{\partial f(x|\boldsymbol{\theta})}{\partial \theta_i}\right)\right\} f(x|\boldsymbol{\theta})\,\mathrm{d}x = \int \left\{\frac{\partial}{\partial \theta_i}\left(\frac{\frac{\partial f(x|\boldsymbol{\theta})}{\partial \theta_i}}{f(x|\boldsymbol{\theta})}\right)\right\} f(x|\boldsymbol{\theta})\,\mathrm{d}x \\
&= \int \frac{\frac{\partial^2 f(x|\boldsymbol{\theta})}{\partial \theta_i^2} f(x|\boldsymbol{\theta}) - \frac{\partial f(x|\boldsymbol{\theta})}{\partial \theta_i}\frac{\partial f(x|\boldsymbol{\theta})}{\partial \theta_i}}{[f(x|\boldsymbol{\theta})]^2} f(x|\boldsymbol{\theta})\,\mathrm{d}x = \int \frac{\frac{\partial^2 f(x|\boldsymbol{\theta})}{\partial \theta_i^2}[f(x|\boldsymbol{\theta})]^2 - \left[\frac{\partial f(x|\boldsymbol{\theta})}{\partial \theta_i}\right]^2 [f(x|\boldsymbol{\theta})]}{[f(x|\boldsymbol{\theta})]^2}\mathrm{d}x \\
&= \int \frac{\partial^2 f(x|\boldsymbol{\theta})}{\partial \theta_i^2}\mathrm{d}x - \int \left(\frac{\frac{\partial f(x|\boldsymbol{\theta})}{\partial \theta_i}}{f(x|\boldsymbol{\theta})}\right)^2 f(x|\boldsymbol{\theta})\,\mathrm{d}x \overset{\begin{cases}\text{Leibnitz rule}\\ X(\Omega) \perp \boldsymbol{\theta}\end{cases}}{=} \frac{\partial^2}{\partial \theta_i^2}\int f(x|\boldsymbol{\theta})\,\mathrm{d}x - \int \left(\frac{\frac{\partial f(x|\boldsymbol{\theta})}{\partial \theta_i}}{f(x|\boldsymbol{\theta})}\right)^2 f(x|\boldsymbol{\theta})\,\mathrm{d}x \\
&= \frac{\partial^2}{\partial \theta_i^2}1 - \int \left(\frac{\frac{\partial f(x|\boldsymbol{\theta})}{\partial \theta_i}}{f(x|\boldsymbol{\theta})}\right)^2 f(x|\boldsymbol{\theta})\,\mathrm{d}x = 0 - \int \left(\frac{\frac{\partial f(x|\boldsymbol{\theta})}{\partial \theta_i}}{f(x|\boldsymbol{\theta})}\right)^2 f(x|\boldsymbol{\theta})\,\mathrm{d}x = -\int \left(\frac{\frac{\partial f(x|\boldsymbol{\theta})}{\partial \theta_i}}{f(x|\boldsymbol{\theta})}\right)^2 f(x|\boldsymbol{\theta})\,\mathrm{d}x \\
&= -\int \left(\frac{\partial \ln f(x|\boldsymbol{\theta})}{\partial \theta_i}\right)^2 f(x|\boldsymbol{\theta})\,\mathrm{d}x = -\mathrm{E}\left[\left(\frac{\partial \ln f(X|\boldsymbol{\theta})}{\partial \theta_i}\right)^2\right]
\end{aligned}
$$

□

$$\begin{cases} \mathrm{E}\left[\dfrac{\partial \ln f(X|\boldsymbol{\theta})}{\partial \theta_i}\right] = & 0 \\ \mathrm{E}\left[\dfrac{\partial^2 \ln f(X|\boldsymbol{\theta})}{\partial \theta_i^2}\right] = & -\mathrm{E}\left[\left(\dfrac{\partial \ln f(X|\boldsymbol{\theta})}{\partial \theta_i}\right)^2\right] \end{cases}, \forall i = 1, \cdots, p$$

exponential family expectation and variance

Theorem 15.15. *exponential family expectation*

$$X \in \mathcal{E}^f = \left\{ f \middle| f = h(x)\, c(\boldsymbol{\theta}) \exp\left(\sum_{j=1}^{k} w_j(\boldsymbol{\theta})\, t_j(x)\right)\right\}$$
$$\Downarrow \forall i = 1, \cdots, p$$
$$\mathrm{E}\left[\sum_{j=1}^{k} \frac{\partial w_j(\boldsymbol{\theta})}{\partial \theta_i} t_j(X)\right] = \frac{-\partial \ln c(\boldsymbol{\theta})}{\partial \theta_i}$$

Proof:

$$\begin{aligned}
f(x|\boldsymbol{\theta}) =& h(x)\, c(\boldsymbol{\theta})\, \mathrm{e}^{\sum_{j=1}^{k} w_j(\boldsymbol{\theta}) t_j(x)} = h(x)\, c(\boldsymbol{\theta}) \exp\left(\sum_{j=1}^{k} w_j(\boldsymbol{\theta})\, t_j(x)\right) \\
\ln f(x|\boldsymbol{\theta}) =& \ln h(x)\, c(\boldsymbol{\theta})\, \mathrm{e}^{\sum_{j=1}^{k} w_j(\boldsymbol{\theta}) t_j(x)} \\
=& \ln h(x) + \ln c(\boldsymbol{\theta}) + \ln \mathrm{e}^{\sum_{j=1}^{k} w_j(\boldsymbol{\theta}) t_j(x)} \\
=& \ln h(x) + \ln c(\boldsymbol{\theta}) + \sum_{j=1}^{k} w_j(\boldsymbol{\theta})\, t_j(x) \\
\ln f(x|\boldsymbol{\theta}) =& \ln h(x) + \ln c(\boldsymbol{\theta}) + \sum_{j=1}^{k} w_j(\boldsymbol{\theta})\, t_j(x) \\
\ln f(X|\boldsymbol{\theta}) =& \ln h(X) + \ln c(\boldsymbol{\theta}) + \sum_{j=1}^{k} w_j(\boldsymbol{\theta})\, t_j(X) \\
\frac{\partial}{\partial \theta_i} \ln f(X|\boldsymbol{\theta}) =& \frac{\partial}{\partial \theta_i}\left[\ln h(X) + \ln c(\boldsymbol{\theta}) + \sum_{j=1}^{k} w_j(\boldsymbol{\theta})\, t_j(X)\right] \\
=& \frac{\partial}{\partial \theta_i} \ln h(X) + \frac{\partial}{\partial \theta_i} \ln c(\boldsymbol{\theta}) + \frac{\partial}{\partial \theta_i} \sum_{j=1}^{k} w_j(\boldsymbol{\theta})\, t_j(X) \\
=& 0 + \frac{\partial}{\partial \theta_i} \ln c(\boldsymbol{\theta}) + \sum_{j=1}^{k} \frac{\partial}{\partial \theta_i}\{w_j(\boldsymbol{\theta})\, t_j(X)\} \\
=& \frac{\partial \ln c(\boldsymbol{\theta})}{\partial \theta_i} + \sum_{j=1}^{k}\left\{\left(\frac{\partial}{\partial \theta_i} w_j(\boldsymbol{\theta})\right) t_j(X) + w_j(\boldsymbol{\theta})\left(\frac{\partial}{\partial \theta_i} t_j(X)\right)\right\} \\
=& \frac{\partial \ln c(\boldsymbol{\theta})}{\partial \theta_i} + \sum_{j=1}^{k}\left\{\left(\frac{\partial}{\partial \theta_i} w_j(\boldsymbol{\theta})\right) t_j(X) + w_j(\boldsymbol{\theta})\, 0\right\} \\
=& \frac{\partial \ln c(\boldsymbol{\theta})}{\partial \theta_i} + \sum_{j=1}^{k}\left\{\left(\frac{\partial}{\partial \theta_i} w_j(\boldsymbol{\theta})\right) t_j(X)\right\} \\
\frac{\partial \ln f(X|\boldsymbol{\theta})}{\partial \theta_i} =& \frac{\partial \ln c(\boldsymbol{\theta})}{\partial \theta_i} + \sum_{j=1}^{k}\left\{\frac{\partial w_j(\boldsymbol{\theta})}{\partial \theta_i} t_j(X)\right\}
\end{aligned}$$

$$
\begin{aligned}
\frac{\partial \ln f(X|\boldsymbol{\theta})}{\partial \theta_i} &= \frac{\partial \ln c(\boldsymbol{\theta})}{\partial \theta_i} + \sum_{j=1}^{k}\left\{\frac{\partial w_j(\boldsymbol{\theta})}{\partial \theta_i} t_j(X)\right\} \\
0 \overset{\text{lemma: } \mathrm{E}\left[\frac{\partial \ln f(X|\boldsymbol{\theta})}{\partial \theta_i}\right]=0}{=} \mathrm{E}\left[\frac{\partial \ln f(X|\boldsymbol{\theta})}{\partial \theta_i}\right] &= \mathrm{E}\left[\frac{\partial \ln c(\boldsymbol{\theta})}{\partial \theta_i} + \sum_{j=1}^{k}\left\{\frac{\partial w_j(\boldsymbol{\theta})}{\partial \theta_i} t_j(X)\right\}\right] \\
&= \mathrm{E}\left[\frac{\partial \ln c(\boldsymbol{\theta})}{\partial \theta_i}\right] + \mathrm{E}\left[\sum_{j=1}^{k}\frac{\partial w_j(\boldsymbol{\theta})}{\partial \theta_i} t_j(X)\right] \\
&\overset{\mathrm{E}[c]=c}{=} \frac{\partial \ln c(\boldsymbol{\theta})}{\partial \theta_i} + \mathrm{E}\left[\sum_{j=1}^{k}\frac{\partial w_j(\boldsymbol{\theta})}{\partial \theta_i} t_j(X)\right] \\
&= \frac{\partial \ln c(\boldsymbol{\theta})}{\partial \theta_i} + \sum_{j=1}^{k}\mathrm{E}\left[\frac{\partial w_j(\boldsymbol{\theta})}{\partial \theta_i} t_j(X)\right] \\
&= \frac{\partial \ln c(\boldsymbol{\theta})}{\partial \theta_i} + \sum_{j=1}^{k}\frac{\partial w_j(\boldsymbol{\theta})}{\partial \theta_i} \mathrm{E}\left[t_j(X)\right] \\
0 &= \frac{\partial \ln c(\boldsymbol{\theta})}{\partial \theta_i} + \mathrm{E}\left[\sum_{j=1}^{k}\frac{\partial w_j(\boldsymbol{\theta})}{\partial \theta_i} t_j(X)\right] \\
\mathrm{E}\left[\sum_{j=1}^{k}\frac{\partial w_j(\boldsymbol{\theta})}{\partial \theta_i} t_j(X)\right] &= \frac{-\partial \ln c(\boldsymbol{\theta})}{\partial \theta_i}
\end{aligned}
$$

□

Theorem 15.16. *exponential family variance*

$$
\begin{gathered}
X \in \mathcal{E}^f = \left\{ f \middle| f = h(x)\, c(\boldsymbol{\theta}) \exp\left(\sum_{j=1}^{k} w_j(\boldsymbol{\theta})\, t_j(x)\right)\right\} \\
\Downarrow \forall i = 1, \cdots, p \\
\mathrm{V}\left[\sum_{j=1}^{k}\frac{\partial w_j(\boldsymbol{\theta})}{\partial \theta_i} t_j(X)\right] = \frac{-\partial^2 \ln c(\boldsymbol{\theta})}{\partial \theta_i^2} - \mathrm{E}\left[\sum_{j=1}^{k}\frac{\partial^2 w_j(\boldsymbol{\theta})}{\partial \theta_i^2} t_j(X)\right]
\end{gathered}
$$

Proof:

same as the above

$$
\begin{aligned}
f(x|\boldsymbol{\theta}) =& h(x)\, c(\boldsymbol{\theta})\, \mathrm{e}^{\sum_{j=1}^{k} w_j(\boldsymbol{\theta}) t_j(x)} = h(x)\, c(\boldsymbol{\theta}) \exp\left(\sum_{j=1}^{k} w_j(\boldsymbol{\theta})\, t_j(x)\right) \\
\ln f(x|\boldsymbol{\theta}) =& \ln h(x)\, c(\boldsymbol{\theta})\, \mathrm{e}^{\sum_{j=1}^{k} w_j(\boldsymbol{\theta}) t_j(x)} \\
=& \ln h(x) + \ln c(\boldsymbol{\theta}) + \ln \mathrm{e}^{\sum_{j=1}^{k} w_j(\boldsymbol{\theta}) t_j(x)} \\
=& \ln h(x) + \ln c(\boldsymbol{\theta}) + \sum_{j=1}^{k} w_j(\boldsymbol{\theta})\, t_j(x) \\
\ln f(x|\boldsymbol{\theta}) =& \ln h(x) + \ln c(\boldsymbol{\theta}) + \sum_{j=1}^{k} w_j(\boldsymbol{\theta})\, t_j(x) \\
\ln f(X|\boldsymbol{\theta}) =& \ln h(X) + \ln c(\boldsymbol{\theta}) + \sum_{j=1}^{k} w_j(\boldsymbol{\theta})\, t_j(X) \\
\frac{\partial}{\partial \theta_i} \ln f(X|\boldsymbol{\theta}) =& \frac{\partial}{\partial \theta_i}\left[\ln h(X) + \ln c(\boldsymbol{\theta}) + \sum_{j=1}^{k} w_j(\boldsymbol{\theta})\, t_j(X)\right] \\
=& \frac{\partial}{\partial \theta_i} \ln h(X) + \frac{\partial}{\partial \theta_i} \ln c(\boldsymbol{\theta}) + \frac{\partial}{\partial \theta_i} \sum_{j=1}^{k} w_j(\boldsymbol{\theta})\, t_j(X) \\
=& 0 + \frac{\partial}{\partial \theta_i} \ln c(\boldsymbol{\theta}) + \sum_{j=1}^{k} \frac{\partial}{\partial \theta_i}\left\{w_j(\boldsymbol{\theta})\, t_j(X)\right\} \\
=& \frac{\partial \ln c(\boldsymbol{\theta})}{\partial \theta_i} + \sum_{j=1}^{k}\left\{\left(\frac{\partial}{\partial \theta_i} w_j(\boldsymbol{\theta})\right) t_j(X) + w_j(\boldsymbol{\theta})\left(\frac{\partial}{\partial \theta_i} t_j(X)\right)\right\} \\
=& \frac{\partial \ln c(\boldsymbol{\theta})}{\partial \theta_i} + \sum_{j=1}^{k}\left\{\left(\frac{\partial}{\partial \theta_i} w_j(\boldsymbol{\theta})\right) t_j(X) + w_j(\boldsymbol{\theta})\, 0\right\} \\
=& \frac{\partial \ln c(\boldsymbol{\theta})}{\partial \theta_i} + \sum_{j=1}^{k}\left\{\left(\frac{\partial}{\partial \theta_i} w_j(\boldsymbol{\theta})\right) t_j(X)\right\} \\
\frac{\partial \ln f(X|\boldsymbol{\theta})}{\partial \theta_i} =& \frac{\partial \ln c(\boldsymbol{\theta})}{\partial \theta_i} + \sum_{j=1}^{k}\left\{\frac{\partial w_j(\boldsymbol{\theta})}{\partial \theta_i} t_j(X)\right\}
\end{aligned}
$$

$$
\begin{aligned}
\frac{\partial \ln f(X|\boldsymbol{\theta})}{\partial \theta_i} =& \frac{\partial \ln c(\boldsymbol{\theta})}{\partial \theta_i} + \sum_{j=1}^{k} \frac{\partial w_j(\boldsymbol{\theta})}{\partial \theta_i} t_j(X) \\
\mathrm{V}\left[\frac{\partial \ln f(X|\boldsymbol{\theta})}{\partial \theta_i}\right] =& \mathrm{V}\left[\frac{\partial \ln c(\boldsymbol{\theta})}{\partial \theta_i} + \sum_{j=1}^{k} \frac{\partial w_j(\boldsymbol{\theta})}{\partial \theta_i} t_j(X)\right] \\
\overset{\mathrm{V}[aX+b]=a^2\mathrm{V}[X]}{=}& \mathrm{V}\left[\sum_{j=1}^{k} \frac{\partial w_j(\boldsymbol{\theta})}{\partial \theta_i} t_j(X)\right] \\
\mathrm{V}\left[\sum_{j=1}^{k} \frac{\partial w_j(\boldsymbol{\theta})}{\partial \theta_i} t_j(X)\right] =& \mathrm{V}\left[\frac{\partial \ln f(X|\boldsymbol{\theta})}{\partial \theta_i}\right]
\end{aligned}
$$

$$
\begin{aligned}
\mathrm{V}\left[\sum_{j=1}^{k}\frac{\partial w_j(\boldsymbol{\theta})}{\partial\theta_i}t_j(X)\right] &= \mathrm{V}\left[\frac{\partial\ln f(X|\boldsymbol{\theta})}{\partial\theta_i}\right] = \mathrm{E}\left[\left(\frac{\partial\ln f(X|\boldsymbol{\theta})}{\partial\theta_i} - \mathrm{E}\left[\frac{\partial\ln f(X|\boldsymbol{\theta})}{\partial\theta_i}\right]\right)^2\right] \\
&\overset{\text{lemma: } \mathrm{E}\left[\frac{\partial\ln f(X|\boldsymbol{\theta})}{\partial\theta_i}\right]=0}{=} \mathrm{E}\left[\left(\frac{\partial\ln f(X|\boldsymbol{\theta})}{\partial\theta_i} - 0\right)^2\right] = \mathrm{E}\left[\left(\frac{\partial\ln f(X|\boldsymbol{\theta})}{\partial\theta_i}\right)^2\right] \\
&\overset{\text{lemma: } \mathrm{E}\left[\frac{\partial^2\ln f(X|\boldsymbol{\theta})}{\partial\theta_i^2}\right]=-\mathrm{E}\left[\left(\frac{\partial\ln f(X|\boldsymbol{\theta})}{\partial\theta_i}\right)^2\right]}{=} - \mathrm{E}\left[\frac{\partial^2\ln f(X|\boldsymbol{\theta})}{\partial\theta_i^2}\right] = -\mathrm{E}\left[\frac{\partial}{\partial\theta_i}\frac{\partial\ln f(X|\boldsymbol{\theta})}{\partial\theta_i}\right] \\
&\overset{\frac{\partial\ln f(X|\boldsymbol{\theta})}{\partial\theta_i}=\frac{\partial\ln c(\boldsymbol{\theta})}{\partial\theta_i}+\sum_{j=1}^{k}\frac{\partial w_j(\boldsymbol{\theta})}{\partial\theta_i}t_j(X)}{=} - \mathrm{E}\left[\frac{\partial}{\partial\theta_i}\left\{\frac{\partial\ln c(\boldsymbol{\theta})}{\partial\theta_i} + \sum_{j=1}^{k}\frac{\partial w_j(\boldsymbol{\theta})}{\partial\theta_i}t_j(X)\right\}\right] \\
&= - \mathrm{E}\left[\frac{\partial^2\ln c(\boldsymbol{\theta})}{\partial\theta_i^2} + \frac{\partial}{\partial\theta_i}\sum_{j=1}^{k}\frac{\partial w_j(\boldsymbol{\theta})}{\partial\theta_i}t_j(X)\right] \\
&= \frac{-\partial^2\ln c(\boldsymbol{\theta})}{\partial\theta_i^2} - \mathrm{E}\left[\sum_{j=1}^{k}\frac{\partial}{\partial\theta_i}\left\{\frac{\partial w_j(\boldsymbol{\theta})}{\partial\theta_i}t_j(X)\right\}\right] \\
&= \frac{-\partial^2\ln c(\boldsymbol{\theta})}{\partial\theta_i^2} - \mathrm{E}\left[\sum_{j=1}^{k}\left\{\left(\frac{\partial}{\partial\theta_i}\frac{\partial w_j(\boldsymbol{\theta})}{\partial\theta_i}\right)t_j(X) + \frac{\partial w_j(\boldsymbol{\theta})}{\partial\theta_i}\frac{\partial}{\partial\theta_i}t_j(X)\right\}\right] \\
&= \frac{-\partial^2\ln c(\boldsymbol{\theta})}{\partial\theta_i^2} - \mathrm{E}\left[\sum_{j=1}^{k}\left\{\frac{\partial^2 w_j(\boldsymbol{\theta})}{\partial\theta_i^2}t_j(X) + \frac{\partial w_j(\boldsymbol{\theta})}{\partial\theta_i}0\right\}\right] \\
&= \frac{-\partial^2\ln c(\boldsymbol{\theta})}{\partial\theta_i^2} - \mathrm{E}\left[\sum_{j=1}^{k}\frac{\partial^2 w_j(\boldsymbol{\theta})}{\partial\theta_i^2}t_j(X)\right] \\
\mathrm{V}\left[\sum_{j=1}^{k}\frac{\partial w_j(\boldsymbol{\theta})}{\partial\theta_i}t_j(X)\right] &= \frac{-\partial^2\ln c(\boldsymbol{\theta})}{\partial\theta_i^2} - \mathrm{E}\left[\sum_{j=1}^{k}\frac{\partial^2 w_j(\boldsymbol{\theta})}{\partial\theta_i^2}t_j(X)\right]
\end{aligned}
$$

$\square$

---

exponential family expectation and variance

sense of downgrading by differentiation instead of integration

$$
X \in \mathcal{E}^f = \left\{ f \,\middle|\, f = h(x)\,c(\boldsymbol{\theta})\exp\left(\sum_{j=1}^{k} w_j(\boldsymbol{\theta})\,t_j(x)\right)\right\}
$$

$$
\Rightarrow \begin{cases} \mathrm{E}\left[\sum_{j=1}^{k}\frac{\partial w_j(\boldsymbol{\theta})}{\partial\theta_i}t_j(X)\right] = \frac{-\partial\ln c(\boldsymbol{\theta})}{\partial\theta_i} \\ \mathrm{V}\left[\sum_{j=1}^{k}\frac{\partial w_j(\boldsymbol{\theta})}{\partial\theta_i}t_j(X)\right] = \frac{-\partial^2\ln c(\boldsymbol{\theta})}{\partial\theta_i^2} - \mathrm{E}\left[\sum_{j=1}^{k}\frac{\partial^2 w_j(\boldsymbol{\theta})}{\partial\theta_i^2}t_j(X)\right] \end{cases}, \forall i = 1, \cdots, p
$$

15.1.3 multivariable distribution

univariable random vector

discrete: PMF equals probability function

$$
f_X(x) = \mathrm{P}(X = x)
$$

continuous: PDF equals probability intensity

$$f_X(x)\,\mathrm{d}x = \mathrm{dP}(X \le x)$$

Given $(X,Y) = (X_1,X_2) = \langle X,Y\rangle = \langle X_1,X_2\rangle \sim f_{XY} = f_{XY}(x,y) = f_{X_1X_2} = f_{X_1X_2}(x_1,x_2)$

discrete:

Definition 15.19. JPMF = joint probability mass function: the JPMF of $(X_1,X_2) = \langle X_1,X_2\rangle$ is

$$f_{X_1X_2} = f_{X_1X_2}(x_1,x_2) = \mathrm{P}(X_1 = x_1, X_2 = x_2)$$

joint vs. marginal

Theorem 15.17. *joint probability mass function can inference marginal probability mass function: The marginal PMF of X_1, $f_{X_1}(x_1) = \mathrm{P}(X_1 = x_1)$ is given by*

$$f_{X_1}(x_1) = \sum_{x_2\in X_2(\Omega)} f_{X_1X_2}(x_1,x_2) = \sum_{x_2\in X_2(\Omega)} \mathrm{P}(X_1 = x_1, X_2 = x_2)$$

Proof:

$$\begin{aligned}
f_{X_1}(x_1) =& \mathrm{P}(X_1 = x_1)\\
=& \mathrm{P}(\{X_1 = x_1\} \cap \{X_2 \in (-\infty,\infty)\})\\
=& \mathrm{P}\left(\bigcup_{x_2\in(-\infty,\infty)} \{X_1 = x_1 \wedge X_2 = x_2\}\right)\\
=& \bigcup_{x_2\in(-\infty,\infty)} \mathrm{P}(X_1 = x_1 \wedge X_2 = x_2) = \bigcup_{x_2\in X_2(\Omega)} \mathrm{P}(X_1 = x_1, X_2 = x_2)\\
=& \bigcup_{x_2\in X_2(\Omega)} \mathrm{P}(X_1 = x_1, X_2 = x_2) = \sum_{x_2\in X_2(\Omega)} f_{X_1X_2}(x_1,x_2)
\end{aligned}$$

□

continuous:

Definition 15.20. JCDF = joint cumulative distribution function: the JCDF of $(X_1,X_2) = \langle X_1,X_2\rangle$ is

$$F_{X_1X_2} = F_{X_1X_2}(x_1,x_2) = \mathrm{P}(X_1 \le x_1, X_2 \le x_2)$$

Definition 15.21. JPDF = joint probability density function: the JPDF of $(X_1,X_2) = \langle X_1,X_2\rangle$ is

$$f_{X_1X_2} = f_{X_1X_2}(x_1,x_2) = \frac{\partial^2}{\partial x_1 \partial x_2} F_{X_1X_2}(x_1,x_2)$$

joint vs. marginal

Theorem 15.18. *joint cumulative distribution function can inference marginal cumulative distribution function: The marginal CDF of X_1 is*

$$F_{X_1}(x_1) = F_{X_1X_2}(x_1,\infty)$$

Proof:

$$\begin{aligned}
F_{X_1}(x_1) =& \mathrm{P}(X_1 \le x_1)\\
=& \mathrm{P}(X_1 \le x_1 \wedge X_2 \le \infty)\\
=& \mathrm{P}(X_1 \le x_1, X_2 \le \infty)\\
=& F_{X_1X_2}(x_1,\infty)
\end{aligned}$$

□

Note:

$F_{X_1}(x_1)$

$$F_{X_1}(x_1) = F_{X_1X_2}(x_1, \infty) = \int_{-\infty}^{\infty}\int_{-\infty}^{x_1} f_{X_1X_2}(u_1, u_2)\,\mathrm{d}u_1\mathrm{d}u_2$$

$$F_{X_1}(x_1) = \int_{-\infty}^{\infty}\int_{-\infty}^{x_1} f_{X_1X_2}(u_1, u_2)\,\mathrm{d}u_1\mathrm{d}u_2$$

$f_{X_1}(x_1)$

According to the fundamental theorem of calculus,

$$f_{X_1}(x_1) = \frac{\mathrm{d}}{\mathrm{d}x_1}F_{X_1}(x_1) = \frac{\mathrm{d}}{\mathrm{d}x_1}F_{X_1X_2}(x_1, \infty) = \frac{\mathrm{d}}{\mathrm{d}x_1}\int_{-\infty}^{\infty}\int_{-\infty}^{x} f_{X_1X_2}(u_1, u_2)\,\mathrm{d}u_1\mathrm{d}u_2 = \int_{-\infty}^{\infty} f_{X_1X_2}(x_1, u_2)\,\mathrm{d}u_2$$

$$f_{X_1}(x_1) = \int_{-\infty}^{\infty} f_{X_1X_2}(x_1, u_2)\,\mathrm{d}u_2$$

$$f_{X_1}(x_1) = \int_{-\infty}^{\infty} f_{X_1X_2}(x_1, x_2)\,\mathrm{d}x_2$$

$$f_{X_1}(x_1) = \begin{cases} \int_{-\infty}^{\infty} f_{X_1X_2}(x_1, x_2)\,\mathrm{d}x_2 & \text{continuous} \\ \sum\limits_{x_2 \in X_2(\Omega)} f_{X_1X_2}(x_1, x_2) & \text{discrete} \end{cases}$$

Theorem 15.19. *A function $f(x, y)$ is a joint PDF/PMF or JPDF/JPMF iff*

$$\begin{cases} f(x, y) \geq 0 & (ne)\ \text{non-negative} \\ \begin{cases} \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f(x, y)\,\mathrm{d}x\mathrm{d}y = 1 & \text{continuous} \\ \sum\limits_{x \in X(\Omega_X)}\sum\limits_{y \in Y(\Omega_Y)} f(x, y) = 1 & \text{discrete} \end{cases} & (1)\ \text{total event} \end{cases}$$

Definition 15.22. expected value: The expected value of a random vector $g(X, Y)$ is

$$\mathrm{E}_{X,Y}[g(X, Y)] = \mathrm{E}[g(X, Y)] = \begin{cases} \int\limits_{-\infty}^{+\infty}\int\limits_{-\infty}^{+\infty} g(x, y) f_{XY}(x, y)\,\mathrm{d}x\mathrm{d}y & \text{continuous} \\ \sum\limits_{x \in X(\Omega_X)}\sum\limits_{y \in Y(\Omega_Y)} g(x, y) f_{XY}(x, y) & \text{discrete} \end{cases}$$

Def: 15.14

$$\mathrm{P}((X, Y) \in E) = \mathrm{E}[1((X, Y) \in E)] = \begin{cases} \int_{y \in Y(\Omega_Y)}\int_{x \in X(\Omega_X)} 1((X, Y) \in E) f_{XY}(x, y)\,\mathrm{d}x\mathrm{d}y & \text{continuous} \\ \sum\limits_{x \in X(\Omega_X)}\sum\limits_{y \in Y(\Omega_Y)} 1((X, Y) \in E)\, g(x, y) f_{XY}(x, y) & \text{discrete} \end{cases}$$
$$= \begin{cases} \iint_{(X,Y) \in E} f_{XY}(x, y)\,\mathrm{d}x\mathrm{d}y & \text{continuous} \\ \sum\sum\limits_{(X,Y) \in E} f_{XY}(x, y) & \text{discrete} \end{cases}$$

---

example

$f_{XY}(x, y) = \mathrm{e}^{-y}1(0 < x < y < \infty) \Rightarrow \mathrm{P}(X + Y \leq 1)$

Check if

$$f_{XY}(x, y) = \mathrm{e}^{-y}1(0 < x < y < \infty)$$

is a JPDF:

$$\begin{cases} f_{XY}(x,y) = \mathrm{e}^{-y} 1(0<x<y<\infty) & \overset{\begin{cases}\mathrm{e}^{-y}>0 \\ 1(0<x<y<\infty) = \begin{cases}0\\1\end{cases}\end{cases}}{\geq} 0 & (ne)\,\text{non-negative} \\ \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f_{XY}(x,y)\,\mathrm{d}x\mathrm{d}y = ? & & (1)\,\text{total event} \end{cases}$$

$$\begin{aligned}
&\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f_{XY}(x,y)\,\mathrm{d}x\mathrm{d}y = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \mathrm{e}^{-y}1(0<x<y<\infty)\,\mathrm{d}x\mathrm{d}y, x\in(0,\infty) \\
&= \int_{-\infty}^{\infty}\int_{0}^{\infty} \mathrm{e}^{-y}1(0<x<y<\infty)\,\mathrm{d}x\mathrm{d}y, y>x\in(0,\infty) \\
&= \int_{0}^{\infty}\int_{0}^{\infty} \mathrm{e}^{-y}1(0<x<y<\infty)\,\mathrm{d}x\mathrm{d}y \overset{\text{Fubini}}{=} \int_{0}^{\infty}\int_{0}^{\infty} \mathrm{e}^{-y}1(0<x<y<\infty)\,\mathrm{d}y\mathrm{d}x, y>x \\
&= \int_{0}^{\infty}\int_{x}^{\infty} \mathrm{e}^{-y}1\mathrm{d}y\mathrm{d}x = \int_{0}^{\infty}\int_{x}^{\infty} \mathrm{e}^{-y}\mathrm{d}y\mathrm{d}x = \int_{0}^{\infty}\left[-\mathrm{e}^{-y}\right]_{y=x}^{\infty}\mathrm{d}x = \int_{0}^{\infty}\left[-\mathrm{e}^{-\infty}-\left(-\mathrm{e}^{-x}\right)\right]\mathrm{d}x \\
&= \int_{0}^{\infty}\left[-0-\left(-\mathrm{e}^{-x}\right)\right]\mathrm{d}x = \int_{0}^{\infty}\mathrm{e}^{-x}\mathrm{d}x = \left[-\mathrm{e}^{-x}\right]_{x=0}^{\infty} = \left[-\mathrm{e}^{-\infty}-\left(-\mathrm{e}^{-0}\right)\right] = [-0-(-1)] = 1
\end{aligned}$$

$$\begin{cases} f_{XY}(x,y) = \mathrm{e}^{-y}1(0<x<y<\infty) \geq 0 & (ne)\,\text{non-negative} \\ \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f_{XY}(x,y)\,\mathrm{d}x\mathrm{d}y = 1 & (1)\,\text{total event} \end{cases}$$

$\mathrm{P}(X+Y\leq 1)$

https://tex.stackexchange.com/questions/75933/how-to-draw-the-region-of-inequality

https://tex.stackexchange.com/questions/352511/how-to-fill-in-inequality-where-all-inequalities-overlap

interpolation dashed lines[13.4.6]

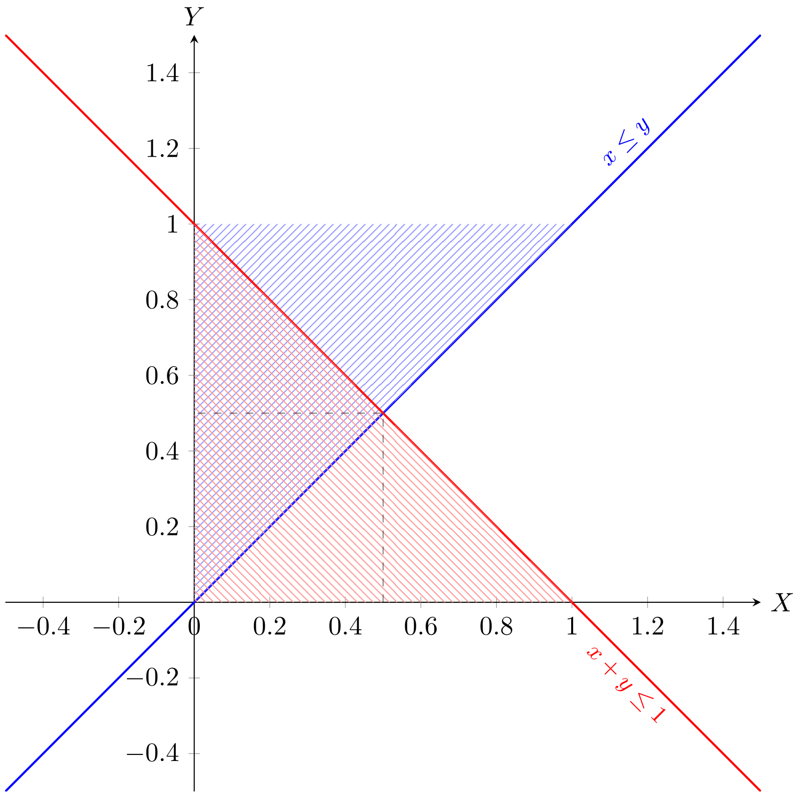


Figure 15.3: $X+Y\leq 1$

$$\begin{aligned}
\mathrm{P}\left(X+Y\leq 1\right)=&\iint_{(X,Y)\in E} f_{XY}\left(x,y\right)\mathrm{d}x\mathrm{d}y, E=\{X+Y\leq 1\}\\
=&\iint_{E}\mathrm{e}^{-y}1\left(0<x<y<\infty\right)\mathrm{d}x\mathrm{d}y, E=\{X+Y\leq 1\}\\
=&\int_0^{0.5}\int_x^{1-x}\mathrm{e}^{-y}\mathrm{d}y\mathrm{d}x=\int_0^{0.5}\left[-\mathrm{e}^{-y}\right]_{y=x}^{1-x}\mathrm{d}x=\int_0^{0.5}\left[-\mathrm{e}^{-(1-x)}-\left(-\mathrm{e}^{-x}\right)\right]\mathrm{d}x\\
=&\int_0^{0.5}\left[\mathrm{e}^{-x}-\mathrm{e}^{x-1}\right]\mathrm{d}x=\left[-\mathrm{e}^{-x}-\mathrm{e}^{x-1}\right]_{x=0}^{0.5}=\left[-\mathrm{e}^{-0.5}-\mathrm{e}^{0.5-1}\right]-\left[-\mathrm{e}^{-0}-\mathrm{e}^{0-1}\right]\\
=&1+\mathrm{e}^{-1}-2\mathrm{e}^{-0.5}\risingdotseq 0.154818\cdots
\end{aligned}$$

□

15.2 Chen, Lin-An

https://www.youtube.com/playlist?list=PLTpF-A8hKVUPXtNAX9lro-leGgEK0OSEW

https://en.wikipedia.org/wiki/Inverse_distance_weighting Shepard interpolation

Chapter 16

covariance matrix

16.1 vector direct product

- scalar = rank-0 tensor
- vector = rank-1 tensor
- matrix = rank-2 tensor
- vector direct product = rank-1 tensor times rank-1 tensor equals rank-2 tensor: increasing rank
- vector inner product = rank-1 tensor times rank-1 tensor equals rank-0 tensor: decreasing rank

scalar = rank-0 tensor

vector = rank-1 tensor

matrix = rank-2 tensor

16.1.1 vector direct product: increasing rank

vector direct product = rank-1 tensor times rank-1 tensor equals rank-2 tensor: increasing rank

$$\begin{aligned} U \otimes V &= UV^\intercal = U_i V_j \\ = \begin{pmatrix} U_1 \\ U_2 \\ U_3 \end{pmatrix} \otimes \begin{pmatrix} V_1 \\ V_2 \\ V_3 \end{pmatrix} &= \begin{pmatrix} U_1 \\ U_2 \\ U_3 \end{pmatrix} \begin{pmatrix} V_1 & V_2 & V_3 \end{pmatrix} = \begin{pmatrix} U_1V_1 & U_1V_2 & U_1V_3 \\ U_2V_1 & U_2V_2 & U_2V_3 \\ U_3V_1 & U_3V_2 & U_3V_3 \end{pmatrix} \\ &= \begin{pmatrix} U_1V^\intercal \\ U_2V^\intercal \\ U_3V^\intercal \end{pmatrix} = \begin{pmatrix} UV_1 & UV_2 & UV_3 \end{pmatrix} \end{aligned}$$

16.1.2 vector inner product: decreasing rank

vector inner product = rank-1 tensor times rank-1 tensor equals rank-0 tensor: decreasing rank

$$\begin{aligned} U \cdot V &= V^\intercal U = V_i U_i \\ = \begin{pmatrix} U_1 \\ U_2 \\ U_3 \end{pmatrix} \cdot \begin{pmatrix} V_1 \\ V_2 \\ V_3 \end{pmatrix} &= \begin{pmatrix} V_1 & V_2 & V_3 \end{pmatrix} \begin{pmatrix} U_1 \\ U_2 \\ U_3 \end{pmatrix} = V_1U_1 + V_2U_2 + V_3U_3 \end{aligned}$$

16.1.3 tensor direct product: increasing rank

$$\begin{aligned}
S \otimes T &= S_{ik} T_{jl} && (16.1)\\
(i,j),(k,l) &\in \{(1,1),(1,2),(2,1),(2,2)\} && (16.2)\\
\begin{pmatrix} S_{11} & S_{12} \\ S_{21} & S_{22} \end{pmatrix} \otimes \begin{pmatrix} T_{11} & T_{12} \\ T_{21} & T_{22} \end{pmatrix} &= \begin{pmatrix} S_{11}T_{11} & S_{11}T_{12} & S_{12}T_{11} & S_{12}T_{12} \\ S_{11}T_{21} & S_{11}T_{22} & S_{12}T_{21} & S_{12}T_{22} \\ S_{21}T_{11} & S_{21}T_{12} & S_{22}T_{11} & S_{22}T_{12} \\ S_{21}T_{21} & S_{21}T_{22} & S_{22}T_{21} & S_{22}T_{22} \end{pmatrix} && (16.3)\\
\begin{pmatrix} S_{11} & S_{12} \\ S_{21} & S_{22} \end{pmatrix} \begin{pmatrix} T_{11} & T_{12} \\ T_{21} & T_{22} \end{pmatrix} &= \begin{pmatrix} S_{11}T & S_{12}T \\ S_{21}T & S_{22}T \end{pmatrix} && (16.4)\\
&= \begin{pmatrix} S_{11}\begin{pmatrix} T_{11} & T_{12} \\ T_{21} & T_{22} \end{pmatrix} & S_{12}\begin{pmatrix} T_{11} & T_{12} \\ T_{21} & T_{22} \end{pmatrix} \\ S_{21}\begin{pmatrix} T_{11} & T_{12} \\ T_{21} & T_{22} \end{pmatrix} & S_{22}\begin{pmatrix} T_{11} & T_{12} \\ T_{21} & T_{22} \end{pmatrix} \end{pmatrix} && (16.5)
\end{aligned}$$

16.2 covariance matrix and its properties

8

$$\begin{aligned}
\mathrm{C}[\boldsymbol{X}] = \mathrm{Cov}[\boldsymbol{X}] = \mathrm{V}[\boldsymbol{X}] =& \mathrm{E}\left[[\boldsymbol{X} - \mathrm{E}(\boldsymbol{X})][\boldsymbol{X} - \mathrm{E}(\boldsymbol{X})]^\intercal\right]\\
=& \mathrm{E}\left[[\boldsymbol{X} - \mathrm{E}(\boldsymbol{X})][\boldsymbol{X}^\intercal - \mathrm{E}(\boldsymbol{X})^\intercal]\right]\\
=& \mathrm{E}\left[\boldsymbol{X}\boldsymbol{X}^\intercal - \mathrm{E}(\boldsymbol{X})\boldsymbol{X}^\intercal - \boldsymbol{X}\mathrm{E}(\boldsymbol{X})^\intercal + \mathrm{E}(\boldsymbol{X})\mathrm{E}(\boldsymbol{X})^\intercal\right]\\
=& \mathrm{E}[\boldsymbol{X}\boldsymbol{X}^\intercal] - \mathrm{E}[\mathrm{E}(\boldsymbol{X})\boldsymbol{X}^\intercal] - \mathrm{E}[\boldsymbol{X}\mathrm{E}(\boldsymbol{X})^\intercal] + \mathrm{E}[\mathrm{E}(\boldsymbol{X})\mathrm{E}(\boldsymbol{X})^\intercal]\\
=& \mathrm{E}[\boldsymbol{X}\boldsymbol{X}^\intercal] - \mathrm{E}(\boldsymbol{X})\mathrm{E}[\boldsymbol{X}^\intercal] - \mathrm{E}[\boldsymbol{X}]\mathrm{E}(\boldsymbol{X})^\intercal + \mathrm{E}(\boldsymbol{X})\mathrm{E}(\boldsymbol{X})^\intercal\\
=& \mathrm{E}[\boldsymbol{X}\boldsymbol{X}^\intercal] - \mathrm{E}(\boldsymbol{X})\mathrm{E}(\boldsymbol{X})^\intercal - \mathrm{E}(\boldsymbol{X})\mathrm{E}(\boldsymbol{X})^\intercal + \mathrm{E}(\boldsymbol{X})\mathrm{E}(\boldsymbol{X})^\intercal\\
=& \mathrm{E}[\boldsymbol{X}\boldsymbol{X}^\intercal] - \mathrm{E}(\boldsymbol{X})\mathrm{E}(\boldsymbol{X})^\intercal
\end{aligned}$$

$$\begin{aligned}
\boldsymbol{X} = [X]_{1\times 1} = X \Rightarrow C(X) = \mathrm{C}[\boldsymbol{X}] =& \mathrm{E}[\boldsymbol{X}\boldsymbol{X}^\intercal] - \mathrm{E}(\boldsymbol{X})\mathrm{E}(\boldsymbol{X})^\intercal\\
=& \mathrm{E}[XX] - \mathrm{E}(X)\mathrm{E}(X)\\
=& \mathrm{E}\left(X^2\right) - [\mathrm{E}(X)]^2 = \mathrm{V}(X)
\end{aligned}$$

16.2.1 $\mathrm{V}[\boldsymbol{X} + \boldsymbol{b}] = \mathrm{V}[\boldsymbol{X}]$

$$\begin{aligned}
\mathrm{V}[\boldsymbol{X} + \boldsymbol{b}] =& \mathrm{E}\left[[(\boldsymbol{X} + \boldsymbol{b}) - \mathrm{E}(\boldsymbol{X} + \boldsymbol{b})][(\boldsymbol{X} + \boldsymbol{b}) - \mathrm{E}(\boldsymbol{X} + \boldsymbol{b})]^\intercal\right]\\
\overset{\mathrm{E}(\boldsymbol{X}+\boldsymbol{b})=\mathrm{E}(\boldsymbol{X})+\boldsymbol{b}}{=}& \mathrm{E}\left[[\boldsymbol{X} + \boldsymbol{b} - \mathrm{E}(\boldsymbol{X}) - \boldsymbol{b}][\boldsymbol{X} + \boldsymbol{b} - \mathrm{E}(\boldsymbol{X}) - \boldsymbol{b}]^\intercal\right]\\
=& \mathrm{E}\left[[\boldsymbol{X} - \mathrm{E}(\boldsymbol{X})][\boldsymbol{X} - \mathrm{E}(\boldsymbol{X})]^\intercal\right] = \mathrm{V}[\boldsymbol{X}]
\end{aligned}$$

16.2.2 $\mathrm{V}[A\boldsymbol{X}] = A\mathrm{V}[\boldsymbol{X}]A^\intercal$

$$\begin{aligned}
\mathrm{V}[A\boldsymbol{X}] =& \mathrm{E}\left[[(A\boldsymbol{X}) - \mathrm{E}(A\boldsymbol{X})][(A\boldsymbol{X}) - \mathrm{E}(A\boldsymbol{X})]^\intercal\right]\\
\overset{\mathrm{E}(A\boldsymbol{X})=A\mathrm{E}(\boldsymbol{X})}{=}& \mathrm{E}\left[[A\boldsymbol{X} - A\mathrm{E}(\boldsymbol{X})][A\boldsymbol{X} - A\mathrm{E}(\boldsymbol{X})]^\intercal\right]\\
=& \mathrm{E}\left[A[\boldsymbol{X} - \mathrm{E}(\boldsymbol{X})][A[\boldsymbol{X} - \mathrm{E}(\boldsymbol{X})]]^\intercal\right]\\
=& \mathrm{E}\left[A[\boldsymbol{X} - \mathrm{E}(\boldsymbol{X})][\boldsymbol{X} - \mathrm{E}(\boldsymbol{X})]^\intercal A^\intercal\right]\\
=& A\mathrm{E}\left[[\boldsymbol{X} - \mathrm{E}(\boldsymbol{X})][\boldsymbol{X} - \mathrm{E}(\boldsymbol{X})]^\intercal\right]A^\intercal = A\mathrm{V}[\boldsymbol{X}]A^\intercal
\end{aligned}$$

16.2.3 $\mathrm{V}[A\boldsymbol{X} + \boldsymbol{b}] = A\mathrm{V}[\boldsymbol{X}]A^\intercal$

$$\mathrm{V}[A\boldsymbol{X} + \boldsymbol{b}] = \mathrm{V}[A\boldsymbol{X}] = A\mathrm{V}[\boldsymbol{X}]A^\intercal$$

Chapter 17

Gosper algorithm

Chapter 18

Lorentz transformation

18.1 Einstein

https://wap.hillpublisher.com/UpFile/202204/20220414165340.pdf

18.2 Bondi k-calculus

https://en.wikipedia.org/wiki/Bondi_k-calculus

18.3 wordline in Minkowski space

18.3.1 Wick rotation

https://ncatlab.org/nlab/show/Wick+rotation

18.3.1.1 Osterwalder-Schrader reconstruction theorem

https://ncatlab.org/nlab/show/Osterwalder-Schrader+theorem

Chapter 19

R

19.1 TonyKuoYJ

郭耀仁 認識 R 的美好

https://bookdown.org/tonykuoyj/eloquentr/getting-started.html

https://bookdown.org/tonykuoyj/eloquentr/easy-installation.html#about-packages

`install.pacakges()`

`library()`

https://bookdown.org/tonykuoyj/eloquentr/getting-started.html

19.1.1 quick intro

`Ctrl + Alt + I` to insert a new code chunk in RStudio

`Ctrl + Enter` to run the current line

`Ctrl + Shift + Enter` to run the current chunk

```
R.version
```

```
##                _
## platform       x86_64-w64-mingw32
## arch           x86_64
## os             mingw32
## crt            ucrt
## system         x86_64, mingw32
## status
## major          4
## minor          2.1
## year           2022
## month          06
## day            23
## svn rev        82513
## language       R
## version.string R version 4.2.1 (2022-06-23 ucrt)
## nickname       Funny-Looking Kid
```

```
a <- 23 # prime
a
```

```
## [1] 23
```

```
combine <- c(11, 13) # twin prime
combine
```

```
## [1] 11 13
```

```r
# ?c
# help(c)
```

`Ctrl + L` to clean R console

path with slash `/` in R, differing backslash `\` in M$ Windows

19.1.1.1 function

```r
add <- function(x, y) {
  return(x + y)
}

add(11, 13)
```

```
## [1] 24
```

$$BMI = \frac{BW\ [\mathrm{Kg}]}{BH\ [\mathrm{m}]^2}$$

```r
get_bmi <- function (bw, bh) {
  return (bw/(bh/100)^2)
}

get_bmi(70, 170)
```

```
## [1] 24.22145
```

19.1.2 R style

https://bookdown.org/tonykuoyj/eloquentr/styleguide.html

snake_case rather than camelCase

19.1.3 data workflow or forward pipe

from *chaining method* in *object-oriented programming* to **functional programming**

19.1.3.1 %>% operator

```r
abs(-5:5)
```

```
##  [1] 5 4 3 2 1 0 1 2 3 4 5
```

```r
# install.packages("magrittr")

library(magrittr)
```

```
##
## Attaching package: 'magrittr'

## The following object is masked _by_ '.GlobalEnv':
##
##     add
```

```r
-5:5 %>% abs()
```

```
##  [1] 5 4 3 2 1 0 1 2 3 4 5
```

```r
# with readability but too many lines
sys_date <- Sys.Date()
sys_date_yr <- format(sys_date, format = "%Y")
sys_date_num <- as.numeric(sys_date_yr)
sys_date_num
```

```
## [1] 2024
```

```r
# less line but also less readability
sys_date_num <- as.numeric(format(Sys.Date(), format = "%Y"))
sys_date_num
```

```
## [1] 2024
```

```r
# use %>% operator to demonstrate data workflow or forward pipe
sys_date_num <- Sys.Date() %>%
   format(format = "%Y") %>%
   as.numeric()
sys_date_num
```

```
## [1] 2024
```

19.1.4 data processing with dplyr

https://bookdown.org/tonykuoyj/eloquentr/dplyr.html

some functions functioning like those in **SQL**

```r
library(dplyr)
```

```
## Warning: package 'dplyr' was built under R version 4.2.3
```

```
##
## Attaching package: 'dplyr'
```

```
## The following objects are masked from 'package:stats':
##
##     filter, lag
```

```
## The following objects are masked from 'package:base':
##
##     intersect, setdiff, setequal, union
```

```r
# install.packages("gapminder")

library(gapminder)
```

```
## Warning: package 'gapminder' was built under R version 4.2.3
```

```r
head(gapminder)
```

```
## # A tibble: 6 x 6
##   country     continent  year lifeExp      pop gdpPercap
##   <fct>       <fct>     <int>   <dbl>    <int>     <dbl>
## 1 Afghanistan Asia       1952    28.8  8425333      779.
## 2 Afghanistan Asia       1957    30.3  9240934      821.
## 3 Afghanistan Asia       1962    32.0 10267083      853.
## 4 Afghanistan Asia       1967    34.0 11537966      836.
## 5 Afghanistan Asia       1972    36.1 13079460      740.
## 6 Afghanistan Asia       1977    38.4 14880372      786.
```

```r
library(gapminder)
library(dplyr)
library(magrittr)

gapminder %>%
  filter(year == 2007)
```

```
## # A tibble: 142 x 6
##    country     continent  year lifeExp       pop gdpPercap
##    <fct>       <fct>     <int>   <dbl>     <int>     <dbl>
##  1 Afghanistan Asia       2007    43.8  31889923      975.
##  2 Albania     Europe     2007    76.4   3600523     5937.
##  3 Algeria     Africa     2007    72.3  33333216     6223.
```

```
## 4 Angola      Africa     2007    42.7  12420476     4797.
## 5 Argentina   Americas   2007    75.3  40301927    12779.
## 6 Australia   Oceania    2007    81.2  20434176    34435.
## 7 Austria     Europe     2007    79.8   8199783    36126.
## 8 Bahrain     Asia       2007    75.6    708573    29796.
## 9 Bangladesh  Asia       2007    64.1 150448339     1391.
## 10 Belgium    Europe     2007    79.4  10392226    33693.
## # i 132 more rows
```

```
library(gapminder)
library(dplyr)
library(magrittr)

gapminder %>%
  filter(year == 2007) %>%
  select(country)
```

```
## # A tibble: 142 x 1
##    country
##    <fct>
##  1 Afghanistan
##  2 Albania
##  3 Algeria
##  4 Angola
##  5 Argentina
##  6 Australia
##  7 Austria
##  8 Bahrain
##  9 Bangladesh
## 10 Belgium
## # i 132 more rows
```

```
library(gapminder)
library(dplyr)
library(magrittr)

gapminder %>%
  mutate(pop_in_thousands = pop / 1000)
```

```
## # A tibble: 1,704 x 7
##    country     continent  year lifeExp      pop gdpPercap pop_in_thousands
##    <fct>       <fct>     <int>   <dbl>    <int>     <dbl>            <dbl>
##  1 Afghanistan Asia       1952    28.8  8425333      779.             8425.
##  2 Afghanistan Asia       1957    30.3  9240934      821.             9241.
##  3 Afghanistan Asia       1962    32.0 10267083      853.            10267.
##  4 Afghanistan Asia       1967    34.0 11537966      836.            11538.
##  5 Afghanistan Asia       1972    36.1 13079460      740.            13079.
##  6 Afghanistan Asia       1977    38.4 14880372      786.            14880.
##  7 Afghanistan Asia       1982    39.9 12881816      978.            12882.
##  8 Afghanistan Asia       1987    40.8 13867957      852.            13868.
##  9 Afghanistan Asia       1992    41.7 16317921      649.            16318.
## 10 Afghanistan Asia       1997    41.8 22227415      635.            22227.
## # i 1,694 more rows
```

```
library(gapminder)
library(dplyr)
library(magrittr)

gapminder %>%
  arrange(year)
```

```
## # A tibble: 1,704 x 6
##    country     continent  year lifeExp      pop gdpPercap
##    <fct>       <fct>     <int>   <dbl>    <int>     <dbl>
```

```
##  1 Afghanistan Asia       1952    28.8  8425333      779.
##  2 Albania     Europe     1952    55.2  1282697     1601.
##  3 Algeria     Africa     1952    43.1  9279525     2449.
##  4 Angola      Africa     1952    30.0  4232095     3521.
##  5 Argentina   Americas   1952    62.5 17876956     5911.
##  6 Australia   Oceania    1952    69.1  8691212    10040.
##  7 Austria     Europe     1952    66.8  6927772     6137.
##  8 Bahrain     Asia       1952    50.9   120447     9867.
##  9 Bangladesh  Asia       1952    37.5 46886859      684.
## 10 Belgium     Europe     1952    68     8730405     8343.
## # i 1,694 more rows
```

total population in the world in 2007

```
library(gapminder)
library(dplyr)
library(magrittr)

gapminder %>%
  filter(year == 2007) %>%
  summarise(ttl_pop = sum(as.numeric(pop)))
```

```
## # A tibble: 1 x 1
##      ttl_pop
##        <dbl>
## 1 6251013179
```

total population group by the continents in 2007

```
library(gapminder)
library(dplyr)
library(magrittr)

gapminder %>%
  filter(year == 2007) %>%
  group_by(continent) %>%
  summarise(ttl_pop = sum(as.numeric(pop)))
```

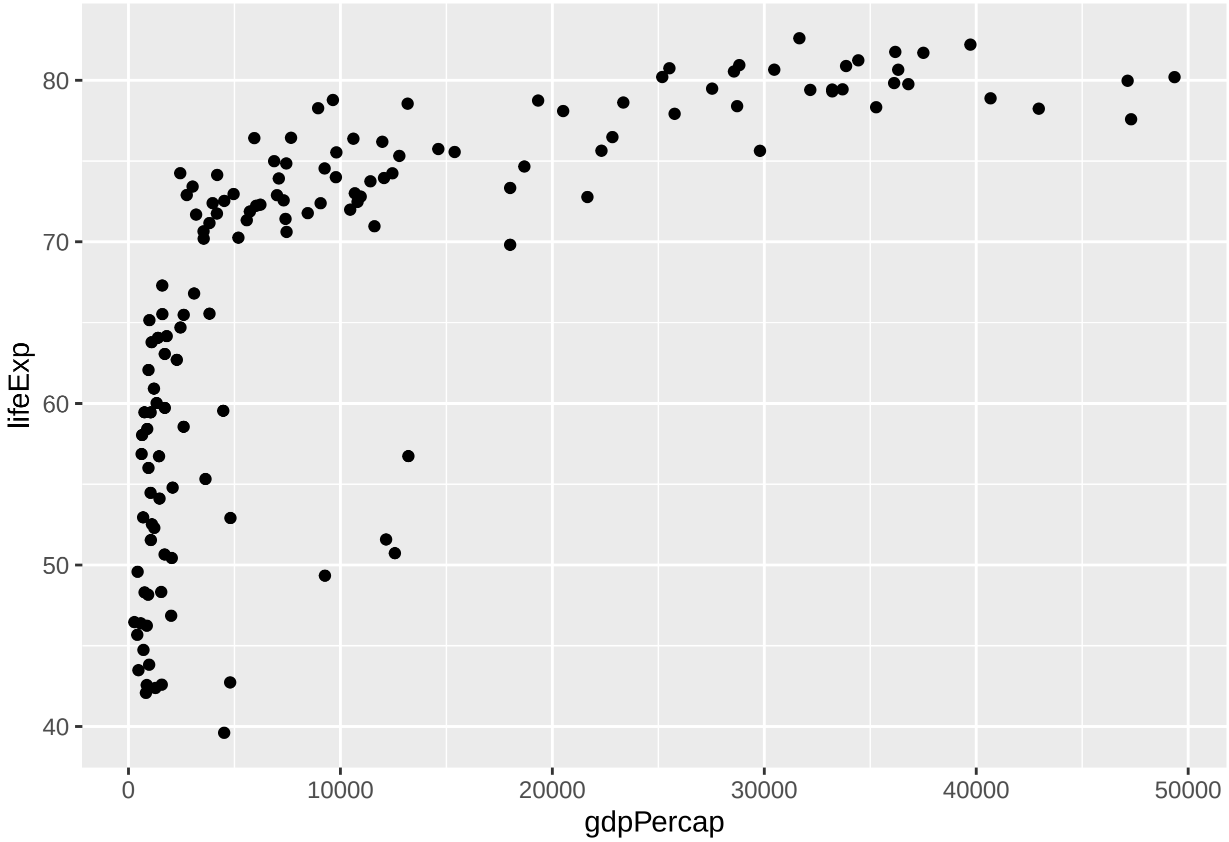
```
## # A tibble: 5 x 2
##   continent    ttl_pop
##   <fct>          <dbl>
## 1 Africa     929539692
## 2 Americas   898871184
## 3 Asia      3811953827
## 4 Europe     586098529
## 5 Oceania     24549947
```

19.1.5 visualization statically with ggplot2

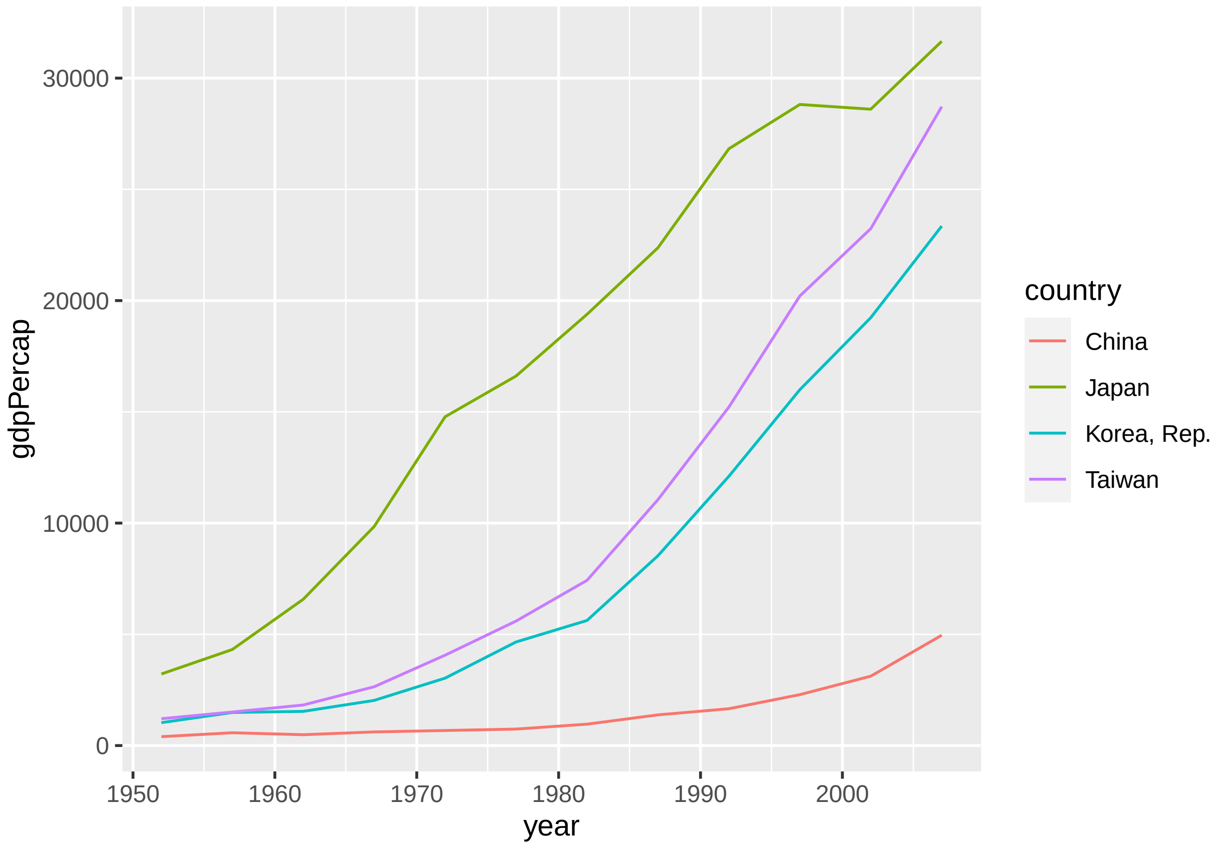
```
library(ggplot2)
```

```
## Warning: package 'ggplot2' was built under R version 4.2.3
```

```
library(gapminder)

gapminder_2007 <- gapminder %>%
  filter(year == 2007)
scatter_plot <- ggplot(gapminder_2007, aes(x = gdpPercap, y = lifeExp)) +
  geom_point()
scatter_plot
```

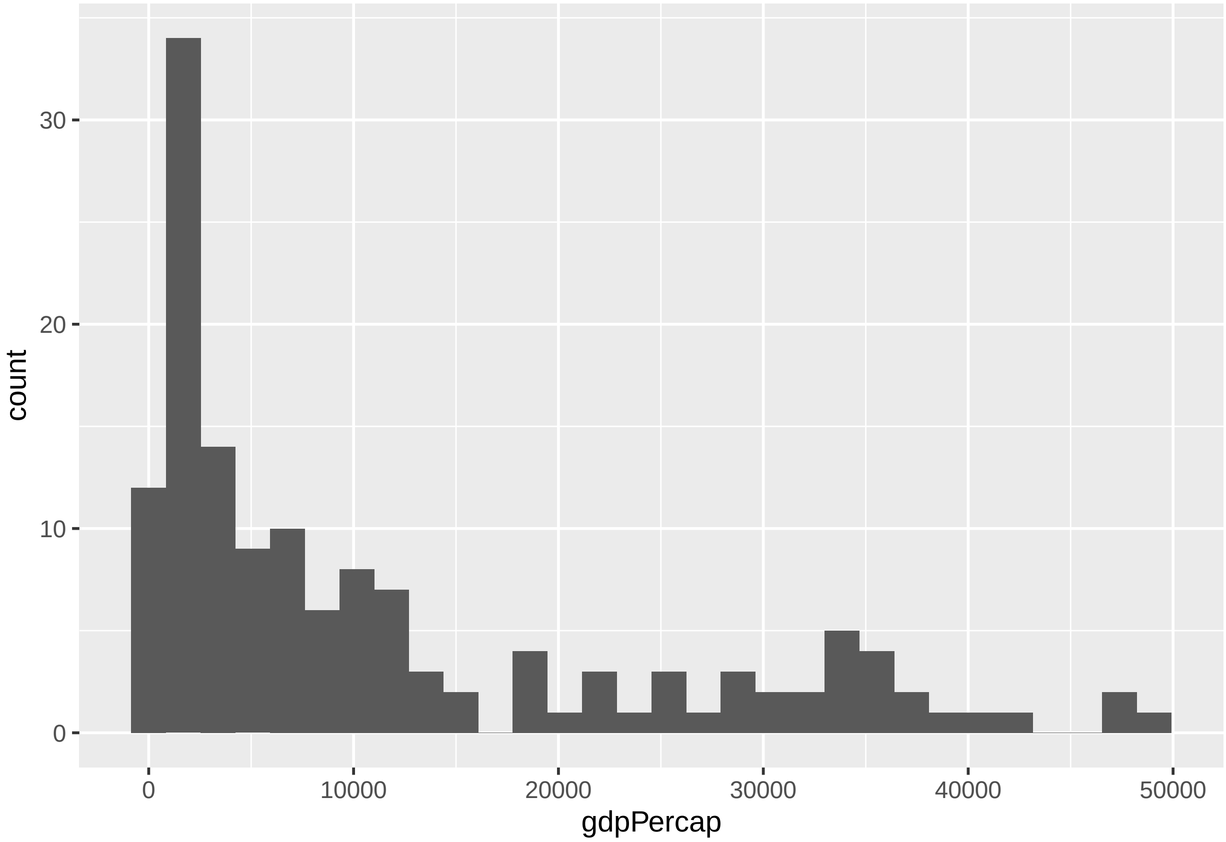



```r
library(ggplot2)
library(gapminder)

north_asia <- gapminder %>%
  filter(country %in% c("China", "Japan", "Taiwan", "Korea, Rep."))
line_plot <- ggplot(north_asia, aes(x = year, y = gdpPercap, colour = country)) +
  geom_line()
line_plot
```

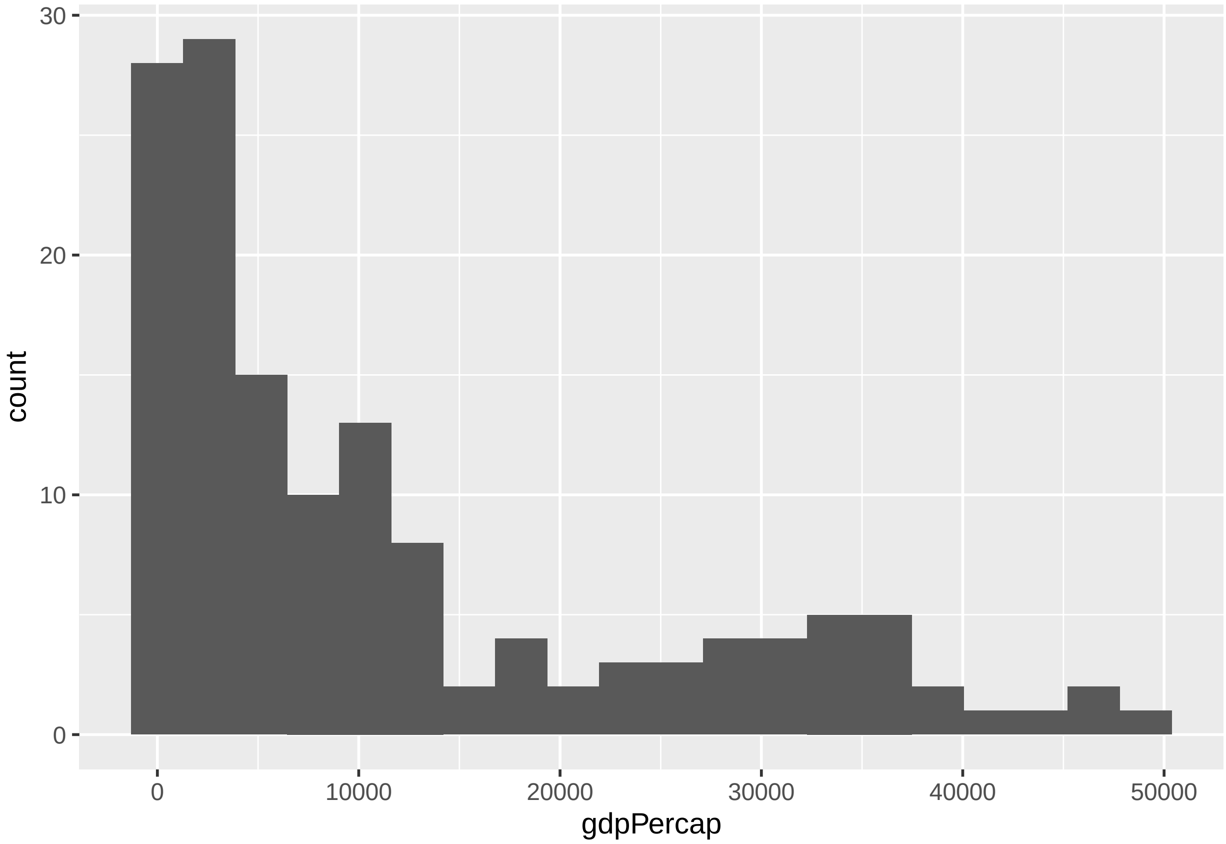



```
library(ggplot2)
library(gapminder)

hist_plot <- ggplot(gapminder_2007, aes(x = gdpPercap)) +
  geom_histogram()
hist_plot
```

```
## `stat_bin()` using `bins = 30`. Pick better value with `binwidth`.
```

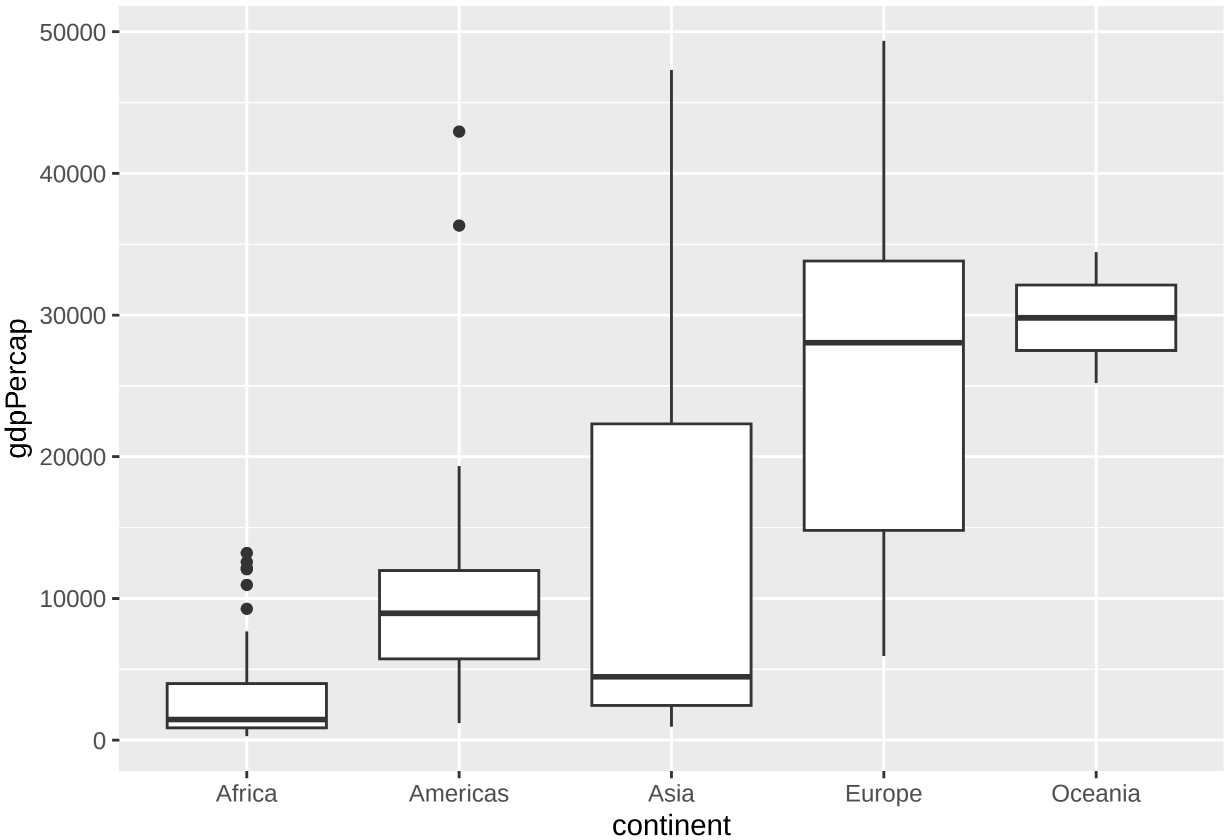



```r
hist_plot <- ggplot(gapminder_2007, aes(x = gdpPercap)) +
  geom_histogram(bins = 20)
hist_plot
```

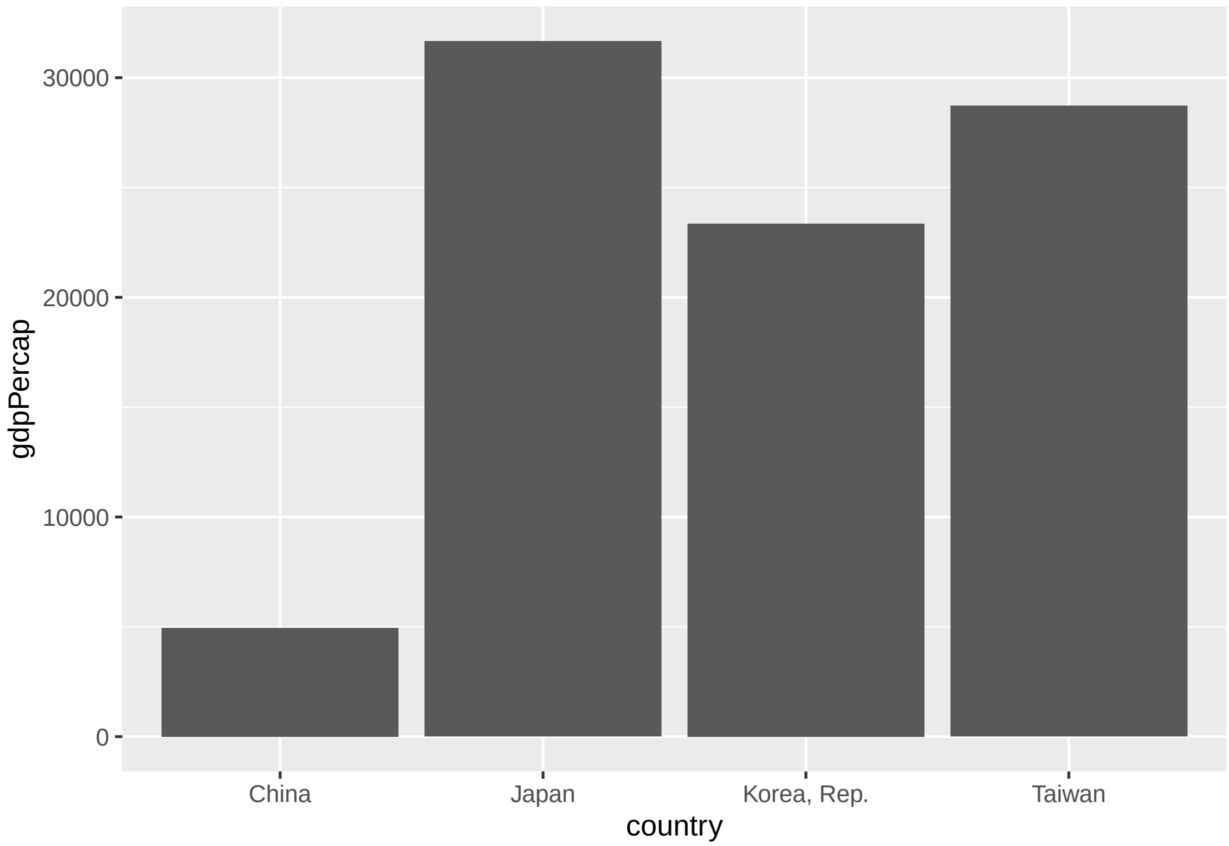



```r
library(ggplot2)
library(gapminder)
```

```
box_plot <- ggplot(gapminder_2007, aes(x = continent, y = gdpPercap)) +
  geom_boxplot()
box_plot
```




```
library(ggplot2)
library(gapminder)

gdpPercap_2007_na <- gapminder %>%
  filter(year == 2007 & country %in% c("China", "Japan", "Taiwan", "Korea, Rep."))
bar_plot <- ggplot(gdpPercap_2007_na, aes(x = country, y = gdpPercap)) +
  geom_bar(stat = "identity")
bar_plot
```



19.1.6 loop

https://bookdown.org/tonykuoyj/eloquentr/for.html

```
month.name
```

```
##  [1] "January"   "February"  "March"     "April"     "May"       "June"
##  [7] "July"      "August"    "September" "October"   "November"  "December"
```

```
month.name[1]
```

```
## [1] "January"
```

```
for (month in month.name) {
  print(month)
}
```

```
## [1] "January"
## [1] "February"
## [1] "March"
## [1] "April"
## [1] "May"
## [1] "June"
## [1] "July"
## [1] "August"
## [1] "September"
## [1] "October"
## [1] "November"
## [1] "December"
```

19.1.7 variable type

https://bookdown.org/tonykuoyj/eloquentr/variable-types.html

https://www.w3schools.com/r/r_data_types.asp

- numeric
- integer

- complex = complex number
- character
- logical = boolean

```r
class(2L)
```

```
## [1] "integer"
```

```r
class(2.0L)
```

```
## [1] "integer"
```

```r
class(2.3L)
```

```
## [1] "numeric"
```

time: POSIXct POSIXt

```r
class(Sys.time())
```

```
## [1] "POSIXct" "POSIXt"
```

```r
0 %in% -5:5
```

```
## [1] TRUE
```

19.1.7.1 date

1970-01-01 = 0L

```r
date_of_origin <- as.Date("1970-01-01")
as.integer(date_of_origin)
```

```
## [1] 0
```

check if type of `x` is `Date`

`inherits(x, what = "Date")`

convert `character` to `Date`

`as.Date("01-01-1970", format = "%m-%d-%Y")`

19.1.7.2 time

1970-01-01 00:00:00 GMT = 0L

tz = time zone

```r
time_of_origin <- as.POSIXct("1970-01-01 00:00:00", tz = "GMT")
as.integer(time_of_origin)
```

```
## [1] 0
```

check if type of `x` is time

`inherits(x, what = "POSIXct")`

convert `character` to time

`as.POSIXct("1970-01-01 00:00:00", tz = "GMT")`

19.1.7.3 quotient `%/%` operator

https://www.w3schools.com/r/r_operators.asp

```r
7 %/% 3
```

```
## [1] 2
```

19.1.8 data type

https://bookdown.org/tonykuoyj/eloquentr/vector-factor.html

- 1D
  - vector[19.1.8.1]
  - factor[19.1.8.2]
- 2D
  - matrix[19.1.8.3]
  - data frame[19.1.8.4]
- nD
  - array[19.2.6.1]
  - list[19.2.6.2]

19.1.8.1 vector

```
four_seasons <- c("spring", "summer", "autumn", "winter")
four_seasons
```

```
## [1] "spring" "summer" "autumn" "winter"
```

```
favorite_season <- four_seasons[3]
favorite_season
```

```
## [1] "autumn"
```

```
favorite_seasons <- four_seasons[c(-2, -4)]
favorite_seasons
```

```
## [1] "spring" "autumn"
```

only one variable type for a vector

```
lucky_numbers <- c(7L, 24)
class(lucky_numbers[1])
```

```
## [1] "numeric"
```

```
lucky_numbers <- c(7L, FALSE)
lucky_numbers
```

```
## [1] 7 0
```

```
class(lucky_numbers[2])
```

```
## [1] "integer"
```

```
mixed_vars <- c(TRUE, 7L, 24, "spring")
mixed_vars
```

```
## [1] "TRUE"   "7"      "24"     "spring"
```

```
class(mixed_vars[1])
```

```
## [1] "character"
```

```
class(mixed_vars[2])
```

```
## [1] "character"
```

```
class(mixed_vars[3])
```

```
## [1] "character"
```

```

four_seasons <- c("spring", "summer", "autumn", "winter")
my_favorite_seasons <- four_seasons == "spring" | four_seasons == "autumn"
four_seasons[my_favorite_seasons]
```

19.1.8.1.1 logic

```
## [1] "spring" "autumn"
```

```r
rep(7L, times = 8)
```

19.1.8.1.2 rep repeat

```
## [1] 7 7 7 7 7 7 7 7
```

```r
rep("R", times = 10)
```

```
##  [1] "R" "R" "R" "R" "R" "R" "R" "R" "R" "R"
```

```r
seq(from = 7, to = 77, by = 7)
```

19.1.8.1.3 seq sequence

```
##  [1]  7 14 21 28 35 42 49 56 63 70 77
```

```r
11:20
```

```
##  [1] 11 12 13 14 15 16 17 18 19 20
```

19.1.8.2 factor

https://bookdown.org/tonykuoyj/eloquentr/vector-factor.html#factor

```r
four_seasons <- c("spring", "summer", "autumn", "winter")
four_seasons
```

```
## [1] "spring" "summer" "autumn" "winter"
```

```r
four_seasons_factor <- factor(four_seasons)
four_seasons_factor
```

```
## [1] spring summer autumn winter
## Levels: autumn spring summer winter
```

```r
four_seasons <- c("spring", "summer", "autumn", "winter")
four_seasons_factor <- factor(four_seasons, ordered = TRUE, levels = c("summer", "winter", "spring",
↪  "autumn"))
four_seasons_factor
```

```
## [1] spring summer autumn winter
## Levels: summer < winter < spring < autumn
```

```r
temperatures <- c("warm", "hot", "cold")
temp_factors <- factor(temperatures, ordered = TRUE, levels = c("cold", "warm", "hot"))
temp_factors
```

```
## [1] warm hot  cold
## Levels: cold < warm < hot
```

if no levels specified, the levels will be specified alphabetically, sometimes not really true

```r
temperatures <- c("warm", "hot", "cold")
temp_factors <- factor(temperatures, ordered = TRUE)
temp_factors
```

```
## [1] warm hot  cold
## Levels: cold < hot < warm
```

19.1.8.3 matrix

https://bookdown.org/tonykuoyj/eloquentr/matrix-dataframe-more.html

```r
my_mat <- matrix(1:6, nrow = 2)
my_mat
```

```
##      [,1] [,2] [,3]
## [1,]    1    3    5
## [2,]    2    4    6
```

```r
class(my_mat)
```

```
## [1] "matrix" "array"
```

```r
my_mat2 <- matrix(1:6, nrow = 2, byrow = TRUE)
my_mat2
```

```
##      [,1] [,2] [,3]
## [1,]    1    2    3
## [2,]    4    5    6
```

```r
my_mat2[2, 3]
```

```
## [1] 6
```

```r
my_mat2[2, ]
```

```
## [1] 4 5 6
```

```r
my_mat2[, 3]
```

```
## [1] 3 6
```

```r
filter <- my_mat2 < 6 & my_mat2 > 1
my_mat2[filter]
```

```
## [1] 4 2 5 3
```

boolean will become value in a matrix, like vector

```r
my_mat3 <- matrix(c(1, 2, TRUE, FALSE, 3, 4), nrow = 2)
my_mat3
```

```
##      [,1] [,2] [,3]
## [1,]    1    1    3
## [2,]    2    0    4
```

```r
class(my_mat3[, 2])
```

```
## [1] "numeric"
```

19.1.8.4 data frame

- variable: column
- observation: row
- value: cell

```r
team_name <- c("Chicago Bulls", "Golden State Warriors")
wins <- c(72, 73)
losses <- c(10, 9)
is_champion <- c(TRUE, FALSE)
season <- c("1995-96", "2015-16")

great_nba_teams <- data.frame(team_name, wins, losses, is_champion, season)
great_nba_teams
```

```
##               team_name wins losses is_champion  season
## 1         Chicago Bulls   72     10        TRUE 1995-96
## 2 Golden State Warriors   73      9       FALSE 2015-16
```

```r
great_nba_teams[1, 1]
```

```
## [1] "Chicago Bulls"
```

```r
great_nba_teams[1, ]
```

```
##       team_name wins losses is_champion  season
## 1 Chicago Bulls   72     10        TRUE 1995-96
```

```r
great_nba_teams[, 1]
```

```
## [1] "Chicago Bulls"         "Golden State Warriors"
```

stringsAsFactors = TRUE

```r
team_name <- c("Chicago Bulls", "Golden State Warriors")
wins <- c(72, 73)
losses <- c(10, 9)
is_champion <- c(TRUE, FALSE)
season <- c("1995-96", "2015-16")

great_nba_teams <- data.frame(team_name, wins, losses, is_champion, season, stringsAsFactors = TRUE)
great_nba_teams[, 1]
```

```
## [1] Chicago Bulls         Golden State Warriors
## Levels: Chicago Bulls Golden State Warriors
```

stringsAsFactors = FALSE

```r
team_name <- c("Chicago Bulls", "Golden State Warriors")
wins <- c(72, 73)
losses <- c(10, 9)
is_champion <- c(TRUE, FALSE)
season <- c("1995-96", "2015-16")

great_nba_teams <- data.frame(team_name, wins, losses, is_champion, season, stringsAsFactors = FALSE)
great_nba_teams[, 1]
```

```
## [1] "Chicago Bulls"         "Golden State Warriors"
```

```r

great_nba_teams$team_name
```

19.1.8.4.1 selecting variable or column

```
## [1] "Chicago Bulls"         "Golden State Warriors"
```

```r
great_nba_teams[, "team_name"]
```

```
## [1] "Chicago Bulls"         "Golden State Warriors"
```

```r

filter <- great_nba_teams$is_champion == TRUE
great_nba_teams[filter, ]
```

19.1.8.4.2 filtering observation or row

```
##       team_name wins losses is_champion  season
## 1 Chicago Bulls   72     10        TRUE 1995-96
```

```r

str(great_nba_teams)
```

19.1.8.4.3 check mixed data type

```
## 'data.frame':    2 obs. of  5 variables:
##  $ team_name  : chr  "Chicago Bulls" "Golden State Warriors"
##  $ wins       : num  72 73
```

```
## $ losses     : num  10 9
## $ is_champion: logi  TRUE FALSE
## $ season     : chr  "1995-96" "2015-16"
```

19.2 W3School

https://www.w3schools.com/r/default.asp

19.2.1 same multiple variable

https://www.w3schools.com/r/r_variables_multiple.asp

```
# Assign the same value to multiple variables in one line
var1 <- var2 <- var3 <- "Orange"

# Print variable values
var1
```

```
## [1] "Orange"
```

```
var2
```

```
## [1] "Orange"
```

```
var3
```

```
## [1] "Orange"
```

19.2.2 legal variable name

https://www.w3schools.com/r/r_variables_name.asp

```
# Legal variable names:
myvar <- "John"
my_var <- "John"
myVar <- "John"
MYVAR <- "John"
myvar2 <- "John"
.myvar <- "John"

## Illegal variable names:
# 2myvar <- "John"
# my-var <- "John"
# my var <- "John"
# _my_var <- "John"
# my_v@ar <- "John"
# TRUE <- "John"
```

19.2.3 complex number

https://www.w3schools.com/r/r_data_types.asp

https://www.w3schools.com/r/r_numbers.asp

19.2.4 escape character

https://www.w3schools.com/r/r_strings_esc.asp

19.2.5 global assignment <<-

```
my_function <- function() {
txt <<- "fantastic"
  paste("R is", txt)
}
```

```r
my_function()
```

```
## [1] "R is fantastic"
```

```r
print(txt)
```

```
## [1] "fantastic"
```

```r
txt <- "awesome"
my_function <- function() {
  txt <<- "fantastic"
  paste("R is", txt)
}

my_function()
```

```
## [1] "R is fantastic"
```

```r
paste("R is", txt)
```

```
## [1] "R is fantastic"
```

19.2.6 data type

19.2.6.1 array

https://www.w3schools.com/r/r_arrays.asp

```r
# An array with one dimension with values ranging from 1 to 24
thisarray <- c(1:24)
thisarray
```

```
##  [1]  1  2  3  4  5  6  7  8  9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24
```

```r
# An array with more than one dimension
multiarray <- array(thisarray, dim = c(4, 3, 2))
multiarray
```

```
## , , 1
## 
##      [,1] [,2] [,3]
## [1,]    1    5    9
## [2,]    2    6   10
## [3,]    3    7   11
## [4,]    4    8   12
## 
## , , 2
## 
##      [,1] [,2] [,3]
## [1,]   13   17   21
## [2,]   14   18   22
## [3,]   15   19   23
## [4,]   16   20   24
```

```r
multiarray[2, 3, 2]
```

```
## [1] 22
```

19.2.6.2 list

https://www.w3schools.com/r/r_lists.asp

19.3 Apan Liao

R 演習室

https://www.youtube.com/playlist?list=PL5AC0ADBF65924EAD

19.3.1 data input

https://www.youtube.com/watch?v=STcIxf_vUWY&list=PL5AC0ADBF65924EAD&index=1

- `scan()`
- read
  - `read.table()`
  - `read.csv()`

19.3.2 descriptive statistics

https://www.youtube.com/watch?v=GL3Wv_45LaU&list=PL5AC0ADBF65924EAD&index=2

Chapter 20

Laplace transform

Chapter 21

conic section

conic section 圓錐曲線 / 圓錐截痕

https://en.wikipedia.org/wiki/Conic_section

https://tex.stackexchange.com/questions/222882/drawing-minimal-xy-axis

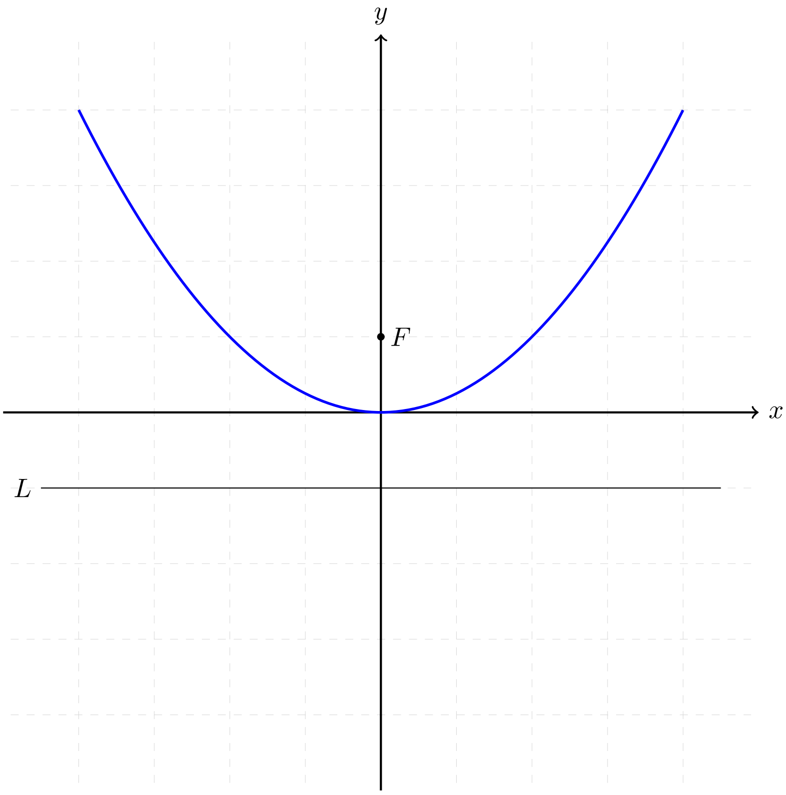


Figure 21.1: parabola defined by focus, directrix, eccentricity

21.1 Cartesian coordinate: focus, directrix, eccentricity

focus, directrix, eccentricity 焦點, 準線, 離心率

$$\begin{cases} F = (0, y_F) \\ L = y - y_L = 0 \\ \epsilon = \dfrac{\overline{PF}}{d\,(P, L)} = \dfrac{\|(x, y) - (0, y_F)\|}{\|y - y_L\|} \end{cases} \qquad \begin{array}{l} F : \text{focus} \\ L : \text{directrix} \\ \begin{cases} P = (x, y) \\ \epsilon : \text{eccentricity} \end{cases} \end{array}$$

$$0 \le \epsilon = \frac{\overline{PF}}{d(P,L)} = \frac{\overline{PF}}{\overline{PP'}} = \frac{\|(x,y)-(0,y_F)\|}{\|(x,y)-(x,y_L)\|} = \frac{\|(x,y-y_F)\|}{\|(0,y-y_L)\|} = \frac{\sqrt{x^2+(y-y_F)^2}}{\sqrt{(y-y_L)^2}} \tag{21.1}$$

$$\epsilon^2 = \frac{x^2+(y-y_F)^2}{(y-y_L)^2} = \frac{x^2+y^2-2y_Fy+y_F^2}{y^2-2y_Ly+y_L^2} \tag{21.2}$$

$$0 = x^2+\left(1-\epsilon^2\right)y^2-2\left(y_F-\epsilon^2y_L\right)y+\left(y_F^2-\epsilon^2y_L^2\right) \tag{21.3}$$

$$\stackrel{\epsilon\neq 1}{=} x^2+\left(1-\epsilon^2\right)\left[y^2-\frac{2\left(y_F-\epsilon^2y_L\right)}{1-\epsilon^2}y+\frac{y_F^2-\epsilon^2y_L^2}{1-\epsilon^2}\right] \tag{21.4}$$

$$= x^2+\left(1-\epsilon^2\right) \tag{21.5}$$

$$\left[y^2-\frac{2\left(y_F-\epsilon^2y_L\right)}{1-\epsilon^2}y+\left(\frac{y_F-\epsilon^2y_L}{1-\epsilon^2}\right)^2-\left(\frac{y_F-\epsilon^2y_L}{1-\epsilon^2}\right)^2+\frac{y_F^2-\epsilon^2y_L^2}{1-\epsilon^2}\right] \tag{21.6}$$

$$= x^2+\left(1-\epsilon^2\right)\left[\left(y-\frac{y_F-\epsilon^2y_L}{1-\epsilon^2}\right)^2+\frac{\left(y_F^2-\epsilon^2y_L^2\right)\left(1-\epsilon^2\right)-\left(y_F-\epsilon^2y_L\right)^2}{\left(1-\epsilon^2\right)^2}\right] \tag{21.7}$$

$$= x^2+\left(1-\epsilon^2\right)\left(y-\frac{y_F-\epsilon^2y_L}{1-\epsilon^2}\right)^2+\frac{\left(y_F^2-\epsilon^2y_L^2\right)\left(1-\epsilon^2\right)-\left(y_F-\epsilon^2y_L\right)^2}{1-\epsilon^2} \tag{21.8}$$

$$\begin{aligned}&\left(y_F^2-\epsilon^2y_L^2\right)\left(1-\epsilon^2\right)-\left(y_F-\epsilon^2y_L\right)^2\\ =&\left(1-\epsilon^2\right)y_F^2-\left(\epsilon^2-\epsilon^4\right)y_L^2-y_F^2+2\epsilon^2y_Fy_L-\epsilon^4y_L^2\\ =&-\epsilon^2y_F^2-\epsilon^2y_L^2+2\epsilon^2y_Fy_L=-\epsilon^2\left(y_F-y_L\right)^2\end{aligned}$$

$$\frac{\epsilon^2\left(y_F-y_L\right)^2}{1-\epsilon^2} \stackrel{\epsilon\neq 1}{=} x^2+\left(1-\epsilon^2\right)\left(y-\frac{y_F-\epsilon^2y_L}{1-\epsilon^2}\right)^2$$

$$1 \stackrel{\epsilon\neq 0,1}{=} \begin{cases} \left(\dfrac{x-0}{\frac{\epsilon\left(y_F-y_L\right)}{\sqrt{1-\epsilon^2}}}\right)^2+\left(\dfrac{y-\frac{y_F-\epsilon^2y_L}{1-\epsilon^2}}{\frac{\epsilon\left(y_F-y_L\right)}{1-\epsilon^2}}\right)^2 & 1-\epsilon^2>0 \stackrel{\epsilon\geq 0}{\Rightarrow} 0<\epsilon<1 \\ -\left(\dfrac{x-0}{\frac{\epsilon\left(y_F-y_L\right)}{\sqrt{\epsilon^2-1}}}\right)^2+\left(\dfrac{y-\frac{y_F-\epsilon^2y_L}{1-\epsilon^2}}{\frac{\epsilon\left(y_F-y_L\right)}{1-\epsilon^2}}\right)^2 & 1-\epsilon^2<0 \stackrel{\epsilon\geq 0}{\Rightarrow} \epsilon>1 \end{cases}$$

$\epsilon = 0$ or $\lim_{|y_L|\to\infty} \epsilon = 0$

$$r = \overline{PF} = \|(x,y)-(0,y_F)\| = \|(x,y-y_F)\| = \sqrt{x^2+(y-y_F)^2}$$

$$\epsilon = \frac{r}{d(P,L)} = \frac{\overline{PF}}{\overline{PP'}} = \frac{\|(x,y)-(0,y_F)\|}{\|(x,y)-(x,y_L)\|} = \frac{\|(x,y-y_F)\|}{\|(0,y-y_L)\|} = \frac{\sqrt{x^2+(y-y_F)^2}}{|y-y_L|}$$

$$\lim_{|y_L|\to\infty}\epsilon = \lim_{|y_L|\to\infty}\frac{r}{d(P,L)} = \lim_{|y_L|\to\infty}\frac{\sqrt{x^2+(y-y_F)^2}}{|y-y_L|} = 0$$

$\epsilon = 1$

$$\begin{aligned}0 =& x^2+\left(1-\epsilon^2\right)y^2-2\left(y_F-\epsilon^2y_L\right)y+\left(y_F^2-\epsilon^2y_L^2\right)\\ \stackrel{\epsilon=1}{=}& x^2+\left(1-1^2\right)y^2-2\left(y_F-1^2y_L\right)y+\left(y_F^2-1^2y_L^2\right)\\ =& x^2-2\left(y_F-y_L\right)y+\left(y_F^2-y_L^2\right)\\ =& x^2-2\left(y_F-y_L\right)y+\left(y_F+y_L\right)\left(y_F-y_L\right)\\ x^2 =& 2\left(y_F-y_L\right)\left(y-\frac{y_F+y_L}{2}\right)\end{aligned}$$

Let one curve vertex $P = V = (0, 0)$ on the curve, and fix the directrix L or y_L,

$\epsilon \neq 1$

$$\begin{aligned}
1 &\overset{P(x,y)=V(0,0)}{=} 0 + \left(\frac{0 - \frac{y_F - \epsilon^2 y_L}{1-\epsilon^2}}{\frac{\epsilon(y_F - y_L)}{1-\epsilon^2}} \right)^2 \\
&\Rightarrow y_F - \epsilon^2 y_L = \pm\epsilon (y_F - y_L) \\
&\Rightarrow \begin{cases} (1-\epsilon) y_F = \epsilon(\epsilon - 1) y_L & + \\ (1+\epsilon) y_F = \epsilon(\epsilon+1) y_L & - \end{cases} \\
&\Rightarrow y_F = \begin{cases} -\epsilon y_L & + \\ \epsilon y_L & - \end{cases}
\end{aligned}$$

$\epsilon = 1$

$$\begin{aligned}
x^2 &= 2(y_F - y_L)\left(y - \frac{y_F + y_L}{2}\right) \\
\overset{P(x,y)=V(0,0)}{\Rightarrow} 0^2 &= 2(y_F - y_L)\left(0 - \frac{y_F + y_L}{2}\right) \\
\Rightarrow 0 &= (y_F - y_L)(y_F + y_L) \\
\Rightarrow y_F &= \mp y_L
\end{aligned}$$

or by definition of eccentricity (21.1)

$$\begin{aligned}
0 \leq \epsilon = \frac{\overline{PF}}{d(P,L)} = \frac{\overline{PF}}{\overline{PP'}} &= \frac{\|(x,y) - (0,y_F)\|}{\|(x,y) - (x,y_L)\|} = \frac{\|(x, y - y_F)\|}{\|(0, y - y_L)\|} = \frac{\sqrt{x^2 + (y - y_F)^2}}{\sqrt{(y - y_L)^2}} \\
&\overset{P(x,y)=V(0,0)}{=} \frac{\sqrt{0^2 + (0 - y_F)^2}}{\sqrt{(0 - y_L)^2}} = \sqrt{\left(\frac{y_F}{y_L}\right)^2} \\
\epsilon^2 = \left(\frac{y_F}{y_L}\right)^2 &\Rightarrow y_F = \mp\epsilon y_L
\end{aligned}$$

actually,

$$y_F = -\epsilon y_L$$

21.2 two-definition equivalence for ellipse and hyperbola

https://math.stackexchange.com/questions/1833973/prove-that-the-directrix-focus-and-focus-focus-definitions-are-equivalent

https://www.geogebra.org/calculator/zkppuxwp

$$\begin{cases} P = (x,y) \\ F = (x_F, y_F) = (\alpha, \varphi) & F' = (x_{F'}, y_{F'}) = (\chi, \psi) \\ L = A'x + B'y + C' = 0 \end{cases}$$

21.2.1 first definition for conic sections including ellipses and hyperbolas

distance from a point to a line[^22^]

$$0 \leq \epsilon = \frac{\overline{PF}}{d(P,L)} = \frac{\sqrt{(x - x_F)^2 + (y - y_F)^2}}{\frac{|A'x + B'y + C'|}{\sqrt{A'^2 + B'^2}}} = \frac{\sqrt{(x-\alpha)^2 + (y - \varphi)^2}}{|Ax + By + C|}, \begin{cases} A = \frac{A'}{\sqrt{A'^2 + B'^2}} \\ B = \frac{B'}{\sqrt{A'^2 + B'^2}} \\ C = \frac{C'}{\sqrt{A'^2 + B'^2}} \end{cases}$$

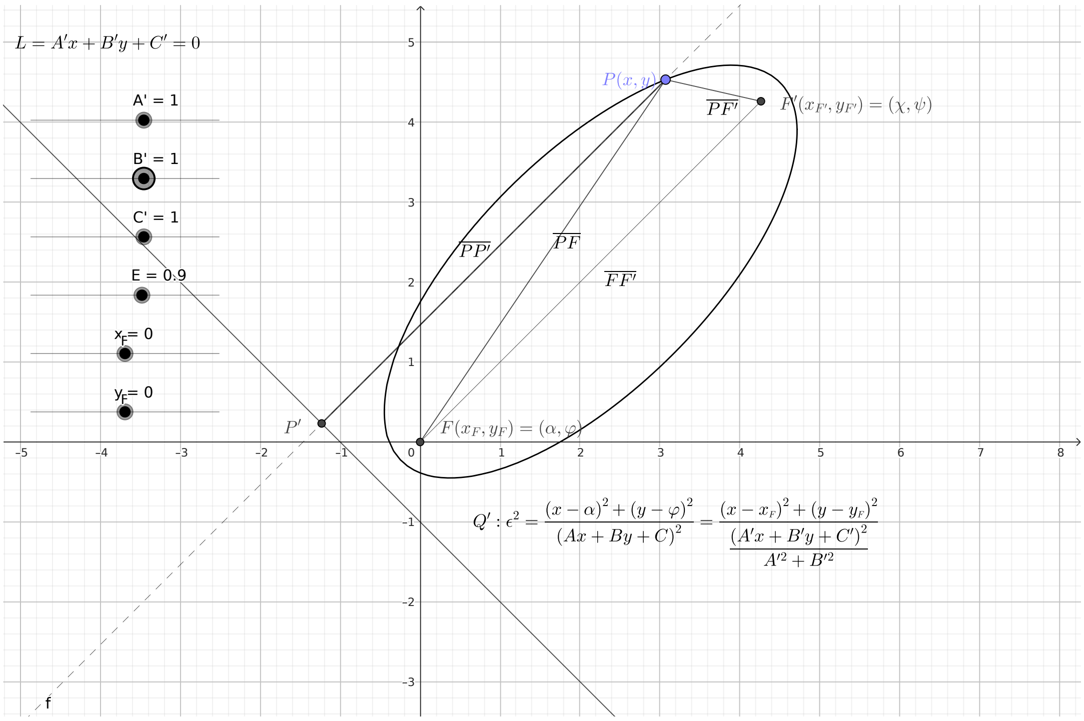


Figure 21.2: conic sections

$$A^2 + B^2 = \left(\frac{A'}{\sqrt{A'^2 + B'^2}}\right)^2 + \left(\frac{B'}{\sqrt{A'^2 + B'^2}}\right)^2 = 1$$

or allowing $\epsilon < 0$ by squaring the definition

$$\epsilon^2 = \frac{(x-\alpha)^2 + (y-\varphi)^2}{(Ax+By+C)^2} = \frac{(x-x_F)^2+(y-y_F)^2}{\frac{(A'x+B'y+C')^2}{A'^2+B'^2}}$$

$$(x-\alpha)^2 + (y-\varphi)^2 = [\epsilon\,(Ax+By+C)]^2$$

21.2.2 second definition for ellipses and hyperbolas

$$\begin{aligned}2c = \overline{FF'} &= \|(x_F, y_F) - (x_{F'}, y_{F'})\| = \|(\alpha, \varphi) - (\chi, \psi)\| \\ &= \sqrt{(\alpha-\chi)^2 + (\chi-\psi)^2}\end{aligned}$$

$$\begin{aligned}D &= \begin{cases}\sqrt{(x-x_F)^2+(y-y_F)^2} + \sqrt{(x-x_{F'})^2+(y-y_{F'})^2} & \text{ellipse} \\ \sqrt{(x-x_F)^2+(y-y_F)^2} - \sqrt{(x-x_{F'})^2+(y-y_{F'})^2} & \text{hyperbola}\end{cases} \\ &= \sqrt{(x-x_F)^2+(y-y_F)^2} \pm \sqrt{(x-x_{F'})^2+(y-y_{F'})^2} \\ &= \sqrt{(x-\alpha)^2+(y-\varphi)^2} \pm \sqrt{(x-\chi)^2+(y-\psi)^2}\end{aligned}$$

$$\begin{aligned}(x-\alpha)^2 + (y-\varphi)^2 &= \left(D \mp \sqrt{(x-\chi)^2+(y-\psi)^2}\right)^2 \\ &= D^2 \mp 2D\sqrt{(x-\chi)^2+(y-\psi)^2} \\ &\quad + (x-\chi)^2 + (y-\psi)^2\end{aligned}$$

$$\begin{aligned}
D^2 =& (x-\alpha)^2 + (y-\varphi)^2 + (x-\chi)^2 + (y-\psi)^2 \\
&\pm 2\sqrt{\left[(x-\alpha)^2 + (y-\varphi)^2\right]\left[(x-\chi)^2 + (y-\psi)^2\right]} \\
&(x-\alpha)^2 + (y-\varphi)^2 + (x-\chi)^2 + (y-\psi)^2 - D^2 \\
=& \mp 2\sqrt{\left[(x-\alpha)^2 + (y-\varphi)^2\right]\left[(x-\chi)^2 + (y-\psi)^2\right]} \\
&\left[(x-\alpha)^2 + (y-\varphi)^2 + (x-\chi)^2 + (y-\psi)^2\right]^2 + D^4 \\
&- 2D^2\left[(x-\alpha)^2 + (y-\varphi)^2 + (x-\chi)^2 + (y-\psi)^2\right] \\
=&4\left[(x-\alpha)^2 + (y-\varphi)^2\right]\left[(x-\chi)^2 + (y-\psi)^2\right] \\
&\left[(x-\alpha)^2 + (y-\varphi)^2\right]^2 + \left[(x-\chi)^2 + (y-\psi)^2\right]^2 \\
&+ 2\left[(x-\alpha)^2 + (y-\varphi)^2\right]\left[(x-\chi)^2 + (y-\psi)^2\right] + D^4 \\
&- 2D^2\left[(x-\alpha)^2 + (y-\varphi)^2 + (x-\chi)^2 + (y-\psi)^2\right] \\
=&4\left[(x-\alpha)^2 + (y-\varphi)^2\right]\left[(x-\chi)^2 + (y-\psi)^2\right] \\
0 =& \left[(x-\alpha)^2 + (y-\varphi)^2\right]^2 + \left[(x-\chi)^2 + (y-\psi)^2\right]^2 \\
&- 2\left[(x-\alpha)^2 + (y-\varphi)^2\right]\left[(x-\chi)^2 + (y-\psi)^2\right] + D^4 \\
&- 2D^2\left[(x-\alpha)^2 + (y-\varphi)^2 + (x-\chi)^2 + (y-\psi)^2\right] \\
0 =& \left\{\left[(x-\alpha)^2 + (y-\varphi)^2\right] - \left[(x-\chi)^2 + (y-\psi)^2\right]\right\}^2 + D^4 \\
&- 2D^2\left\{\left[(x-\alpha)^2 + (y-\varphi)^2\right] + \left[(x-\chi)^2 + (y-\psi)^2\right]\right\} \\
0 =& \left\{\left[(x-\chi)^2 + (y-\psi)^2\right] - \left[(x-\alpha)^2 + (y-\varphi)^2\right]\right\}^2 + D^4 \\
&- 2D^2\left\{\left[(x-\chi)^2 + (y-\psi)^2\right] - \left[(x-\alpha)^2 + (y-\varphi)^2\right]\right\} \\
&- 4D^2\left[(x-\alpha)^2 + (y-\varphi)^2\right] \\
&(2D)^2\left[(x-\alpha)^2 + (y-\varphi)^2\right] \\
=& \left\{\left[(x-\chi)^2 + (y-\psi)^2\right] - \left[(x-\alpha)^2 + (y-\varphi)^2\right] - D^2\right\}^2 \\
=& \left\{\left[(x-\chi)^2 - (x-\alpha)^2\right] + \left[(y-\psi)^2 - (y-\varphi)^2\right] - D^2\right\}^2 \\
=& \left\{(2x-\chi-\alpha)(\alpha-\chi) + (2y-\psi-\varphi)(\varphi-\psi) - D^2\right\}^2 \\
=& \left\{2(\alpha-\chi)x - \left(\alpha^2-\chi^2\right) + 2(\varphi-\psi)y - \left(\varphi^2-\psi^2\right) - D^2\right\}^2 \\
=& \left\{2(\alpha-\chi)x + 2(\varphi-\psi)y - \left[\left(\alpha^2-\chi^2\right) + \left(\varphi^2-\psi^2\right) + D^2\right]\right\}^2 \\
D \neq 0& \\
&(x-\alpha)^2 + (y-\varphi)^2 \\
=& \left[\frac{\alpha-\chi}{D}x + \frac{\varphi-\psi}{D}y - \left(\frac{\alpha^2-\chi^2}{2D} + \frac{\varphi^2-\psi^2}{2D} + \frac{D}{2}\right)\right]^2
\end{aligned}$$

$$\begin{cases}
(x-\alpha)^2 + (y-\varphi)^2 = \left[\epsilon\left(Ax+By+C\right)\right]^2 \\
(x-\alpha)^2 + (y-\varphi)^2 = \left[\frac{\alpha-\chi}{D}x + \frac{\varphi-\psi}{D}y - \left(\frac{\alpha^2-\chi^2}{2D} + \frac{\varphi^2-\psi^2}{2D} + \frac{D}{2}\right)\right]^2
\end{cases}$$

$$(A,B,C) \rightleftarrows (\chi,\psi,D)$$

$$\begin{cases} \epsilon A = \pm \dfrac{\alpha - \chi}{D} & \chi \pm \epsilon AD = \alpha \\ \epsilon B = \pm \dfrac{\varphi - \psi}{D} & \psi \pm \epsilon BD = \varphi \\ \epsilon C = \mp \left(\dfrac{\alpha^2 - \chi^2}{2D} + \dfrac{\varphi^2 - \psi^2}{2D} + \dfrac{D}{2} \right) \end{cases}$$

$$\begin{aligned} 2\epsilon C &= \mp \left(\frac{\alpha - \chi}{D} (\alpha + \chi) + \frac{\varphi - \psi}{D} (\varphi + \psi) + D \right) \\ &= \mp \left(\pm \epsilon A (\alpha + \chi) \pm \epsilon B (\varphi + \psi) + D \right) \\ \mp \epsilon (A\alpha + B\varphi + 2C) &= \pm \epsilon A\chi \pm \epsilon B\psi + D \end{aligned}$$

$$\begin{pmatrix} 1 & 0 & \pm\epsilon A \\ 0 & 1 & \pm\epsilon B \\ \pm\epsilon A & \pm\epsilon B & 1 \end{pmatrix} \begin{pmatrix} \chi \\ \psi \\ D \end{pmatrix} = \begin{pmatrix} \alpha \\ \varphi \\ \mp\epsilon (A\alpha + B\varphi + 2C) \end{pmatrix}$$

$$\begin{pmatrix} 1 & 0 & \pm\epsilon A & \alpha \\ 0 & 1 & \pm\epsilon B & \varphi \\ 0 & \pm\epsilon B & 1 \mp \epsilon^2 A^2 & \mp\epsilon (2A\alpha + B\varphi + 2C) \end{pmatrix}$$

$$\begin{pmatrix} 1 & 0 & \pm\epsilon A & \alpha \\ 0 & 1 & \pm\epsilon B & \varphi \\ 0 & 0 & 1 \mp \epsilon^2 A^2 \mp \epsilon^2 B^2 & \mp\epsilon (2A\alpha + 2B\varphi + 2C) \end{pmatrix}$$

$$\begin{pmatrix} 1 & 0 & \pm\epsilon A & \alpha \\ 0 & 1 & \pm\epsilon B & \varphi \\ 0 & 0 & 1 & \dfrac{\mp 2\epsilon (A\alpha + B\varphi + C)}{1 \mp \epsilon^2 (A^2 + B^2)} \end{pmatrix}$$

$$A^2 + B^2 = \left(\frac{A'}{\sqrt{A'^2 + B'^2}} \right)^2 + \left(\frac{B'}{\sqrt{A'^2 + B'^2}} \right)^2 = 1$$

$$\begin{cases} \chi = \alpha \mp \epsilon AD = \alpha \mp \epsilon \dfrac{A'}{\sqrt{A'^2 + B'^2}} D \\ \psi = \varphi \mp \epsilon BD = \varphi \mp \epsilon \dfrac{B'}{\sqrt{A'^2 + B'^2}} D \\ D = \dfrac{\mp 2\epsilon (A\alpha + B\varphi + C)}{1 \mp \epsilon^2 (A^2 + B^2)} = \dfrac{\mp 2\epsilon}{1 \mp \epsilon^2} \dfrac{A'\alpha + B'\varphi + C'}{\sqrt{A'^2 + B'^2}} \quad A^2 + B^2 = 1 \end{cases}$$

actually, only one of two solutions is true

$$\begin{cases} \chi = \alpha - \epsilon AD = \alpha - \epsilon \dfrac{A'}{\sqrt{A'^2 + B'^2}} D = \alpha - \dfrac{2\epsilon^2}{\epsilon^2 - 1} \dfrac{A'^2\alpha + A'B'\varphi + A'C'}{A'^2 + B'^2} \\ \psi = \varphi - \epsilon BD = \varphi - \epsilon \dfrac{B'}{\sqrt{A'^2 + B'^2}} D = \varphi - \dfrac{2\epsilon^2}{\epsilon^2 - 1} \dfrac{A'B'\alpha + B'^2\varphi + B'C'}{A'^2 + B'^2} \\ D = \dfrac{-2\epsilon (A\alpha + B\varphi + C)}{1 - \epsilon^2 (A^2 + B^2)} = \dfrac{-2\epsilon}{1 - \epsilon^2} \dfrac{A'\alpha + B'\varphi + C'}{\sqrt{A'^2 + B'^2}} = \dfrac{2\epsilon}{\epsilon^2 - 1} \dfrac{A'\alpha + B'\varphi + C'}{\sqrt{A'^2 + B'^2}} \end{cases}$$

$$\begin{cases} \chi = \dfrac{\left(\epsilon^2 - 1\right) \left(A'^2 + B'^2\right) \alpha - 2\epsilon^2 \left(A'^2\alpha + A'B'\varphi + A'C'\right)}{\left(\epsilon^2 - 1\right) \left(A'^2 + B'^2\right)} \\ \psi = \dfrac{\left(\epsilon^2 - 1\right) \left(A'^2 + B'^2\right) \varphi - 2\epsilon^2 \left(A'B'\alpha + B'^2\varphi + B'C'\right)}{\left(\epsilon^2 - 1\right) \left(A'^2 + B'^2\right)} \\ \left| \dfrac{D}{d(F, L)} \right| = \left| \dfrac{2\epsilon}{1 - \epsilon^2} \right| \Rightarrow \left(\dfrac{D}{d(F, L)} \right)^2 = \left(\dfrac{2\epsilon}{1 - \epsilon^2} \right)^2 \end{cases}$$

$$\begin{aligned} &\left(\epsilon^2 - 1\right) \left(A'^2 + B'^2\right) \alpha - 2\epsilon^2 \left(A'^2\alpha + A'B'\varphi + A'C'\right) \\ &= \left(-\left(\epsilon^2 + 1\right) A'^2 + \left(\epsilon^2 - 1\right) B'^2 \right) \alpha - 2\epsilon^2 \left(A'B'\varphi + A'C'\right) \\ &= \left(-\left(\epsilon^2 + 1\right) A'^2 + \left(\epsilon^2 - 1\right) B'^2 \right) \alpha - 2\epsilon^2 \left(A'B'\varphi + A'C'\right) \end{aligned}$$

Can the above be more simplified?

$$\begin{aligned}\overline{FF'}^2 &= (\alpha-\chi)^2 + (\varphi-\psi)^2 \\ &= (\alpha-(\alpha-\epsilon AD))^2 + (\varphi-(\varphi-\epsilon BD))^2 \\ &= (\epsilon D)^2\left(A^2+B^2\right) \\ &= (\epsilon D)^2\end{aligned}$$

21.2.3 eccentricity and its equivalent representation

$$\left(\frac{c}{a}\right)^2 = \left(\frac{\overline{PF}}{d(P,L)}\right)^2 = \epsilon^2 = \left(\frac{\overline{FF'}}{D}\right)^2 = \left(\frac{2c}{D}\right)^2 \Rightarrow D = 2a$$

$$\left(\frac{D}{d(F,L)}\right)^2 = \left(\frac{2\epsilon}{1-\epsilon^2}\right)^2$$

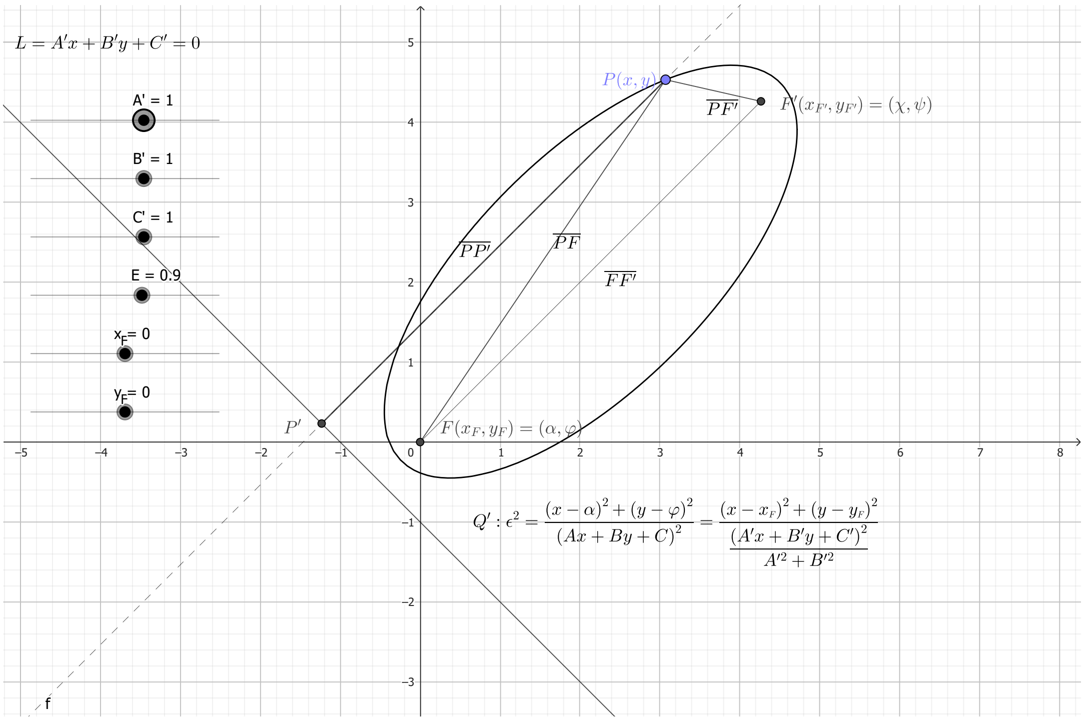


Figure 21.3: conic sections: ellipse

21.3 Cartesian coordinate: standard form / standard equation

circle	$\left(\frac{y-k}{a}\right)^2 + \left(\frac{x-h}{a}\right)^2$	$=1$		$b=a$
ellipse	$\left(\frac{y-k}{b}\right)^2 + \left(\frac{x-h}{a}\right)^2$	$=1$	vertical	$b>a$
	$\left(\frac{y-k}{b}\right)^2 + \left(\frac{x-h}{a}\right)^2$	$=1$	horizontal	$a>b$
parabola	$(y-k) - 4c(x-h)^2$	$=0$	vertical	
	$-4c(y-k)^2 + (x-h)$	$=0$	horizontal	
hyperbola	$\left(\frac{y-k}{b}\right)^2 - \left(\frac{x-h}{a}\right)^2$	$=1$	vertical	$\frac{x-h}{a}=0 \Rightarrow \frac{y-k}{b}=\pm 1$
	$-\left(\frac{y-k}{b}\right)^2 + \left(\frac{x-h}{a}\right)^2$	$=1$	horizontal	$\frac{y-k}{b}=0 \Rightarrow \frac{x-h}{a}=\pm 1$

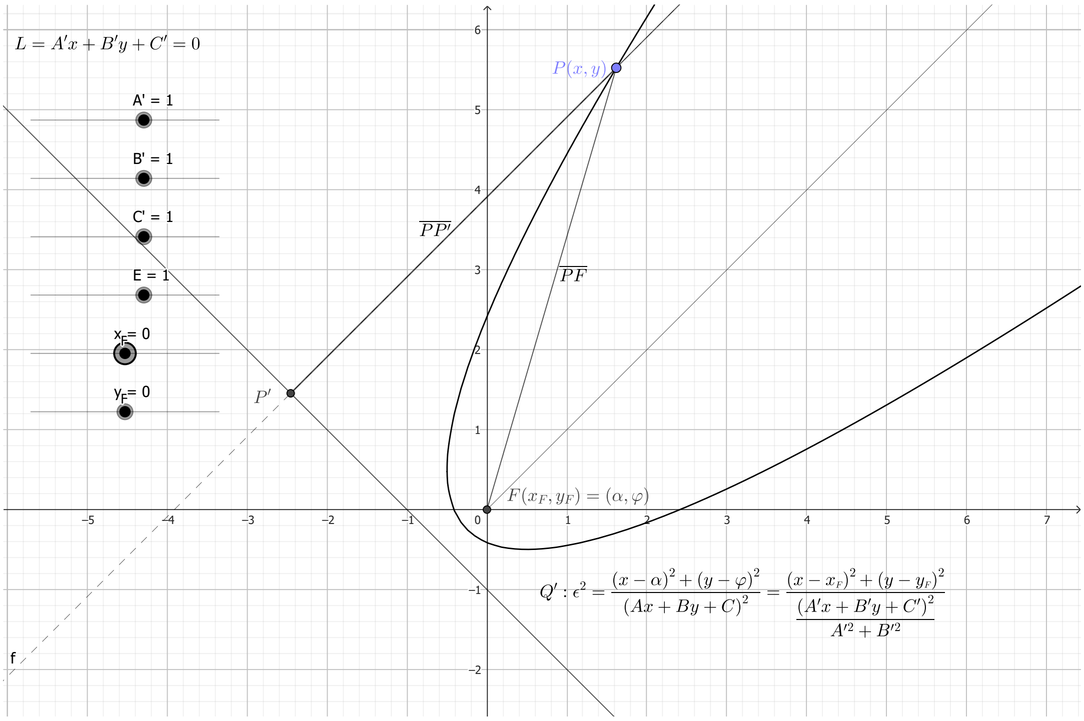


Figure 21.4: conic sections: parabola

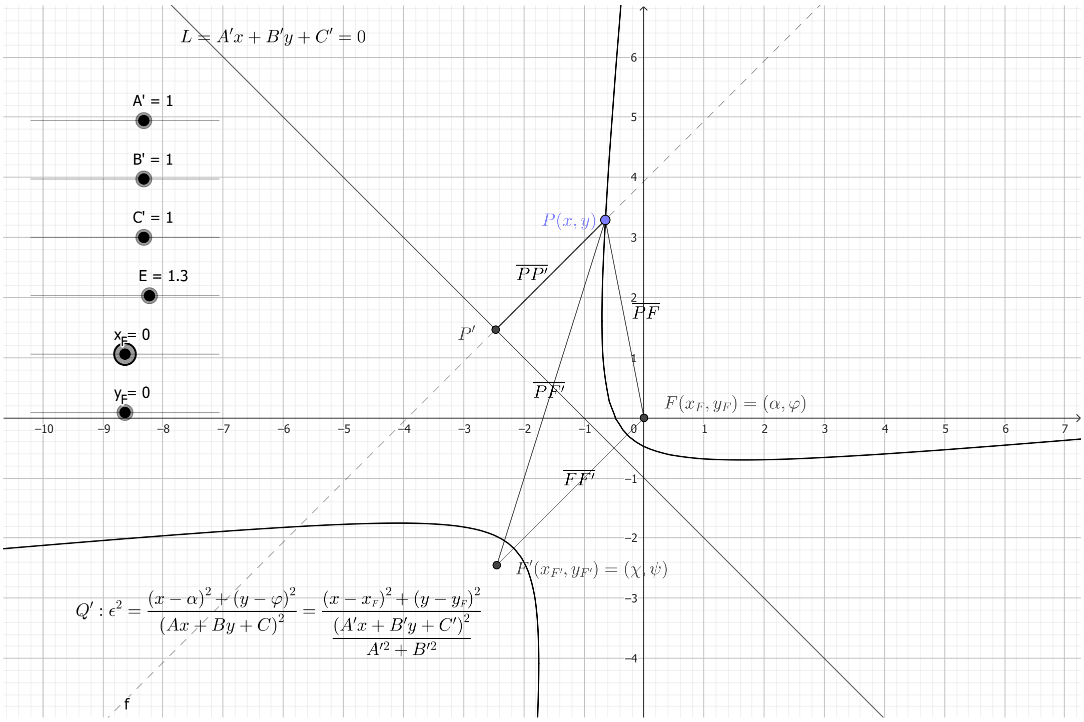


Figure 21.5: conic sections: hyperbola

21.4 parametric equation

$$\begin{array}{llll}
\text{circle} & \left(\frac{y-k}{a}\right)^2+\left(\frac{x-h}{a}\right)^2 & =1 & \begin{pmatrix}x\\y\\1\end{pmatrix}=\begin{pmatrix}a&0&h\\0&a&k\\0&0&1\end{pmatrix}\begin{pmatrix}\cos t\\\sin t\\1\end{pmatrix}=\begin{pmatrix}\cos t&0&h\\0&\sin t&k\\0&0&1\end{pmatrix}\begin{pmatrix}a\\a\\1\end{pmatrix}\\
\text{ellipse} & \left(\frac{y-k}{b}\right)^2+\left(\frac{x-h}{a}\right)^2 & =1 & \begin{pmatrix}x\\y\\1\end{pmatrix}=\begin{pmatrix}a&0&h\\0&b&k\\0&0&1\end{pmatrix}\begin{pmatrix}\cos t\\\sin t\\1\end{pmatrix}=\begin{pmatrix}\cos t&0&h\\0&\sin t&k\\0&0&1\end{pmatrix}\begin{pmatrix}a\\b\\1\end{pmatrix}\\
\text{parabola} & (y-k)-4c(x-h)^2 & =0 & \begin{pmatrix}x\\y\\1\end{pmatrix}=\begin{pmatrix}1&0&h\\0&4c&k\\0&0&1\end{pmatrix}\begin{pmatrix}t\\t^2\\1\end{pmatrix}=\begin{pmatrix}t&0&h\\0&t^2&k\\0&0&1\end{pmatrix}\begin{pmatrix}1\\4c\\1\end{pmatrix}\\
 & -4c(y-k)^2+(x-h) & =0 & \begin{pmatrix}x\\y\\1\end{pmatrix}=\begin{pmatrix}4c&0&h\\0&1&k\\0&0&1\end{pmatrix}\begin{pmatrix}t^2\\t\\1\end{pmatrix}=\begin{pmatrix}t^2&0&h\\0&t&k\\0&0&1\end{pmatrix}\begin{pmatrix}4c\\1\\1\end{pmatrix}\\
\text{hyperbola} & \left(\frac{y-k}{b}\right)^2-\left(\frac{x-h}{a}\right)^2 & =1 & \begin{pmatrix}x\\y\\1\end{pmatrix}=\begin{pmatrix}a&0&h\\0&b&k\\0&0&1\end{pmatrix}\begin{pmatrix}\pm\cosh t\\\sinh t\\1\end{pmatrix}=\begin{pmatrix}\tan t&0&h\\0&\sec t&k\\0&0&1\end{pmatrix}\begin{pmatrix}a\\b\\1\end{pmatrix}\\
 & -\left(\frac{y-k}{b}\right)^2+\left(\frac{x-h}{a}\right)^2 & =1 & \begin{pmatrix}x\\y\\1\end{pmatrix}=\begin{pmatrix}a&0&h\\0&b&k\\0&0&1\end{pmatrix}\begin{pmatrix}\sinh t\\\pm\cosh t\\1\end{pmatrix}=\begin{pmatrix}\sec t&0&h\\0&\tan t&k\\0&0&1\end{pmatrix}\begin{pmatrix}a\\b\\1\end{pmatrix}
\end{array}$$

tangent half-angle formula[24]

21.5 polar coordinate

$$(x-\alpha)^2+(y-\varphi)^2=[\epsilon(Ax+By+C)]^2$$

$\begin{cases}x=r\cos\theta\\y=r\sin\theta\end{cases}$

$$(r\cos\theta-\alpha)^2+(r\sin\theta-\varphi)^2=[\epsilon(Ar\cos\theta+Br\sin\theta+C)]^2$$

If $\begin{cases}F=(x_F,y_F)=(\alpha,\varphi)=(0,0)\\L=Ax+By+C=x+p=0\end{cases}$

$$\begin{aligned}(r\cos\theta)^2+(r\sin\theta)^2 &=[\epsilon(r\cos\theta+p)]^2\\
r^2 &=\\
r &=\pm\epsilon(r\cos\theta+p)\\
&=\pm(r\epsilon\cos\theta+\epsilon p)\\
r(1\mp\epsilon\cos\theta) &=\epsilon p\\
r &=\frac{\epsilon p}{1\mp\epsilon\cos\theta}\end{aligned}$$

https://www.geogebra.org/calculator/azksjxbq

$r=\frac{\epsilon p}{1-\epsilon\cos\theta}$ will not cross $L=x+p=0$ on graphs, so maybe it is a more correct solution

$$r=\frac{\epsilon p}{1-\epsilon\cos\theta}$$

21.6 Cartesian coordinate: general form / quadratic equation

https://ccjou.wordpress.com/2013/05/24/圓錐曲線/

https://en.wikipedia.org/wiki/Matrix_representation_of_conic_sections

$$ax^2+bxy+cy^2+dx+ey+f=0$$

$$\begin{pmatrix}x&y\end{pmatrix}\begin{pmatrix}a&b/2\\b/2&c\end{pmatrix}\begin{pmatrix}x\\y\end{pmatrix}=\begin{pmatrix}x&y\end{pmatrix}\begin{pmatrix}ax+(b/2)y\\(b/2)x+cy\end{pmatrix}=ax^2+bxy+cy^2$$

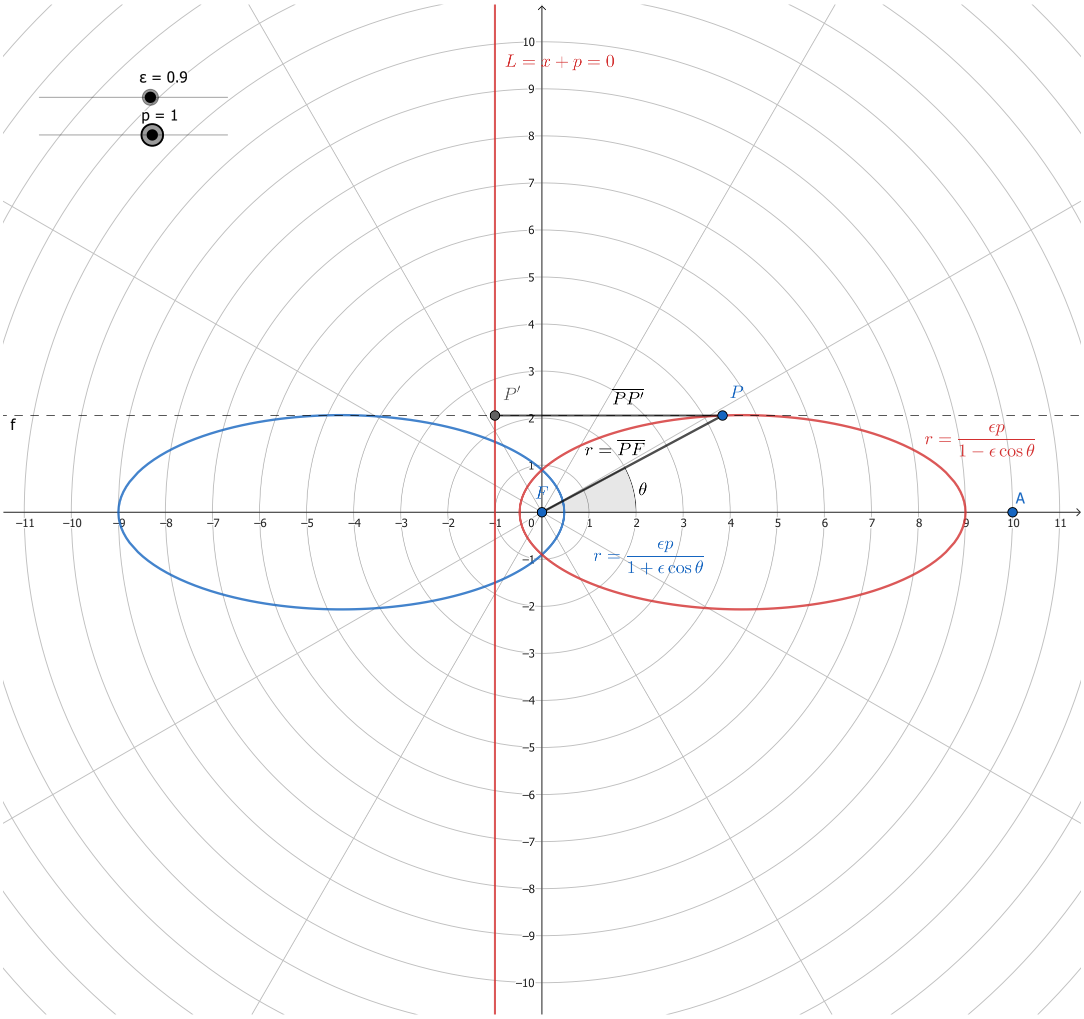


Figure 21.6: polar conic sections: ellipse

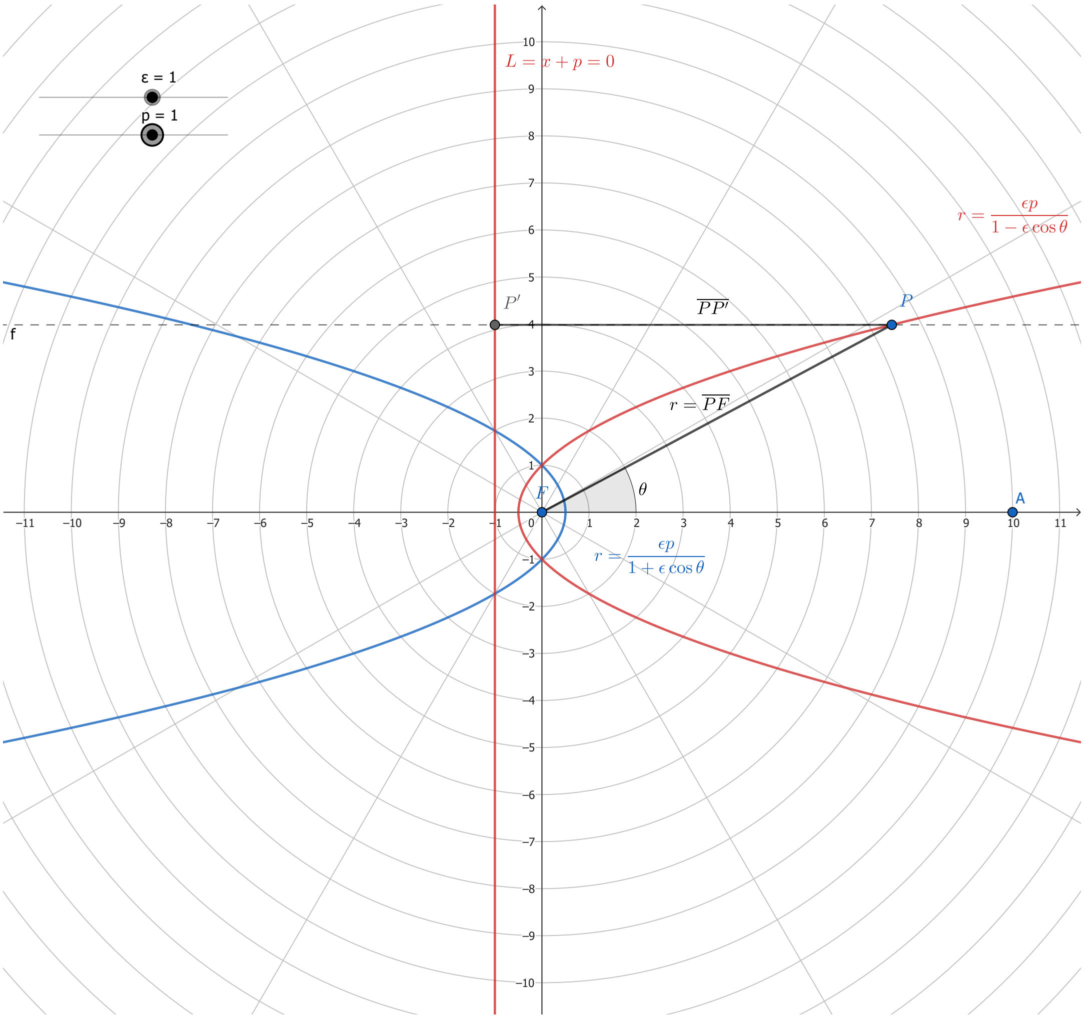


Figure 21.7: polar conic sections: parabola

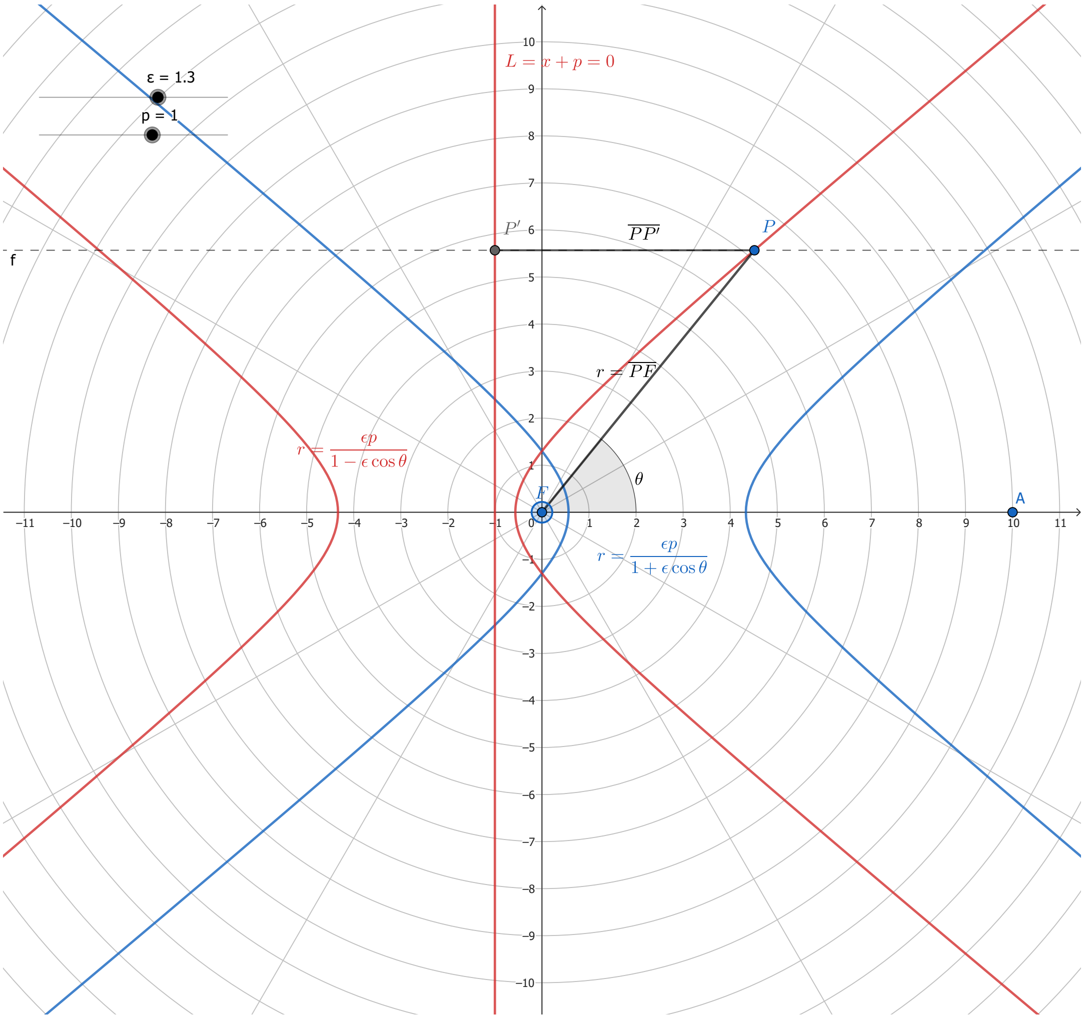


Figure 21.8: polar conic sections: hyperbola

$$
\begin{aligned}
0 &= \begin{pmatrix} x & y \end{pmatrix} \begin{pmatrix} a & b/2 \\ b/2 & c \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} + \begin{pmatrix} d & e \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} + f \\
&= \boldsymbol{x}^\intercal A \boldsymbol{x} + \boldsymbol{b}^\intercal \boldsymbol{x} + f, \begin{cases} A = \begin{pmatrix} a & b/2 \\ b/2 & c \end{pmatrix} & A \text{ real symmetric} \\ \boldsymbol{b} = \begin{pmatrix} d \\ e \end{pmatrix} \\ \boldsymbol{x} = \begin{pmatrix} x \\ y \end{pmatrix} \end{cases}
\end{aligned}
$$

real symmetric matrix diagonalizable[23]

21.7 homogeneous coordinate

X homogeneous coordinate

homogeneous coordinate O: HTML, X: PDF becoming web link

O homogeneous coordinate[25]

X homogeneous coordinate

X homogeneous coordinate[21.7]

https://ccjou.wordpress.com/2013/05/24/圓錐曲線/

$$
\begin{pmatrix} x & y \end{pmatrix} \begin{pmatrix} a & b/2 \\ b/2 & c \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} x & y & 1 \end{pmatrix} \begin{pmatrix} a & b/2 & \\ b/2 & c & \\ & & \end{pmatrix} \begin{pmatrix} x \\ y \\ 1 \end{pmatrix} = \begin{pmatrix} x & y & 1 \end{pmatrix} \begin{pmatrix} a & b/2 & 0 \\ b/2 & c & 0 \\ 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} x \\ y \\ 1 \end{pmatrix}
$$

$$
\begin{aligned}
\begin{pmatrix} d & e \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} &= \begin{pmatrix} x & y & 1 \end{pmatrix} \begin{pmatrix} \alpha & \beta & \gamma \\ \delta & \epsilon & \zeta \\ \eta & \theta & \kappa \end{pmatrix} \begin{pmatrix} x \\ y \\ 1 \end{pmatrix} = \begin{pmatrix} x & y & 1 \end{pmatrix} \begin{pmatrix} \alpha x + \beta y + \gamma \\ \delta x + \epsilon y + \zeta \\ \eta x + \theta y + \kappa \end{pmatrix}, \begin{cases} \gamma + \eta = d \\ \zeta + \theta = e \end{cases} \\
&= \begin{pmatrix} x & y & 1 \end{pmatrix} \begin{pmatrix} 0 & 0 & \gamma \\ 0 & 0 & \zeta \\ \eta & \theta & 0 \end{pmatrix} \begin{pmatrix} x \\ y \\ 1 \end{pmatrix} = \begin{pmatrix} x & y & 1 \end{pmatrix} \begin{pmatrix} 0 & 0 & d/2 \\ 0 & 0 & e/2 \\ d/2 & e/2 & 0 \end{pmatrix} \begin{pmatrix} x \\ y \\ 1 \end{pmatrix}
\end{aligned}
$$

$$
\begin{aligned}
0 =& ax^2 + bxy + cy^2 + dx + ey + f \\
=& \begin{pmatrix} x & y \end{pmatrix} \begin{pmatrix} a & b/2 \\ b/2 & c \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} + \begin{pmatrix} d & e \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} + f = \boldsymbol{x}^\intercal A \boldsymbol{x} + \boldsymbol{b}^\intercal \boldsymbol{x} + f \\
=& \begin{pmatrix} x & y & 1 \end{pmatrix} \begin{pmatrix} a & b/2 & d/2 \\ b/2 & c & e/2 \\ d/2 & e/2 & f \end{pmatrix} \begin{pmatrix} x \\ y \\ 1 \end{pmatrix} = \begin{pmatrix} \boldsymbol{x}^\intercal & 1 \end{pmatrix} M \begin{pmatrix} \boldsymbol{x} \\ 1 \end{pmatrix}, M = \begin{pmatrix} a & b/2 & d/2 \\ b/2 & c & e/2 \\ d/2 & e/2 & f \end{pmatrix}
\end{aligned}
$$

$$
\begin{aligned}
0 =& ax^2 + bxy + cy^2 + dx + ey + f \\
=& \begin{pmatrix} x & y \end{pmatrix} \begin{pmatrix} a & b/2 \\ b/2 & c \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} + \begin{pmatrix} d & e \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} + f = \boldsymbol{x}^\intercal A \boldsymbol{x} + \boldsymbol{b}^\intercal \boldsymbol{x} + f \\
=& \begin{pmatrix} x & y & 1 \end{pmatrix} \begin{pmatrix} a & b/2 & d/2 \\ b/2 & c & e/2 \\ d/2 & e/2 & f \end{pmatrix} \begin{pmatrix} x \\ y \\ 1 \end{pmatrix} = \begin{pmatrix} \boldsymbol{x}^\intercal & 1 \end{pmatrix} M \begin{pmatrix} \boldsymbol{x} \\ 1 \end{pmatrix}, M = \begin{pmatrix} a & b/2 & d/2 \\ b/2 & c & e/2 \\ d/2 & e/2 & f \end{pmatrix}
\end{aligned}
$$

https://en.wikipedia.org/wiki/Matrix_representation_of_conic_sections

$$
\begin{aligned}
0 = Q =& Ax^2 + Bxy + Cy^2 + Dx + Ey + F \\
=& \begin{bmatrix} x & y & 1 \end{bmatrix} \begin{bmatrix} A & B/2 & D/2 \\ B/2 & C & E/2 \\ D/2 & E/2 & F \end{bmatrix} \begin{bmatrix} x \\ y \\ 1 \end{bmatrix} = \boldsymbol{x}_{\mathrm{h}}^\intercal A_Q \boldsymbol{x}_{\mathrm{h}} \\
=& \begin{bmatrix} x & y \end{bmatrix} \begin{bmatrix} A & B/2 \\ B/2 & C \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} + \begin{bmatrix} D & E \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} + F = \boldsymbol{x}^\intercal A_{Q,33} \boldsymbol{x} + \boldsymbol{b}^\intercal \boldsymbol{x} + F
\end{aligned}
$$

Chapter 22

distance from a point to a line

點到直線距離

Theorem 22.1.

$$\begin{cases} P = P(x_0, y_0) \\ L = L(x, y) = Ax + By + C = 0, A^2 + B^2 \neq 0 \end{cases}$$
$$\Downarrow$$
$$d(P, L) = \frac{|Ax_0 + By_0 + C|}{\sqrt{A^2 + B^2}}$$

https://en.wikipedia.org/wiki/Distance_from_a_point_to_a_line

https://highscope.ch.ntu.edu.tw/wordpress/?p=47407

https://web.math.sinica.edu.tw/mathmedia/HTMLarticle18.jsp?mID=40312

Proofs:

22.1 by shortest $\overline{PP'}$

$$\begin{aligned} P' = P'(x, y) &\in L = Ax + By + C = 0 \\ \Rightarrow y &= \frac{-1}{B}(Ax + C) \end{aligned}$$

$$\begin{aligned} \overline{PP'}^2(x, y) &= (x_0 - x)^2 + (y_0 - y)^2 \\ &= (x_0 - x)^2 + \left(y_0 - \frac{-1}{B}(Ax + C)\right)^2 \\ &= (x - x_0)^2 + \left(\frac{A}{B}x + \frac{C}{B} + y_0\right)^2 = \overline{PP'}^2(x) \end{aligned}$$

$$\begin{aligned} 0 = \frac{\partial}{\partial x}\overline{PP'}^2(x) =& 2(x - x_0) + 2\left(\frac{A}{B}x + \frac{C}{B} + y_0\right)\frac{A}{B} \\ =& \frac{2}{B^2}\left(B^2(x - x_0) + A^2x + AC + ABy_0\right) \\ =& \frac{2}{B^2}\left[\left(A^2 + B^2\right)x - \left(B^2x_0 - ABy_0 - AC\right)\right] \\ x =& \frac{B^2x_0 - ABy_0 - AC}{A^2 + B^2} \end{aligned}$$

or by completing the square to find x.

$$
\begin{aligned}
&\overline{PP'}^2\left(x=\frac{B^2x_0-ABy_0-AC}{A^2+B^2}\right)\\
=&\left(\frac{B^2x_0-ABy_0-AC}{A^2+B^2}-x_0\right)^2+\left(\frac{A}{B}\frac{B^2x_0-ABy_0-AC}{A^2+B^2}+\frac{C}{B}+y_0\right)^2\\
=&\left(\frac{-A^2x_0-ABy_0-AC}{A^2+B^2}\right)^2+\left(\frac{A\left(B^2x_0-ABy_0-AC\right)+C\left(A^2+B^2\right)+B\left(A^2+B^2\right)y_0}{B\left(A^2+B^2\right)}\right)^2\\
=&\left(\frac{-A\left(Ax_0+By_0+C\right)}{A^2+B^2}\right)^2+\left(\frac{AB^2x_0+B^3y_0+B^2C}{B\left(A^2+B^2\right)}\right)^2\\
=&\frac{A^2\left(Ax_0+By_0+C\right)^2}{\left(A^2+B^2\right)^2}+\frac{B^2\left(Ax_0+By_0+C\right)^2}{\left(A^2+B^2\right)^2}\\
=&\frac{\left(Ax_0+By_0+C\right)^2}{A^2+B^2}\\
\overline{PP'}=&\overline{PP'}\left(x=\frac{B^2x_0-ABy_0-AC}{A^2+B^2}\right)=\frac{\left|Ax_0+By_0+C\right|}{\sqrt{A^2+B^2}}
\end{aligned}
$$

22.2 by perpendicular foot

$$y=\frac{-A}{B}x-\frac{C}{B}=\frac{-1}{B}\left(Ax+C\right),\ \text{if } B\neq 0$$

$$L_\perp:\left(y=\frac{B}{A}x+K\right)\perp\left(y=\frac{-A}{B}x-\frac{C}{B}\right):L$$

$$L_\perp=L_\perp\left(x,y\right)=Bx-Ay+K=0$$

$$P=P\left(x_0,y_0\right)\in L_\perp=B\left(x-x_0\right)-A\left(y-y_0\right)=0$$

$$L_\perp=Bx-Ay-\left(Bx_0-Ay_0\right)=0$$

perpendicular foot = foot of the perpendicular P'

$$
\begin{aligned}
P'\in\left(L_\perp\cap L\right)=&\begin{cases}L=Ax+By+C=0\\ L_\perp=Bx-Ay-\left(Bx_0-Ay_0\right)=0\end{cases}\\
=&\begin{cases}Ax+By=-C\\ Bx-Ay=Bx_0-Ay_0\end{cases}\\
P'=P'\left(x,y\right)=&\left(\frac{\begin{vmatrix}-C & B\\ Bx_0-Ay_0 & -A\end{vmatrix}}{\begin{vmatrix}A & B\\ B & -A\end{vmatrix}},\frac{\begin{vmatrix}A & -C\\ B & Bx_0-Ay_0\end{vmatrix}}{\begin{vmatrix}A & B\\ B & -A\end{vmatrix}}\right)\\
=&\left(\frac{\begin{vmatrix}C & B\\ -Bx_0+Ay_0 & -A\end{vmatrix}}{\begin{vmatrix}A & -B\\ B & A\end{vmatrix}},\frac{\begin{vmatrix}A & C\\ B & -Bx_0+Ay_0\end{vmatrix}}{\begin{vmatrix}A & -B\\ B & A\end{vmatrix}}\right)\\
=&\left(\frac{B^2x_0-ABy_0-AC}{A^2+B^2},\frac{-ABx_0+A^2y_0-BC}{A^2+B^2}\right)
\end{aligned}
$$

$$
\begin{aligned}
& d(P,L) = \overline{PP'} \\
=& \left\| (x_0, y_0) - \left(\frac{B^2 x_0 - ABy_0 - AC}{A^2+B^2}, \frac{-ABx_0 + A^2 y_0 - BC}{A^2+B^2} \right) \right\| \\
=& \sqrt{\left(x_0 - \frac{B^2 x_0 - ABy_0 - AC}{A^2+B^2} \right)^2 + \left(y_0 - \frac{-ABx_0 + A^2 y_0 - BC}{A^2+B^2} \right)^2} \\
=& \sqrt{\left(\frac{A^2 x_0 + ABy_0 + AC}{A^2+B^2} \right)^2 + \left(\frac{ABx_0 + B^2 y_0 + BC}{A^2+B^2} \right)^2} \\
=& \sqrt{\frac{A^2 (Ax_0 + By_0 + C)^2 + B^2 (Ax_0 + By_0 + C)^2}{(A^2+B^2)^2}} = \sqrt{\frac{(Ax_0 + By_0 + C)^2}{A^2+B^2}} \\
=& \frac{|Ax_0 + By_0 + C|}{\sqrt{A^2+B^2}}
\end{aligned}
$$

22.3 by normal vector

$$
\begin{cases}
\overrightharpoon{n} = (A,B) \perp L = Ax + By + C = 0 \\
\overrightharpoon{PP'} = P' - P = (x - x_0, y - y_0)
\end{cases}
$$

P到L的距離$d(P,L)$即為L線上一點P'對應之$\overrightharpoon{PP'}$在L法向量$\overrightharpoon{n}$方向上的投影長

$$
\begin{aligned}
\overrightharpoon{PP'} \cdot \overrightharpoon{n} &= \left\| \overrightharpoon{PP'} \right\| \left\| \overrightharpoon{n} \right\| \cos\theta \\
\left| \overrightharpoon{PP'} \cdot \overrightharpoon{n} \right| &= \left\| \overrightharpoon{PP'} \right\| \left\| \overrightharpoon{n} \right\| |\cos\theta| \\
\left\| \overrightharpoon{PP'} \right\| |\cos\theta| = \left| \overrightharpoon{PP'} \cdot \hat{\boldsymbol{n}} \right| &= \frac{\left| \overrightharpoon{PP'} \cdot \overrightharpoon{n} \right|}{\left\| \overrightharpoon{n} \right\|} = \frac{|(x - x_0, y - y_0) \cdot (A,B)|}{\|(A,B)\|} \\
&= \frac{|A(x - x_0) + B(y - y_0)|}{\sqrt{A^2+B^2}} = \frac{|-Ax_0 - By_0 + Ax + By|}{\sqrt{A^2+B^2}} \\
&\overset{Ax+By+C=0}{\underset{Ax+By=-C}{=}} \frac{|-Ax_0 - By_0 - C|}{\sqrt{A^2+B^2}} = \frac{|Ax_0 + By_0 + C|}{\sqrt{A^2+B^2}}
\end{aligned}
$$

PDF LaTeX `\usepackage{fdsymbol}` to have `\overrightharpoon` vector; however, there are too many side effects, including ugly mathptmx $\sum$, ...

```
\usepackage{fdsymbol} % vector over accent, but will use mathptmx
% replace the rather ugly mathptmx \sum operator with the equivalent Computer Modern one
\let\sum\relax
\DeclareSymbolFont{CMlargesymbols}{OMX}{cmex}{m}{n}
\DeclareMathSymbol{\sum}{\mathop}{CMlargesymbols}{"50}
```

22.4 by Cauchy inequality

$$
\begin{aligned}
Ax + By + C &= 0 \\
Ax + By &= -C \\
(Ax + By) - (Ax_0 + By_0) &= -C - (Ax_0 + By_0) \\
A(x - x_0) + B(y - y_0) &= -(Ax_0 + By_0 + C)
\end{aligned}
$$

$$
\begin{aligned}
\overline{PP'}^2 &= (x_0 - x)^2 + (y_0 - y)^2 \\
\left[A^2 + B^2\right] \overline{PP'}^2 &= \left[A^2 + B^2\right] \left[(x_0 - x)^2 + (y_0 - y)^2\right] \\
&\geq \left[A\,(x - x_0) + B\,(y - y_0)\right]^2 \\
&= \left[-\left(Ax_0 + By_0 + C\right)\right]^2 = (Ax_0 + By_0 + C)^2 \\
\overline{PP'}^2 &\geq \frac{(Ax_0 + By_0 + C)^2}{A^2 + B^2} \\
\overline{PP'} &\geq \frac{|Ax_0 + By_0 + C|}{\sqrt{A^2 + B^2}}
\end{aligned}
$$

Chapter 23

real symmetric matrix diagonalizable

https://ccjou.wordpress.com/2011/02/09/實對稱矩陣可正交對角化的證明/

https://tex.stackexchange.com/questions/30619/what-is-the-best-symbol-for-vector-matrix-transpose

Theorem 23.1.

實對稱矩陣的特徵值皆是實數，且對應特徵向量是實向量。

$$\begin{cases}\begin{cases}A \in \mathcal{M}_{n\times n}(\mathbb{R}) & \text{real matrix} \\ A^\intercal = A & \text{symmetric matrix}\end{cases} & \text{real symmetric matrix} \\ A\boldsymbol{x} = \lambda \boldsymbol{x} & \begin{cases}\lambda \in \mathbb{C} & \text{complex eigenvalue} \\ \mathbf{0} \neq \boldsymbol{x} \in \mathbb{C}^n & \text{complex eigenvector}\end{cases}\end{cases}$$
$$\Downarrow$$
$$\begin{cases}\lambda \in \mathbb{R} & \text{real eigenvalue (1)} \\ \boldsymbol{x} \in \mathbb{R}^n & \text{real eigenvector (2)}\end{cases}$$

Proof. (1)

$$\begin{aligned}
A\boldsymbol{x} =& \lambda\boldsymbol{x} \\
\overline{A}\overline{\boldsymbol{x}} = \overline{A\boldsymbol{x}} =& \overline{\lambda\boldsymbol{x}} = \overline{\lambda}\overline{\boldsymbol{x}} \\
\overline{\boldsymbol{x}}^\intercal\overline{A}^\intercal = \left(\overline{A}\overline{\boldsymbol{x}}\right)^\intercal = \left(\overline{\lambda}\overline{\boldsymbol{x}}\right)^\intercal =& \overline{\lambda}\overline{\boldsymbol{x}}^\intercal \\
\overline{\boldsymbol{x}}^\intercal A \overset{\text{symmetric}}{=} \overline{\boldsymbol{x}}^\intercal A^\intercal \overset{\text{real}}{=} & \\
\overline{\boldsymbol{x}}^\intercal A =& \overline{\lambda}\overline{\boldsymbol{x}}^\intercal \\
\lambda\overline{\boldsymbol{x}}^\intercal\boldsymbol{x} = \overline{\boldsymbol{x}}^\intercal(\lambda\boldsymbol{x}) \underset{A\boldsymbol{x}=\lambda\boldsymbol{x}}{\overset{\cdot\boldsymbol{x}}{=}} \overline{\boldsymbol{x}}^\intercal A\boldsymbol{x} =& \overline{\lambda}\overline{\boldsymbol{x}}^\intercal\boldsymbol{x} \\
\lambda\overline{\boldsymbol{x}}^\intercal\boldsymbol{x} =& \overline{\lambda}\overline{\boldsymbol{x}}^\intercal\boldsymbol{x} \\
\left(\lambda - \overline{\lambda}\right)\overline{\boldsymbol{x}}^\intercal\boldsymbol{x} =& 0 \wedge \begin{cases}\overline{\boldsymbol{x}}^\intercal\boldsymbol{x} = \sum_{i=1}^{n}|x_i|^2 \\ \boldsymbol{x} \neq \mathbf{0}\end{cases} \Rightarrow \overline{\boldsymbol{x}}^\intercal\boldsymbol{x} \neq 0 \\
\lambda - \overline{\lambda} =& 0 \\
\lambda =& \overline{\lambda} \Leftrightarrow \lambda \in \mathbb{R}
\end{aligned}$$

□

Proof. (1) fast concept

$$
\begin{aligned}
\left(\overline{A\boldsymbol{x}}\right)^{\intercal}\boldsymbol{x} = \left(\overline{\boldsymbol{x}}^{\intercal}\overline{A}^{\intercal}\right)\boldsymbol{x} &\overset{\text{symmetric}}{=} \left(\overline{\boldsymbol{x}}^{\intercal}\overline{A}\right)\boldsymbol{x} = \overline{\boldsymbol{x}}^{\intercal}\left(\overline{A}\boldsymbol{x}\right) \\
(L) = \left(\overline{A\boldsymbol{x}}\right)^{\intercal}\boldsymbol{x} &= \overline{\boldsymbol{x}}^{\intercal}\left(\overline{A}\boldsymbol{x}\right) = (R) \\
(L) = \left(\overline{A\boldsymbol{x}}\right)^{\intercal}\boldsymbol{x} &\overset{A\boldsymbol{x}=\lambda\boldsymbol{x}}{=} \left(\overline{\lambda\boldsymbol{x}}\right)^{\intercal}\boldsymbol{x} = \overline{\lambda}\overline{\boldsymbol{x}}^{\intercal}\boldsymbol{x} \\
(R) = \overline{\boldsymbol{x}}^{\intercal}\left(\overline{A}\boldsymbol{x}\right) &\overset{\text{real}}{=} \overline{\boldsymbol{x}}^{\intercal}\left(A\boldsymbol{x}\right) \overset{A\boldsymbol{x}=\lambda\boldsymbol{x}}{=} \overline{\boldsymbol{x}}^{\intercal}\left(\lambda\boldsymbol{x}\right) = \lambda\overline{\boldsymbol{x}}^{\intercal}\boldsymbol{x} \\
\overline{\lambda}\overline{\boldsymbol{x}}^{\intercal}\boldsymbol{x} = \left(\overline{A\boldsymbol{x}}\right)^{\intercal}\boldsymbol{x} &= \overline{\boldsymbol{x}}^{\intercal}\left(\overline{A}\boldsymbol{x}\right) = \lambda\overline{\boldsymbol{x}}^{\intercal}\boldsymbol{x} \\
\overline{\lambda}\overline{\boldsymbol{x}}^{\intercal}\boldsymbol{x} &= \lambda\overline{\boldsymbol{x}}^{\intercal}\boldsymbol{x}
\end{aligned}
$$

□

Proof. (2)

???

推論特徵空間 $N(A-\lambda I)$ ($A-\lambda I$ 的零空間) 為 $\mathbb{R}^n$ 的子空間，故 $\boldsymbol{x} \in N(A-\lambda I)$ 是一個非零實向量。

□

Theorem 23.2.

實對稱矩陣對應相異特徵值的特徵向量互為正交。

$$
\begin{cases}
\begin{cases} A \in \mathcal{M}_{n\times n}(\mathbb{R}) & \text{real matrix} \\ A^{\intercal} = A & \text{symmetric matrix} \end{cases} & \text{real symmetric matrix} \\
A\boldsymbol{x} = \lambda\boldsymbol{x} & 23.1 \begin{cases} \lambda \in \mathbb{R} & \text{real eigenvalue} \\ \boldsymbol{x} \in \mathbb{R}^n & \text{real eigenvector} \end{cases} \\
\begin{cases} A\boldsymbol{x}_1 = \lambda_1\boldsymbol{x}_1 & (e_1) \\ A\boldsymbol{x}_2 = \lambda_2\boldsymbol{x}_2 & (e_2) \end{cases} & \lambda_1 \neq \lambda_2
\end{cases}
$$

$$\Downarrow$$

$$\boldsymbol{x}_1^{\intercal}\boldsymbol{x}_2 = 0 \Leftrightarrow \boldsymbol{x}_1 \perp \boldsymbol{x}_2$$

Proof. (1)

$$
\begin{aligned}
A\boldsymbol{x}_2 &= \lambda_2\boldsymbol{x}_2 \\
\boldsymbol{x}_1^{\intercal}A\boldsymbol{x}_2 &\overset{\boldsymbol{x}_1^{\intercal}\cdot}{=} \boldsymbol{x}_1^{\intercal}\lambda_2\boldsymbol{x}_2 = \lambda_2\boldsymbol{x}_1^{\intercal}\boldsymbol{x}_2 = (1) \\
A\boldsymbol{x}_1 &= \lambda_1\boldsymbol{x}_1 \\
\boldsymbol{x}_1^{\intercal}A^{\intercal} = (A\boldsymbol{x}_1)^{\intercal} &= (\lambda_1\boldsymbol{x}_1)^{\intercal} = \lambda_1\boldsymbol{x}_1^{\intercal} \\
\boldsymbol{x}_1^{\intercal}A^{\intercal} &= \lambda_1\boldsymbol{x}_1^{\intercal} \\
\boldsymbol{x}_1^{\intercal}A\boldsymbol{x}_2 \overset{\text{symmetric}}{=} \boldsymbol{x}_1^{\intercal}A^{\intercal}\boldsymbol{x}_2 &\overset{\cdot\boldsymbol{x}_2}{=} \lambda_1\boldsymbol{x}_1^{\intercal}\boldsymbol{x}_2 = (2) \\
\lambda_2\boldsymbol{x}_1^{\intercal}\boldsymbol{x}_2 \overset{(1)}{=} \boldsymbol{x}_1^{\intercal}A\boldsymbol{x}_2 &\overset{(2)}{=} \lambda_1\boldsymbol{x}_1^{\intercal}\boldsymbol{x}_2 \\
\lambda_2\boldsymbol{x}_1^{\intercal}\boldsymbol{x}_2 &= \lambda_1\boldsymbol{x}_1^{\intercal}\boldsymbol{x}_2 \\
(\lambda_2-\lambda_1)\,\boldsymbol{x}_1^{\intercal}\boldsymbol{x}_2 &= 0 \wedge \lambda_1 \neq \lambda_2 \\
\boldsymbol{x}_1^{\intercal}\boldsymbol{x}_2 &= 0
\end{aligned}
$$

□

Proof. (1) fast concept

$$
\begin{aligned}
(A\boldsymbol{x}_1)^{\intercal}\boldsymbol{x}_2 = \left(\boldsymbol{x}_1^{\intercal}A^{\intercal}\right)\boldsymbol{x}_2 &\overset{\text{symmetric}}{=} \left(\boldsymbol{x}_1^{\intercal}A\right)\boldsymbol{x}_2 = \boldsymbol{x}_1^{\intercal}(A\boldsymbol{x}_2) \\
(L) = (A\boldsymbol{x}_1)^{\intercal}\boldsymbol{x}_2 &= \boldsymbol{x}_1^{\intercal}(A\boldsymbol{x}_2) = (R) \\
(L) = (A\boldsymbol{x}_1)^{\intercal}\boldsymbol{x}_2 &\overset{(e_1)}{=} (\lambda_1\boldsymbol{x}_1)^{\intercal}\boldsymbol{x}_2 = \lambda_1\boldsymbol{x}_1^{\intercal}\boldsymbol{x}_2 \\
(R) = \boldsymbol{x}_1^{\intercal}(A\boldsymbol{x}_2) &\overset{(e_2)}{=} \boldsymbol{x}_1^{\intercal}(\lambda_2\boldsymbol{x}_2) = \lambda_2\boldsymbol{x}_1^{\intercal}\boldsymbol{x}_2 \\
\lambda_1\boldsymbol{x}_1^{\intercal}\boldsymbol{x}_2 = (A\boldsymbol{x}_1)^{\intercal}\boldsymbol{x}_2 &= \boldsymbol{x}_1^{\intercal}(A\boldsymbol{x}_2) = \lambda_2\boldsymbol{x}_1^{\intercal}\boldsymbol{x}_2 \\
\lambda_1\boldsymbol{x}_1^{\intercal}\boldsymbol{x}_2 &= \lambda_2\boldsymbol{x}_1^{\intercal}\boldsymbol{x}_2
\end{aligned}
$$

□

Theorem 23.3.

$$\begin{cases} \begin{cases} A \in \mathcal{M}_{n\times n}(\mathbb{R}) & \text{real matrix} \\ A^\intercal = A & \text{symmetric matrix} \end{cases} & \text{real symmetric matrix} \\ A\boldsymbol{x}_1 = \lambda \boldsymbol{x}_1 & (e) \\ \boldsymbol{x}_2^\intercal \boldsymbol{x}_1 = 0 \Leftrightarrow \boldsymbol{x}_2 \perp \boldsymbol{x}_1 & (o) \end{cases}$$

$$\Downarrow$$

$$A\boldsymbol{x}_2 \perp \boldsymbol{x}_1 \Leftrightarrow (A\boldsymbol{x}_2)^\intercal \boldsymbol{x}_1 = 0$$

Proof.

$$\begin{aligned} (A\boldsymbol{x}_2)^\intercal \boldsymbol{x}_1 &= (\boldsymbol{x}_2^\intercal A^\intercal)\, \boldsymbol{x}_1 \overset{\text{symmetric}}{=} (\boldsymbol{x}_2^\intercal A)\, \boldsymbol{x}_1 \\ &= \boldsymbol{x}_2^\intercal (A\boldsymbol{x}_1) \overset{(e)}{=} \boldsymbol{x}_2^\intercal (\lambda \boldsymbol{x}_1) \\ &= \lambda \boldsymbol{x}_2^\intercal \boldsymbol{x}_1 \overset{(o)}{=} \lambda \cdot 0 = 0 \\ (A\boldsymbol{x}_2)^\intercal \boldsymbol{x}_1 &= 0 \Leftrightarrow A\boldsymbol{x}_2 \perp \boldsymbol{x}_1 \end{aligned}$$

□

Chapter 24

tangent half-angle formula

https://en.wikipedia.org/wiki/Tangent_half-angle_formula

https://zh.wikipedia.org/zh-tw/正切半角公式

正切半形公式又稱萬能公式

以切表弦公式，簡稱以切表弦

Chapter 25

homogeneous coordinate

https://youtu.be/EKN7dTJ4ep8?si=8woajZxbqPfEXhdK&t=2263

https://youtu.be/1z1S2kQKXDs?si=71o339yBtIQYhWtj&t=3082

Chapter 26

Archimedean property

26.1 integer Archimedean property

26.2 rational Archimedean property

https://math.stackexchange.com/questions/3699023/proof-the-the-field-of-rational-numbers-has-the-archimedean-property

https://math.stackexchange.com/questions/1919829/proving-the-archimedean-properties-of-rational-numbers

26.3 real Archimedean property

Chapter 27

MatPlotLib / matplotlib

- tikzplotlib[13.5]: Python[12] matplotlib[27] export to TikZ[13] .tex

27.1 Timothy H. Wu

巫孟叡

- API = application programming interface
  - functional[27.1.1]
  - object-oriented[27.1.2]
    * figure
    * axes
    * subplot

https://matplotlib.org/stable/tutorials/introductory/quick_start.html

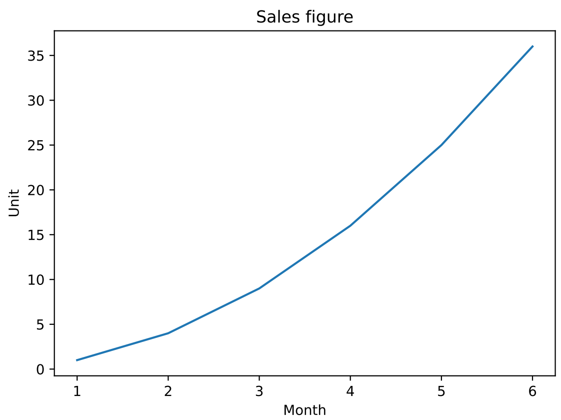
https://pbpython.com/effective-matplotlib.html

https://tex.stackexchange.com/questions/84847/can-i-use-webp-images-in-latex

You probably need to convert the image to png.

27.1.1 funcitonal API

```
import matplotlib.pyplot as plt

a1 = list(range(1, 7)) # [1, 2, 3, 4, 5, 6]
a2 = [1, 4, 9, 16, 25, 36]
plt.title('Sales figure')
plt.xlabel('Month')
plt.ylabel('Unit')
plt.plot(a1, a2) # this doesn't actually show
↪  the plot.
# plt.show() This is automatically called for
↪  Jupyter notebook.
```



To plot a scatterplot, call `scatter()` instead of `plot()`.

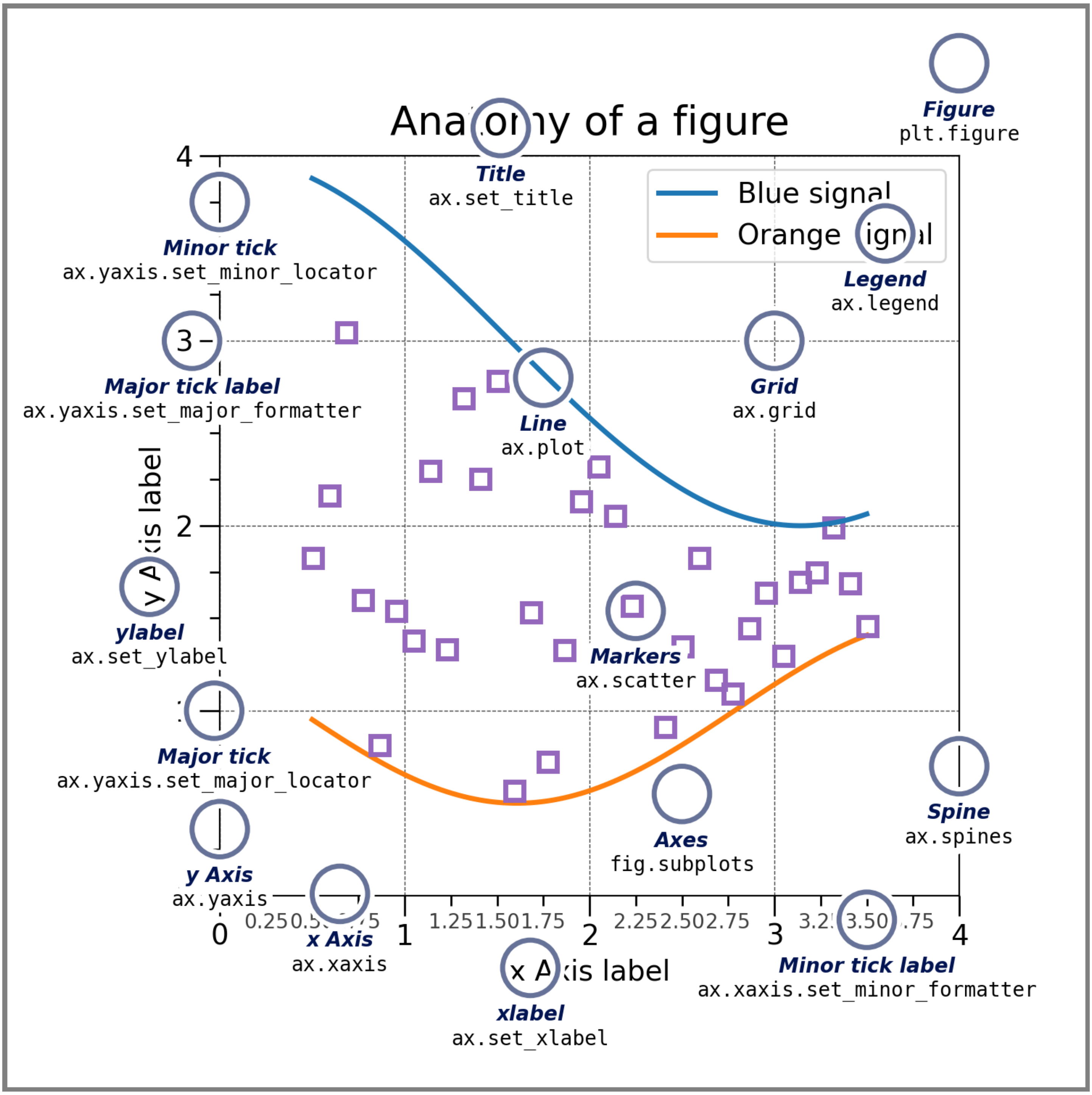


Figure 27.1: matplotlib figure anatomy

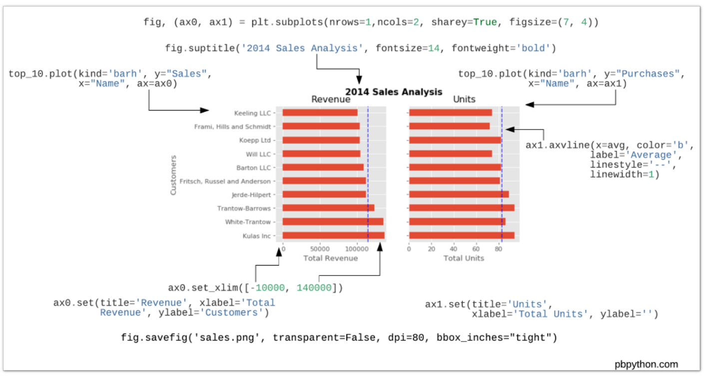
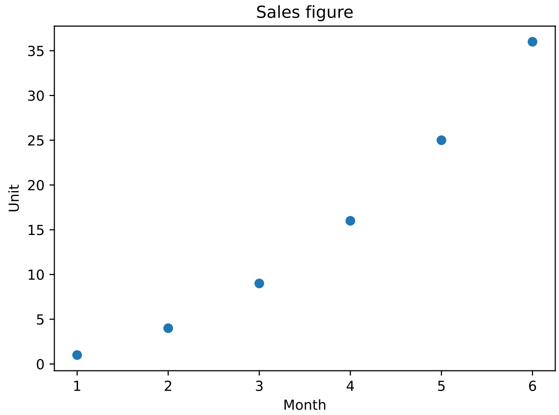
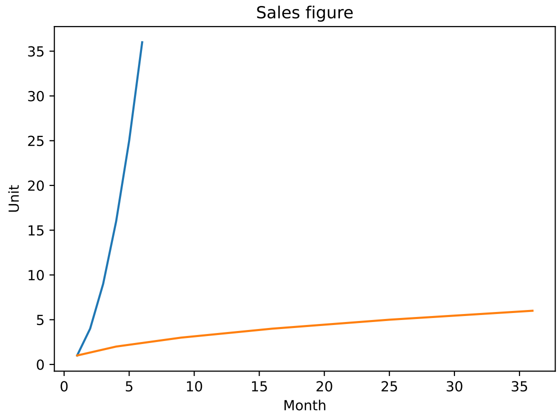


Figure 27.2: matplotlib subplot anatomy

```
import matplotlib.pyplot as plt

a1 = list(range(1, 7)) # [1, 2, 3, 4, 5, 6]
a2 = [1, 4, 9, 16, 25, 36]
plt.title('Sales figure')
plt.xlabel('Month')
plt.ylabel('Unit')
plt.scatter(a1, a2) # instead of plot.plot(),
↪  use scatter() to show scatter plot
```

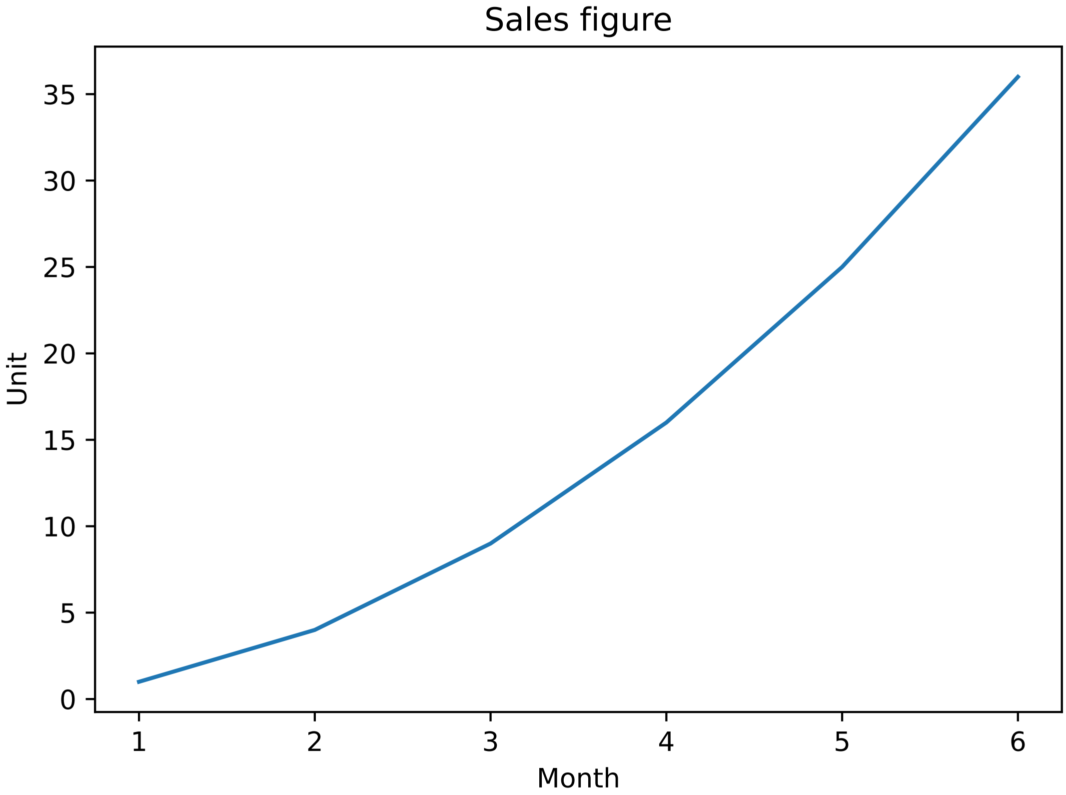


```
import matplotlib.pyplot as plt

a1 = list(range(1, 7)) # [1, 2, 3, 4, 5, 6]
a2 = [1, 4, 9, 16, 25, 36]
plt.title('Sales figure')
plt.xlabel('Month')
plt.ylabel('Unit')
plt.plot(a1, a2)
plt.plot(a2, a1)
```

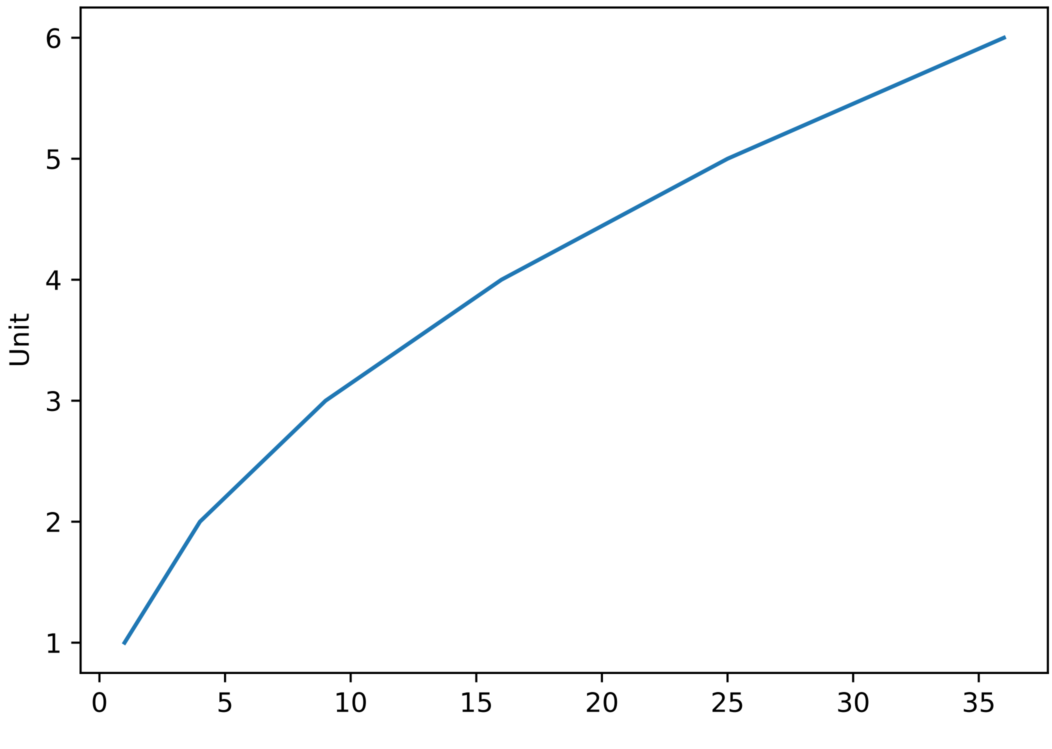


The behavior of the functional API is stateful. What's stateful? An example is when you read a text file. When you `open()` a text file to read, the library read the next line every time you call `readline()`. It rememember where you left off, despite the fact that you do not give it the position to read from. This behavior of the library is called **stateful**. The way we've used Matplotlib is also **stateful**. And everytime, `plot.show()` is called (and it automatically gets called on cell ends), some state about plots is reset. We can see that here:

```
import matplotlib.pyplot as plt

a1 = list(range(1, 7)) # [1, 2, 3, 4, 5, 6]
a2 = [1, 4, 9, 16, 25, 36]
plt.title('Sales figure')
plt.xlabel('Month')
plt.ylabel('Unit')
plt.plot(a1, a2)
```




```
plt.plot(a2, a1)
plt.ylabel('Unit')
```



It makes two graphs instead of one. Also note that `ylabel()` was called after `plot()`, and it is still shown before `plot.show()` but `Sales Figure` plot title and other labels don't show up on this graph. Because every time `plot.show()` is called, things are reset. This is a `stateful API` we're using. The functional APIs are used when you plot Matplotlib by calling on `pyplot` module level API (module level functions).

27.1.2 object-oriented API

In object-orinted API, we're getting two type of objects. One is `Figure`, the other one is `Axes`. Figure is the *canvas* of the plot. In English, axes is the plural form of axis. We're talking about the axis in x axis and y axis. Since one **plot** consists of both axis, in MatPlotlib the object that represents one plot is called `Axes`. Since it's an object. We'll call it "a" axes.

```
import matplotlib.pyplot as plt

a1 = list(range(1, 7)) # [1, 2, 3, 4, 5, 6]
a2 = [1, 4, 9, 16, 25, 36]
fig = plt.figure()
print("fig is of type:", type(fig))
```

```
## fig is of type: <class 'matplotlib.figure.Figure'>
```

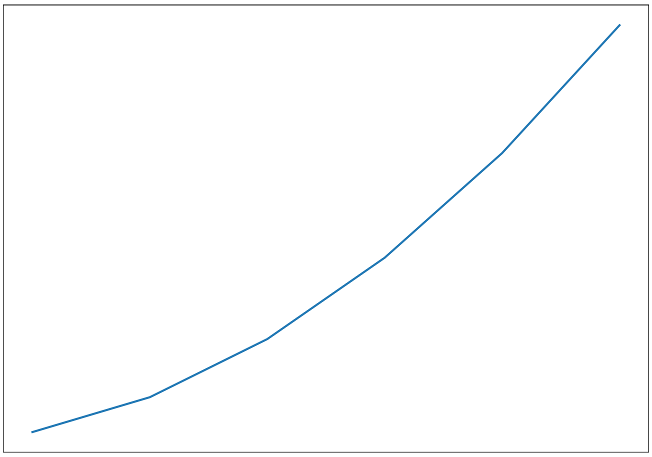
```
ax1 = fig.add_axes([0, 0, 1, 1]) # [left, bottom, width, height]
print("ax1 is of type:", type(ax1))
# ax1.plot(a1, a2)
```

```
## ax1 is of type: <class 'matplotlib.axes._axes.Axes'>
```

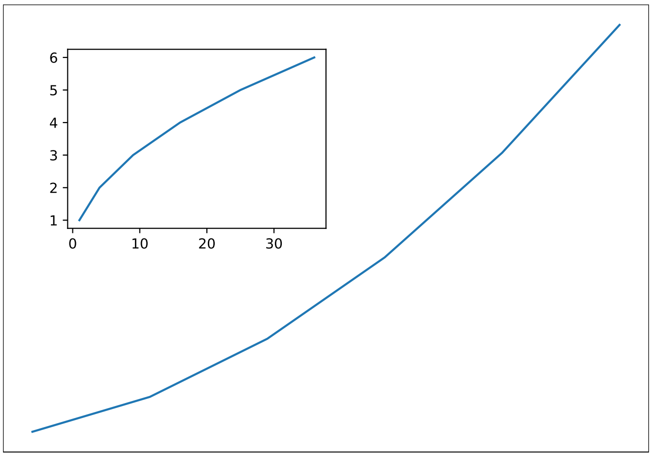
1. Call `figure()` to get a Figure type
2. Call `add_axes()` to get a ax1 type `ax1 = fig.add_axes([0, 0, 1, 1]) # [left, bottom, width, height]`

- The list given to `add_axes()` is the rectangular region of where to show the plot:
  - Bottom left corner at x=0, y=0, width and height of both 1, 1

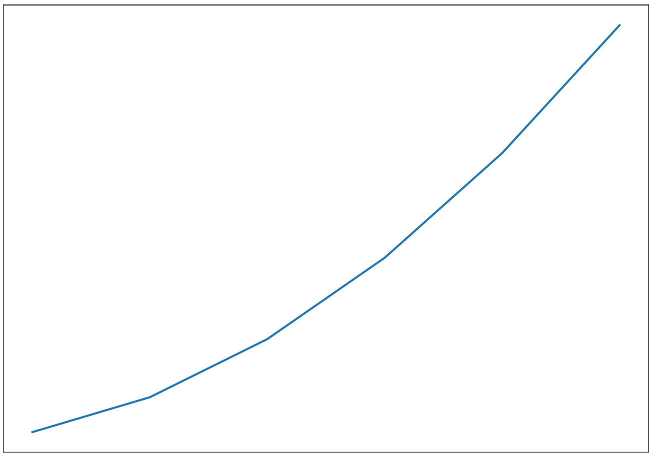
```
import matplotlib.pyplot as plt

a1 = list(range(1, 7)) # [1, 2, 3, 4, 5, 6]
a2 = [1, 4, 9, 16, 25, 36]
fig = plt.figure()
# print("fig is of type:", type(fig))
ax1 = fig.add_axes([0, 0, 1, 1]) # [left,
↪  bottom, width, height]
# print("ax1 is of type:", type(ax1))
ax1.plot(a1, a2)
```



```
import matplotlib.pyplot as plt

a1 = list(range(1, 7)) # [1, 2, 3, 4, 5, 6]
a2 = [1, 4, 9, 16, 25, 36]
fig = plt.figure()
ax1 = fig.add_axes([0, 0, 1, 1])
ax1.plot(a1, a2)
ax2 = fig.add_axes([0.1, 0.5, .4, .4])
ax2.plot(a2, a1)
```



```
import matplotlib.pyplot as plt

a1 = list(range(1, 7)) # [1, 2, 3, 4, 5, 6]
a2 = [1, 4, 9, 16, 25, 36]
fig = plt.figure()
ax = fig.add_axes([0, 0, 1, 1])
ax.plot(a1, a2)
ax.set_xlabel('Time')
ax.set_ylabel('Unit')
ax.set_title('Sales figure')
# alternatively:
# ax.set(xlabel='Time', ylabel='Unit',
↪   title='Sales figure')
```



27.1.2.1 Configure the figure size and DPI

Get image size for the figure object. 6 by 4 is the default.

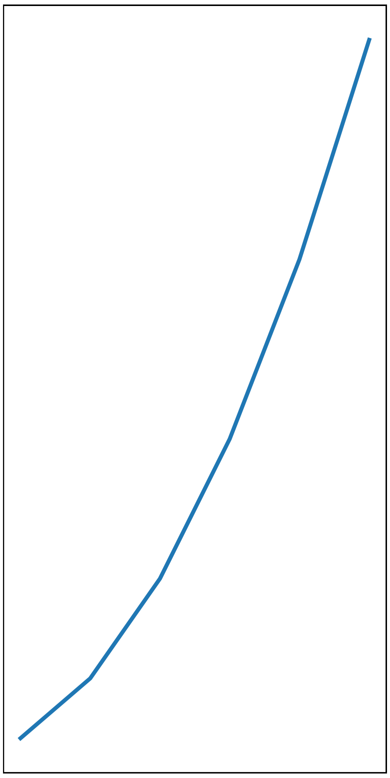
```
fig.get_size_inches()
```

```
import matplotlib.pyplot as plt

fig = plt.figure()
fig.get_size_inches()
```

```
## array([6.5, 4.5])
```

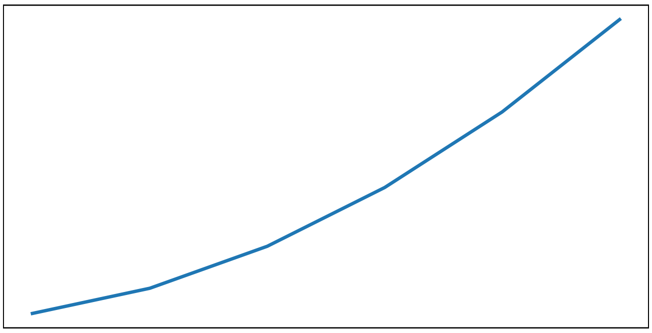
```
fig = plt.figure(figsize=(2, 4))
```

```
import matplotlib.pyplot as plt

a1 = list(range(1, 7)) # [1, 2, 3, 4, 5, 6]
a2 = [1, 4, 9, 16, 25, 36]
fig = plt.figure(figsize=(2, 4))
# you can also set after getting the figure
# fig.set_size_inches((12, 2))
ax = fig.add_axes([0, 0, 1, 1])
ax.plot(a1, a2)
```



```
fig.set_size_inches((4, 2))
```

```
import matplotlib.pyplot as plt

a1 = list(range(1, 7)) # [1, 2, 3, 4, 5, 6]
a2 = [1, 4, 9, 16, 25, 36]
fig = plt.figure()
# you can also set after getting the figure
fig.set_size_inches((4, 2))
ax = fig.add_axes([0, 0, 1, 1])
ax.plot(a1, a2)
```



DPI = dots per inch

Get image DPI for the figure object. 100 is the default here.

fig.get_dpi()

```
import matplotlib.pyplot as plt

fig = plt.figure()
fig.get_dpi()
```

```
## 100.0
```

```
import matplotlib.pyplot as plt

a1 = list(range(1, 7)) # [1, 2, 3, 4, 5, 6]
a2 = [1, 4, 9, 16, 25, 36]
fig = plt.figure()
# you can also set after getting the figure
fig.set_size_inches((12, 2))
ax = fig.add_axes([0, 0, 1, 1])
ax.plot(a1, a2)
plt.title('Sales figure')
fig.get_dpi()
```

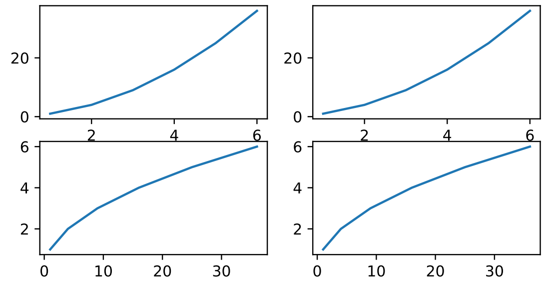
```
## 100.0
```

27.1.2.2 subplot

```
import matplotlib.pyplot as plt

# Subplots handles add_axes for you according to the number of rows and columns
fig, axes = plt.subplots(nrows=2, ncols=2, figsize=(6, 3))
# axes is a numpy array, you can use it like using a list.
print(axes)
```

```
## [[<Axes: > <Axes: >]
##  [<Axes: > <Axes: >]]
```

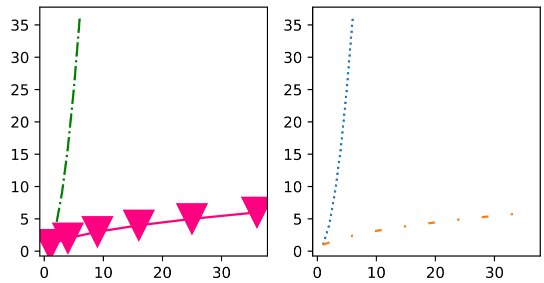
```
import matplotlib.pyplot as plt

a1 = list(range(1, 7)) # [1, 2, 3, 4, 5, 6]
a2 = [1, 4, 9, 16, 25, 36]
# Subplots handles add_axes for you according
↪  to the number of rows and columns
fig, axes = plt.subplots(nrows=2, ncols=2,
↪  figsize=(6, 3))
# axes is a numpy array, you can use it like
↪  using a list.
# print(axes)
axes[0][0].plot(a1, a2)
axes[0][1].plot(a1, a2)
axes[1][0].plot(a2, a1)
axes[1][1].plot(a2, a1)
```



27.1.2.3 color and linestyle

- Color
  - https://matplotlib.org/stable/tutorials/colors/colors.html
- line-style (ls)
  - https://matplotlib.org/stable/gallery/lines_bars_and_markers/linestyles.html
- marker
  - https://matplotlib.org/stable/api/markers_api.html
- linewidth (lw)

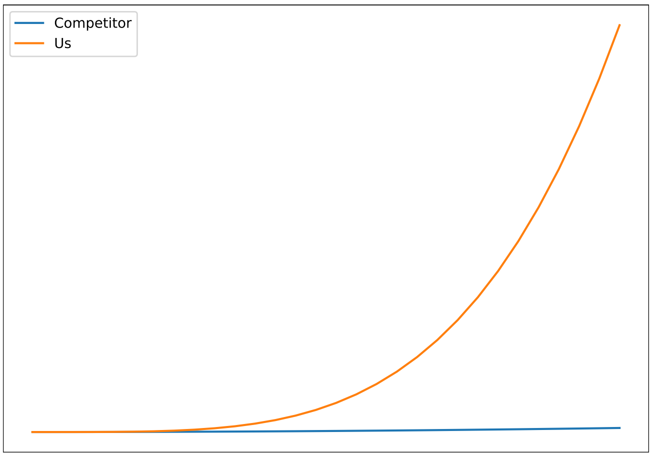
```
import matplotlib.pyplot as plt

a1 = list(range(1, 7)) # [1, 2, 3, 4, 5, 6]
a2 = [1, 4, 9, 16, 25, 36]
# Subplots handles add_axes for you according
↪  to the number of rows and columns
fig, axes = plt.subplots(nrows=1, ncols=2,
↪  figsize=(6, 3))
# axes is a numpy array
axes[0].plot(a1, a2, color='green',
↪  linestyle='-.')
axes[0].plot(a2, a1, color=(1, 0, 0.5),
↪  marker='v', markersize=20)
axes[1].plot(a1, a2, linestyle='dotted')
axes[1].plot(a2, a1, linestyle=(0, (3, 10, 1,
↪  10)))
```



27.1.2.4 other inputs

- NumPy array
- Pandas series

27.1.2.5 legend




```
import matplotlib.pyplot as plt
import numpy as np

x = np.linspace(0, 10, 30)
y = x * x
fig = plt.figure()
ax = fig.add_axes([0, 0, 1, 1])
line1 = ax.plot(x, y, label="Competitor")
line2 = ax.plot(x, y**2, label="Us")
ax.legend()
```

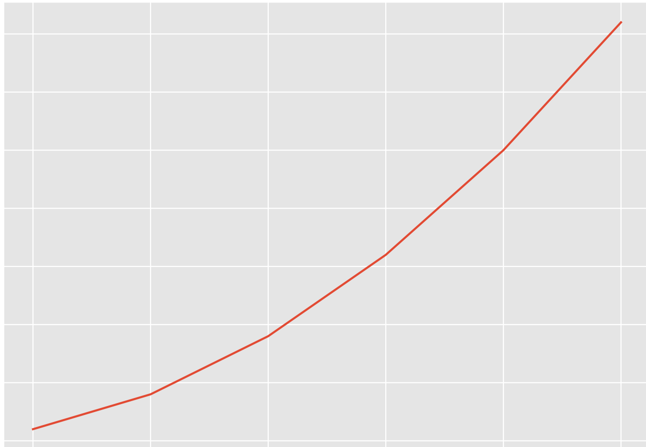
27.1.2.6 customize style

predefined styles

```
import matplotlib.pyplot as plt

print(plt.style.available)
```

```
## ['Solarize_Light2', '_classic_test_patch', '_mpl-gallery', '_mpl-gallery-nogrid', 'bmh', 'classic', 'dark_
```

```
import matplotlib.pyplot as plt

a1 = list(range(1, 7)) # [1, 2, 3, 4, 5, 6]
a2 = [1, 4, 9, 16, 25, 36]
fig = plt.figure()
plt.style.use('ggplot')
ax = fig.add_axes([0, 0, 1, 1])
ax.plot(a1, a2)
```



restore default style

```
import matplotlib.pyplot as plt

plt.style.use('default') # But strangely enough figure size gets changed still
```

```
import matplotlib.pyplot as plt

plt.rcParams["figure.figsize"] = (6, 4)
plt.rcParams["figure.dpi"] = 100
```

27.1.2.7 save to file

> `savefig` saves image to file. We also set the `bbox_inches` parameter to `tight` to make sure the image doesn't get out of the image bound.

save to .png

```
import matplotlib.pyplot as plt

a1 = list(range(1, 7)) # [1, 2, 3, 4, 5, 6]
a2 = [1, 4, 9, 16, 25, 36]
fig = plt.figure()
ax = fig.add_axes([0, 0, 1, 1])
ax.plot(a1, a2)
fig.savefig('out.png', bbox_inches = 'tight')
```

save to .pdf

```
import matplotlib.pyplot as plt

a1 = list(range(1, 7)) # [1, 2, 3, 4, 5, 6]
a2 = [1, 4, 9, 16, 25, 36]
fig = plt.figure()
ax = fig.add_axes([0, 0, 1, 1])
ax.plot(a1, a2)
fig.savefig('out.pdf', bbox_inches = 'tight')
```

27.1.3 Seaborn

- Kimberly Fessel

  - visually explained
  - Seaborn
  - MatPlotLib
  - Pandas
  - iPyWidgets

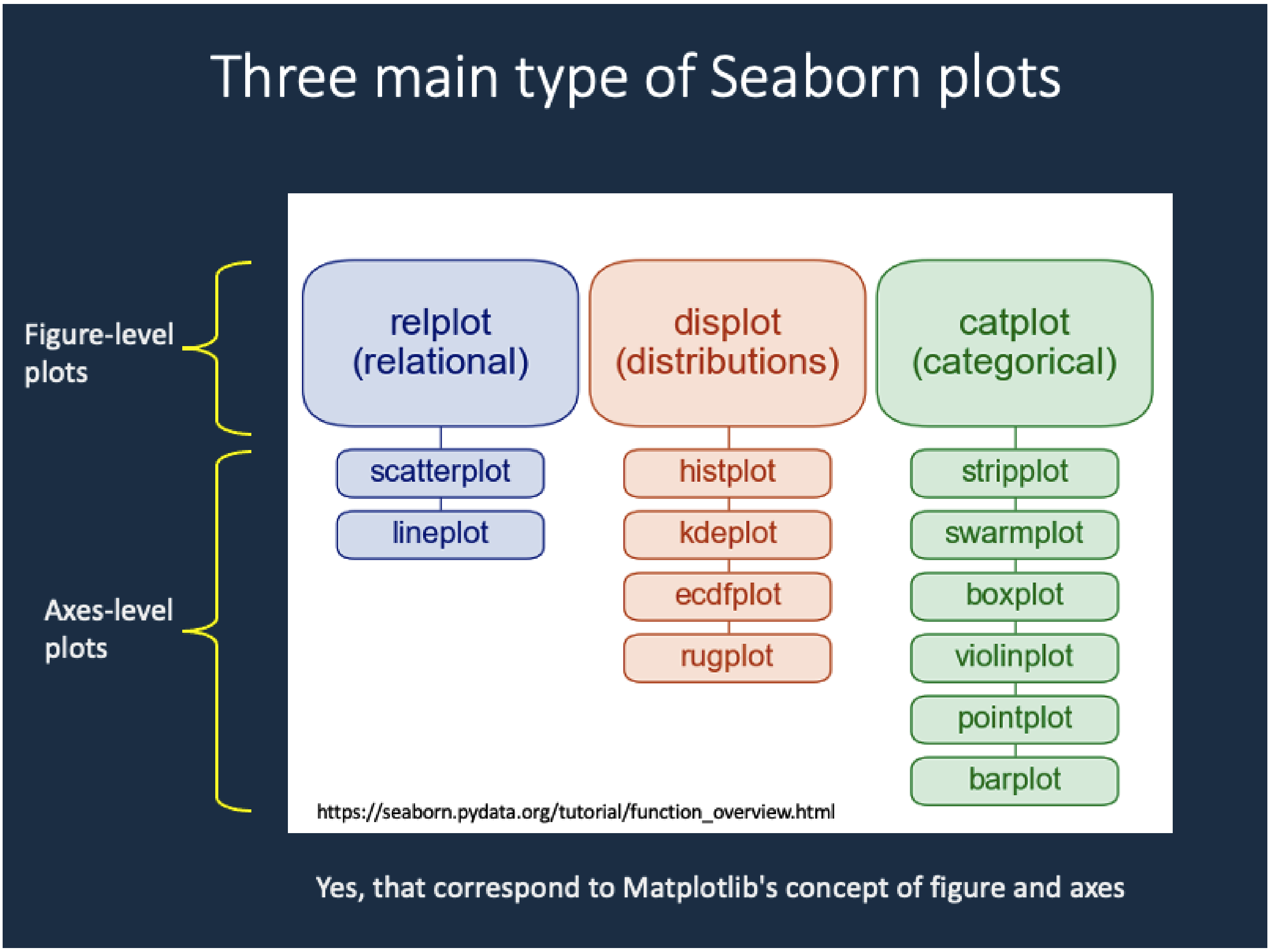


Figure 27.3: seaborn plot type

- Seaborn
  - figure-level plot
  - axes-level plot

R or RStudio run Python with installing packages or modules by using `reticulate`, R package

and directly using Anaconda `conda` environment for convenience, instead of `virtualenv`

https://rstudio.github.io/reticulate/articles/python_packages.html

```
library(reticulate)
```

```
## Warning: package 'reticulate' was built under R version 4.2.3
```

```
conda_list()
```

```
##                 name                                    python
## 1               base                  D:\\Anaconda3/python.exe
## 2           fiftyone D:\\Anaconda3\\envs\\fiftyone/python.exe
## 3              keras    D:\\Anaconda3\\envs\\keras/python.exe
## 4            labelme  D:\\Anaconda3\\envs\\labelme/python.exe
## 5              manim    D:\\Anaconda3\\envs\\manim/python.exe
```

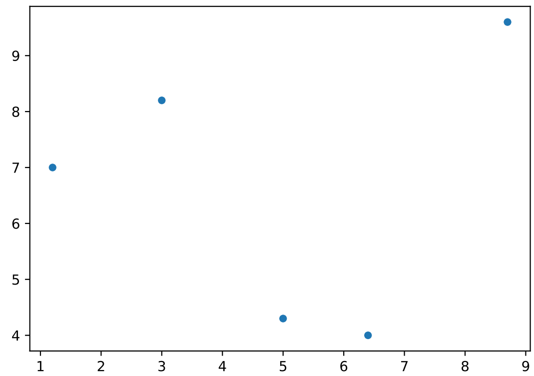
```
## 6              mmyolo            D:\\Anaconda3\\envs\\mmyolo/python.exe
## 7        r-reticulate      D:\\Anaconda3\\envs\\r-reticulate/python.exe
## 8  rsconnect-jupyter D:\\Anaconda3\\envs\\rsconnect-jupyter/python.exe
## 9             sandbox           D:\\Anaconda3\\envs\\sandbox/python.exe
## 10        sandbox-3.9       D:\\Anaconda3\\envs\\sandbox-3.9/python.exe
```

```
use_condaenv(condaenv = 'sandbox-3.9')

## install Seaborn
# conda_install("r-reticulate", "seaborn")

## import Seaborn (it will be automatically discovered in "r-reticulate")
seaborn <- import("seaborn")
```

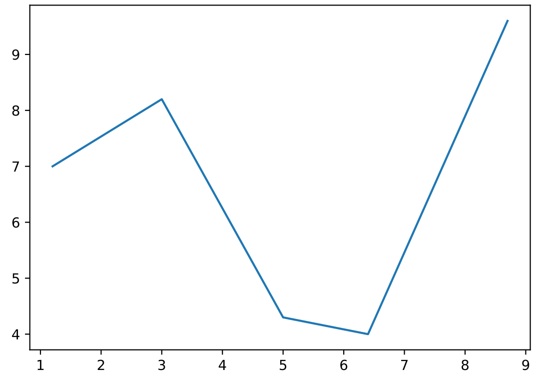
27.1.3.1 basic

```
import matplotlib.pyplot as plt  # need it sometimes
import seaborn as sns

sns.set_theme() # set the default theme
```

```
import seaborn as sns

x = [3, 5, 1.2, 8.7, 6.4]
y = [8.2, 4.3, 7, 9.6, 4]
sns.scatterplot(x=x, y=y)
```



```
import seaborn as sns

x = [3, 5, 1.2, 8.7, 6.4]
y = [8.2, 4.3, 7, 9.6, 4]
sns.lineplot(x=x, y=y)
```



27.1.3.2 data frame

```
sns.load_dataset
```

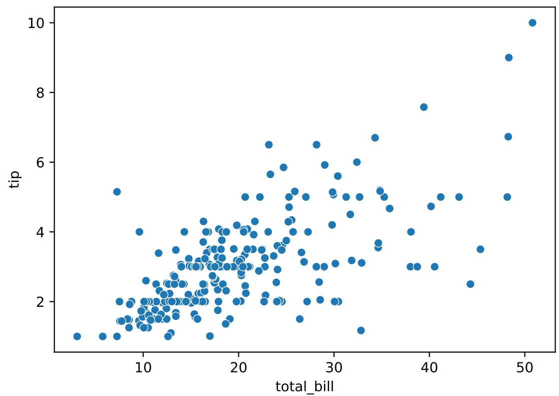
```
import seaborn as sns

tips = sns.load_dataset("tips")
tips
```

```
##      total_bill   tip     sex smoker   day    tim
## 0         16.99  1.01  Female     No   Sun  Dinne
## 1         10.34  1.66    Male     No   Sun  Dinne
## 2         21.01  3.50    Male     No   Sun  Dinne
## 3         23.68  3.31    Male     No   Sun  Dinne
## 4         24.59  3.61  Female     No   Sun  Dinne
## ..          ...   ...     ...    ...   ...    ..
## 239       29.03  5.92    Male     No   Sat  Dinne
## 240       27.18  2.00  Female    Yes   Sat  Dinne
## 241       22.67  2.00    Male    Yes   Sat  Dinne
## 242       17.82  1.75    Male     No   Sat  Dinne
## 243       18.78  3.00  Female     No  Thur  Dinne
##
## [244 rows x 7 columns]
```

---

data

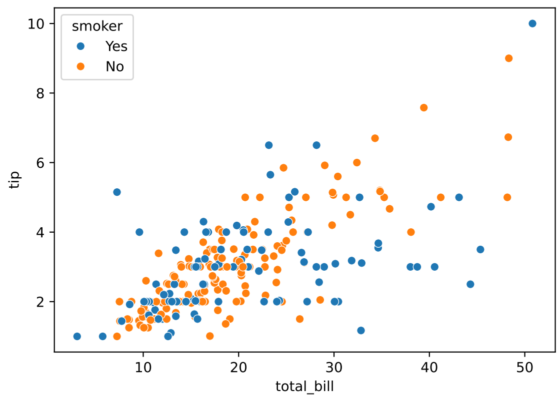
```
import seaborn as sns

tips = sns.load_dataset("tips")
sns.scatterplot(data=tips,
                x='total_bill',
                y='tip')
```



---

hue

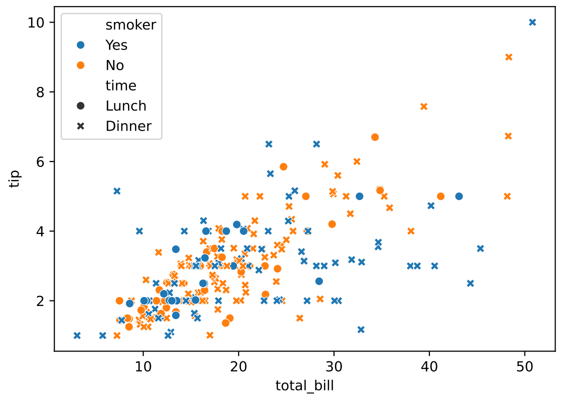
```
import seaborn as sns

tips = sns.load_dataset("tips")
sns.scatterplot(data=tips,
                x='total_bill',
                y='tip',
                hue='smoker')
```



---

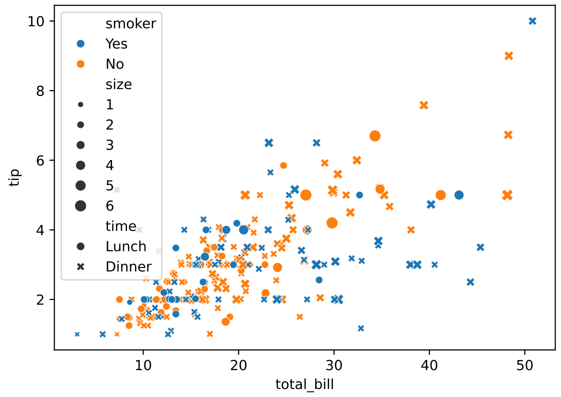
style

```
import seaborn as sns

tips = sns.load_dataset("tips")
sns.scatterplot(data=tips,
                x='total_bill',
                y='tip',
                hue='smoker',
                style='time')
```



`size`

```
import seaborn as sns

tips = sns.load_dataset("tips")
sns.scatterplot(data=tips,
                x='total_bill',
                y='tip',
                hue='smoker',
                style='time',
                size='size')
```



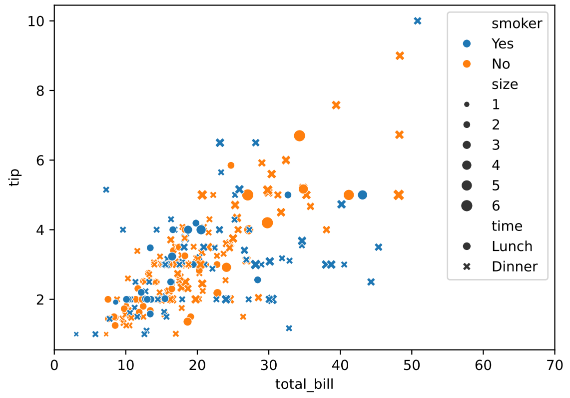
- Legend is covering up the graph, it's getting out of hand. Let's tune the range of x
- `sns.scatterplot()` actually returns something that resembles Matplotlib axis. So we use a Matplotlib axis function:

```
ax = sns.scatterplot(...)
```

https://stackoverflow.com/questions/26597116/seaborn-plots-not-showing-up

```
import matplotlib.pyplot as plt
import seaborn as sns

tips = sns.load_dataset("tips")
ax = sns.scatterplot(data=tips,
                x='total_bill',
                y='tip',
                hue='smoker',
                style='time',
                size='size')
ax.set_xlim((0, 70))
# alternatively:
# ax.set(xlim=(0, 70))
plt.show()
```

```
## (0.0, 70.0)
```



27.1.3.3 axis-level plot and figure-level plot

`sns.relplot`

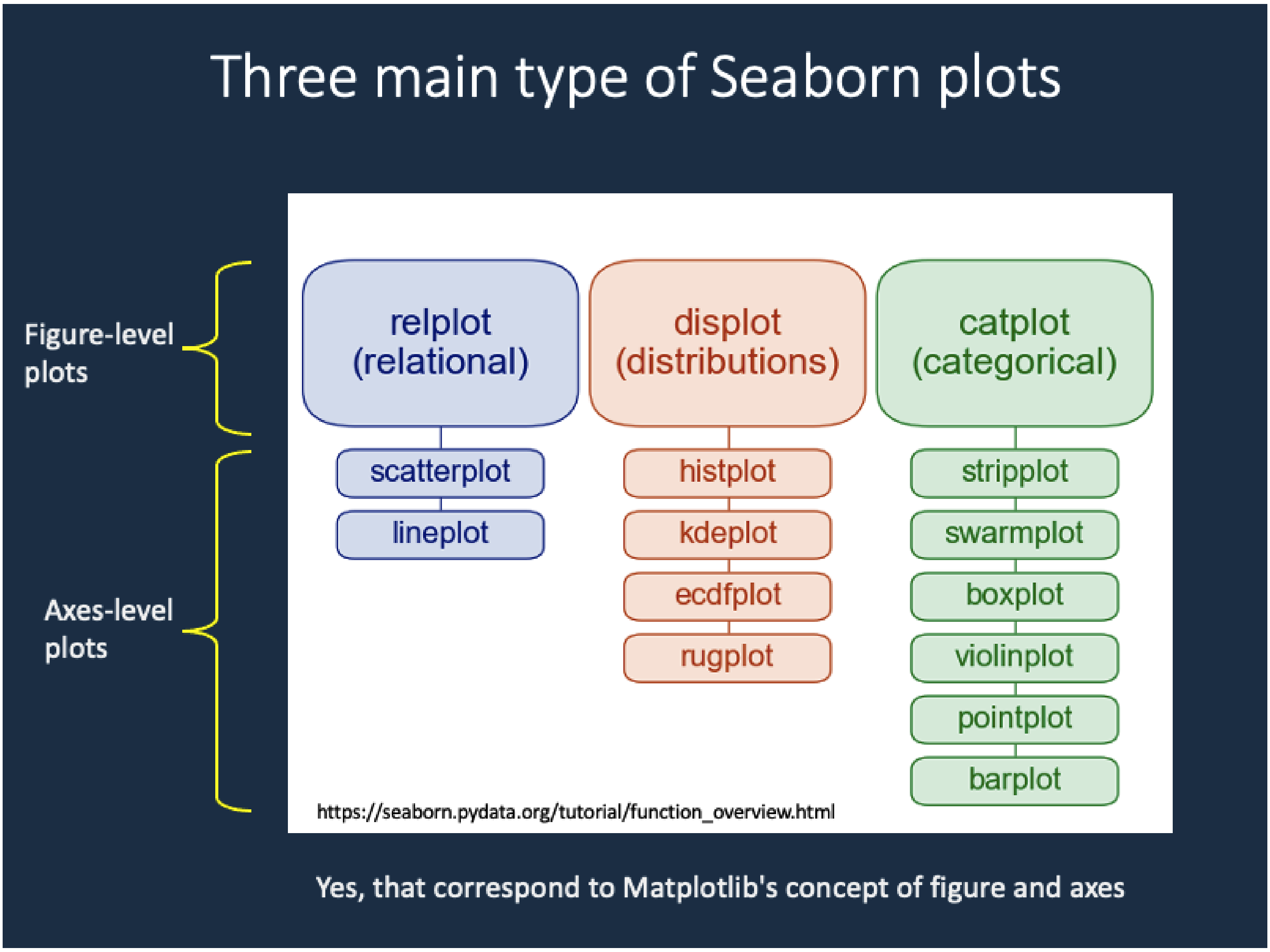
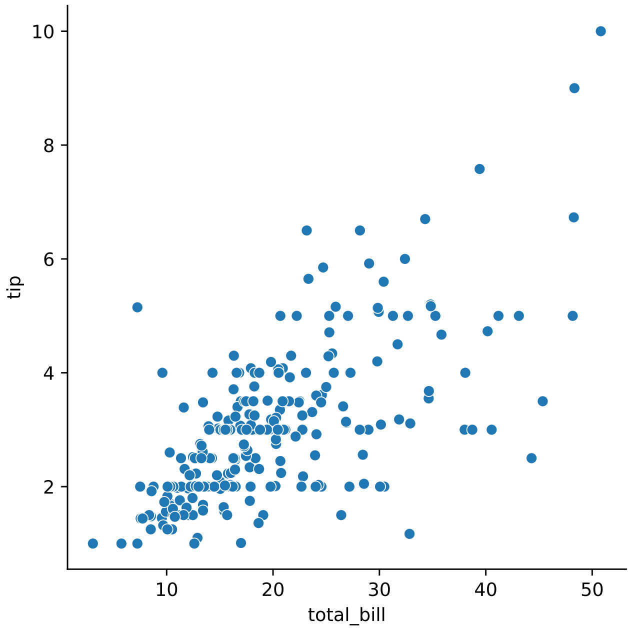
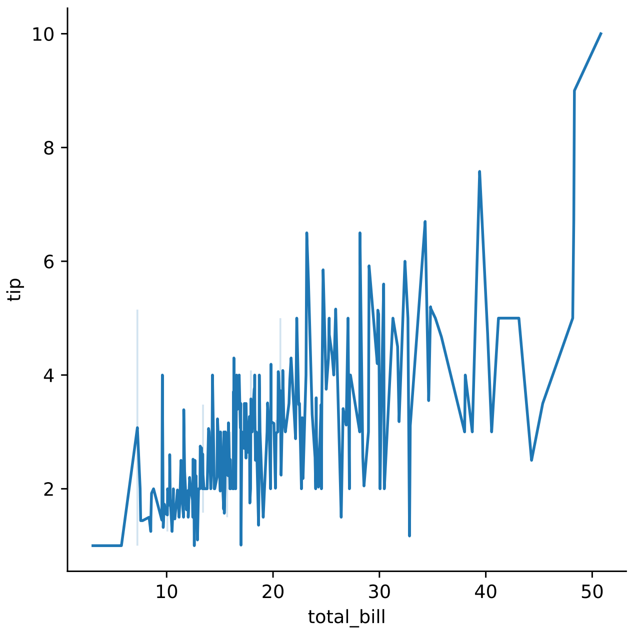


Figure 27.4: seaborn plot type

```
import seaborn as sns

tips = sns.load_dataset("tips")
sns.relplot(kind='scatter',
            data=tips,
            x='total_bill',
            y='tip')
```




```
import seaborn as sns

tips = sns.load_dataset("tips")
sns.relplot(kind='line',
            data=tips,
            x='total_bill',
            y='tip')
```

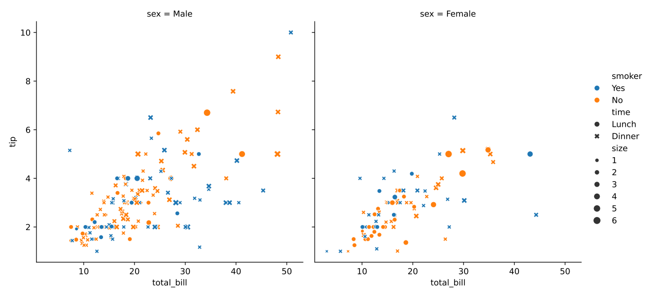


With figure-level plot, we can draw more than one plot (one `axes`).

Here we specify that different `sex` be on different column by specifying `col=sex`.

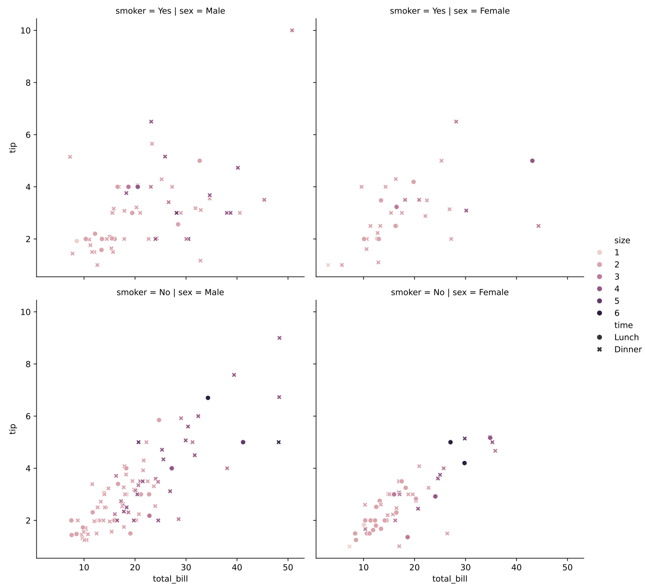
`col=sex`

```
import seaborn as sns

tips = sns.load_dataset("tips")
sns.relplot(kind='scatter',
            data=tips,
            x='total_bill',
            y='tip',
            hue='smoker',
            style='time',
            size='size',
            col='sex')
```



`col=sex row=smoker`

```
import seaborn as sns

tips = sns.load_dataset("tips")
sns.relplot(kind='scatter',
            data=tips,
            x='total_bill',
            y='tip',
            style='time',
            hue='size',
            row='smoker',
            col='sex')
```

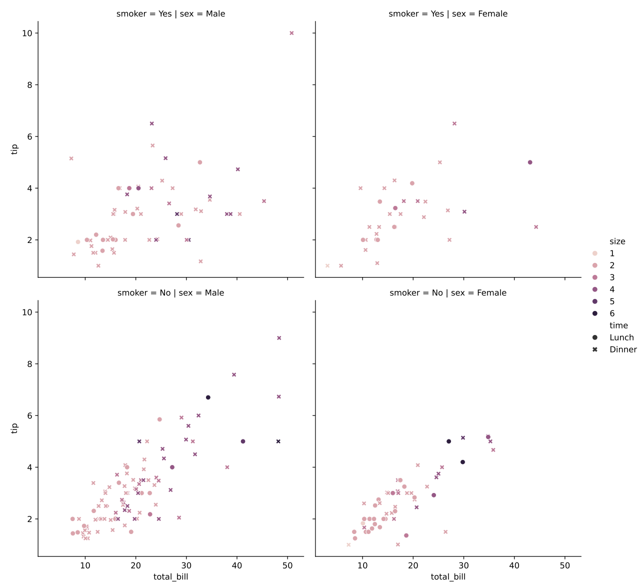


27.1.3.4 accessing fgure and axes objects

Recall that Seaborn uses Matplotlib to draw the graphics. So underneath the Seaborn libray, you can still access Matplotlib's figure object and axes objects if necessary. The call to figure-level plot returns an object.

```
import seaborn as sns

tips = sns.load_dataset("tips")
g = sns.relplot(kind='scatter',
                data=tips,
                x='total_bill',
                y='tip',
                style='time',
                hue='size',
                row='smoker',
                col='sex')
print(type(g))
print(type(g.fig)) # g.fig gets you the Figure
g.fig
```

```
## D:\Anaconda3\envs\sandbox-3.9\lib\site-packages\s
##   self._figure.tight_layout(*args, **kwargs)
## <class 'seaborn.axisgrid.FacetGrid'>
## <class 'matplotlib.figure.Figure'>
```

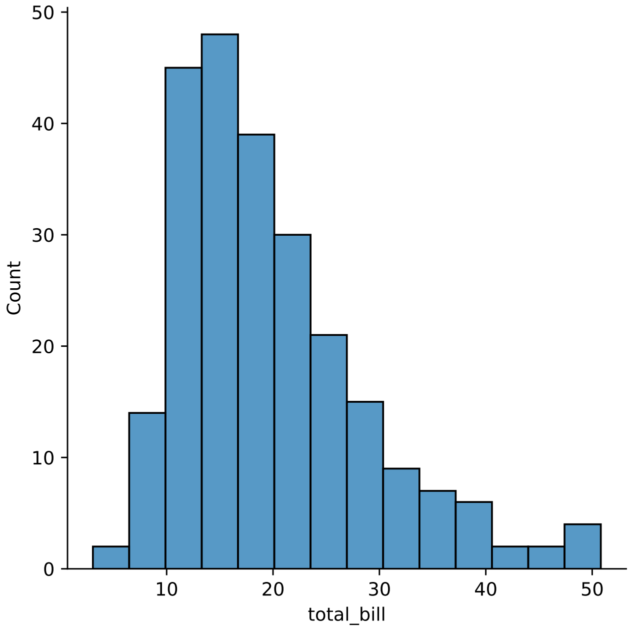


27.1.3.5 distribution plot

```
sns.displot
```

27.1.3.5.1 histogram

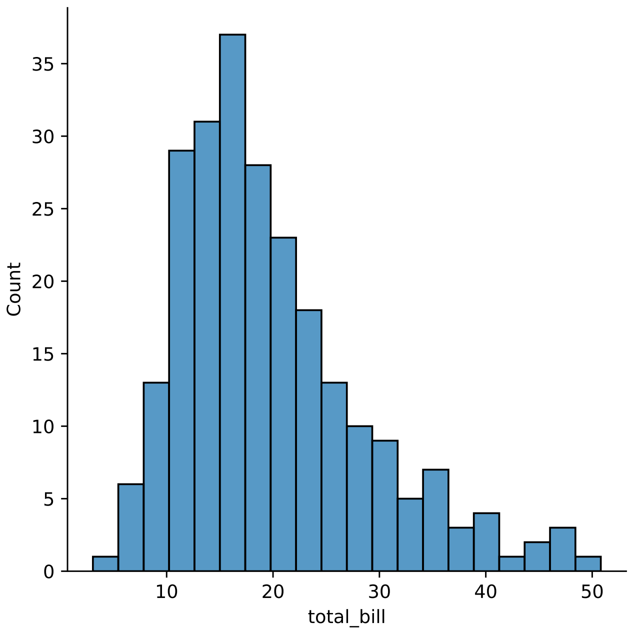
```
import seaborn as sns

tips = sns.load_dataset("tips")
sns.displot(kind='hist',
            data=tips,
            x='total_bill')
```



---

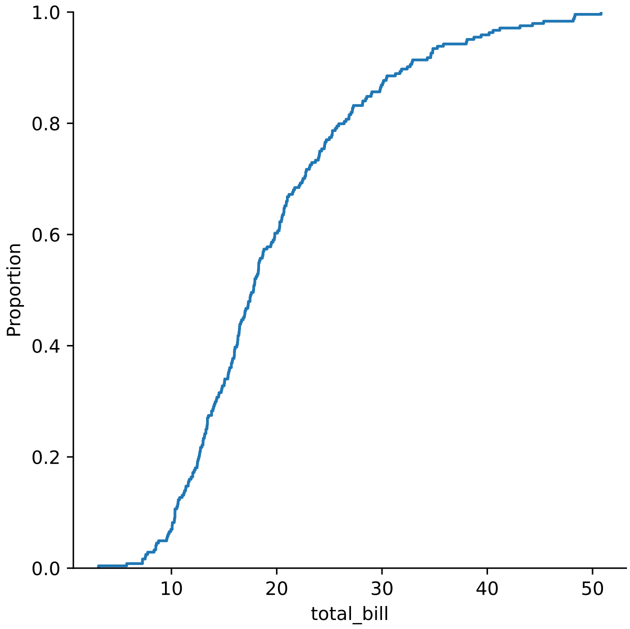
bins

```
import seaborn as sns

tips = sns.load_dataset("tips")
sns.displot(kind='hist', bins=20,
            data=tips,
            x='total_bill')
```



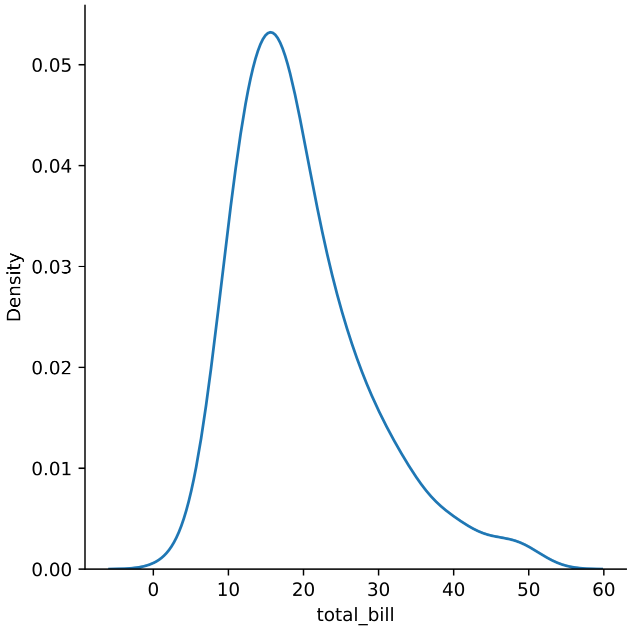
27.1.3.5.2 ECDF = empirical cumulative distrutive function

```
import seaborn as sns

tips = sns.load_dataset("tips")
sns.displot(kind='ecdf',
            data=tips,
            x='total_bill')
```



27.1.3.5.3 KDE = kernel density estimation

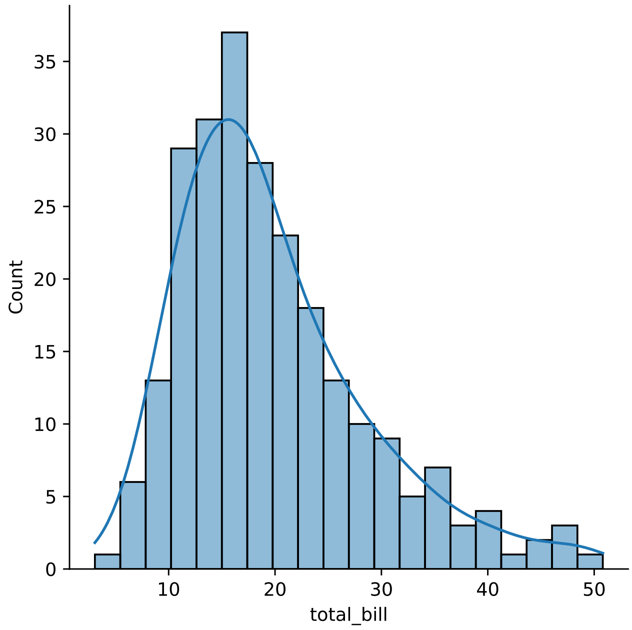
```
import seaborn as sns

tips = sns.load_dataset("tips")
sns.displot(kind='kde',
            data=tips,
            x='total_bill')
```



---

histogram with KDE

```python
import seaborn as sns

tips = sns.load_dataset("tips")
sns.displot(kind='hist', bins=20, kde='true',
            data=tips,
            x='total_bill')
```




```
## '% This file was created with tikzplotlib v0.10.1
```

TikZ / `tikzpicture` with PGFplots `axis` by transforming Matplotlib-based Seaborn plot to .tex via Python package `tikzplotlib`

`xlabel={total_bill},` changed to `xlabel={total bill},` wihout _ in text

X `ylabel={Count},` changed to `ylabel={$\text{Count}$},`: not necessoary to be changed, the problem is _ in xlabel name

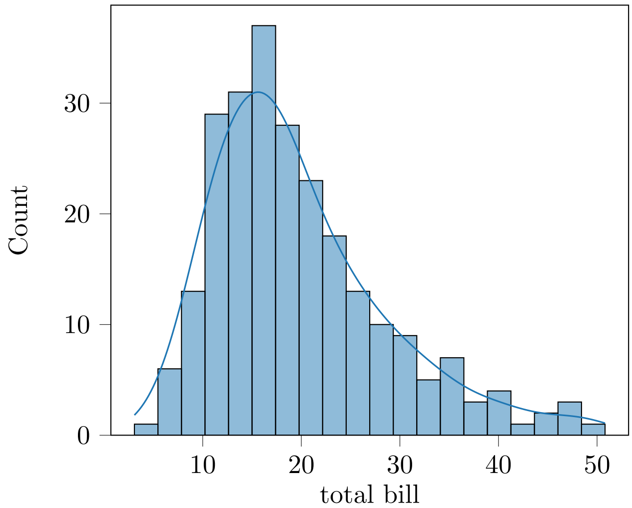
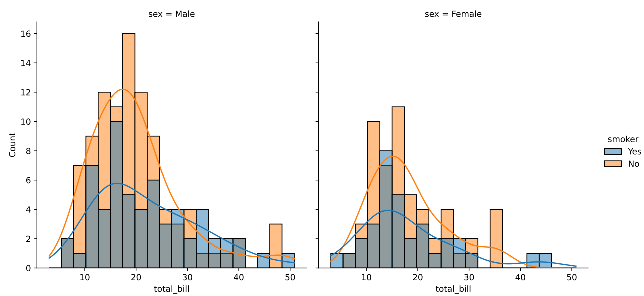


Figure 27.5: tikzplotlib

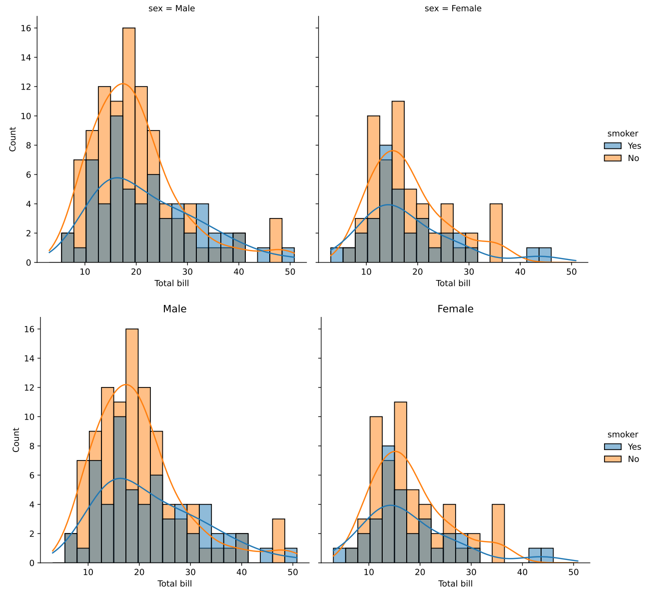
```python
import seaborn as sns

tips = sns.load_dataset("tips")
sns.displot(kind='hist', bins=20, kde='true',
            data=tips,
            x='total_bill',
            hue='smoker',
            col='sex'
            )
```



Let's cutomize the label and title.

```
import seaborn as sns

tips = sns.load_dataset("tips")
g = sns.displot(kind='hist', bins=20, kde='true',
                data=tips,
                x='total_bill',
                hue='smoker',
                col='sex'
                )
g.set(xlabel='Total bill')
g.axes[0][0].set(title='Male')
g.axes[0][1].set(title='Female')
g.fig
```

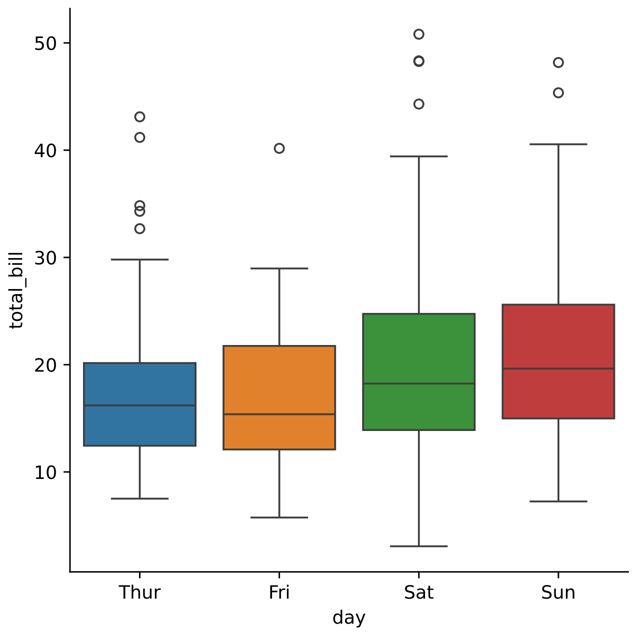
```
## D:\Anaconda3\envs\sandbox-3.9\lib\site-packages\s
##   self._figure.tight_layout(*args, **kwargs)
```



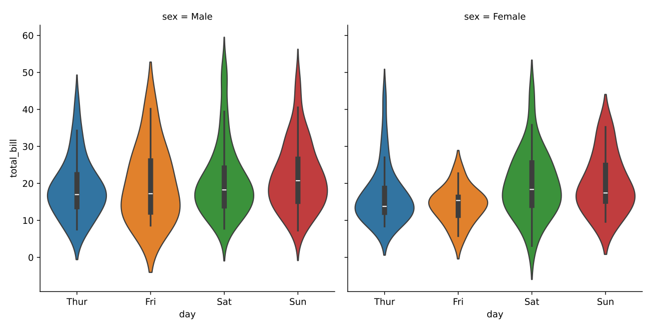
27.1.3.6 categorical plot

`sns.catplot`

```
import seaborn as sns

tips = sns.load_dataset("tips")
sns.catplot(kind='box',
            data=tips,
            x="day", hue="day",
            y="total_bill"
            )
```



```
import seaborn as sns

tips = sns.load_dataset("tips")
sns.catplot(kind='violin',
            data=tips,
            x="day", hue="day",
            y="total_bill",
            col='sex'
            )
```

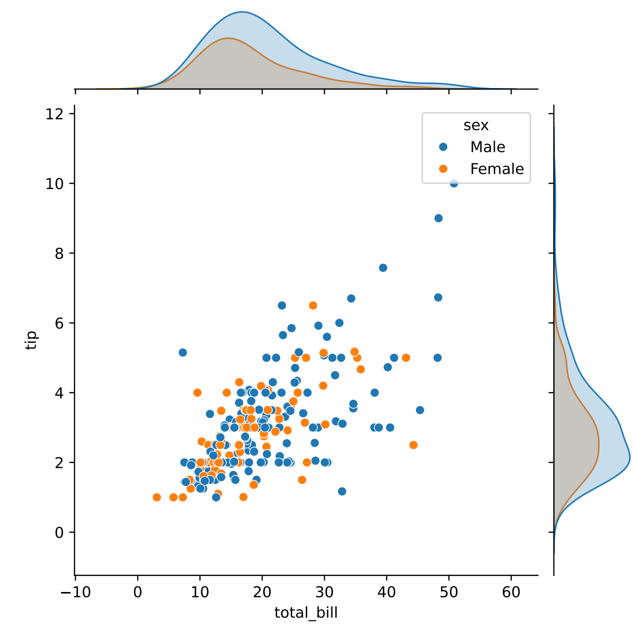


27.1.3.7 joint plot

`sns.jointplot`

```
import matplotlib.pyplot as plt
import seaborn as sns

tips = sns.load_dataset("tips")
sns.jointplot(data=tips,
              x="total_bill",
              y="tip",
              hue='sex'
              )
plt.show()
```

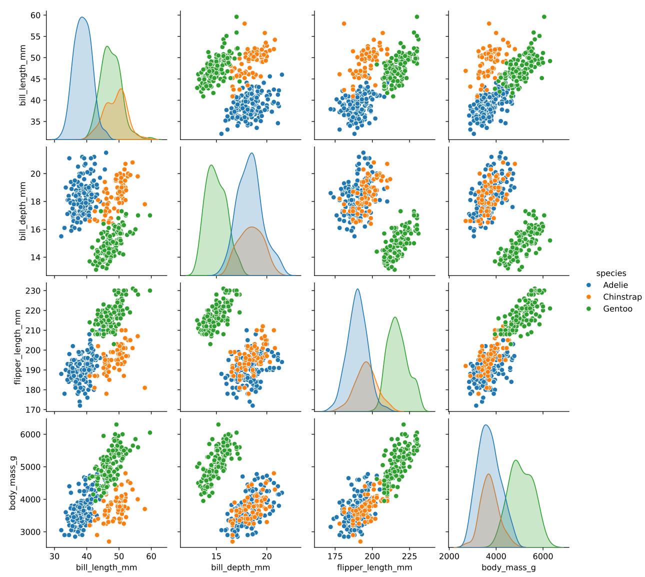
```
## <seaborn.axisgrid.JointGrid object at 0x000001EE7
```



27.1.3.8 pair plot

sns.pairplot

```
import seaborn as sns

penguins = sns.load_dataset("penguins")
sns.pairplot(data=penguins,
             hue='species')
```

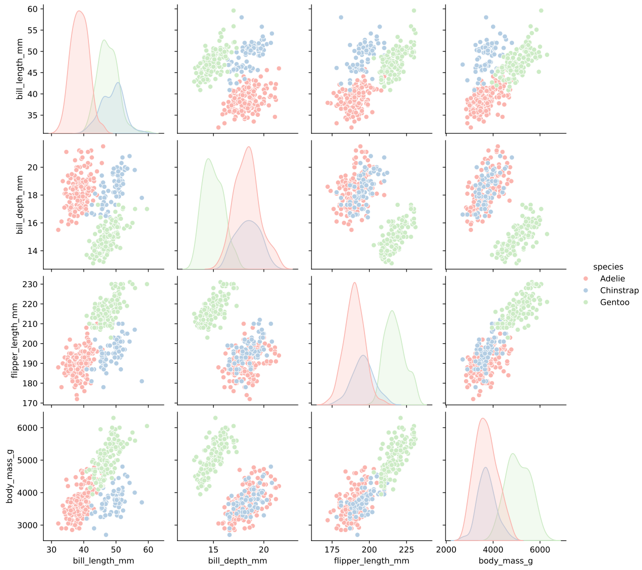


27.1.3.9 palette

Changing the color with Matplotlib's color

A list of color: https://matplotlib.org/stable/tutorials/colors/colormaps.html

```python
import seaborn as sns

penguins = sns.load_dataset("penguins")
sns.pairplot(data=penguins,
             hue='species',
             palette='Pastel1')
```



27.1.3.10 volcano plot

27.1.3.11 heatmap

27.2 export

27.2.1 .svg

https://stackoverflow.com/questions/24525111/how-can-i-get-the-output-of-a-matplotlib-plot-as-an-svg

```python
import matplotlib.pyplot as plt
import seaborn as sns

penguins = sns.load_dataset("penguins")
sns.pairplot(data=penguins,
             hue='species')

plt.savefig("test.svg")
# plt.savefig("test.svg", dpi=1200)
```

27.2.2 .eps

https://stackoverflow.com/questions/16183462/saving-images-in-python-at-a-very-high-quality

The PostScript backend does not support transparency; partially transparent artists will be rendered opaque.

```python
import matplotlib.pyplot as plt
import seaborn as sns

penguins = sns.load_dataset("penguins")
sns.pairplot(data=penguins,
             hue='species')

plt.savefig("test.eps")
# plt.savefig("test.eps", dpi=1200)
```

Chapter 28

survial analysis

28.1 Python package `tableone`

28.2 Python package `lifelines`

Chapter 29

Manim

29.1 VSCode extension: Manim Sideview

https://marketplace.visualstudio.com/items?itemName=Rickaym.manim-sideview

`ffmpeg.exe` placed in the same folder with `.py`

VSCode Ctrl + Shift + P: `open Mobject gallery`

29.2 installation

https://docs.manim.community/en/stable/installation.html

29.2.1 Conda

`conda install -c conda-forge manim`

29.3 quickstart

https://docs.manim.community/en/stable/tutorials/quickstart.html

https://www.w3schools.com/tags/att_video_autoplay.asp

https://www.w3schools.com/tags/att_video_loop.asp

```
from manim import *

class CreateCircle(Scene):
    def construct(self):
        circle = Circle()  # create a circle
        circle.set_fill(PINK, opacity=0.5)  #
↪   set the color and transparency
        self.play(Create(circle))  # show the
            ↪   circle on screen
```

`manim -pql scene.py CreateCircle`

Chapter 30

ggplot2

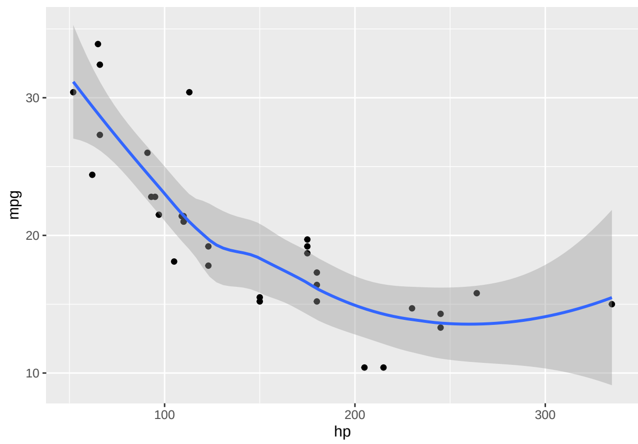
https://bookdown.org/xiangyun/msg/system.html#chap:system

Modern Statistical Graphics section 5.1

- https://www.rdocumentation.org/ to search function
- ggplot2
  - https://ggplot2.tidyverse.org/index.html to search ggplot2 function
  - panel = layer
    - geom = geometric objects / geometry = element
      - element
    - statistic
    - scale
    - coordinate system
    - facet

```r
library(ggplot2)

p <- ggplot(aes(x = hp, y = mpg), data =
↪  mtcars) +
  geom_point() # layer of scatterplot
p + geom_smooth(method = "loess") # add layer
↪  of smooth
```

```
## `geom_smooth()` using formula = 'y ~ x'
```

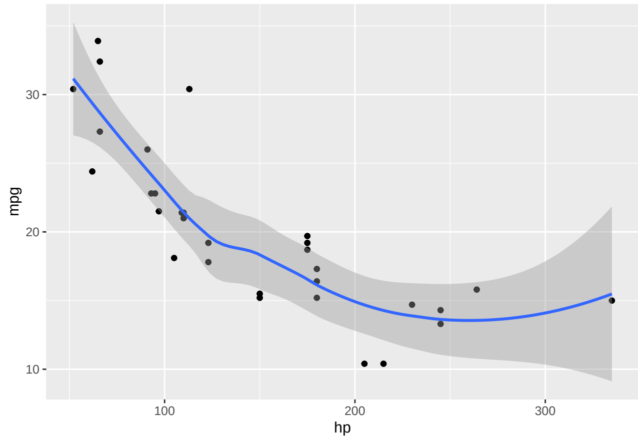


30.1 geom

https://bookdown.org/xiangyun/msg/system.html#section-13

```r
library(ggplot2)

ggplot(aes(x = hp, y = mpg), data = mtcars) +
  geom_point() +
  geom_smooth(method = "loess")
```

```
## `geom_smooth()` using formula = 'y ~ x'
```



points https://ggplot2.tidyverse.org/reference/geom_point.html?q=geom_point#null

`geom_point`

smoothed conditional means https://ggplot2.tidyverse.org/reference/geom_smooth.html?q=geom_sm#null

> Aids the eye in seeing patterns in the presence of overplotting. `geom_smooth()` and `stat_smooth()` are effectively aliases: they both use the same arguments. Use `stat_smooth()` if you want to display the results with a non-standard geom.

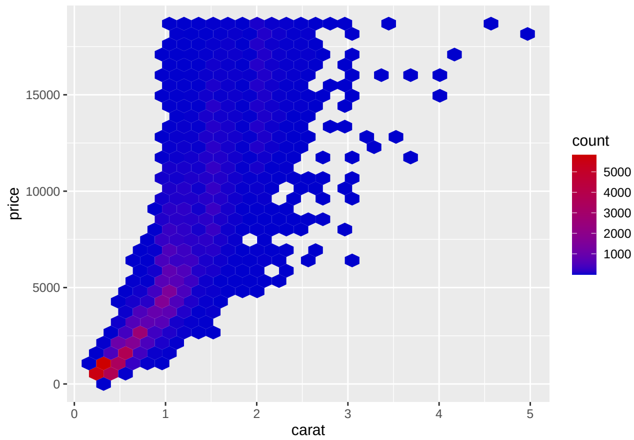
`geom_smooth`

`stat_smooth`

> `method` Smoothing method (function) to use, accepts either NULL or a character vector, e.g. `"lm"`, `"glm"`, `"gam"`, `"loess"` or a function, e.g. `MASS::rlm` or `mgcv::gam`, `stats::lm`, or `stats::loess`. `"auto"` is also accepted for backwards compatibility. It is equivalent to `NULL`.

```r
# install.packages("hexbin")
```

```r
library(ggplot2)

ggplot(aes(x = carat, y = price), data =
  diamonds) +
  geom_hex() +
  scale_fill_gradient(low = "blue3", high =
    "red3")
```

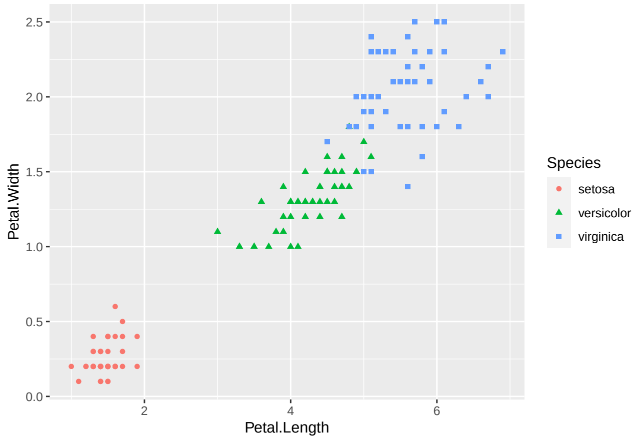


gradient color scales https://ggplot2.tidyverse.org/reference/scale_gradient.html?q=scale_fill_gradient#ref-usage

https://ggplot2.tidyverse.org/reference/scale_gradient.html?q=scale_fill_gradient#ref-examples

---

```r
library(ggplot2)

ggplot(aes(x = Petal.Length, y = Petal.Width),
↪  data = iris) +
  geom_point(aes(color = Species, shape =
  ↪  Species))
```

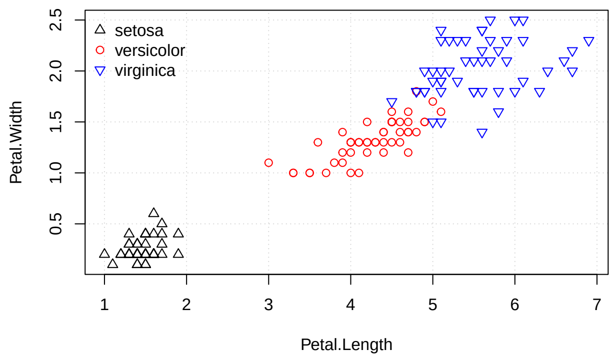


---

basic plot system

https://bookdown.org/xiangyun/msg/elements.html#sec:points

```r
# iris species converted to type integer 1, 2,
↪  3 for further using vectors
idx <- as.integer(iris[["Species"]])
plot(iris[, 3:4],
  pch = c(24, 21, 25)[idx],
  col = c("black", "red", "blue")[idx],
  ↪  panel.first = grid()
)
legend("topleft",
  legend = levels(iris[["Species"]]),
  col = c("black", "red", "blue"), pch = c(24,
  ↪  21, 25), bty = "n"
)
```



`plot` https://www.rdocumentation.org/packages/graphics/versions/3.6.2/topics/plot.default

`pch` a vector of **plotting characters** or symbols: see `points`.

`col` The **colors** for lines and points. Multiple colors can be specified so that each point can be given its own color. If there are fewer colors than points they are recycled in the standard fashion. Lines will all be plotted in the first colour specified.

`panel.first` an 'expression' to be evaluated after the plot axes are set up but before any plotting takes place. This can be useful for drawing **background grids** or **scatterplot smooths**. Note that this works by lazy evaluation: passing this argument from other plot methods may well not work since it may be evaluated too early.

`legend` https://www.rdocumentation.org/packages/graphics/versions/3.6.2/topics/legend

`bty` the **type of box** to be drawn around the legend. The allowed values are "o" (the default) and "n".

https://stackoverflow.com/questions/10108073/plot-legends-without-border-and-with-white-background

Use option `bty = "n"` in legend to remove the box around the legend.

`legend` a character or expression vector of length ≥ 1 to appear in the legend. Other objects will be coerced by `as.graphicsAnnot`

30.1.1 basic plot system decomposition

```
##    Sepal.Length Sepal.Width Petal.Length Petal.W
## 1           5.1         3.5          1.4
## 2           4.9         3.0          1.4
## 3           4.7         3.2          1.3
## 4           4.6         3.1          1.5
## 5           5.0         3.6          1.4
## 6           5.4         3.9          1.7
## 7           4.6         3.4          1.4
## 8           5.0         3.4          1.5
## 9           4.4         2.9          1.4
## 10          4.9         3.1          1.5
## 11          5.4         3.7          1.5
## 12          4.8         3.4          1.6
## 13          4.8         3.0          1.4
## 14          4.3         3.0          1.1
## 15          5.8         4.0          1.2
## 16          5.7         4.4          1.5
## 17          5.4         3.9          1.3
## 18          5.1         3.5          1.4
## 19          5.7         3.8          1.7
## 20          5.1         3.8          1.5
## 21          5.4         3.4          1.7
## 22          5.1         3.7          1.5
## 23          4.6         3.6          1.0
## 24          5.1         3.3          1.7
## 25          4.8         3.4          1.9
## 26          5.0         3.0          1.6
## 27          5.0         3.4          1.6
## 28          5.2         3.5          1.5
## 29          5.2         3.4          1.4
## 30          4.7         3.2          1.6
## 31          4.8         3.1          1.6
## 32          5.4         3.4          1.5
## 33          5.2         4.1          1.5
## 34          5.5         4.2          1.4
## 35          4.9         3.1          1.5
## 36          5.0         3.2          1.2
## 37          5.5         3.5          1.3
## 38          4.9         3.6          1.4
## 39          4.4         3.0          1.3
## 40          5.1         3.4          1.5
## 41          5.0         3.5          1.3
## 42          4.5         2.3          1.3
## 43          4.4         3.2          1.3
## 44          5.0         3.5          1.6
## 45          5.1         3.8          1.9
## 46          4.8         3.0          1.4
## 47          5.1         3.8          1.6
## 48          4.6         3.2          1.4
## 49          5.3         3.7          1.5
## 50          5.0         3.3          1.4
## 51          7.0         3.2          4.7
## 52          6.4         3.2          4.5
## 53          6.9         3.1          4.9
## 54          5.5         2.3          4.0
## 55          6.5         2.8          4.6
## 56          5.7         2.8          4.5
## 57          6.3         3.3          4.7
## 58          4.9         2.4          3.3
## 59          6.6         2.9          4.6
## 60          5.2         2.7          3.9
## 61          5.0         2.0          3.5
## 62          5.9         3.0          4.2
## 63          6.0         2.2          4.0
## 64          6.1         2.9          4.7
## 65          5.6         2.9          3.6
```

```r
iris$Species
```

```
##   [1] setosa     setosa     setosa     setosa
##   [7] setosa     setosa     setosa     setosa
##  [13] setosa     setosa     setosa     setosa
##  [19] setosa     setosa     setosa     setosa
##  [25] setosa     setosa     setosa     setosa
##  [31] setosa     setosa     setosa     setosa
##  [37] setosa     setosa     setosa     setosa
##  [43] setosa     setosa     setosa     setosa
##  [49] setosa     setosa     versicolor versicolor
##  [55] versicolor versicolor versicolor versicolor
##  [61] versicolor versicolor versicolor versicolor
##  [67] versicolor versicolor versicolor versicolor
##  [73] versicolor versicolor versicolor versicolor
##  [79] versicolor versicolor versicolor versicolor
##  [85] versicolor versicolor versicolor versicolor
##  [91] versicolor versicolor versicolor versicolor
##  [97] versicolor versicolor versicolor versicolor
## [103] virginica  virginica  virginica  virginica
## [109] virginica  virginica  virginica  virginica
## [115] virginica  virginica  virginica  virginica
## [121] virginica  virginica  virginica  virginica
## [127] virginica  virginica  virginica  virginica
## [133] virginica  virginica  virginica  virginica
## [139] virginica  virginica  virginica  virginica
## [145] virginica  virginica  virginica  virginica
## Levels: setosa versicolor virginica
```

```r
iris[["Species"]]
```

```
##   [1] setosa     setosa     setosa     setosa
##   [7] setosa     setosa     setosa     setosa
##  [13] setosa     setosa     setosa     setosa
##  [19] setosa     setosa     setosa     setosa
##  [25] setosa     setosa     setosa     setosa
##  [31] setosa     setosa     setosa     setosa
##  [37] setosa     setosa     setosa     setosa
##  [43] setosa     setosa     setosa     setosa
##  [49] setosa     setosa     versicolor versicolor
##  [55] versicolor versicolor versicolor versicolor
##  [61] versicolor versicolor versicolor versicolor
##  [67] versicolor versicolor versicolor versicolor
##  [73] versicolor versicolor versicolor versicolor
##  [79] versicolor versicolor versicolor versicolor
##  [85] versicolor versicolor versicolor versicolor
##  [91] versicolor versicolor versicolor versicolor
##  [97] versicolor versicolor versicolor versicolor
## [103] virginica  virginica  virginica  virginica
## [109] virginica  virginica  virginica  virginica
## [115] virginica  virginica  virginica  virginica
## [121] virginica  virginica  virginica  virginica
## [127] virginica  virginica  virginica  virginica
## [133] virginica  virginica  virginica  virginica
## [139] virginica  virginica  virginica  virginica
## [145] virginica  virginica  virginica  virginica
## Levels: setosa versicolor virginica
```

```r
as.integer(iris[["Species"]])
```

```
##   [1] 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
##  [38] 1 1 1 1 1 1 1 1 1 1 1 1 1 2 2 2 2 2 2 2 2 2
##  [75] 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2
## [112] 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3
## [149] 3 3
```

```r
levels(iris[["Species"]])
```

```
## [1] "setosa"     "versicolor" "virginica"
```

```r
idx <- as.integer(iris[["Species"]])
idx
```

```
##   [1] 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
##  [38] 1 1 1 1 1 1 1 1 1 1 1 1 1 2 2 2 2 2 2 2 2
##  [75] 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2
## [112] 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3
## [149] 3 3
```

```r
idx <- as.integer(iris[["Species"]])
c(24, 21, 25)[idx]
```

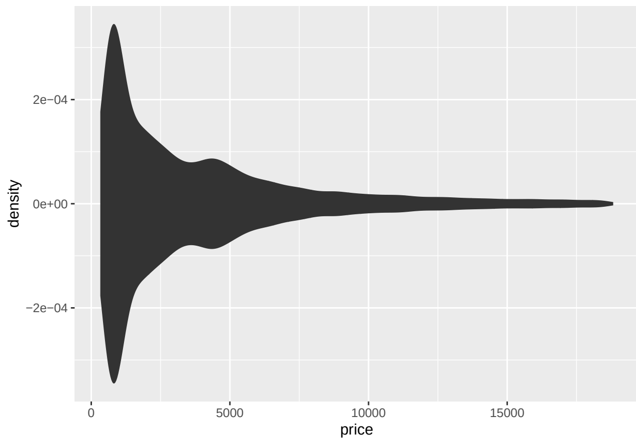
```
##   [1] 24 24 24 24 24 24 24 24 24 24 24 24 24 24 2
##  [26] 24 24 24 24 24 24 24 24 24 24 24 24 24 24 2
##  [51] 21 21 21 21 21 21 21 21 21 21 21 21 21 2
##  [76] 21 21 21 21 21 21 21 21 21 21 21 21 21 2
## [101] 25 25 25 25 25 25 25 25 25 25 25 25 25 25 2
## [126] 25 25 25 25 25 25 25 25 25 25 25 25 25 25 2
```

30.2 statistic

https://bookdown.org/xiangyun/msg/system.html#section-14

```r
library(ggplot2)

ggplot(diamonds, aes(x = price)) +
  stat_density(aes(ymax = ..density.., ymin =
  ↪  -..density..),
    geom = "ribbon", position = "identity"
  )
```

```
## Warning: The dot-dot notation (`..density..`) was
## i Please use `after_stat(density)` instead.
## This warning is displayed once every 8 hours.
## Call `lifecycle::last_lifecycle_warnings()` to se
## generated.
```



smoothed density estimates https://ggplot2.tidyverse.org/reference/geom_density.html

> Computes and draws kernel density estimate, which is a smoothed version of the histogram. This is a useful alternative to the histogram for continuous data that comes from an underlying smooth distribution.

`geom_density`

`stat_density`

https://ggplot2.tidyverse.org/reference/geom_density.html#ref-examples

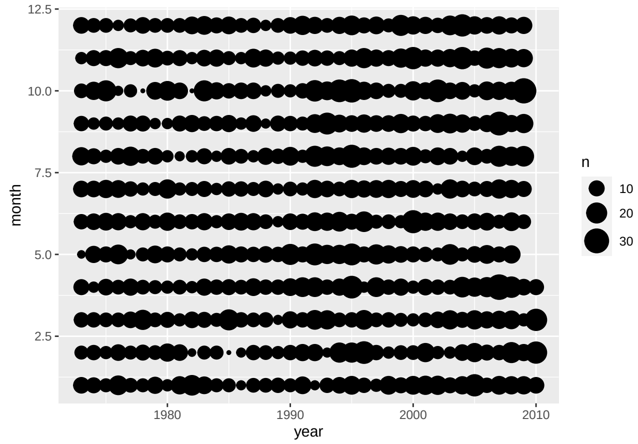
30.3 scale

ggplot2

https://bookdown.org/xiangyun/msg/system.html#section-15

```r
library(ggplot2)

data(quake6, package = "MSG")
ggplot(quake6, aes(x = year, y = month)) +
  stat_sum(aes(size = ..n..)) +
  ↪  scale_size(range = c(1, 8))
```

```
## Warning: package 'ggplot2' was built under R vers
## Warning: The dot-dot notation (`..n..`) was depre
## i Please use `after_stat(n)` instead.
## This warning is displayed once every 8 hours.
## Call `lifecycle::last_lifecycle_warnings()` to se
## generated.
```



count overlapping points https://ggplot2.tidyverse.org/reference/geom_count.html?q=stat_sum#ref-usage

> This is a variant `geom_point()` that counts the number of observations at each location, then maps the count to point area. It useful when you have discrete data and overplotting.

`geom_count`

`stat_sum`

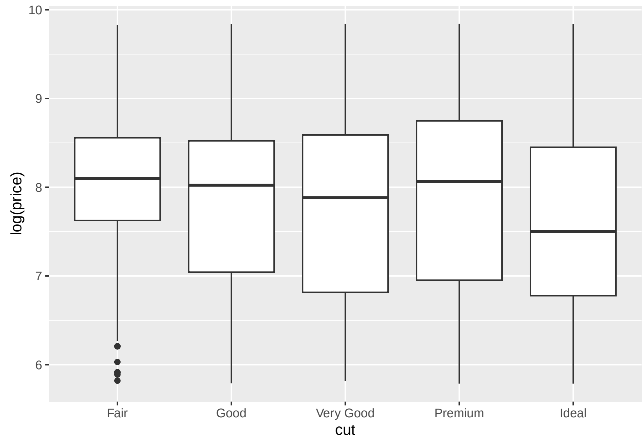
https://ggplot2.tidyverse.org/reference/geom_count.html?q=stat_sum#ref-examples

scales for area or radius https://ggplot2.tidyverse.org/reference/scale_size.html?q=scale_size#null

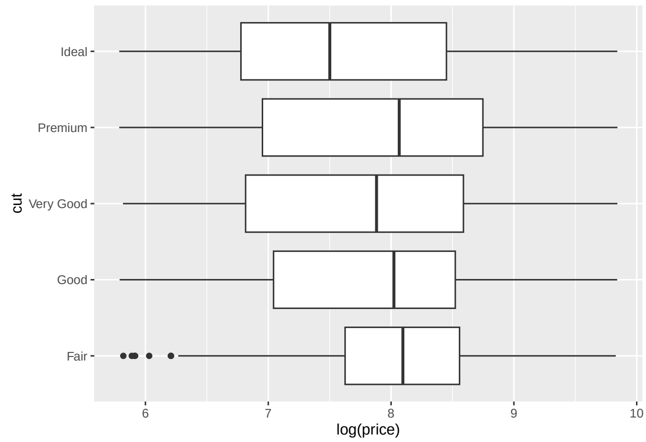
`scale_size`

30.4 coordinate system

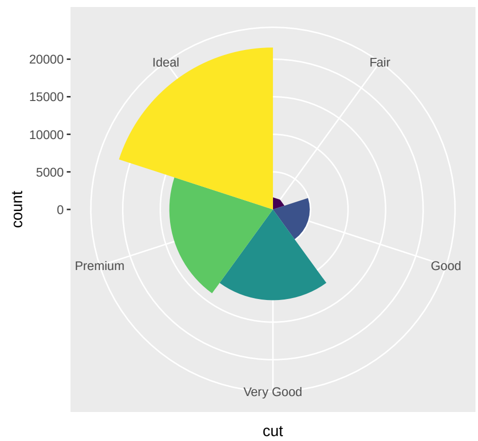
https://bookdown.org/xiangyun/msg/system.html#section-16




```
library(ggplot2)

p <- ggplot(aes(x = cut, y = log(price)), data
↪ = diamonds) +
  geom_boxplot()
p
p + coord_flip()
```



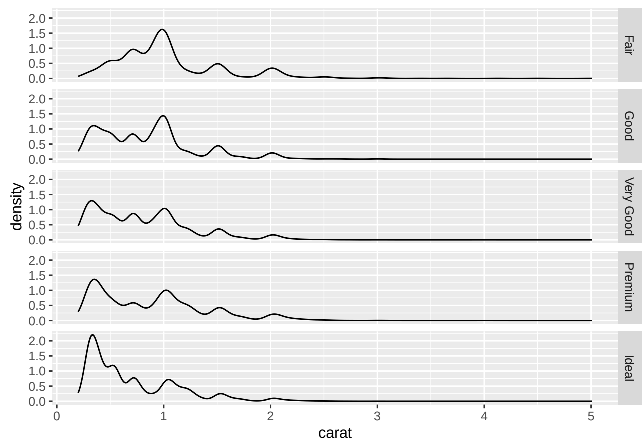

```
library(ggplot2)

ggplot(aes(x = cut, fill = cut), data =
↪ diamonds) +
  coord_polar() +
  geom_bar(width = 1, show.legend = FALSE)
```



30.5 facet

https://bookdown.org/xiangyun/msg/system.html#subsec:facet

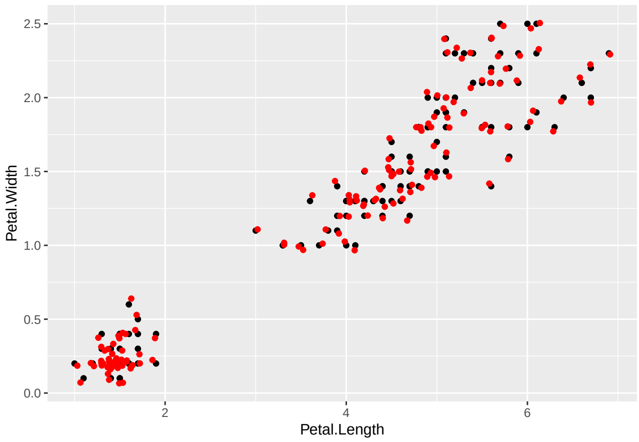
```
library(ggplot2)

ggplot(aes(x = carat), data = diamonds) +
  geom_density() +
  facet_grid(cut ~ .)
```



30.6 jitter

https://bookdown.org/xiangyun/msg/system.html#section-17

```
library(ggplot2)

ggplot(aes(x = Petal.Length, y = Petal.Width),
↪   data = iris) +
  geom_point() +
  geom_jitter(color = "red")
```



30.7 font

https://bookdown.org/xiangyun/msg/system.html#subsec:font

Chapter 31

notational system for design

[9] p.51

NSD = notational system for design

[5]

31.1 graphic notation

[9] p.51

- X: treatment or exposure to an agent or an event of interest
- P: **placebo**, i.e. blank treatment or exposure, or standard treatment, or exposure as an active control
- O: **observation** or process of measurement
- R: **randomization**, i.e. random assignment of research subjects to separate treatment or exposure groups
- subscript
  - $_g$: **groups**
  - $_k$: **kinds** of treatments, exposures, or placebos
  - $_t$: **time** or sequential order

https://tex.stackexchange.com/questions/591882/citation-within-a-latex-figure-caption-in-rmarkdown

`(ref:rudolph) *nice* cite: [@Lam94].`

`(ref:campbell1963) *nice* cite: [@campbell1963].`

`(ref:campbell1963) ([@campbell1963]`

`(ref:campbell1963) \ [@campbell1963]`

[5] [5]

31.2 pre-experimental design

[5] p.6

31.2.1 one-shot case study

$$X \quad O$$

31.2.2 one-group pretest-posttest design

$$O \quad X \quad O$$

paired t test

[9] p.62

$$\begin{array}{ccc} O & X & O \\ O_t & X & O_t \\ O_0 & X & O_1 \end{array}$$

Sources of Invalidity

	Internal								External			
	History	Maturation	Testing	Instrumentation	Regression	Selection	Mortality	Interaction of Selection and Maturation, etc.	Interaction of Testing and X	Interaction of Selection and X	Reactive Arrangements	Multiple-X Interference
Pre-Experimental Designs: 1. One-Shot Case Study X O	−	−				−	−			−		
2. One-Group Pretest-Posttest Design O X O	−	−	−	−	?	+	+	−	−	−	?	
3. Static-Group Comparison X O / O	+	?	+	+	+	−	−	−		−		
True Experimental Designs: 4. Pretest-Posttest Control Group Design R O X O / R O O	+	+	+	+	+	+	+	+	−	?	?	
5. Solomon Four-Group Design R O X O / R O O / R X O / R O	+	+	+	+	+	+	+	+	+	?	?	
6. Posttest-Only Control Group Design R X O / R O	+	+	+	+	+	+	+	+	+	?	?	

Figure 31.1: pre- and true experimental designs ([5] p.8)

$$\begin{array}{ccc} O & X & O \\ O_{gt} & X_g & O_{gt} \\ O_{10} & X & O_{11} \end{array}$$

31.2.3 static-group comparison

$$\begin{array}{cc} X & O \\ X_g & O_{gt} \\ X & O_{11} \\ & O_{21} \end{array}$$

$$\begin{array}{cc} X & O \\ X_g & O_{gt} \\ X & O_{11} \\ & O_{01} \end{array}$$

31.3 true experimental design

[5] p.13

31.3.1 posttest-only control group design

basic experimental design

two-sample t test

[9] p.53

$$\begin{array}{ccc} R & X & O \\ R & & O \end{array}$$

	Internal								External			
	History	Maturation	Testing	Instrumentation	Regression	Selection	Mortality	Interaction of Selection and Maturation, etc.	Interaction of Testing and X	Interaction of Selection and X	Reactive Arrangements	Multiple-X Interference
Quasi-Experimental Designs:												
7. Time Series O O O OXO O O O	−	+	+	?	+	+	+	+	−	?	?	
8. Equivalent Time Samples Design X_1O X_0O X_1O X_0O, etc.	+	+	+	+	+	+	+	+	−	?	−	−
9. Equivalent Materials Samples Design M_aX_1O M_bX_0O M_cX_1O M_dX_0O, etc.	+	+	+	+	+	+	+	+	−	?	?	−
10. Nonequivalent Control Group Design O X O - - - - - - - - O O	+	+	+	+	?	+	+	−	−	?	?	
11. Counterbalanced Designs X_1O X_2O X_3O X_4O X_2O X_4O X_1O X_3O X_3O X_1O X_4O X_2O X_4O X_3O X_2O X_1O	+	+	+	+	+	+	+	?	?	?	?	−
12. Separate-Sample Pretest-Posttest Design R O (X) R X O	−	−	+	?	+	+	−	−	+	+	+	
12a. R O (X) R X O - - - - - - - - R O (X) R X O	+	−	+	?	+	+	−	+	+	+	+	
12b. R O_1 (X) R O_2 (X) R X O_3	−	+	+	?	+	+	−	?	+	+	+	
12c. R O_1 X O_2 R X O_3	−	−	+	?	+	+	+	−	+	+	+	

Figure 31.2: quasi-experimental designs ([5] p.40)

	Sources of Invalidity: Internal								Sources of Invalidity: External			
	History	Maturation	Testing	Instrumentation	Regression	Selection	Mortality	Interaction of Selection and Maturation, etc.	Interaction of Testing and X	Interaction of Selection and X	Reactive Arrangements	Multiple-X Interference
Quasi-Experimental Designs Continued:												
13. Separate-Sample Pretest-Posttest Control Group Design R O (X) R X O - - - - - - - R O R O	+	+	+	+	+	+	+	−	+	+	+	
13a. R O (X) R X O - - - - - - - R': R O (X) R X O - - - - - - - R O (X) R X O R O R O - - - - - - - R': R O R O - - - - - - - R O R O	+	+	+	+	+	+	+	+	+	+	+	
14. Multiple Time-Series O O OXO O O - - - - - - - O O O O O O	+	+	+	+	+	+	+	+	−	−	?	
15. Institutional Cycle Design Class A X O_1 - - - - - - - Class B_1 RO_2 X O_3 Class B_2 R X O_4 - - - - - - - Class C O_5 X - - - - - - - [a]Gen. Pop. Con. Cl. B O_6 [a]Gen. Pop. Con. Cl. C O_7												
$O_2 < O_1$, $O_5 < O_4$	+	−	+	+	?	−	?		+	?	+	
$O_2 < O_3$	−	−	−	?	?	+	+		−	?	+	
$O_2 < O_4$	−	−	+	?	?	+	?		+	?	?	
$O_6 = O_7$, $O_{2y} = O_{2o}$		+						−				
16. Regression Discontinuity	+	+	+	?	+	+	?	+	+	−	+	+

[a] General Population Controls for Class B, etc.

Figure 31.3: quasi-experimental designs continued ([5] p.56)

or, with a placebo or an active control,

$$\begin{array}{ccc} R & X & O \\ R & P & O \end{array}$$

$$R \quad X_g \quad O_{gt}$$

$$\begin{array}{lll} R & X_g = X_1 = X & O_{gt} = O_{11} \\ R & X_g = X_2 = \emptyset & O_{gt} = O_{21} \end{array}$$

or, with a placebo or an active control

$$\begin{array}{lll} R & X_g = X_1 = X & O_{gt} = O_{11} \\ R & X_g = X_2 = P & O_{gt} = O_{21} \end{array}$$

$$\begin{array}{ccc} R & X & O_{11} \\ R & & O_{21} \end{array}$$

or, with a placebo or an active control

$$\begin{array}{ccc} R & X & O_{11} \\ R & P & O_{21} \end{array}$$

$$\begin{array}{ccc} R & X_g & O_{gt} \\ R & X & O_{11} \\ R & P & O_{21} \end{array}$$

$$\begin{array}{ccc} R & X & O \\ R & X_g & O_{gt} \\ R & X & O_{11} \\ R & P & O_{21} \end{array}$$

31.3.2 pretest-posttest control group design

$$\begin{array}{cccc} R & O & X & O \\ R & O_{gt} & X_g & O_{gt} \\ R & O_{10} & X & O_{11} \\ R & O_{20} & & O_{21} \end{array}$$

31.3.3 Solomon four-group design

[9] p.52

Solomon 4-group design = pretest-posttest + posttest-only control group design

$$\begin{array}{cccc} R & O_{gt} & X_g & O_{gt} \\ R & O_{10} & X & O_{11} \\ R & O_{20} & & O_{21} \\ R & & X & O_{31} \\ R & & & O_{41} \end{array}$$

$$\begin{array}{cccc} R & O & X & O \\ R & O_{gt} & X_g & O_{gt} \\ R & O_{10} & X & O_{11} \\ R & O_{20} & & O_{21} \\ R & & X & O_{31} \\ R & & & O_{41} \end{array}$$

31.4 quasi-experimental design

[5] p.34

31.5 correlational and ex post facto designs

[5] p.64

31.6 graphic notation, advanced

[9] p.74

- X: treatment or exposure to an agent or an event of interest
- P: **placebo**, i.e. blank treatment or exposure, or standard treatment, or exposure as an active control
- O: **observation** or process of measurement
- R: **randomization**, i.e. random assignment of research subjects to separate treatment or exposure groups
- subscript
  - $_g$: **groups**
  - $_k$: **kinds** of treatments, exposures, or placebos
  - $_t$: **time** or sequential order
- V: **variable(s)**
  - $B(V)$: **blocking** by the variable(s)
  - $M(V)$: **matching** by the variable(s)
  - $S(V)$: **stratifying** by the variable(s)
  - $L(V/L)$: **limiting** to the level(s) of the variable(s)
- M^*: research **material(s)** selected
- $-$: cohort

Chapter 32

design of experiment

experimental design = experiment design = design of experiments = DoE

[9] p.72

question-design-analysis loop

32.1 notational system for design[31]

graphic notation, advanced[31.6]

32.2 terminology

- population
  - sample
    - subsample
- unit
  - experimental unit
    - response
    - block: group of similar experimental unit ([10] p.74)
  - observational unit / measurement unit [1]
- replication ([9] p.76): an independent observation of the treatment ([10] p.74)
  - treatment replication: experimental-unit-to-experimental-unit variation
  - measurement replication = subsample: measurement-to-measurement variation
- replicate
  - experimental replicate
  - biological replicate
  - technical replicate

---

[10] p.73

Y_{ij}: the response observed from the j^{th} experimental unit assigned to the i^{th} treatment

μ_i: the mean response to the i^{th} treatment

$\mathcal{E}_{ij}$: the noise from other possible natural variation or nonrandom and random error

$$Y_{ij} = \mu_i + \mathcal{E}_{ij}, \begin{cases} i \in \mathbb{N} \cap [1, n_i] & \mathbb{N} \ni n_i \text{ treatments} \\ j \in \mathbb{N} \cap [1, n_j] & \mathbb{N} \ni n_j \text{ experimental units per treatment} \end{cases}$$

Each treatment has n_j experimental units, so there are totally $n_i n_j$ experimental units.

If experimental units cannot be homogeneous, we can try to

- stratify them
- group them, and measure group to group variation
- block them

[1] https://passel2.unl.edu/view/lesson/2e09f0055f13/6

here n_j blocks each with n_i experimental units where **each treatment occurs once in each block**

$$\begin{aligned} Y_{ij} =& \mu_i + \mathcal{E}_{ij} \\ =& \mu_i + b_j + \mathcal{E}^*_{ij}, \begin{cases} i \in \mathbb{N} \cap [1, n_i] & \mathbb{N} \ni n_i \text{ experimental units per block} \\ j \in \mathbb{N} \cap [1, n_j] & \mathbb{N} \ni n_j \text{ blocks} \end{cases} \end{aligned}$$

where

$$\mathcal{E}_{ij} = b_j + \mathcal{E}^*_{ij}$$

i.e. the variation between groups or blocks of experimental units has been identified and isolated from $\mathcal{E}^*_{ij}$, which represents the variability of experimental units within a block. By isolating the block effect from the experimental units, the within-block variation can be used to compare treatment effects, which involves computing the estimated standard errors of contrasts of the treatments.

$$\begin{aligned} Y_{ij} - Y_{i'j} &= \left(\mu_i + b_j + \mathcal{E}^*_{ij}\right) \\ &\quad - \left(\mu_{i'} + b_j + \mathcal{E}^*_{i'j}\right) \\ &= \left(\mu_i - \mu_{i'}\right) + \left(\mathcal{E}^*_{ij} - \mathcal{E}^*_{i'j}\right) \end{aligned}$$

which does not depend on the block effect b_j or free of block effects. The result of this difference is that the variance of the difference of two treatment responses within a block depends on the within-block variation among the experimental units and not the between-block variation.

32.2.1 replication vs. subsample

It is very important to distinguish between a **subsample** and a **replication** since the error variance estimated from between subsamples is in general considerably smaller than the error variance estimated from replications or between experimental units. ([10] p.77)

https://www.researchgate.net/post/What-is-Experimental-Unit-Replicate-Total-sample-size-treatment-size

32.2.2 replication vs. repeated measurements

32.2.3 replication, replicate

32.2.3.1 technical replicate, biological replicate

https://www.youtube.com/watch?v=c_cpl5YsBV8

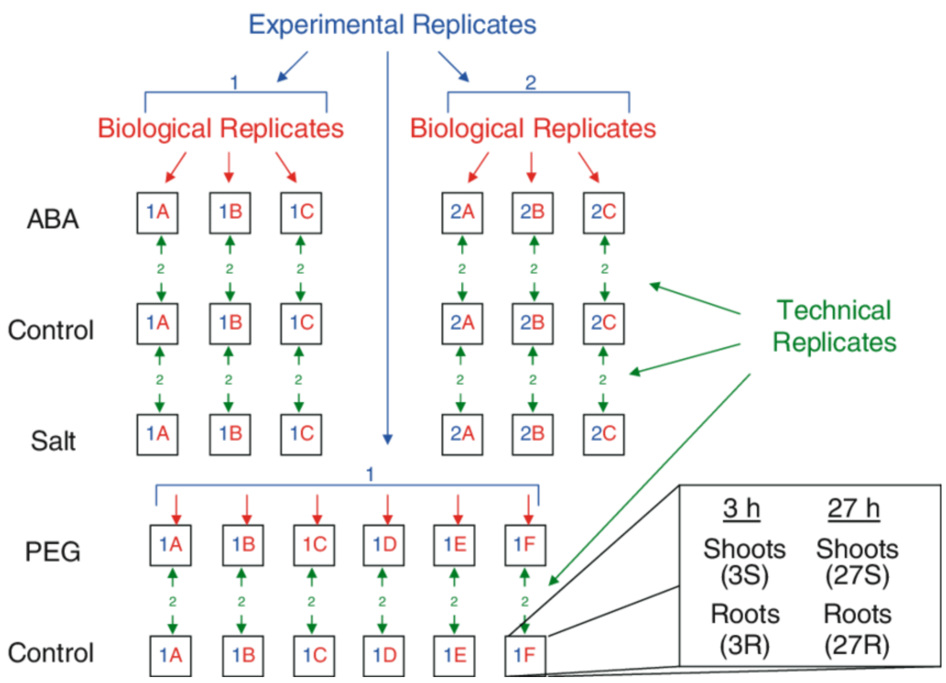


Figure 32.1: experimental, biological, technical replicates ([11])

32.2.4 Latin square design

LSD = Latin square design

[6] p.505~507

https://tex.stackexchange.com/questions/501671/how-to-get-math-mode-curly-braces-in-tikz

`\\usepackage{pgfplots}` in `engine.opts=list(extra.preamble=c("\\usepackage{pgfplots}"))`

`\usetikzlibrary{decorations}`

$$\underbrace{p = 4 \text{ rows}}\left\{\begin{array}{|c|c|c|c|} \multicolumn{4}{c}{\overbrace{\qquad\qquad\qquad\qquad}^{p=4\text{ columns}}} \\ \hline A & B & C & D \\ \hline B & C & D & A \\ \hline C & D & A & B \\ \hline D & A & B & C \\ \hline \end{array}\right.$$

$p = |\{A, B, C, D\}| = 4$ treatments

Figure 32.2: Latin square example

$$Y_{ijk} = \mu + \alpha_i + \beta_j + \gamma_k + \mathcal{E}_{ijk}, \begin{cases} i \in \mathbb{N} \cap [1, p] & \mathbb{N} \ni p \text{ treatments} \\ j \in \mathbb{N} \cap [1, p] & \mathbb{N} \ni p \text{ rows} \\ k \in \mathbb{N} \cap [1, p] & \mathbb{N} \ni p \text{ columns} \end{cases}$$

$$\mathcal{E}_{ijk} \overset{\text{i.i.d.}}{\sim} \mathrm{n}\left(0, \sigma^2\right)$$

ρ_i: i^{th} row

κ_j: j^{th} column

τ_k: k^{th} treatment

$$\begin{aligned} Y_{ijk} =& \mu + \alpha_i + \beta_j + \gamma_k + \mathcal{E}_{ijk}, \begin{cases} i \in \mathbb{N} \cap [1, p] & \mathbb{N} \ni p \text{ treatments} \\ j \in \mathbb{N} \cap [1, p] & \mathbb{N} \ni p \text{ rows} \\ k \in \mathbb{N} \cap [1, p] & \mathbb{N} \ni p \text{ columns} \end{cases} \\ =& \mu + \rho_i + \kappa_j + \tau_k + \mathcal{E}_{ijk}, \begin{cases} i \in \mathbb{N} \cap [1, p] & \mathbb{N} \ni p \text{ rows} \\ j \in \mathbb{N} \cap [1, p] & \mathbb{N} \ni p \text{ columns} \\ k \in \mathbb{N} \cap [1, p] & \mathbb{N} \ni p \text{ treatments} \end{cases} \end{aligned}$$

32.2.5 model assumption and experimental unit, measurement/observational unit

---

$$Y_{ijk} = \mu + \tau_i + \beta_j + \epsilon_{ij} + \mathit{\Delta}_{ijk}$$

32.3 experiment structure

32.3.1 treatment structure

[10] p.77

- 1-way treatment structure
- 2-way treatment structure
- factorial arrangement treatment structure

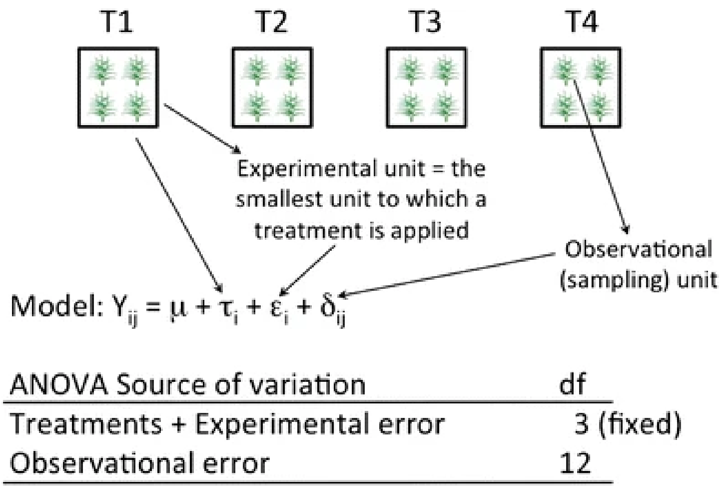


Figure 32.3: model assumption and experimental unit 1 ([12] fig.1)

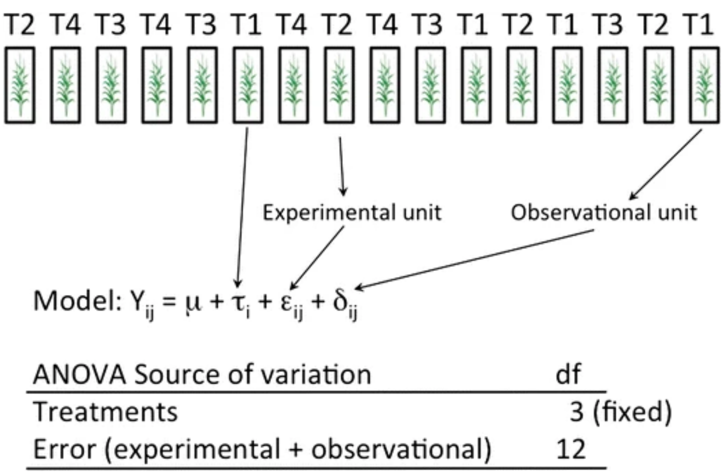


Figure 32.4: model assumption and experimental unit 2 ([12] fig.2)

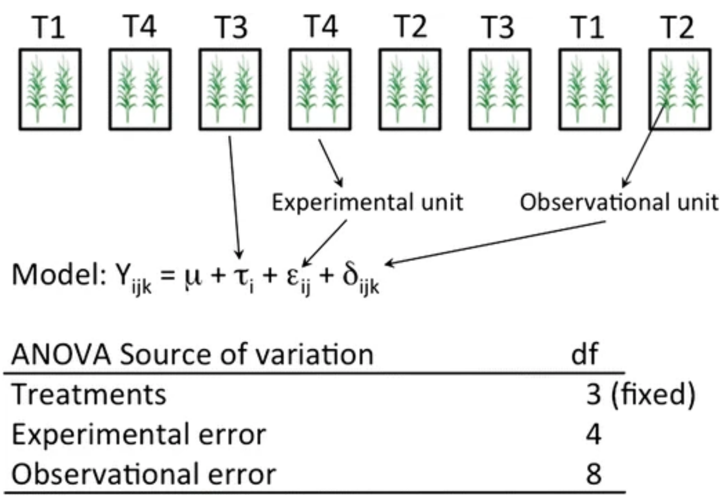


Figure 32.5: model assumption and experimental unit 3 ([12] fig.3)

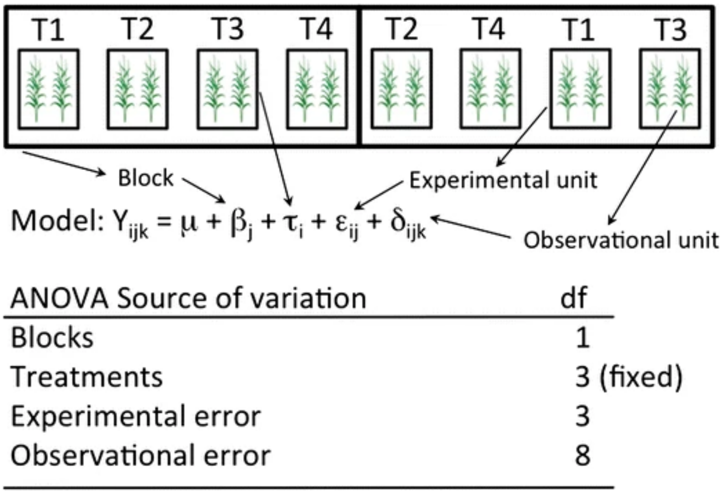


Figure 32.6: model assumption and experimental unit 4 ([12] fig.4)

- *fractional* factorial arrangement treatment structure
- factorial arrangement with one or more controls

32.3.2 design structure

[10] p.77

- CRD = completely randomized design
- RCBD = randomized complete block design
  - ? why not called CRBD = completely randomized block design
- LSD = Latin square design[32.2.4]
- IBD = incomplete block design
  - BIBD = balanced IBD
- various combinations and generalizations

32.3.3 size of experimental unit

- split-plot design
  - split-split-plot design
  - split-split-split-plot design
- repeated measures design
  - cross-over design
  - change-over design
- nested design = hierarchical design
- variations and combinations
  - SSEU = several sizes of experimental units

32.4 approach to experimentation

[9] p.75

- approach to experimentation
  - best-guess approach
  - one-factor-at-a-time approach = OFAT
  - factorial approach

32.5 sample size estimation

32.6 statistical analysis plan

32.7 protocol

[9] p.95

- study objective
- study endpoint
  - primary endpoint
  - secondary endpoint(s)
- experimental unit(s)
- treatment structure[32.3.1]
- design structure[32.3.2]
- potential confounder(s)
- randomization
- blinding
- chance reduction
- sample size estimation[32.5]
- data collection
- data management system
- statistical analysis plan[32.6]
- DSMB / DSMC = data and safety monitoring board / committee

32.8 DoE course with six sigma and Mintab

https://zhuanlan.zhihu.com/p/265914617

https://www.zhihu.com/question/416312693/answer/1426399810

32.8.1 evolution

- Fisher
- Rao

Chapter 33

quine

```
s = 's = %r\nprint(s%%s)'
print(s%s)
```

```
## s = 's = %r\nprint(s%%s)'
## print(s%s)
```

This snippet is a clever example of a quine. A quine is a computer program that takes no input and produces a copy of its own source code as its output. The given code in Python is written to print its own source when executed. Let's break it down:

`s = 's = %r\nprint(s%%s)'`: This line defines a string `s` that contains a format string. `%r` is a placeholder that gets replaced with the `repr()` of the argument provided to the `%` operator, which in this case will be the string s itself. This means it will insert the string representation of `s` into the format string at `%r`.

`print(s%s)`: This line prints the result of `s%s`. Here, the `%` operator is used to format the string `s` with itself. The `%s` inside the print statement is replaced by the string `s`, leading to the entire string being printed out, including the print statement itself.

This is because the format operation replaces %r with the representation of the string `s`, and `%%` is a way to escape the `%` sign in format strings, resulting in a single `%` in the output. This output is exactly the same as the source code, making it a quine.

33.1 %r

The `%r` in Python string formatting represents the "representation" of a value, which is typically the way you would see it if you were to type it into a Python interpreter. It uses the `repr()` function to convert the value to a string. This is useful for debugging, among other things, because it shows strings with quotes around them and escapes special characters. Essentially, `%r` gives you the "developer's view" of what a variable looks like.

Here's a simple example to illustrate %r versus %s in string formatting:

```
my_str = "Hello, World!\nNew line character is represented with \\n"
print("Using %%s: %s" % my_str)
```

```
## Using %s: Hello, World!
## New line character is represented with \n
```

```
print("Using %%r: %r" % my_str)
```

```
## Using %r: 'Hello, World!\nNew line character is represented with \\n'
```

In this example:

The `%s` specifier tells Python to convert the object using `str()`, which is designed to be readable and outputs the string `"Hello, World!\nNew line character is represented with \n"`, interpreting the escape character `\n` as a newline.

The `%r` specifier tells Python to convert the object using `repr()`, which aims to generate output that could be used to recreate the object, outputting the string `'Hello, World!\nNew line character is represented with \\n'`, preserving the actual escape characters in the output.

Notice how `%r` preserves the string exactly as it is, including the quotes and escaped characters, making it clear it's a string and showing the escape sequence explicitly.

Chapter 34

quaternion

34.1 TaylorCatAlice

https://en.wikipedia.org/wiki/Blackboard_bold

34.1.1 complex / dionion / bionion

$$
\begin{aligned}
& c = a + b\mathrm{i} = a + \mathrm{i}b, \begin{cases} c \in \mathbb{C} \\ \mathrm{i}^2 = -1 \\ a, b \in \mathbb{R} \quad \Leftrightarrow \langle a, b \rangle \in \mathbb{R}^2 \end{cases} \\
= & z = x + y\mathrm{i} = x + \mathrm{i}y, \begin{cases} z \in \mathbb{C} \\ \mathrm{i}^2 = -1 \\ x, y \in \mathbb{R} \quad \Leftrightarrow \langle x, y \rangle \in \mathbb{R}^2 \end{cases} \\
= & \sqrt{x^2 + y^2} \left(\frac{x}{\sqrt{x^2 + y^2}} + \frac{y}{\sqrt{x^2 + y^2}} \mathrm{i} \right) = r (\cos\theta + \mathrm{i} \sin\theta) = r\mathrm{e}^{\mathrm{i}\theta}
\end{aligned}
$$

Also see complex group representation[37.4].

34.1.2 trionion / triernion / triplex / ternion

https://zh.wikipedia.org/zh-tw/%E4%B8%89%E5%85%83%E6%95%B8

https://math.stackexchange.com/questions/1784166/why-are-there-no-triernions-3-dimensional-analogue-of-complex-numbers-quate

https://math.stackexchange.com/questions/32100/is-there-a-third-dimension-of-numbers/4453131

$$
\begin{aligned}
& t = a + b\mathrm{i} + c\mathrm{j} = a + \mathrm{i}b + \mathrm{j}c, \begin{cases} t \in \mathbb{T} \\ \mathrm{i}^2 = -1 \\ \mathrm{j}^2 = -1 \end{cases} \\
= & w = x + y\mathrm{i} + z\mathrm{j} = x + \mathrm{i}y + \mathrm{j}z, \begin{cases} w \in \mathbb{T} \\ \mathrm{i}^2 = -1 \\ \mathrm{j}^2 = -1 \end{cases} \\
= & \sqrt{x^2 + y^2 + z^2} \left(\frac{x}{\sqrt{x^2 + y^2 + z^2}} + \frac{y}{\sqrt{x^2 + y^2 + z^2}} \mathrm{i} + \frac{z}{\sqrt{x^2 + y^2 + z^2}} \mathrm{j} \right) = ?
\end{aligned}
$$

$$
\begin{cases} A(BC) = (AB)C & (a)\,\text{associativity} \\ A(B + C) = AB + AC & (d)\,\text{distributivity} \end{cases}
$$

$$
\begin{aligned}
\mathbb{T} \ni \mathrm{ij} =& X + Y\mathrm{i} + Z\mathrm{j} \in \mathbb{T} \\
-\mathrm{j} = \left(\mathrm{i}^2\right)\mathrm{j} \overset{(a)}{=} \mathrm{i}\,(\mathrm{ij}) =& \mathrm{i}\,(X + Y\mathrm{i} + Z\mathrm{j}) \overset{(d)}{=} -Y + X\mathrm{i} + Z\mathrm{ij} \\
\mathrm{ij} =& \frac{Y}{Z} - \frac{X}{Z}\mathrm{i} - \frac{1}{Z}\mathrm{j} \Rightarrow \begin{cases} X = \dfrac{Y}{Z} \\ Y = -\dfrac{X}{Z} \\ Z = -\dfrac{1}{Z} & \Rightarrow Z^2 = -1 \Rightarrow Z \notin \mathbb{R} \end{cases} \\
-\mathrm{i} = \mathrm{i}\left(\mathrm{j}^2\right) \overset{(a)}{=} (\mathrm{ij})\,\mathrm{j} =& (X + Y\mathrm{i} + Z\mathrm{j})\,\mathrm{j} \overset{(d)}{=} -Z + X\mathrm{j} + Y\mathrm{ij} \\
\mathrm{ij} =& \frac{Z}{Y} - \frac{1}{Y}\mathrm{i} - \frac{X}{Y}\mathrm{j} \Rightarrow \begin{cases} X = \dfrac{Z}{Y} \\ Y = -\dfrac{1}{Y} & \Rightarrow Y^2 = -1 \Rightarrow Y \notin \mathbb{R} \\ Z = -\dfrac{X}{Y} \end{cases}
\end{aligned}
$$

34.1.3 quaternion

https://en.wikipedia.org/wiki/Quaternion

$$
\begin{aligned}
q = a + b\mathrm{i} + c\mathrm{j} + d\mathrm{k} = a + \mathrm{i}b + \mathrm{j}c + \mathrm{k}d,& \begin{cases} q \in \mathbb{H} \\ a, b, c, d \in \mathbb{R} & \Leftrightarrow \langle a, b, c, d\rangle \in \mathbb{R}^4 \end{cases} \\
= w = t + x\mathrm{i} + y\mathrm{j} + z\mathrm{k} = t + \mathrm{i}x + \mathrm{j}y + \mathrm{k}z,& \begin{cases} w \in \mathbb{H} \\ t, x, y, z \in \mathbb{R} & \Leftrightarrow \langle t, x, y, z\rangle \in \mathbb{R}^4 \end{cases} \\
=?&
\end{aligned}
$$

$$
\begin{aligned}
q = a + b\mathrm{i} + c\mathrm{j} + d\mathrm{k} = a + \mathrm{i}b + \mathrm{j}c + \mathrm{k}d,& \begin{cases} q \in \mathbb{H} \\ \mathrm{i}^2 = \mathrm{j}^2 = \mathrm{k}^2 = -1 = \mathrm{ijk} & \Rightarrow \mathrm{ij} = \mathrm{k} \\ a, b, c, d \in \mathbb{R} & \Leftrightarrow \langle a, b, c, d\rangle \in \mathbb{R}^4 \end{cases} \\
= w = t + x\mathrm{i} + y\mathrm{j} + z\mathrm{k} = t + \mathrm{i}x + \mathrm{j}y + \mathrm{k}z,& \begin{cases} w \in \mathbb{H} \\ \mathrm{i}^2 = \mathrm{j}^2 = \mathrm{k}^2 = -1 = \mathrm{ijk} & \Rightarrow \mathrm{ij} = \mathrm{k} \\ t, x, y, z \in \mathbb{R} & \Leftrightarrow \langle t, x, y, z\rangle \in \mathbb{R}^4 \end{cases} \\
= t + \begin{pmatrix} \mathrm{i} & \mathrm{j} & \mathrm{k} \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = x_0 + \begin{pmatrix} \boldsymbol{i} & \boldsymbol{j} & \boldsymbol{k} \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} &= x_0 + \begin{pmatrix} \boldsymbol{e}_1 & \boldsymbol{e}_2 & \boldsymbol{e}_3 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = x_0 + \boldsymbol{x}, \begin{cases} \boldsymbol{e}_1 = \boldsymbol{i} = \mathrm{i} \\ \boldsymbol{e}_2 = \boldsymbol{j} = \mathrm{j} \\ \boldsymbol{e}_3 = \boldsymbol{k} = \mathrm{k} \end{cases} \\
= t + \frac{\mathrm{i}x + \mathrm{j}y + \mathrm{k}z}{r} r,& \begin{cases} r^2 = x^2 + y^2 + z^2 & |q|^2 = t^2 + r^2 \\ \|\boldsymbol{n}\|^2 = \left(\dfrac{\mathrm{i}x + \mathrm{j}y + \mathrm{k}z}{r}\right)^2 = -1 \end{cases} \\
= \sqrt{t^2 + r^2}\left(\frac{t}{\sqrt{t^2 + r^2}} + \boldsymbol{n}\frac{r}{\sqrt{t^2 + r^2}}\right) &= |q|\left(\cos\frac{\theta}{2} + \boldsymbol{n}\sin\frac{\theta}{2}\right) = |q|\,\mathrm{e}^{\boldsymbol{n}\frac{\theta}{2}}
\end{aligned}
$$

The quaternion set is denoted $\mathbb{H}$ for Sir R.W. **Hamilton**, because he suddenly and strikingly realized

$$
\begin{cases} \mathrm{ij} = \mathrm{k} \\ \mathrm{k} \in \mathbb{H} \end{cases} \Rightarrow \mathrm{ij} \in \mathbb{H} \text{ for closure property}
$$

---

for the sake of rigorosity, see group theory[37]

$$\mathrm{k}^2 = -1$$

$$
\begin{aligned}
\mathrm{ij} &= \mathrm{k} \\
\mathrm{ijk} = \mathrm{i}\,(\mathrm{jk}) &\overset{(a)}{=} (\mathrm{ij})\,\mathrm{k} = \mathrm{kk} = \mathrm{k}^2 = -1 \\
\mathrm{kij} = (\mathrm{ki})\,\mathrm{j} &\overset{(a)}{=} \mathrm{k}\,(\mathrm{ij}) = \mathrm{kk} = \mathrm{k}^2 = -1
\end{aligned}
$$

$$\begin{aligned}\mathrm{ij} &= \mathrm{k}\\ -\mathrm{j} = \left(\mathrm{i}^2\right)\mathrm{j} &\overset{(a)}{=} \mathrm{i}\,(\mathrm{ij}) = \mathrm{ik}\\ -\mathrm{i} = \mathrm{i}\left(\mathrm{j}^2\right) &\overset{(a)}{=} (\mathrm{ij})\,\mathrm{j} = \mathrm{kj}\end{aligned}$$

$$\begin{aligned}-\mathrm{j} &= \left(\mathrm{i}^2\right)\mathrm{j} \overset{(a)}{=} \mathrm{i}\,(\mathrm{ij}) = \mathrm{ik}\\ 1 &= -\mathrm{j}^2 = \mathrm{j}\,(-\mathrm{j}) = \mathrm{j}\,(\mathrm{ik}) \overset{(a)}{=} (\mathrm{ji})\,\mathrm{k}\\ \mathrm{k} &= [1]\,\mathrm{k} = [(\mathrm{ji})\,\mathrm{k}]\,\mathrm{k} \overset{(a)}{=} (\mathrm{ji})\left(\mathrm{k}^2\right) = (\mathrm{ji})\,(-1)\\ -\mathrm{k} &= \mathrm{ji}\\ -\mathrm{i} &= \mathrm{i}\left(\mathrm{j}^2\right) \overset{(a)}{=} (\mathrm{ij})\,\mathrm{j} = \mathrm{kj}\\ 1 &= (-\mathrm{i})\,\mathrm{i} = (\mathrm{kj})\,\mathrm{i} \overset{(a)}{=} \mathrm{k}\,(\mathrm{ji}) = \mathrm{kji}\\ 1 &= \mathrm{kji}\end{aligned}$$

There is no more commutativity[34.1.3.2.1], i.e.

$$AB \not\equiv BA$$

but

$$AB + BA = 0 \Leftrightarrow AB = -BA$$

satisfying anticommutativity[34.1.3.2.2].

$$\begin{cases}\mathrm{ij} = \mathrm{k}\\ \mathrm{ji} = -\mathrm{k}\end{cases} \begin{aligned}&\Leftrightarrow \mathrm{ji} = -\mathrm{k} = -\mathrm{ij}\\ &\Rightarrow \mathrm{ji} = -\mathrm{ij}\\ &\Leftrightarrow \mathrm{ij} + \mathrm{ji} = 0\end{aligned}$$

$$\begin{cases}\mathrm{kij} = -1\\ \mathrm{kji} = 1\end{cases} \begin{aligned}&\Leftrightarrow \mathrm{kij} = -1 = -\mathrm{kji}\\ &\Rightarrow \mathrm{kij} = -\mathrm{kji}\\ &\Leftrightarrow \mathrm{kij} + \mathrm{kji} = 0\end{aligned}$$

$$\begin{aligned}q = a + b\mathrm{i} + c\mathrm{j} + d\mathrm{k} = a + \mathrm{i}b + \mathrm{j}c + \mathrm{k}d, &\begin{cases} q \in \mathbb{H} \\ \mathrm{i}^2 = \mathrm{j}^2 = \mathrm{k}^2 = -1 = \mathrm{ijk} & \Rightarrow \mathrm{ij} = \mathrm{k} \\ a, b, c, d \in \mathbb{R} & \Leftrightarrow \langle a, b, c, d\rangle \in \mathbb{R}^4 \end{cases}\\ = w = t + x\mathrm{i} + y\mathrm{j} + z\mathrm{k} = t + \mathrm{i}x + \mathrm{j}y + \mathrm{k}z, &\begin{cases} w \in \mathbb{H} \\ \mathrm{i}^2 = \mathrm{j}^2 = \mathrm{k}^2 = -1 = \mathrm{ijk} & \Rightarrow \mathrm{ij} = \mathrm{k} \\ t, x, y, z \in \mathbb{R} & \Leftrightarrow \langle t, x, y, z\rangle \in \mathbb{R}^4 \end{cases}\\ = t + \begin{pmatrix}\mathrm{i} & \mathrm{j} & \mathrm{k}\end{pmatrix}\begin{pmatrix}x\\y\\z\end{pmatrix} = x_0 + \begin{pmatrix}\boldsymbol{i} & \boldsymbol{j} & \boldsymbol{k}\end{pmatrix}\begin{pmatrix}x\\y\\z\end{pmatrix} &= x_0 + \begin{pmatrix}\boldsymbol{e}_1 & \boldsymbol{e}_2 & \boldsymbol{e}_3\end{pmatrix}\begin{pmatrix}x_1\\x_2\\x_3\end{pmatrix} = x_0 + \boldsymbol{x}, \begin{cases}\boldsymbol{e}_1 = \boldsymbol{i} = \mathrm{i}\\ \boldsymbol{e}_2 = \boldsymbol{j} = \mathrm{j}\\ \boldsymbol{e}_3 = \boldsymbol{k} = \mathrm{k}\end{cases}\end{aligned}$$

·	1	i	j	k
1	1	i	j	k
i	i	−1	k	−j
j	j	−k	−1	i
k	k	j	−i	−1

$\cdot$	1	i	j	k	-1	$-$i	$-$j	$-$k
1	1	i	j	k				
i	i	-1	k	$-$j				
j	j	$-$k	-1	i				
k	k	j	$-$i	-1				
-1								
$-$i								
$-$j								
$-$k								

Figure 34.1: quaternion basis group table

$\cdot$	1	-1	i	$-$i	j	$-$j	k	$-$k
1								
-1								
i								
$-$i								
j								
$-$j								
k								
$-$k								

Figure 34.2: quaternion basis group table 2

34.1.3.1 true origin of (dot product & cross product) / (inner product & outer product)

product of two pure imaginary quaternions

$$\begin{aligned}
\boldsymbol{x}_1\boldsymbol{x}_2 =& \left(x_{11}\mathrm{i} + x_{12}\mathrm{j} + x_{13}\mathrm{k}\right)\left(x_{21}\mathrm{i} + x_{22}\mathrm{j} + x_{23}\mathrm{k}\right) \\
=& x_{11}x_{21}\mathrm{i}^2 + x_{11}x_{22}\mathrm{ij} + x_{11}x_{23}\mathrm{ik} \\
&+ x_{12}x_{21}\mathrm{ji} + x_{12}x_{22}\mathrm{j}^2 + x_{12}x_{23}\mathrm{jk} \\
&+ x_{13}x_{21}\mathrm{ki} + x_{13}x_{22}\mathrm{kj} + x_{13}x_{23}\mathrm{k}^2 \\
=& -\left(x_{11}x_{21} + x_{12}x_{22} + x_{13}x_{23}\right) \\
&+ \left(x_{12}x_{23} - x_{13}x_{22}\right)\mathrm{jk} + \left(x_{13}x_{21} - x_{11}x_{23}\right)\mathrm{ki} + \left(x_{11}x_{22} - x_{12}x_{21}\right)\mathrm{ij} \\
=& -\left(x_{11}x_{21} + x_{12}x_{22} + x_{13}x_{23}\right) \\
&+ \left(x_{12}x_{23} - x_{13}x_{22}\right)\mathrm{i} + \left(x_{13}x_{21} - x_{11}x_{23}\right)\mathrm{j} + \left(x_{11}x_{22} - x_{12}x_{21}\right)\mathrm{k} \\
=& -\left(\boldsymbol{x}_1 \cdot \boldsymbol{x}_2\right) + \left(\boldsymbol{x}_1 \times \boldsymbol{x}_2\right), \begin{cases} \boldsymbol{x}_1 \cdot \boldsymbol{x}_2 = x_{11}x_{21} + x_{12}x_{22} + x_{13}x_{23} \\ \boldsymbol{x}_1 \times \boldsymbol{x}_2 = \begin{vmatrix} \mathrm{i} & \mathrm{j} & \mathrm{k} \\ x_{11} & x_{12} & x_{13} \\ x_{21} & x_{22} & x_{23} \end{vmatrix} \end{cases} \\
=& -\boldsymbol{x}_1 \cdot \boldsymbol{x}_2 + \boldsymbol{x}_1 \times \boldsymbol{x}_2
\end{aligned}$$

product of two general quaternions / ordinary quaternions = Grassmann product

$$\begin{aligned}
q_1q_2 &= (q_{10}+q_{11}\mathrm{i}+q_{12}\mathrm{j}+q_{13}\mathrm{k})\,(q_{20}+q_{21}\mathrm{i}+q_{22}\mathrm{j}+q_{23}\mathrm{k})\\
&= (x_{10}+\boldsymbol{x}_1)\,(x_{20}+\boldsymbol{x}_2)\,, \begin{cases} x_{i\mu}=q_{i\mu} & \mu\in\{0\}\cup(\mathbb{N}\cap[1,3]) \\ \boldsymbol{x}_i = x_{ij}\boldsymbol{e}_j & i,j\in\mathbb{N}\cap[1,3]\,, \begin{cases}\boldsymbol{e}_1=\mathrm{i}\\ \boldsymbol{e}_2=\mathrm{j}\\ \boldsymbol{e}_3=\mathrm{k}\end{cases}\end{cases}\\
&= x_{10}x_{20}+x_{10}\boldsymbol{x}_2+x_{20}\boldsymbol{x}_1+\boldsymbol{x}_1\boldsymbol{x}_2\\
&= x_{10}x_{20}+x_{10}\boldsymbol{x}_2+x_{20}\boldsymbol{x}_1-\boldsymbol{x}_1\cdot\boldsymbol{x}_2+\boldsymbol{x}_1\times\boldsymbol{x}_2\\
&= (x_{10}x_{20}-\boldsymbol{x}_1\cdot\boldsymbol{x}_2)+(x_{10}\boldsymbol{x}_2+x_{20}\boldsymbol{x}_1+\boldsymbol{x}_1\times\boldsymbol{x}_2)
\end{aligned}$$

$$\begin{aligned}
x_{10}x_{20}-\boldsymbol{x}_1\cdot\boldsymbol{x}_2 &= \begin{pmatrix} q_{10} & q_{11} & q_{12} & q_{13}\end{pmatrix}\begin{pmatrix}1&&&\\&-1&&\\&&-1&\\&&&-1\end{pmatrix}\begin{pmatrix}q_{20}\\q_{21}\\q_{22}\\q_{23}\end{pmatrix} = q_1^{\mu}\eta_{\mu\nu}q_2^{\nu}\\
&= \begin{pmatrix} q_{10} & q_{11} & q_{12} & q_{13}\end{pmatrix}\begin{pmatrix}1&0&0&0\\0&-1&0&0\\0&0&-1&0\\0&0&0&-1\end{pmatrix}\begin{pmatrix}q_{20}\\q_{21}\\q_{22}\\q_{23}\end{pmatrix} = \boldsymbol{q}_1^{\intercal}H\boldsymbol{q}_2, H=[\eta_{\mu\nu}]_{4\times4}=\eta_{\mu\nu}
\end{aligned}$$

$$\begin{aligned}
q_1q_2 &= (x_{10}x_{20}-\boldsymbol{x}_1\cdot\boldsymbol{x}_2)+(x_{10}\boldsymbol{x}_2+x_{20}\boldsymbol{x}_1+\boldsymbol{x}_1\times\boldsymbol{x}_2)\\
QP &= (Q_0P_0-\boldsymbol{Q}\cdot\boldsymbol{P})+(Q_0\boldsymbol{P}+P_0\boldsymbol{Q}+\boldsymbol{Q}\times\boldsymbol{P})
\end{aligned}$$

$$ab=(a_0b_0-\boldsymbol{a}\cdot\boldsymbol{b})+(a_0\boldsymbol{b}+b_0\boldsymbol{a}+\boldsymbol{a}\times\boldsymbol{b})$$

Minkowski metric tensor

$$\eta=H=[\eta_{\mu\nu}]_{4\times4}=\begin{pmatrix}1&0&0&0\\0&-1&0&0\\0&0&-1&0\\0&0&0&-1\end{pmatrix}=\begin{pmatrix}1&&&\\&-1&&\\&&-1&\\&&&-1\end{pmatrix}=\eta_{\mu\nu}$$

and quaternions as 4-vectors or four-vectors

$$\boldsymbol{q}_1^{\intercal}=\begin{pmatrix} q_{10} & q_{11} & q_{12} & q_{13}\end{pmatrix}$$

$$\boldsymbol{q}_2=\begin{pmatrix}q_{20}\\q_{21}\\q_{22}\\q_{23}\end{pmatrix}$$

34.1.3.2 commutativity vs. anticommutativity

34.1.3.2.1 commutativity 交換律 = 交換性 = 對易性

$$AB=BA\Leftrightarrow AB-BA=0$$

$$AB=BA\Rightarrow AB\equiv BA$$

34.1.3.2.2 anticommutativity 反交換律 = 反交換性 = 反對易性

$$AB+BA=0\Leftrightarrow AB=-BA$$

$$AB=-BA\Rightarrow AB\not\equiv BA$$

34.1.3.3 bracket

34.1.3.3.1 self-invented bracket 自創括號 = 自創括

34.1.3.3.1.1 commutative bracket 交換括號 = 對易式

$$[X,Y]=\frac{XY-YX}{2}$$

34.1.3.3.1.2 anticommutative bracket 反交換括號 = 反對易式

$$\{X,Y\}=\frac{XY+YX}{2}$$

34.1.3.3.2 Poisson bracket https://en.wikipedia.org/wiki/Poisson_bracket

34.1.3.3.3 Lagrange bracket

34.1.3.3.4 Lie bracket

34.1.3.4 triple product

product = double product = Grassmann product

$$ab=(a_0b_0-\boldsymbol{a}\cdot\boldsymbol{b})+(a_0\boldsymbol{b}+b_0\boldsymbol{a}+\boldsymbol{a}\times\boldsymbol{b})$$

pure imaginary

$$\begin{aligned} ab&=(a_0b_0-\boldsymbol{a}\cdot\boldsymbol{b})+(a_0\boldsymbol{b}+b_0\boldsymbol{a}+\boldsymbol{a}\times\boldsymbol{b})\\ \begin{cases}a_0=0\\ b_0=0\end{cases}&=(00-\boldsymbol{a}\cdot\boldsymbol{b})+(0\boldsymbol{b}+0\boldsymbol{a}+\boldsymbol{a}\times\boldsymbol{b})\\ &=-\boldsymbol{a}\cdot\boldsymbol{b}+\boldsymbol{a}\times\boldsymbol{b}\end{aligned}$$

pure imaginary product can get both (real & imaginary) / (scalar & vector) parts

$$ab=\boldsymbol{ab}=-\boldsymbol{a}\cdot\boldsymbol{b}+\boldsymbol{a}\times\boldsymbol{b},\begin{cases}a=0+\boldsymbol{a}=\boldsymbol{a}\\ b=0+\boldsymbol{b}=\boldsymbol{b}\end{cases}$$

---

triple product

https://en.wikipedia.org/wiki/Triple_product

pure imaginary

$$\begin{cases}a=0+\boldsymbol{a}=\boldsymbol{a}\\ b=0+\boldsymbol{b}=\boldsymbol{b}\\ c=0+\boldsymbol{c}=\boldsymbol{c}\end{cases}$$

$$\begin{aligned}abc=\boldsymbol{abc}&=(\boldsymbol{ab})\,\boldsymbol{c}=(-\boldsymbol{a}\cdot\boldsymbol{b}+\boldsymbol{a}\times\boldsymbol{b})\,\boldsymbol{c}\\ &=\boldsymbol{a}\,(\boldsymbol{bc})=\boldsymbol{a}\,(-\boldsymbol{b}\cdot\boldsymbol{c}+\boldsymbol{b}\times\boldsymbol{c})\end{aligned}$$

$$\begin{aligned}\boldsymbol{abc}=(\boldsymbol{ab})\,\boldsymbol{c}&=(-\boldsymbol{a}\cdot\boldsymbol{b}+\boldsymbol{a}\times\boldsymbol{b})\,\boldsymbol{c}\\ &=-(\boldsymbol{a}\cdot\boldsymbol{b})\,\boldsymbol{c}+(\boldsymbol{a}\times\boldsymbol{b})\,\boldsymbol{c}\\ &=-(\boldsymbol{a}\cdot\boldsymbol{b})\,\boldsymbol{c}+(-(\boldsymbol{a}\times\boldsymbol{b})\cdot\boldsymbol{c}+(\boldsymbol{a}\times\boldsymbol{b})\times\boldsymbol{c})\\ &=[-(\boldsymbol{a}\times\boldsymbol{b})\cdot\boldsymbol{c}]+[(\boldsymbol{a}\times\boldsymbol{b})\times\boldsymbol{c}-(\boldsymbol{a}\cdot\boldsymbol{b})\,\boldsymbol{c}]\\ =\boldsymbol{a}\,(\boldsymbol{bc})&=\boldsymbol{a}\,(-\boldsymbol{b}\cdot\boldsymbol{c}+\boldsymbol{b}\times\boldsymbol{c})\\ &=-\boldsymbol{a}\,(\boldsymbol{b}\cdot\boldsymbol{c})+\boldsymbol{a}\,(\boldsymbol{b}\times\boldsymbol{c})\\ &=-\boldsymbol{a}\,(\boldsymbol{b}\cdot\boldsymbol{c})+(-\boldsymbol{a}\cdot(\boldsymbol{b}\times\boldsymbol{c})+\boldsymbol{a}\times(\boldsymbol{b}\times\boldsymbol{c}))\\ &=[-\boldsymbol{a}\cdot(\boldsymbol{b}\times\boldsymbol{c})]+[\boldsymbol{a}\times(\boldsymbol{b}\times\boldsymbol{c})-\boldsymbol{a}\,(\boldsymbol{b}\cdot\boldsymbol{c})]\end{aligned}$$

by comparing (real & imaginary) / (scalar & vector) parts,

$$\begin{cases} -(\boldsymbol{a}\times\boldsymbol{b})\cdot\boldsymbol{c}=-\boldsymbol{a}\cdot(\boldsymbol{b}\times\boldsymbol{c}) \\ (\boldsymbol{a}\times\boldsymbol{b})\times\boldsymbol{c}-(\boldsymbol{a}\cdot\boldsymbol{b})\,\boldsymbol{c}=\boldsymbol{a}\times(\boldsymbol{b}\times\boldsymbol{c})-\boldsymbol{a}\,(\boldsymbol{b}\cdot\boldsymbol{c}) \end{cases}$$

$$\Rightarrow\begin{cases} (\boldsymbol{a}\times\boldsymbol{b})\cdot\boldsymbol{c}=\boldsymbol{a}\cdot(\boldsymbol{b}\times\boldsymbol{c}) & (s) \\ (\boldsymbol{a}\times\boldsymbol{b})\times\boldsymbol{c}-(\boldsymbol{a}\cdot\boldsymbol{b})\,\boldsymbol{c}=\boldsymbol{a}\times(\boldsymbol{b}\times\boldsymbol{c})-\boldsymbol{a}\,(\boldsymbol{b}\cdot\boldsymbol{c}) & (v) \end{cases}$$

---

permutation

$$\sigma=\begin{pmatrix} x_1 & x_2 & \cdots \\ \sigma(x_1) & \sigma(x_2) & \cdots \end{pmatrix}$$

$$\begin{pmatrix} a & b & c \\ a & b & c \end{pmatrix},\begin{pmatrix} a & b & c \\ b & c & a \end{pmatrix},\begin{pmatrix} a & b & c \\ c & a & b \end{pmatrix},\begin{pmatrix} a & b & c \\ a & c & b \end{pmatrix},\begin{pmatrix} a & b & c \\ b & a & c \end{pmatrix},\begin{pmatrix} a & b & c \\ c & b & a \end{pmatrix}$$

$$(s)\Rightarrow\begin{cases} (\boldsymbol{a}\times\boldsymbol{b})\cdot\boldsymbol{c}=\boldsymbol{a}\cdot(\boldsymbol{b}\times\boldsymbol{c}) & \begin{pmatrix} a & b & c \\ a & b & c \end{pmatrix},s_1 \\ (\boldsymbol{b}\times\boldsymbol{c})\cdot\boldsymbol{a}=\boldsymbol{b}\cdot(\boldsymbol{c}\times\boldsymbol{a}) & \begin{pmatrix} a & b & c \\ b & c & a \end{pmatrix},s_2 \\ (\boldsymbol{c}\times\boldsymbol{a})\cdot\boldsymbol{b}=\boldsymbol{c}\cdot(\boldsymbol{a}\times\boldsymbol{b}) & \begin{pmatrix} a & b & c \\ c & a & b \end{pmatrix},s_3 \\ (\boldsymbol{b}\times\boldsymbol{a})\cdot\boldsymbol{c}=\boldsymbol{b}\cdot(\boldsymbol{a}\times\boldsymbol{c}) & \begin{pmatrix} a & b & c \\ b & a & c \end{pmatrix},s_4 \\ (\boldsymbol{a}\times\boldsymbol{c})\cdot\boldsymbol{b}=\boldsymbol{a}\cdot(\boldsymbol{c}\times\boldsymbol{b}) & \begin{pmatrix} a & b & c \\ a & c & b \end{pmatrix},s_5 \\ (\boldsymbol{c}\times\boldsymbol{b})\cdot\boldsymbol{a}=\boldsymbol{c}\cdot(\boldsymbol{b}\times\boldsymbol{a}) & \begin{pmatrix} a & b & c \\ c & b & a \end{pmatrix},s_6 \end{cases}$$

$$\begin{aligned} &\overset{\cdot\text{ commutative}}{\Rightarrow}(\boldsymbol{a}\times\boldsymbol{b})\cdot\boldsymbol{c}\overset{s_1}{=}(\boldsymbol{b}\times\boldsymbol{c})\cdot\boldsymbol{a}\overset{s_2}{=}(\boldsymbol{c}\times\boldsymbol{a})\cdot\boldsymbol{b}\overset{s_3}{=}(\boldsymbol{a}\times\boldsymbol{b})\cdot\boldsymbol{c} \\ &\overset{\times\text{ anticommutative}}{=}-(\boldsymbol{b}\times\boldsymbol{a})\cdot\boldsymbol{c}\overset{s_6}{=}-(\boldsymbol{c}\times\boldsymbol{b})\cdot\boldsymbol{a}\overset{s_5}{=}-(\boldsymbol{a}\times\boldsymbol{c})\cdot\boldsymbol{b}\overset{s_4}{=}-(\boldsymbol{b}\times\boldsymbol{a})\cdot\boldsymbol{c} \\ &\Leftrightarrow\boldsymbol{a}\cdot(\boldsymbol{b}\times\boldsymbol{c})=\boldsymbol{b}\cdot(\boldsymbol{c}\times\boldsymbol{a})=\boldsymbol{c}\cdot(\boldsymbol{a}\times\boldsymbol{b}) \\ &=-\boldsymbol{a}\cdot(\boldsymbol{c}\times\boldsymbol{b})=-\boldsymbol{b}\cdot(\boldsymbol{a}\times\boldsymbol{c})=-\boldsymbol{c}\cdot(\boldsymbol{b}\times\boldsymbol{a}) \end{aligned}$$

$$(v)\Rightarrow\begin{cases}(\boldsymbol{a}\times\boldsymbol{b})\times\boldsymbol{c}-(\boldsymbol{a}\cdot\boldsymbol{b})\,\boldsymbol{c}=\boldsymbol{a}\times(\boldsymbol{b}\times\boldsymbol{c})-\boldsymbol{a}\,(\boldsymbol{b}\cdot\boldsymbol{c}) & \begin{pmatrix}a&b&c\\a&b&c\end{pmatrix},v_1\\ (\boldsymbol{b}\times\boldsymbol{c})\times\boldsymbol{a}-(\boldsymbol{b}\cdot\boldsymbol{c})\,\boldsymbol{a}=\boldsymbol{b}\times(\boldsymbol{c}\times\boldsymbol{a})-\boldsymbol{b}\,(\boldsymbol{c}\cdot\boldsymbol{a}) & \begin{pmatrix}a&b&c\\b&c&a\end{pmatrix},v_2\\ (\boldsymbol{c}\times\boldsymbol{a})\times\boldsymbol{b}-(\boldsymbol{c}\cdot\boldsymbol{a})\,\boldsymbol{b}=\boldsymbol{c}\times(\boldsymbol{a}\times\boldsymbol{b})-\boldsymbol{c}\,(\boldsymbol{a}\cdot\boldsymbol{b}) & \begin{pmatrix}a&b&c\\c&a&b\end{pmatrix},v_3\\ (\boldsymbol{b}\times\boldsymbol{a})\times\boldsymbol{c}-(\boldsymbol{b}\cdot\boldsymbol{a})\,\boldsymbol{c}=\boldsymbol{b}\times(\boldsymbol{a}\times\boldsymbol{c})-\boldsymbol{b}\,(\boldsymbol{a}\cdot\boldsymbol{c}) & \begin{pmatrix}a&b&c\\b&a&c\end{pmatrix},v_4\\ (\boldsymbol{a}\times\boldsymbol{c})\times\boldsymbol{b}-(\boldsymbol{a}\cdot\boldsymbol{c})\,\boldsymbol{b}=\boldsymbol{a}\times(\boldsymbol{c}\times\boldsymbol{b})-\boldsymbol{a}\,(\boldsymbol{c}\cdot\boldsymbol{b}) & \begin{pmatrix}a&b&c\\a&c&b\end{pmatrix},v_5\\ (\boldsymbol{c}\times\boldsymbol{b})\times\boldsymbol{a}-(\boldsymbol{c}\cdot\boldsymbol{b})\,\boldsymbol{a}=\boldsymbol{c}\times(\boldsymbol{b}\times\boldsymbol{a})-\boldsymbol{c}\,(\boldsymbol{b}\cdot\boldsymbol{a}) & \begin{pmatrix}a&b&c\\c&b&a\end{pmatrix},v_6\end{cases}$$

$$\Rightarrow\begin{cases}-Z-C=X-A & v_1\\ -X-A=Y-B & v_2\\ -Y-B=Z-C & v_3\\ Z-C=-Y-B & v_4\\ Y-B=-X-A & v_5\\ X-A=-Z-C & v_6\end{cases},\begin{cases}\cdot\text{ and scalar-vector product} & \text{commutative}\\ \times & \text{anticommutative}\end{cases},$$

$$\begin{cases}X=\boldsymbol{a}\times(\boldsymbol{b}\times\boldsymbol{c})=-(\boldsymbol{b}\times\boldsymbol{c})\times\boldsymbol{a}=-\boldsymbol{a}\times(\boldsymbol{c}\times\boldsymbol{b})=(\boldsymbol{c}\times\boldsymbol{b})\times\boldsymbol{a}\\ Y=\boldsymbol{b}\times(\boldsymbol{c}\times\boldsymbol{a})=-(\boldsymbol{c}\times\boldsymbol{a})\times\boldsymbol{b}=-\boldsymbol{b}\times(\boldsymbol{a}\times\boldsymbol{c})=(\boldsymbol{a}\times\boldsymbol{c})\times\boldsymbol{b}\\ Z=\boldsymbol{c}\times(\boldsymbol{a}\times\boldsymbol{b})=-(\boldsymbol{a}\times\boldsymbol{b})\times\boldsymbol{c}=-\boldsymbol{c}\times(\boldsymbol{b}\times\boldsymbol{a})=(\boldsymbol{b}\times\boldsymbol{a})\times\boldsymbol{c}\\ A=\boldsymbol{a}\,(\boldsymbol{b}\cdot\boldsymbol{c})=\boldsymbol{a}\,(\boldsymbol{c}\cdot\boldsymbol{b})=(\boldsymbol{c}\cdot\boldsymbol{b})\,\boldsymbol{a}=(\boldsymbol{b}\cdot\boldsymbol{c})\,\boldsymbol{a}\\ B=\boldsymbol{b}\,(\boldsymbol{c}\cdot\boldsymbol{a})=\boldsymbol{b}\,(\boldsymbol{a}\cdot\boldsymbol{c})=(\boldsymbol{a}\cdot\boldsymbol{c})\,\boldsymbol{b}=(\boldsymbol{c}\cdot\boldsymbol{a})\,\boldsymbol{b}\\ C=\boldsymbol{c}\,(\boldsymbol{a}\cdot\boldsymbol{b})=\boldsymbol{c}\,(\boldsymbol{b}\cdot\boldsymbol{a})=(\boldsymbol{b}\cdot\boldsymbol{a})\,\boldsymbol{c}=(\boldsymbol{a}\cdot\boldsymbol{b})\,\boldsymbol{c}\end{cases}$$

$$\Rightarrow\begin{cases}-Z-C=X-A & v_1=v_6\\ -X-A=Y-B & v_2=v_5\\ -Y-B=Z-C & v_3=v_4\end{cases}\Leftrightarrow\begin{cases}Z+X=A-C\\ X+Y=B-A\\ Y+Z=C-B\end{cases}\Leftrightarrow\begin{cases}X+Y=B-A\\ Y+Z=C-B\\ Z+X=A-C\end{cases}$$

$$\Leftrightarrow\begin{cases}2\,(X+Y+Z)=0 & \Rightarrow X+Y+Z=0\\ Y+Z=C-B & \Rightarrow X=B-C\Leftrightarrow\boldsymbol{a}\times(\boldsymbol{b}\times\boldsymbol{c})=\boldsymbol{b}\,(\boldsymbol{c}\cdot\boldsymbol{a})-\boldsymbol{c}\,(\boldsymbol{a}\cdot\boldsymbol{b})\ \text{"back cab"}\\ Z+X=A-C & \Rightarrow Y=C-A\Leftrightarrow\boldsymbol{b}\times(\boldsymbol{c}\times\boldsymbol{a})=\boldsymbol{c}\,(\boldsymbol{a}\cdot\boldsymbol{b})-\boldsymbol{a}\,(\boldsymbol{b}\cdot\boldsymbol{c})\\ X+Y=B-A & \Rightarrow Z=A-B\Leftrightarrow\boldsymbol{c}\times(\boldsymbol{a}\times\boldsymbol{b})=\boldsymbol{a}\,(\boldsymbol{b}\cdot\boldsymbol{c})-\boldsymbol{b}\,(\boldsymbol{c}\cdot\boldsymbol{a})\end{cases}$$

34.1.3.5 differential operator

https://en.wikipedia.org/wiki/Differential_operator

34.1.3.5.1 4-differential operator

4-differential operator / four-differential operator = d'Alembert operator

$$\begin{aligned}\mathrm{D}=&\frac{\partial}{\partial t}+\mathrm{i}\frac{\partial}{\partial x}+\mathrm{j}\frac{\partial}{\partial y}+\mathrm{k}\frac{\partial}{\partial z}=\partial_t+\mathrm{i}\partial_x+\mathrm{j}\partial_y+\mathrm{k}\partial_z\\ =&\frac{\partial}{\partial t}+\boldsymbol{i}\frac{\partial}{\partial x}+\boldsymbol{j}\frac{\partial}{\partial y}+\boldsymbol{k}\frac{\partial}{\partial z}=\partial_t+\boldsymbol{i}\partial_x+\boldsymbol{j}\partial_y+\boldsymbol{k}\partial_z\\ =&\frac{\partial}{\partial x_0}+\boldsymbol{e}_1\frac{\partial}{\partial x_1}+\boldsymbol{e}_2\frac{\partial}{\partial x_2}+\boldsymbol{e}_3\frac{\partial}{\partial x_3}=\partial_0+\boldsymbol{e}_i\partial_i=\partial_0+\boldsymbol{\nabla}\end{aligned}$$

$$\mathrm{D}=\partial_0+\mathrm{i}\partial_1+\mathrm{j}\partial_2+\mathrm{k}\partial_3=\partial_0+\boldsymbol{\nabla}=\mathrm{D}_0+\mathbf{D}$$

34.1.3.5.2 nabla

nabla = spatial differential operator = 3-differential operator / three-differential operator

$$\boldsymbol{\nabla} = \boldsymbol{e}_i \partial_i = \sum_{i=1}^{3} \boldsymbol{e}_i \partial_i = \sum_{i=1}^{3} \boldsymbol{e}_i \frac{\partial}{\partial x_i} = \begin{pmatrix} \frac{\partial}{\partial x} & \frac{\partial}{\partial y} & \frac{\partial}{\partial z} \end{pmatrix}^{\mathsf{T}} = \begin{pmatrix} \frac{\partial}{\partial x} \\ \frac{\partial}{\partial y} \\ \frac{\partial}{\partial z} \end{pmatrix}$$

34.1.3.5.3 Laplace operator Laplace operator = Laplacian

$$\triangle = \boldsymbol{\nabla}^2 = \boldsymbol{\nabla} \cdot \boldsymbol{\nabla} = \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2}$$

34.1.3.5.4 d'Alembert operator

$$\Box = \Box_c = \frac{1}{c^2}\frac{\partial^2}{\partial t^2} - \frac{\partial^2}{\partial x^2} - \frac{\partial^2}{\partial y^2} - \frac{\partial^2}{\partial z^2} = \frac{1}{c^2}\frac{\partial^2}{\partial t^2} - \triangle = \frac{1}{c^2}\frac{\partial^2}{\partial t^2} - \boldsymbol{\nabla}^2$$

$$\Box_1 = \Box_{c=1} = \frac{\partial^2}{\partial t^2} - \frac{\partial^2}{\partial x^2} - \frac{\partial^2}{\partial y^2} - \frac{\partial^2}{\partial z^2} = \frac{\partial^2}{\partial t^2} - \triangle = \frac{\partial^2}{\partial t^2} - \boldsymbol{\nabla}^2$$

34.1.3.6 electromagnetism

Maxwell

34.1.3.6.1 4-potential electromagnetic 4-potential / four-potential

$$A = A_0 + +\mathrm{i}A_1 + \mathrm{j}A_2 + \mathrm{k}A_3 = A_0 + \boldsymbol{A}$$

4-differential operator[34.1.3.5.1]

$$\mathrm{D} = \partial_0 + \mathrm{i}\partial_1 + \mathrm{j}\partial_2 + \mathrm{k}\partial_3 = \partial_0 + \boldsymbol{\nabla} = \mathrm{D}_0 + \mathbf{D}$$

$$QP = (Q_0P_0 - \boldsymbol{Q} \cdot \boldsymbol{P}) + (Q_0\boldsymbol{P} + P_0\boldsymbol{Q} + \boldsymbol{Q} \times \boldsymbol{P})$$

commutative bracket[34.1.3.3.1.1]

$$\begin{aligned}
[\mathrm{D}, A] =& \frac{\mathrm{D}A - A\mathrm{D}}{2} \\
2\,[\mathrm{D}, A] =& \mathrm{D}A - A\mathrm{D} \\
=& \left(\partial_0 + \mathrm{i}\partial_1 + \mathrm{j}\partial_2 + \mathrm{k}\partial_3\right)\left(A_0 + +\mathrm{i}A_1 + \mathrm{j}A_2 + \mathrm{k}A_3\right) \\
&- \left(A_0 + +\mathrm{i}A_1 + \mathrm{j}A_2 + \mathrm{k}A_3\right)\left(\partial_0 + \mathrm{i}\partial_1 + \mathrm{j}\partial_2 + \mathrm{k}\partial_3\right) \\
\mathrm{D}A =& (\mathrm{D}_0A_0 - \mathbf{D} \cdot \boldsymbol{A}) + (\mathrm{D}_0\boldsymbol{A} + A_0\mathbf{D} + \mathbf{D} \times \boldsymbol{A}) \\
=& (\mathrm{D}_0A_0 - \mathbf{D} \cdot \boldsymbol{A}) + (\mathrm{D}_0\boldsymbol{A} + \mathbf{D}A_0 + \mathbf{D} \times \boldsymbol{A}) \\
=& \mathrm{D}_0\left(A_0 + \boldsymbol{A}\right) - \mathbf{D} \cdot \boldsymbol{A} + \mathbf{D}A_0 + \mathbf{D} \times \boldsymbol{A} \\
=& \mathrm{D}_0A - \mathbf{D} \cdot \boldsymbol{A} + \mathbf{D}A_0 + \mathbf{D} \times \boldsymbol{A} \\
=& \partial_0A - \boldsymbol{\nabla} \cdot \boldsymbol{A} + \boldsymbol{\nabla}A_0 + \boldsymbol{\nabla} \times \boldsymbol{A} \\
A\mathrm{D} =& (A_0\mathrm{D}_0 - \boldsymbol{A} \cdot \mathbf{D}) + (A_0\mathbf{D} + \mathrm{D}_0\boldsymbol{A} + \boldsymbol{A} \times \mathbf{D}) \\
=& (\mathrm{D}_0A_0 - \mathbf{D} \cdot \boldsymbol{A}) + (\mathbf{D}A_0 + \mathrm{D}_0\boldsymbol{A} - \mathbf{D} \times \boldsymbol{A}) \\
=& (\mathrm{D}_0A_0 - \mathbf{D} \cdot \boldsymbol{A}) + (\mathrm{D}_0\boldsymbol{A} + \mathbf{D}A_0 - \mathbf{D} \times \boldsymbol{A}) \\
=& \mathrm{D}_0\left(A_0 + \boldsymbol{A}\right) - \mathbf{D} \cdot \boldsymbol{A} + \mathbf{D}A_0 - \mathbf{D} \times \boldsymbol{A} \\
=& \mathrm{D}_0A - \mathbf{D} \cdot \boldsymbol{A} + \mathbf{D}A_0 - \mathbf{D} \times \boldsymbol{A} \\
=& \partial_0A - \boldsymbol{\nabla} \cdot \boldsymbol{A} + \boldsymbol{\nabla}A_0 - \boldsymbol{\nabla} \times \boldsymbol{A}
\end{aligned}$$

$$\begin{aligned}
\mathrm{D}A =& \mathrm{D}_0A - \mathbf{D} \cdot \boldsymbol{A} + \mathbf{D}A_0 + \mathbf{D} \times \boldsymbol{A} = \partial_0A - \boldsymbol{\nabla} \cdot \boldsymbol{A} + \boldsymbol{\nabla}A_0 + \boldsymbol{\nabla} \times \boldsymbol{A} \\
A\mathrm{D} =& \mathrm{D}_0A - \mathbf{D} \cdot \boldsymbol{A} + \mathbf{D}A_0 - \mathbf{D} \times \boldsymbol{A} = \partial_0A - \boldsymbol{\nabla} \cdot \boldsymbol{A} + \boldsymbol{\nabla}A_0 - \boldsymbol{\nabla} \times \boldsymbol{A}
\end{aligned}$$

$$\begin{aligned}\mathrm{D}A - A\mathrm{D} &= 2\mathbf{D} \times \boldsymbol{A} = 2\boldsymbol{\nabla} \times \boldsymbol{A} \\ [\mathrm{D}, A] =& \frac{\mathrm{D}A - A\mathrm{D}}{2} = \mathbf{D} \times \boldsymbol{A} = \boldsymbol{\nabla} \times \boldsymbol{A}\end{aligned}$$

anticommutative bracket[34.1.3.3.1.2]

$$\begin{aligned}\mathrm{D}A + A\mathrm{D} &= 2\left(\mathrm{D}_0 A - \mathbf{D} \cdot \boldsymbol{A} + \mathbf{D} A_0\right) = 2\left(\partial_0 A - \boldsymbol{\nabla} \cdot \boldsymbol{A} + \boldsymbol{\nabla} A_0\right) \\ \{\mathrm{D}, A\} =& \frac{\mathrm{D}A + A\mathrm{D}}{2} = \mathrm{D}_0 A - \mathbf{D} \cdot \boldsymbol{A} + \mathbf{D} A_0 = \partial_0 A - \boldsymbol{\nabla} \cdot \boldsymbol{A} + \boldsymbol{\nabla} A_0\end{aligned}$$

commutation and anticommutation on differential operator and any quaternion

$$\begin{aligned}[\mathrm{D}, Q] =& \boldsymbol{\nabla} \times \boldsymbol{Q} \\ \{\mathrm{D}, Q\} =& \partial_0 Q - \boldsymbol{\nabla} \cdot \boldsymbol{Q} + \boldsymbol{\nabla} Q_0\end{aligned}$$

or more evident

$$\begin{aligned}[\mathrm{D}, Q] =& \boldsymbol{\nabla} \times \boldsymbol{Q} \\ \{\mathrm{D}, Q\} =& \partial_0 Q - \boldsymbol{\nabla} \cdot \boldsymbol{Q} + \boldsymbol{\nabla} Q_0 \\ =& \partial_0 \left(Q_0 + \boldsymbol{Q}\right) - \boldsymbol{\nabla} \cdot \boldsymbol{Q} + \boldsymbol{\nabla} Q_0 \\ =& \left(\partial_0 Q_0 - \boldsymbol{\nabla} \cdot \boldsymbol{Q}\right) + \left(\partial_0 \boldsymbol{Q} + \boldsymbol{\nabla} Q_0\right) \\ =& \left(\frac{\partial Q_0}{\partial t} - \boldsymbol{\nabla} \cdot \boldsymbol{Q}\right) + \left(\frac{\partial \boldsymbol{Q}}{\partial t} + \boldsymbol{\nabla} Q_0\right)\end{aligned}$$

34.1.3.6.2 Maxwell compromise for both quaternion and 3-vector electric potential and vector potential

$$A = A_0 + +\mathrm{i}A_1 + \mathrm{j}A_2 + \mathrm{k}A_3 = A_0 + \boldsymbol{A} = U + \boldsymbol{A}$$

electric quaternion and electric field

$$\begin{aligned}E =& -\{\mathrm{D}, A\} \\ =& -\left(\partial_0 A - \boldsymbol{\nabla} \cdot \boldsymbol{A} + \boldsymbol{\nabla} A_0\right) \\ =& -\partial_0 A + \boldsymbol{\nabla} \cdot \boldsymbol{A} - \boldsymbol{\nabla} A_0 \\ =& -\partial_t \left(U + \boldsymbol{A}\right) + \boldsymbol{\nabla} \cdot \boldsymbol{A} - \boldsymbol{\nabla} U \\ =& -\frac{\partial U}{\partial t} + \boldsymbol{\nabla} \cdot \boldsymbol{A} - \boldsymbol{\nabla} U - \frac{\partial \boldsymbol{A}}{\partial t} \\ =& E_0 + \boldsymbol{E}, \begin{cases} E_0 = -\dfrac{\partial U}{\partial t} + \boldsymbol{\nabla} \cdot \boldsymbol{A} \\ \boldsymbol{E} = -\boldsymbol{\nabla} U - \dfrac{\partial \boldsymbol{A}}{\partial t} \qquad \text{electric field 3-vector} \end{cases}\end{aligned}$$

magnetic field

$$B = [\mathrm{D}, A] = \boldsymbol{\nabla} \times \boldsymbol{A} = \begin{vmatrix} \boldsymbol{i} & \boldsymbol{j} & \boldsymbol{k} \\ \dfrac{\partial}{\partial x} & \dfrac{\partial}{\partial y} & \dfrac{\partial}{\partial z} \\ A_1 & A_2 & A_3 \end{vmatrix} = \boldsymbol{B}$$

Work on time? Yes.

$$qE = qE_0 + q\boldsymbol{E} = qE_0 + \boldsymbol{F}_E$$

force equivalent on time

$$qE_0$$

$$\begin{cases} E = -\{\mathrm{D}, A\} = E_0 + \boldsymbol{E} \\ B = +[\mathrm{D}, A] = B_0 + \boldsymbol{B} = 0 + \boldsymbol{B} = \boldsymbol{B} \quad B_0 = 0 \end{cases}$$

---

for any quaternion commutating and anticommutating with differential operator

$$\begin{aligned} [\mathrm{D}, Q] =& \boldsymbol{\nabla} \times \boldsymbol{Q} \\ \{\mathrm{D}, Q\} =& \partial_0 Q - \boldsymbol{\nabla} \cdot \boldsymbol{Q} + \boldsymbol{\nabla} Q_0 \\ =& \partial_0 \left(Q_0 + \boldsymbol{Q}\right) - \boldsymbol{\nabla} \cdot \boldsymbol{Q} + \boldsymbol{\nabla} Q_0 \\ =& \left(\partial_0 Q_0 - \boldsymbol{\nabla} \cdot \boldsymbol{Q}\right) + \left(\partial_0 \boldsymbol{Q} + \boldsymbol{\nabla} Q_0\right) \\ =& \left(\frac{\partial Q_0}{\partial t} - \boldsymbol{\nabla} \cdot \boldsymbol{Q}\right) + \left(\frac{\partial \boldsymbol{Q}}{\partial t} + \boldsymbol{\nabla} Q_0\right) \end{aligned}$$

$$\begin{cases} [\mathrm{D}, E] = \boldsymbol{\nabla} \times \boldsymbol{E} = (0) + (\boldsymbol{\nabla} \times \boldsymbol{E}) \\ \{\mathrm{D}, E\} = (\partial_0 E_0 - \boldsymbol{\nabla} \cdot \boldsymbol{E}) + (\partial_0 \boldsymbol{E} + \boldsymbol{\nabla} E_0) = \left(\frac{\partial E_0}{\partial t} - \boldsymbol{\nabla} \cdot \boldsymbol{E}\right) + \left(\frac{\partial \boldsymbol{E}}{\partial t} + \boldsymbol{\nabla} E_0\right) \\ [\mathrm{D}, B] = \boldsymbol{\nabla} \times \boldsymbol{B} = (0) + (\boldsymbol{\nabla} \times \boldsymbol{B}) \\ \{\mathrm{D}, B\} = (\partial_0 B_0 - \boldsymbol{\nabla} \cdot \boldsymbol{B}) + (\partial_0 \boldsymbol{B} + \boldsymbol{\nabla} B_0) \overset{B_0=0}{=} -\boldsymbol{\nabla} \cdot \boldsymbol{B} + \partial_0 \boldsymbol{B} = (-\boldsymbol{\nabla} \cdot \boldsymbol{B}) + \left(\frac{\partial \boldsymbol{B}}{\partial t}\right) \end{cases}$$

by comparing (real & imaginary) / (scalar & vector) parts,

Maxwell equations without source terms

$$\begin{cases} [\mathrm{D}, B] = +\{\mathrm{D}, E\} & \Leftrightarrow (0) + (\boldsymbol{\nabla} \times \boldsymbol{B}) = \left(\frac{\partial E_0}{\partial t} - \boldsymbol{\nabla} \cdot \boldsymbol{E}\right) + \left(\frac{\partial \boldsymbol{E}}{\partial t} + \boldsymbol{\nabla} E_0\right) \\ [\mathrm{D}, E] = -\{\mathrm{D}, B\} & \Leftrightarrow (0) + (\boldsymbol{\nabla} \times \boldsymbol{E}) = (-\boldsymbol{\nabla} \cdot \boldsymbol{B}) + \left(\frac{\partial \boldsymbol{B}}{\partial t}\right) \end{cases}$$

$$\Leftrightarrow \begin{cases} \frac{\partial E_0}{\partial t} - \boldsymbol{\nabla} \cdot \boldsymbol{E} = 0 & \Leftrightarrow \boldsymbol{\nabla} \cdot \boldsymbol{E} = \frac{\partial E_0}{\partial t} \\ \frac{\partial \boldsymbol{E}}{\partial t} + \boldsymbol{\nabla} E_0 = \boldsymbol{\nabla} \times \boldsymbol{B} & \Leftrightarrow \boldsymbol{\nabla} \times \boldsymbol{B} = \frac{\partial \boldsymbol{E}}{\partial t} + \boldsymbol{\nabla} E_0 \text{ 動電生磁} \\ -\boldsymbol{\nabla} \cdot \boldsymbol{B} = 0 & \Leftrightarrow \boldsymbol{\nabla} \cdot \boldsymbol{B} = 0 \\ \frac{\partial \boldsymbol{B}}{\partial t} = \boldsymbol{\nabla} \times \boldsymbol{E} & \Leftrightarrow \boldsymbol{\nabla} \times \boldsymbol{E} = \frac{\partial \boldsymbol{B}}{\partial t} \text{ 動磁生電} \end{cases}$$

34.1.3.7 Joule heat vs. Thomson heat (Kelvin heat?)

The Lord Kelvin = William Thomson

34.1.3.7.1 thermoelectric effect thermoelectric effect = Seebeck effect = Peltier effect = Thomson effect

34.1.3.8 source term

$$J = J_0 + \mathrm{i} J_1 + \mathrm{j} J_2 + \mathrm{k} J_3 = J_0 + \boldsymbol{J} = \rho + \boldsymbol{J}$$

Maxwell equations with source terms

$$\begin{cases} [\mathrm{D}, B] = J + \{\mathrm{D}, E\} & \Leftrightarrow (0) + (\boldsymbol{\nabla} \times \boldsymbol{B}) = \left(\rho + \frac{\partial E_0}{\partial t} - \boldsymbol{\nabla} \cdot \boldsymbol{E}\right) + \left(\boldsymbol{J} + \frac{\partial \boldsymbol{E}}{\partial t} + \boldsymbol{\nabla} E_0\right) \\ [\mathrm{D}, E] = 0 - \{\mathrm{D}, B\} & \Leftrightarrow (0) + (\boldsymbol{\nabla} \times \boldsymbol{E}) = (-\boldsymbol{\nabla} \cdot \boldsymbol{B}) + \left(\frac{\partial \boldsymbol{B}}{\partial t}\right) \end{cases}$$

$$\Leftrightarrow \begin{cases} \rho + \frac{\partial E_0}{\partial t} - \boldsymbol{\nabla} \cdot \boldsymbol{E} = 0 & \Leftrightarrow \boldsymbol{\nabla} \cdot \boldsymbol{E} = \rho + \frac{\partial E_0}{\partial t} \\ \boldsymbol{J} + \frac{\partial \boldsymbol{E}}{\partial t} + \boldsymbol{\nabla} E_0 = \boldsymbol{\nabla} \times \boldsymbol{B} & \Leftrightarrow \boldsymbol{\nabla} \times \boldsymbol{B} = \boldsymbol{J} + \frac{\partial \boldsymbol{E}}{\partial t} + \boldsymbol{\nabla} E_0 \text{ 動電生磁} \\ -\boldsymbol{\nabla} \cdot \boldsymbol{B} = 0 & \Leftrightarrow \boldsymbol{\nabla} \cdot \boldsymbol{B} = 0 \\ \frac{\partial \boldsymbol{B}}{\partial t} = \boldsymbol{\nabla} \times \boldsymbol{E} & \Leftrightarrow \boldsymbol{\nabla} \times \boldsymbol{E} = \frac{\partial \boldsymbol{B}}{\partial t} \text{ 動磁生電} \end{cases}$$

34.1.4 quaternion group

https://en.wikipedia.org/wiki/Quaternion_group

group theory[37]

or please first see quaternion group representation[37.5].

34.1.4.1 2D rotation

34.1.4.1.1 matrix

$$\boldsymbol{r}=(x,y)=\langle x,y\rangle=\begin{pmatrix}x\\y\end{pmatrix}=\begin{pmatrix}r\cos\alpha\\r\sin\alpha\end{pmatrix}$$

$$\boldsymbol{r}'=(x',y')=\langle x',y'\rangle=\begin{pmatrix}x'\\y'\end{pmatrix}=\begin{pmatrix}r\cos(\alpha+\theta)\\r\sin(\alpha+\theta)\end{pmatrix}$$

$$\begin{aligned}\boldsymbol{r}'=&\begin{pmatrix}r\cos(\alpha+\theta)\\r\sin(\alpha+\theta)\end{pmatrix}=r\begin{pmatrix}\cos(\alpha+\theta)\\\sin(\alpha+\theta)\end{pmatrix}\\=&r\begin{pmatrix}\cos\alpha\cos\theta-\sin\alpha\sin\theta\\\sin\alpha\cos\theta+\cos\alpha\sin\theta\end{pmatrix}=r\begin{pmatrix}\cos\theta\cos\alpha-\sin\theta\sin\alpha\\\sin\theta\cos\alpha+\cos\theta\sin\alpha\end{pmatrix}\\=&r\begin{pmatrix}\cos\theta&-\sin\theta\\\sin\theta&\cos\theta\end{pmatrix}\begin{pmatrix}\cos\alpha\\\sin\alpha\end{pmatrix}=\begin{pmatrix}\cos\theta&-\sin\theta\\\sin\theta&\cos\theta\end{pmatrix}\begin{pmatrix}r\cos\alpha\\r\sin\alpha\end{pmatrix}\\=&R\boldsymbol{r},\begin{cases}R=\begin{pmatrix}\cos\theta&-\sin\theta\\\sin\theta&\cos\theta\end{pmatrix}=R(\theta)=R_\theta\\\boldsymbol{r}=\begin{pmatrix}r\cos\alpha\\r\sin\alpha\end{pmatrix},\boldsymbol{r}'=\begin{pmatrix}r\cos(\alpha+\theta)\\r\sin(\alpha+\theta)\end{pmatrix}\end{cases}\end{aligned}$$

orthonormal matrix

$$\boldsymbol{r}'=O\boldsymbol{r}$$

$$\begin{aligned}|\boldsymbol{r}'|^2=&|\boldsymbol{r}|^2\\\boldsymbol{r}'\cdot\boldsymbol{r}'=&\boldsymbol{r}\cdot\boldsymbol{r}\\\boldsymbol{r}'^{\mathsf{T}}\boldsymbol{r}'=&\boldsymbol{r}^{\mathsf{T}}\boldsymbol{r}\\(O\boldsymbol{r})^{\mathsf{T}}(O\boldsymbol{r})=&\\\boldsymbol{r}^{\mathsf{T}}O^{\mathsf{T}}O\boldsymbol{r}=&\\\boldsymbol{r}^{\mathsf{T}}O^{\mathsf{T}}O\boldsymbol{r}=&\boldsymbol{r}^{\mathsf{T}}\boldsymbol{r}\\O^{\mathsf{T}}O=&1=I=I_2\end{aligned}$$

$$\begin{aligned}R^{\mathsf{T}}R=&\begin{pmatrix}\cos\theta&-\sin\theta\\\sin\theta&\cos\theta\end{pmatrix}^{\mathsf{T}}\begin{pmatrix}\cos\theta&-\sin\theta\\\sin\theta&\cos\theta\end{pmatrix}\\=&\begin{pmatrix}\cos\theta&\sin\theta\\-\sin\theta&\cos\theta\end{pmatrix}\begin{pmatrix}\cos\theta&-\sin\theta\\\sin\theta&\cos\theta\end{pmatrix}\\=&\begin{pmatrix}\cos^2\theta+\sin^2\theta&-\cos\theta\sin\theta+\sin\theta\cos\theta\\-\sin\theta\cos\theta+\cos\theta\sin\theta&\sin^2\theta+\cos^2\theta\end{pmatrix}\\=&\begin{pmatrix}1&0\\0&1\end{pmatrix}=1=I\\R^{\mathsf{T}}R=&1\Rightarrow R\in\{O|O^{\mathsf{T}}O=1\}\end{aligned}$$

https://en.wikipedia.org/wiki/Transformation_matrix#Affine_transformations

reflection matrix

$$\begin{cases}P_x=\begin{pmatrix}-1&0\\0&1\end{pmatrix}&P_x^{\mathsf{T}}P_x=P_x^2=\begin{pmatrix}-1&0\\0&1\end{pmatrix}^2=\begin{pmatrix}1&0\\0&1\end{pmatrix}=1\Rightarrow P_x\in\{O|O^{\mathsf{T}}O=1\}\\P_y=\begin{pmatrix}1&0\\0&-1\end{pmatrix}&P_y^{\mathsf{T}}P_y=P_y^2=\begin{pmatrix}1&0\\0&-1\end{pmatrix}^2=\begin{pmatrix}1&0\\0&1\end{pmatrix}=1\Rightarrow P_y\in\{O|O^{\mathsf{T}}O=1\}\end{cases}$$

translation matrix?

$$\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = 1$$

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

$O(2)$ group

$$\begin{aligned} O(2) &= \{1, R, P_x, P_y\} \\ &= \{I_2, R_\theta, P_x, P_y\} \subseteq \{O | O^\intercal O = 1\} \end{aligned}$$

$$\begin{aligned} 1 =& O^\intercal O \\ 1 =& \det 1 = \det I = \det (I_2) \\ =& \det (O^\intercal O) = (\det O^\intercal)(\det O) = (\det O)(\det O) = (\det O)^2 \\ 1 =& (\det O)^2 \\ \det O =& \pm 1 \end{aligned}$$

$$\det R = \det R_\theta = \begin{vmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{vmatrix} = \cos^2\theta - \left(-\sin^2\theta\right) = \cos^2\theta + \sin^2\theta = 1$$

$$\begin{cases} P_x = \begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix} & \det P_x = -1 \\ P_y = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} & \det P_y = -1 \end{cases}$$

special orthonormal group of degree 2

$$\begin{aligned} SO(2) = \{1, R\} = \{I_2, R_\theta\} \subseteq \left\{ O \middle| \begin{cases} O^\intercal O = 1 & O \\ \det O = 1 & S \end{cases} \right\} \\ \subset \{1, R, P_x, P_y\} = O(2) \subseteq \{O | O^\intercal O = 1\} \end{aligned}$$

34.1.4.1.2 complex

$$z = r(\cos\alpha + \mathrm{i}\sin\alpha) = r\mathrm{e}^{\mathrm{i}\alpha}$$

$$z' = r(\cos(\alpha+\theta) + \mathrm{i}\sin(\alpha+\theta)) = r\mathrm{e}^{\mathrm{i}(\alpha+\theta)}$$

$$\begin{aligned} z' =& z_\theta z \\ z_\theta =& \frac{z'}{z} = \frac{r'\mathrm{e}^{\mathrm{i}(\alpha+\theta)}}{r\mathrm{e}^{\mathrm{i}\alpha}} = \frac{r'}{r}\mathrm{e}^{\mathrm{i}\theta} = \frac{r'}{r}(\cos\theta + \mathrm{i}\sin\theta) \end{aligned}$$

$$\begin{aligned} z_\theta z =& \left[\frac{r'}{r}(\cos\theta + \mathrm{i}\sin\theta)\right][r(\cos\alpha + \mathrm{i}\sin\alpha)] \\ =& r'[(\cos\theta\cos\alpha - \sin\theta\sin\alpha) + \mathrm{i}(\sin\theta\cos\alpha + \cos\theta\sin\alpha)] \\ =& r'[\cos(\alpha+\theta) + \mathrm{i}\sin(\alpha+\theta)] = z' \end{aligned}$$

$$\hat{z}_\theta = z_\theta\left(\frac{r'}{r} = 1\right) = \mathrm{e}^{\mathrm{i}\theta} = \cos\theta + \mathrm{i}\sin\theta$$

$$\hat{z}_\theta^* = \overline{\hat{z}_\theta} = \mathrm{e}^{-\mathrm{i}\theta} = \cos\theta - \mathrm{i}\sin\theta$$

$$\hat{z}_\theta^*\hat{z}_\theta = \mathrm{e}^{\mathrm{i}\theta}\mathrm{e}^{-\mathrm{i}\theta} = \mathrm{e}^{\mathrm{i}\theta+(-\mathrm{i}\theta)} = \mathrm{e}^{\mathrm{i}0} = \mathrm{e}^0 = 1$$

unitary group of degree 1

$$U(1) = \{1, \hat{z}_\theta\} = \left\{e^{i0}, e^{i\theta}\right\}$$

34.1.4.1.3 $SO(2) \cong U(1)$ $\mathbb{C} \leftrightarrow \mathcal{M}_{2\times 2}(\mathbb{R}) = \mathcal{M}_2(\mathbb{R})$ complex group representation[37.4]

$$x + yi \leftrightarrow \begin{pmatrix} x & -y \\ y & x \end{pmatrix} = xI + yJ$$

$$\begin{aligned}
U(1) &= \{1, \hat{z}_\theta\} = \left\{e^{i0}, e^{i\theta}\right\} \\
&= \{\cos 0 + i\sin 0, \cos\theta + i\sin\theta\} \\
&\leftrightarrow \left\{\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}\cos 0 + \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}\sin 0, \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}\cos\theta + \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}\sin\theta\right\} \\
&= \left\{\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}1 + \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}0, \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}\cos\theta + \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}\sin\theta\right\} \\
&= \left\{\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}\right\} = \{I_2, R_\theta\} = \{1, R\} = SO(2) \\
U(1) &\cong SO(2) \Leftrightarrow SO(2) \cong U(1)
\end{aligned}$$

unitary group of degree 1 and special orthonormal group of degree 2 are isomorphism

34.1.4.2 3D rotation

34.1.4.2.1 matrix

34.1.4.2.1.1 construction with 2D rotation matrix

$$\begin{aligned}
\boldsymbol{r}' &= \begin{pmatrix} r\cos(\alpha+\theta) \\ r\sin(\alpha+\theta) \\ z' \end{pmatrix} \overset{z'=z}{=} \begin{pmatrix} r\cos(\alpha+\theta) \\ r\sin(\alpha+\theta) \\ z \end{pmatrix} = \begin{pmatrix} R(\theta) & \\ & 1 \end{pmatrix}\begin{pmatrix} r\cos\alpha \\ r\sin\alpha \\ z \end{pmatrix} \\
&= \begin{pmatrix} \cos\theta & -\sin\theta & 0 \\ \sin\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{pmatrix}\begin{pmatrix} r\cos\alpha \\ r\sin\alpha \\ z \end{pmatrix} = R_z(\theta)\,\boldsymbol{r}, \begin{cases} R_z(\theta) = \begin{pmatrix} \cos\theta & -\sin\theta & 0 \\ \sin\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{pmatrix} \\ \boldsymbol{r} = \begin{pmatrix} r\cos\alpha \\ r\sin\alpha \\ z \end{pmatrix} \end{cases}
\end{aligned}$$

$$\begin{aligned}
\boldsymbol{r}' &= \begin{pmatrix} x' \\ r\cos(\alpha+\theta) \\ r\sin(\alpha+\theta) \end{pmatrix} \overset{x'=x}{=} \begin{pmatrix} x \\ r\cos(\alpha+\theta) \\ r\sin(\alpha+\theta) \end{pmatrix} = \begin{pmatrix} 1 & \\ & R(\theta) \end{pmatrix}\begin{pmatrix} x \\ r\cos\alpha \\ r\sin\alpha \end{pmatrix} \\
&= \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos\theta & -\sin\theta \\ 0 & \sin\theta & \cos\theta \end{pmatrix}\begin{pmatrix} x \\ r\cos\alpha \\ r\sin\alpha \end{pmatrix} = R_x(\theta)\,\boldsymbol{r}, \begin{cases} R_x(\theta) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos\theta & -\sin\theta \\ 0 & \sin\theta & \cos\theta \end{pmatrix} \\ \boldsymbol{r} = \begin{pmatrix} x \\ r\cos\alpha \\ r\sin\alpha \end{pmatrix} \end{cases}
\end{aligned}$$

$$\begin{aligned}
\boldsymbol{r}' &= \begin{pmatrix} r\sin(\alpha+\theta) \\ y' \\ r\cos(\alpha+\theta) \end{pmatrix} \overset{y'=y}{=} \begin{pmatrix} r\sin(\alpha+\theta) \\ y \\ r\cos(\alpha+\theta) \end{pmatrix} \\
&= \begin{pmatrix} \cos\theta & 0 & \sin\theta \\ 0 & 1 & 0 \\ -\sin\theta & 0 & \cos\theta \end{pmatrix}\begin{pmatrix} r\sin\alpha \\ y \\ r\cos\alpha \end{pmatrix} = R_y(\theta)\,\boldsymbol{r}, \begin{cases} R_y(\theta) = \begin{pmatrix} \cos\theta & 0 & \sin\theta \\ 0 & 1 & 0 \\ -\sin\theta & 0 & \cos\theta \end{pmatrix} \\ \boldsymbol{r} = \begin{pmatrix} r\sin\alpha \\ y \\ r\cos\alpha \end{pmatrix} \end{cases}
\end{aligned}$$

34.1.4.2.1.2 Euler angle $z \to x \to z : \alpha \to \beta \to \gamma$

$$
\begin{aligned}
&R_z(\gamma) R_x(\beta) R_z(\alpha) \\
&= \begin{pmatrix} \cos\gamma & -\sin\gamma & 0 \\ \sin\gamma & \cos\gamma & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos\beta & -\sin\beta \\ 0 & \sin\beta & \cos\beta \end{pmatrix} \begin{pmatrix} \cos\alpha & -\sin\alpha & 0 \\ \sin\alpha & \cos\alpha & 0 \\ 0 & 0 & 1 \end{pmatrix} \\
&= \begin{pmatrix} \cos\gamma & -\sin\gamma & 0 \\ \sin\gamma & \cos\gamma & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} \cos\alpha & -\sin\alpha & 0 \\ \cos\beta\sin\alpha & \cos\beta\cos\alpha & -\sin\beta \\ \sin\beta\sin\alpha & \sin\beta\cos\alpha & \cos\beta \end{pmatrix} \\
&= \begin{pmatrix} \cos\gamma\cos\alpha - \sin\gamma\cos\beta\sin\alpha & -\cos\gamma\sin\alpha - \sin\gamma\cos\beta\cos\alpha & \sin\gamma\sin\beta \\ \sin\gamma\cos\alpha + \cos\gamma\cos\beta\sin\alpha & -\sin\gamma\sin\alpha - \cos\gamma\cos\beta\cos\alpha & -\cos\gamma\sin\beta \\ \sin\beta\sin\alpha & \sin\beta\cos\alpha & \cos\beta \end{pmatrix}
\end{aligned}
$$

$x \to y \to z : \alpha \to \beta \to \gamma$

$$
\begin{aligned}
&R_z(\gamma) R_y(\beta) R_x(\alpha) \\
&= \begin{pmatrix} \cos\gamma & -\sin\gamma & 0 \\ \sin\gamma & \cos\gamma & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} \cos\beta & 0 & \sin\beta \\ 0 & 1 & 0 \\ -\sin\beta & 0 & \cos\beta \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos\alpha & -\sin\alpha \\ 0 & \sin\alpha & \cos\alpha \end{pmatrix} \\
&= \begin{pmatrix} \cos\gamma & -\sin\gamma & 0 \\ \sin\gamma & \cos\gamma & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} \cos\beta & \sin\beta\sin\alpha & \sin\beta\cos\alpha \\ 0 & \cos\alpha & -\sin\alpha \\ -\sin\beta & \cos\beta\sin\alpha & \cos\beta\cos\alpha \end{pmatrix} \\
&= \begin{pmatrix} \cos\gamma\cos\beta & \cos\gamma\sin\beta\sin\alpha - \sin\gamma\cos\alpha & \cos\gamma\sin\beta\cos\alpha + \sin\gamma\sin\alpha \\ \sin\gamma\cos\beta & \sin\gamma\sin\beta\sin\alpha + \cos\gamma\cos\alpha & \sin\gamma\sin\beta\cos\alpha - \cos\gamma\sin\alpha \\ -\sin\beta & \cos\beta\sin\alpha & \cos\beta\cos\alpha \end{pmatrix}
\end{aligned}
$$

34.1.4.2.1.3 3D rotation about an arbitrary axis https://math.stackexchange.com/questions/4550704/rotation-around-an-arbitrary-axis

spherical coordinate unit vector

$$
\begin{cases}
\hat{x} = r\sin\theta\cos\phi & \overset{r=1}{=} \sin\theta\cos\phi \\
\hat{y} = r\sin\theta\sin\phi & \overset{r=1}{=} \sin\theta\sin\phi \\
\hat{z} = r\cos\theta & \overset{r=1}{=} \cos\theta
\end{cases}
$$

although I prefer θ and ϕ switched back to be compatible with 2D coordinate

$$
\begin{cases}
\hat{x} = r\sin\phi\cos\theta & \overset{r=1}{=} \sin\phi\cos\theta \\
\hat{y} = r\sin\phi\sin\theta & \overset{r=1}{=} \sin\phi\sin\theta \\
\hat{z} = r\cos\phi & \overset{r=1}{=} \cos\phi
\end{cases}
$$

or cos first in x, y-plane

$$
\begin{cases}
\hat{x} = r\cos\phi\cos\theta & \overset{r=1}{=} \cos\phi\cos\theta \\
\hat{y} = r\cos\phi\sin\theta & \overset{r=1}{=} \cos\phi\sin\theta \\
\hat{z} = r\sin\phi & \overset{r=1}{=} \sin\phi
\end{cases}
$$

---

still use the most convention

$$
\hat{\boldsymbol{n}} = \begin{pmatrix} \hat{x} \\ \hat{y} \\ \hat{z} \end{pmatrix} = \begin{pmatrix} \sin\theta\cos\phi \\ \sin\theta\sin\phi \\ \cos\theta \end{pmatrix}
$$

$$\begin{cases} \hat{\boldsymbol{n}} = \begin{pmatrix} \sin\theta\cos\phi \\ \sin\theta\sin\phi \\ \cos\theta \end{pmatrix} \\ \hat{\boldsymbol{u}} = \begin{pmatrix} \cos\theta\cos\phi \\ \cos\theta\sin\phi \\ -\sin\theta \end{pmatrix} \Leftarrow \cos\theta\sin\theta - \cos\theta\sin\theta = 0 \Rightarrow \qquad \hat{\boldsymbol{u}}\cdot\hat{\boldsymbol{n}} = 0 \Leftrightarrow \hat{\boldsymbol{u}} \perp \hat{\boldsymbol{n}} \\ \hat{\boldsymbol{v}} = \hat{\boldsymbol{n}} \times \hat{\boldsymbol{u}} = \begin{vmatrix} \boldsymbol{i} & \boldsymbol{j} & \boldsymbol{k} \\ \sin\theta\cos\phi & \sin\theta\sin\phi & \cos\theta \\ \cos\theta\cos\phi & \cos\theta\sin\phi & -\sin\theta \end{vmatrix} = \begin{pmatrix} -\sin\phi \\ \cos\phi \\ 0 \end{pmatrix} \Rightarrow \begin{cases} \hat{\boldsymbol{v}} \perp \hat{\boldsymbol{n}} \\ \hat{\boldsymbol{v}} \perp \hat{\boldsymbol{u}} \end{cases} \end{cases}$$

$$S = \{\hat{\boldsymbol{n}}, \hat{\boldsymbol{u}}, \hat{\boldsymbol{v}}\} = \{\hat{\boldsymbol{u}}, \hat{\boldsymbol{v}}, \hat{\boldsymbol{n}}\} \text{ is a basis of the spherical coordinate}$$

$$[\boldsymbol{V}]_S = \begin{pmatrix} u \\ v \\ n \end{pmatrix}$$

$$\begin{aligned} \boldsymbol{V} &= \begin{pmatrix} \hat{\boldsymbol{u}} & \hat{\boldsymbol{v}} & \hat{\boldsymbol{n}} \end{pmatrix} \begin{pmatrix} u \\ v \\ n \end{pmatrix} = u\hat{\boldsymbol{u}} + v\hat{\boldsymbol{v}} + n\hat{\boldsymbol{n}} \\ &= \begin{pmatrix} \cos\theta\cos\phi & -\sin\phi & \sin\theta\cos\phi \\ \cos\theta\sin\phi & \cos\phi & \sin\theta\sin\phi \\ -\sin\theta & 0 & \cos\theta \end{pmatrix} \begin{pmatrix} u \\ v \\ n \end{pmatrix} = S\,[\boldsymbol{V}]_S \\ \begin{pmatrix} u \\ v \\ n \end{pmatrix} &= \begin{pmatrix} \cos\theta\cos\phi & -\sin\phi & \sin\theta\cos\phi \\ \cos\theta\sin\phi & \cos\phi & \sin\theta\sin\phi \\ -\sin\theta & 0 & \cos\theta \end{pmatrix}^{-1} \boldsymbol{V} \\ [\boldsymbol{V}]_S &= S^{-1}\boldsymbol{V} \end{aligned}$$

$$\begin{aligned} S^{-1} &= \begin{pmatrix} \cos\theta\cos\phi & -\sin\phi & \sin\theta\cos\phi \\ \cos\theta\sin\phi & \cos\phi & \sin\theta\sin\phi \\ -\sin\theta & 0 & \cos\theta \end{pmatrix}^{-1}, S \in \{O | O^{\intercal}O = 1\} \\ &= \begin{pmatrix} \cos\theta\cos\phi & -\sin\phi & \sin\theta\cos\phi \\ \cos\theta\sin\phi & \cos\phi & \sin\theta\sin\phi \\ -\sin\theta & 0 & \cos\theta \end{pmatrix}^{\intercal} \\ &= \begin{pmatrix} \cos\theta\cos\phi & \cos\theta\sin\phi & -\sin\theta \\ -\sin\phi & \cos\phi & 0 \\ \sin\theta\cos\phi & \sin\theta\sin\phi & \cos\theta \end{pmatrix} = \begin{pmatrix} \hat{\boldsymbol{u}}^{\intercal} \\ \hat{\boldsymbol{v}}^{\intercal} \\ \hat{\boldsymbol{n}}^{\intercal} \end{pmatrix} \\ S^{-1}S &= \begin{pmatrix} \cos\theta\cos\phi & \cos\theta\sin\phi & -\sin\theta \\ -\sin\phi & \cos\phi & 0 \\ \sin\theta\cos\phi & \sin\theta\sin\phi & \cos\theta \end{pmatrix} \begin{pmatrix} \cos\theta\cos\phi & -\sin\phi & \sin\theta\cos\phi \\ \cos\theta\sin\phi & \cos\phi & \sin\theta\sin\phi \\ -\sin\theta & 0 & \cos\theta \end{pmatrix} \\ &= \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} = 1 \end{aligned}$$

$\hat{\boldsymbol{n}}$ as z direction

$$\begin{aligned} [\boldsymbol{V}]_S &= S^{-1}\boldsymbol{V} \\ [\boldsymbol{V}']_S = R_z(\gamma)\,[\boldsymbol{V}]_S &= R_z(\gamma)\,S^{-1}\boldsymbol{V} \\ \boldsymbol{V}' = S\,[\boldsymbol{V}']_S &= SR_z(\gamma)\,[\boldsymbol{V}]_S = SR_z(\gamma)\,S^{-1}\boldsymbol{V} \\ \boldsymbol{V}' &= SR_z(\gamma)\,S^{-1}\boldsymbol{V} \end{aligned}$$

https://www.symbolab.com/

$$
\begin{aligned}
&SR_z(\gamma)S^{-1}\\
&=\begin{pmatrix}\cos\theta\cos\phi & -\sin\phi & \sin\theta\cos\phi\\ \cos\theta\sin\phi & \cos\phi & \sin\theta\sin\phi\\ -\sin\theta & 0 & \cos\theta\end{pmatrix}\begin{pmatrix}\cos\gamma & -\sin\gamma & 0\\ \sin\gamma & \cos\gamma & 0\\ 0 & 0 & 1\end{pmatrix}\begin{pmatrix}\cos\theta\cos\phi & \cos\theta\sin\phi & -\sin\theta\\ -\sin\phi & \cos\phi & 0\\ \sin\theta\cos\phi & \sin\theta\sin\phi & \cos\theta\end{pmatrix}\\
&=\begin{pmatrix}c_1c_2 & -s_2 & s_1c_2\\ c_1s_2 & c_2 & s_1s_2\\ -s_1 & 0 & c_1\end{pmatrix}\begin{pmatrix}c_3 & -s_3 & 0\\ s_3 & c_3 & 0\\ 0 & 0 & 1\end{pmatrix}\begin{pmatrix}c_1c_2 & c_1s_2 & -s_1\\ -s_2 & c_2 & 0\\ s_1c_2 & s_1s_2 & c_1\end{pmatrix},\begin{cases}c_1=\cos\theta & s_1=\sin\theta\\ c_2=\cos\phi & s_2=\sin\phi\\ c_3=\cos\gamma & s_3=\sin\gamma\end{cases}\\
&=\begin{pmatrix}c_1c_2 & -s_2 & s_1c_2\\ c_1s_2 & c_2 & s_1s_2\\ -s_1 & 0 & c_1\end{pmatrix}\begin{pmatrix}c_3c_1c_2+s_3s_2 & c_3c_1s_2-c_2s_3 & -c_3s_1\\ c_1c_2s_3-c_3s_2 & c_3c_2+c_1s_3s_2 & -s_3s_1\\ c_2s_1 & s_2s_1 & c_1\end{pmatrix}\\
&=\begin{pmatrix}c_1^2c_2^2c_3+c_3s_2^2+c_2^2s_1^2 & c_1s_2\left(c_1c_2c_3-s_2s_3\right)+c_2\left(-c_1c_2s_3-c_3s_2\right)+c_2s_2s_1^2 & c_1c_2s_1-s_1\left(c_1c_2c_3-s_2\right.\\ c_1c_2\left(c_1c_3s_2+c_2s_3\right)-s_2\left(c_2c_3-c_1s_2s_3\right)+c_2s_2s_1^2 & c_1^2c_3s_2^2+s_2^2s_1^2+c_2^2c_3 & c_1s_2s_1-s_1\left(c_1c_3s_2+c_2\right.\\ -c_1c_2c_3s_1-s_2s_1s_3+c_1c_2s_1 & -c_1c_3s_2s_1+c_2s_1s_3+c_1s_2s_1 & c_3s_1^2+c_1^2\end{pmatrix}
\end{aligned}
$$

34.1.4.2.2 quaternion https://math.stackexchange.com/questions/328117/how-does-one-derive-this-rotation-quaternion-formula

$$
\begin{aligned}
q&=a+b\mathrm{i}+c\mathrm{j}+d\mathrm{k}=a+\mathrm{i}b+\mathrm{j}c+\mathrm{k}d,\begin{cases}q\in\mathbb{H}\\ \mathrm{i}^2=\mathrm{j}^2=\mathrm{k}^2=-1=\mathrm{ijk} & \Rightarrow \mathrm{ij}=\mathrm{k}\\ a,b,c,d\in\mathbb{R} & \Leftrightarrow \langle a,b,c,d\rangle\in\mathbb{R}^4\end{cases}\\
&=w=t+x\mathrm{i}+y\mathrm{j}+z\mathrm{k}=t+\mathrm{i}x+\mathrm{j}y+\mathrm{k}z,\begin{cases}w\in\mathbb{H}\\ \mathrm{i}^2=\mathrm{j}^2=\mathrm{k}^2=-1=\mathrm{ijk} & \Rightarrow \mathrm{ij}=\mathrm{k}\\ t,x,y,z\in\mathbb{R} & \Leftrightarrow \langle t,x,y,z\rangle\in\mathbb{R}^4\end{cases}\\
&=t+\begin{pmatrix}\mathrm{i} & \mathrm{j} & \mathrm{k}\end{pmatrix}\begin{pmatrix}x\\ y\\ z\end{pmatrix}=x_0+\begin{pmatrix}\boldsymbol{i} & \boldsymbol{j} & \boldsymbol{k}\end{pmatrix}\begin{pmatrix}x\\ y\\ z\end{pmatrix}=x_0+\begin{pmatrix}\boldsymbol{e}_1 & \boldsymbol{e}_2 & \boldsymbol{e}_3\end{pmatrix}\begin{pmatrix}x_1\\ x_2\\ x_3\end{pmatrix}=x_0+\boldsymbol{x},\begin{cases}\boldsymbol{e}_1=\boldsymbol{i}=\mathrm{i}\\ \boldsymbol{e}_2=\boldsymbol{j}=\mathrm{j}\\ \boldsymbol{e}_3=\boldsymbol{k}=\mathrm{k}\end{cases}
\end{aligned}
$$

$$q_1q_2=\left(x_{10}x_{20}-\boldsymbol{x}_1\cdot\boldsymbol{x}_2\right)+\left(x_{10}\boldsymbol{x}_2+x_{20}\boldsymbol{x}_1+\boldsymbol{x}_1\times\boldsymbol{x}_2\right)$$

$$QP=\left(Q_0P_0-\boldsymbol{Q}\cdot\boldsymbol{P}\right)+\left(Q_0\boldsymbol{P}+P_0\boldsymbol{Q}+\boldsymbol{Q}\times\boldsymbol{P}\right)$$

$$q=q_0+\boldsymbol{q}=q_0+q_1\mathrm{i}+q_2\mathrm{j}+q_3\mathrm{k}=q_0+q_1\mathbf{i}+q_2\mathbf{j}+q_3\mathbf{k}$$

$$
\begin{aligned}
q^*=\overline{q}=&\overline{q_0+\boldsymbol{q}}=q_0-\boldsymbol{q}\\
=&\overline{q_0+q_1\mathrm{i}+q_2\mathrm{j}+q_3\mathrm{k}}=q_0-q_1\mathrm{i}-q_2\mathrm{j}-q_3\mathrm{k}\\
=&q_0-q_1\mathbf{i}-q_2\mathbf{j}-q_3\mathbf{k}
\end{aligned}
$$

$$v=v_0+\boldsymbol{v}=v_0+v_1\mathrm{i}+v_2\mathrm{j}+v_3\mathrm{k}=v_0+v_1\mathbf{i}+v_2\mathbf{j}+v_3\mathbf{k}$$

$$\boldsymbol{v}=v_1\mathrm{i}+v_2\mathrm{j}+v_3\mathrm{k}=v_1\mathbf{i}+v_2\mathbf{j}+v_3\mathbf{k}$$

$$
\begin{aligned}
qv&=\left(q_0v_0-\boldsymbol{q}\cdot\boldsymbol{v}\right)+\left(q_0\boldsymbol{v}+v_0\boldsymbol{q}+\boldsymbol{q}\times\boldsymbol{v}\right)\\
&\overset{v_0=0}{=}\left(q_00-\boldsymbol{q}\cdot\boldsymbol{v}\right)+\left(q_0\boldsymbol{v}+0\boldsymbol{q}+\boldsymbol{q}\times\boldsymbol{v}\right)\\
&=\left(-\boldsymbol{q}\cdot\boldsymbol{v}\right)+\left(q_0\boldsymbol{v}+\boldsymbol{q}\times\boldsymbol{v}\right)\\
q\boldsymbol{v}&=\left(-\boldsymbol{q}\cdot\boldsymbol{v}\right)+\left(q_0\boldsymbol{v}+\boldsymbol{q}\times\boldsymbol{v}\right)
\end{aligned}
$$

$$
\begin{aligned}
v\overline{q} &= (v_0 q_0 - \boldsymbol{v}\cdot\overline{\boldsymbol{q}}) + (v_0\overline{\boldsymbol{q}} + q_0\boldsymbol{v} + \boldsymbol{v}\times\overline{\boldsymbol{q}}) \\
&\overset{v_0=0}{=} (0q_0 - \boldsymbol{v}\cdot\overline{\boldsymbol{q}}) + (0\overline{\boldsymbol{q}} + q_0\boldsymbol{v} + \boldsymbol{v}\times\overline{\boldsymbol{q}}) \\
&= (-\boldsymbol{v}\cdot\overline{\boldsymbol{q}}) + (q_0\boldsymbol{v} + \boldsymbol{v}\times\overline{\boldsymbol{q}}) \\
&= (-\boldsymbol{v}\cdot(-\boldsymbol{q})) + (q_0\boldsymbol{v} + \boldsymbol{v}\times(-\boldsymbol{q})) \\
&= (\boldsymbol{v}\cdot\boldsymbol{q}) + (q_0\boldsymbol{v} - \boldsymbol{v}\times\boldsymbol{q}) \\
\boldsymbol{v}\overline{q} &= (\boldsymbol{v}\cdot\boldsymbol{q}) + (q_0\boldsymbol{v} - \boldsymbol{v}\times\boldsymbol{q})
\end{aligned}
$$

https://math.stackexchange.com/questions/41574/can-eulers-identity-be-extended-to-quaternions

$$
\begin{aligned}
q =& a + b\mathrm{i} + c\mathrm{j} + d\mathrm{k} = a + \mathrm{i}b + \mathrm{j}c + \mathrm{k}d \\
=& a + r\frac{\mathrm{i}b + \mathrm{j}c + \mathrm{k}d}{r}, r^2 = b^2 + c^2 + d^2 \\
=& a + \theta\frac{\mathrm{i}b + \mathrm{j}c + \mathrm{k}d}{\theta}, \theta^2 = b^2 + c^2 + d^2
\end{aligned}
$$

$$
\begin{aligned}
\left(\frac{\mathrm{i}b + \mathrm{j}c + \mathrm{k}d}{r}\right)^2 =& \frac{-b^2 - c^2 - d^2 + bc\,(\mathrm{ij} + \mathrm{ji}) + cd\,(\mathrm{jk} + \mathrm{kj}) + db\,(\mathrm{ki} + \mathrm{ik})}{r^2} \\
=& \frac{-b^2 - c^2 - d^2 + bc\,(\mathrm{k} - \mathrm{k}) + cd\,(\mathrm{i} - \mathrm{i}) + db\,(\mathrm{j} - \mathrm{j})}{r^2} \\
=& \frac{-b^2 - c^2 - d^2 + bc0 + cd0 + db0}{r^2} = \frac{-b^2 - c^2 - d^2}{r^2} \\
=& \frac{-b^2 - c^2 - d^2}{b^2 + c^2 + d^2} = -1 \\
\frac{\mathrm{i}b + \mathrm{j}c + \mathrm{k}d}{r} =& \pm\sqrt{-1}
\end{aligned}
$$

$$
\begin{aligned}
\mathrm{e}^q =& \mathrm{e}^{a + \mathrm{i}b + \mathrm{j}c + \mathrm{k}d} \\
=& \mathrm{e}^{a + r\sqrt{-1}} = \mathrm{e}^a\mathrm{e}^{r\sqrt{-1}} = \mathrm{e}^{a + \theta\sqrt{-1}} = \mathrm{e}^a\mathrm{e}^{\theta\sqrt{-1}} \\
=& \mathrm{e}^a\left(\cos r + \sqrt{-1}\sin r\right) = \mathrm{e}^a\left(\cos\theta + \sqrt{-1}\sin\theta\right) \\
=& \mathrm{e}^a\left(\cos r + \frac{\mathrm{i}b + \mathrm{j}c + \mathrm{k}d}{r}\sin r\right) = \mathrm{e}^a\left[\cos r + (\mathrm{i}b + \mathrm{j}c + \mathrm{k}d)\frac{\sin r}{r}\right] \\
=& \mathrm{e}^a\left(\cos\theta + \frac{\mathrm{i}b + \mathrm{j}c + \mathrm{k}d}{\theta}\sin\theta\right) = \mathrm{e}^a\left[\cos\theta + (\mathrm{i}b + \mathrm{j}c + \mathrm{k}d)\frac{\sin\theta}{\theta}\right]
\end{aligned}
$$

$$
\begin{aligned}
q =& a + b\mathrm{i} + c\mathrm{j} + d\mathrm{k} = a + \mathrm{i}b + \mathrm{j}c + \mathrm{k}d \\
=& a + r\frac{\mathrm{i}b + \mathrm{j}c + \mathrm{k}d}{r}, r^2 = b^2 + c^2 + d^2 \\
=& \sqrt{a^2 + r^2}\left(\frac{a}{\sqrt{a^2 + r^2}} + \frac{\mathrm{i}b + \mathrm{j}c + \mathrm{k}d}{r}\frac{r}{\sqrt{a^2 + r^2}}\right) \\
=& \rho\left(\cos\phi + \sqrt{-1}\sin\phi\right), \begin{cases} \rho = \sqrt{a^2 + r^2} \\ \tan\phi = \dfrac{r}{a} \Leftrightarrow \phi = \arctan\dfrac{r}{a} \end{cases} \\
=& \rho\mathrm{e}^{\phi\sqrt{-1}}
\end{aligned}
$$

$$
q = \rho\mathrm{e}^{\phi\sqrt{-1}}, \begin{cases} q = a + \mathrm{i}b + \mathrm{j}c + \mathrm{k}d \\ \rho = \sqrt{a^2 + r^2} = \sqrt{a^2 + b^2 + c^2 + d^2} \\ \tan\phi = \dfrac{r}{a} \Leftrightarrow \phi = \arctan\dfrac{r}{a} = \arctan\dfrac{\pm\sqrt{b^2 + c^2 + d^2}}{a} \end{cases}
$$

$$
\begin{aligned}
\rho\mathrm{e}^{-\phi\sqrt{-1}} =& \rho\left[\cos(-\phi) + \sqrt{-1}\sin(-\phi)\right] \\
=& \rho\left[\cos\phi - \sqrt{-1}\sin\phi\right] \\
=& \sqrt{a^2 + r^2}\left[\frac{a}{\sqrt{a^2 + r^2}} - \frac{\mathrm{i}b + \mathrm{j}c + \mathrm{k}d}{r}\frac{r}{\sqrt{a^2 + r^2}}\right] \\
=& a - (\mathrm{i}b + \mathrm{j}c + \mathrm{k}d) = a - \mathrm{i}b - \mathrm{j}c - \mathrm{k}d = \overline{q} = q^*
\end{aligned}
$$

34.1.5 octonion

34.2 Krasjet

https://github.com/Krasjet/quaternion

https://krasjet.github.io/quaternion/

https://krasjet.github.io/quaternion/bonus_gimbal_lock.pdf

34.2.1 Rodrigues rotation

$$\boldsymbol{v} \to \boldsymbol{v}'$$

$$\boldsymbol{v} \xrightarrow{\text{rotate about } \boldsymbol{u}} \boldsymbol{v}'$$

$$\boldsymbol{v} \xrightarrow{\text{rotate about } \boldsymbol{n}} \boldsymbol{v}'$$

$$\begin{cases} \boldsymbol{v} = \boldsymbol{v}_{\parallel \boldsymbol{n}} + \boldsymbol{v}_{\perp \boldsymbol{n}} = \boldsymbol{v}_{\parallel} + \boldsymbol{v}_{\perp} & \Rightarrow \boldsymbol{v}_{\perp} = \boldsymbol{v} - \boldsymbol{v}_{\parallel} \\ \boldsymbol{v}' = \boldsymbol{v}'_{\parallel \boldsymbol{n}} + \boldsymbol{v}'_{\perp \boldsymbol{n}} = \boldsymbol{v}'_{\parallel} + \boldsymbol{v}'_{\perp} & \Rightarrow \boldsymbol{v}'_{\perp} = \boldsymbol{v}' - \boldsymbol{v}'_{\parallel} \end{cases}$$

$$\begin{aligned} \boldsymbol{v}_{\parallel \boldsymbol{n}} = \boldsymbol{v}_{\parallel} &= \text{proj}_{\boldsymbol{n}} \boldsymbol{v} = \frac{\boldsymbol{v} \cdot \boldsymbol{n}}{\|\boldsymbol{n}\|} \hat{\boldsymbol{n}} \\ &= \frac{\boldsymbol{v} \cdot \boldsymbol{n}}{\|\boldsymbol{n}\|} \frac{\boldsymbol{n}}{\|\boldsymbol{n}\|} = \frac{\boldsymbol{v} \cdot \boldsymbol{n}}{\|\boldsymbol{n}\|^2} \boldsymbol{n} \end{aligned}$$

$$\begin{aligned} \boldsymbol{v}_{\parallel \hat{\boldsymbol{n}}} = \boldsymbol{v}_{\parallel} &= \text{proj}_{\hat{\boldsymbol{n}}} \boldsymbol{v} = \frac{\boldsymbol{v} \cdot \hat{\boldsymbol{n}}}{\|\hat{\boldsymbol{n}}\|} \hat{\boldsymbol{n}} \overset{\|\hat{\boldsymbol{n}}\|=1}{=} (\boldsymbol{v} \cdot \hat{\boldsymbol{n}})\, \hat{\boldsymbol{n}} = \boldsymbol{v} \cdot \hat{\boldsymbol{n}} \hat{\boldsymbol{n}} \\ &= \frac{\boldsymbol{v} \cdot \hat{\boldsymbol{n}}}{\|\hat{\boldsymbol{n}}\|} \frac{\hat{\boldsymbol{n}}}{\|\hat{\boldsymbol{n}}\|} = \frac{\boldsymbol{v} \cdot \hat{\boldsymbol{n}}}{\|\hat{\boldsymbol{n}}\|^2} \hat{\boldsymbol{n}} \overset{\|\hat{\boldsymbol{n}}\|=1}{=} (\boldsymbol{v} \cdot \hat{\boldsymbol{n}})\, \hat{\boldsymbol{n}} = \boldsymbol{v} \cdot \hat{\boldsymbol{n}} \hat{\boldsymbol{n}} \end{aligned}$$

$$\boldsymbol{n} = \hat{\boldsymbol{n}}$$

$$\begin{aligned} \boldsymbol{v}_{\parallel \boldsymbol{n}} = \boldsymbol{v}_{\parallel} &= \text{proj}_{\boldsymbol{n}} \boldsymbol{v} = \frac{\boldsymbol{v} \cdot \boldsymbol{n}}{\|\boldsymbol{n}\|} \hat{\boldsymbol{n}} \overset{\|\boldsymbol{n}\|=1}{=} (\boldsymbol{v} \cdot \hat{\boldsymbol{n}})\, \hat{\boldsymbol{n}} = \boldsymbol{v} \cdot \hat{\boldsymbol{n}} \hat{\boldsymbol{n}} \\ &= \frac{\boldsymbol{v} \cdot \boldsymbol{n}}{\|\boldsymbol{n}\|} \frac{\boldsymbol{n}}{\|\boldsymbol{n}\|} = \frac{\boldsymbol{v} \cdot \boldsymbol{n}}{\|\boldsymbol{n}\|^2} \boldsymbol{n} \overset{\|\boldsymbol{n}\|=1}{=} (\boldsymbol{v} \cdot \boldsymbol{n})\, \boldsymbol{n} = \boldsymbol{v} \cdot \boldsymbol{n} \boldsymbol{n} \end{aligned}$$

$$\boldsymbol{v}_{\parallel} = (\boldsymbol{v} \cdot \boldsymbol{n})\, \boldsymbol{n}$$

$$\boldsymbol{v}_{\perp} = \boldsymbol{v} - \boldsymbol{v}_{\parallel} = \boldsymbol{v} - (\boldsymbol{v} \cdot \boldsymbol{n})\, \boldsymbol{n}$$

$$\boldsymbol{v}_{\parallel} + \boldsymbol{v}_{\perp} = \boldsymbol{v} \xrightarrow{\text{rotate about } \boldsymbol{n}} \boldsymbol{v}' = \boldsymbol{v}'_{\parallel} + \boldsymbol{v}'_{\perp}$$

$$\boldsymbol{v}_{\parallel} \xrightarrow{\text{rotate about } \boldsymbol{n}} \boldsymbol{v}'_{\parallel}$$

$$\boldsymbol{v}_{\parallel} = \boldsymbol{v}_{\parallel \boldsymbol{n}} = \boldsymbol{v}'_{\parallel \boldsymbol{n}} = \boldsymbol{v}'_{\parallel}$$

$$\boldsymbol{v}_{\parallel} = \boldsymbol{v}'_{\parallel}$$

$$\boldsymbol{v}'_{\parallel} = \boldsymbol{v}_{\parallel}$$

$$\begin{cases} \boldsymbol{u} = \boldsymbol{n} \times \boldsymbol{v}_\perp \\ \\ \|\boldsymbol{u}\| = \|\boldsymbol{n} \times \boldsymbol{v}_\perp\| = \|\boldsymbol{n}\| \|\boldsymbol{v}_\perp\| \sin\frac{\pi}{2} = \|\boldsymbol{v}_\perp\| \end{cases} \begin{cases} \boldsymbol{u} \perp \boldsymbol{n} \\ \boldsymbol{u} \perp \boldsymbol{v}_\perp \end{cases} \begin{cases} \boldsymbol{n} \perp \boldsymbol{v}_\perp \\ \|\boldsymbol{n}\| = 1 \end{cases}$$

$$\begin{aligned} \boldsymbol{v}'_\perp =& \boldsymbol{v}'_{\|\boldsymbol{v}_\perp} + \boldsymbol{v}'_{\|\boldsymbol{u}} \\ =& (\cos\theta)\, \boldsymbol{v}_\perp + (\sin\theta)\, \boldsymbol{u}, \theta = \angle \boldsymbol{v}'_\perp \boldsymbol{v}_\perp \\ =& \cos\theta \boldsymbol{v}_\perp + \sin\theta \boldsymbol{u} \\ =& \cos\theta \boldsymbol{v}_\perp + \sin\theta\, (\boldsymbol{n} \times \boldsymbol{v}_\perp) \\ \boldsymbol{v}'_\perp =& \cos\theta \boldsymbol{v}_\perp + \sin\theta\, (\boldsymbol{n} \times \boldsymbol{v}_\perp) \end{aligned}$$

$$\boldsymbol{v}'_\perp = \cos\theta \boldsymbol{v}_\perp + \sin\theta\, (\boldsymbol{n} \times \boldsymbol{v}_\perp), \theta = \angle \boldsymbol{v}'_\perp \boldsymbol{v}_\perp$$

$$\boldsymbol{v}'_\perp = \cos\theta \boldsymbol{v}_\perp + \sin\theta\, (\boldsymbol{n} \times \boldsymbol{v}_\perp)$$

$$\begin{aligned} \boldsymbol{v}' =& \boldsymbol{v}'_\| + \boldsymbol{v}'_\perp, \begin{cases} \boldsymbol{v}'_\| = \boldsymbol{v}_\| \\ \boldsymbol{v}'_\perp = \cos\theta \boldsymbol{v}_\perp + \sin\theta\, (\boldsymbol{n} \times \boldsymbol{v}_\perp), \theta = \angle \boldsymbol{v}'_\perp \boldsymbol{v}_\perp \end{cases} \\ =& \boldsymbol{v}_\| + \cos\theta \boldsymbol{v}_\perp + \sin\theta\, (\boldsymbol{n} \times \boldsymbol{v}_\perp), \theta = \angle \boldsymbol{v}'_\perp \boldsymbol{v}_\perp \\ =& \boldsymbol{v}_\| + \cos\theta \boldsymbol{v}_\perp + \sin\theta\, (\boldsymbol{n} \times \boldsymbol{v}_\perp), \begin{cases} \boldsymbol{v}_\| = (\boldsymbol{v} \cdot \boldsymbol{n})\, \boldsymbol{n} \\ \boldsymbol{v}_\perp = \boldsymbol{v} - \boldsymbol{v}_\| = \boldsymbol{v} - (\boldsymbol{v} \cdot \boldsymbol{n})\, \boldsymbol{n} \end{cases} \\ =& (\boldsymbol{v} \cdot \boldsymbol{n})\, \boldsymbol{n} + \cos\theta\, [\boldsymbol{v} - (\boldsymbol{v} \cdot \boldsymbol{n})\, \boldsymbol{n}] + \sin\theta\, [\boldsymbol{n} \times (\boldsymbol{v} - \boldsymbol{v}_\|)] \\ =& (\boldsymbol{v} \cdot \boldsymbol{n})\, \boldsymbol{n} + \cos\theta \boldsymbol{v} - \cos\theta\, (\boldsymbol{v} \cdot \boldsymbol{n})\, \boldsymbol{n} + \sin\theta\, [\boldsymbol{n} \times \boldsymbol{v} - \boldsymbol{n} \times \boldsymbol{v}_\|] \\ =& \cos\theta \boldsymbol{v} + (1 - \cos\theta)\, (\boldsymbol{v} \cdot \boldsymbol{n})\, \boldsymbol{n} + \sin\theta\, [\boldsymbol{n} \times \boldsymbol{v} - \boldsymbol{0}] \\ =& \cos\theta \boldsymbol{v} + (1 - \cos\theta)\, (\boldsymbol{v} \cdot \boldsymbol{n})\, \boldsymbol{n} + \sin\theta \boldsymbol{n} \times \boldsymbol{v} \end{aligned}$$

$$\boldsymbol{v}' = \cos\theta \boldsymbol{v} + (1 - \cos\theta)\, (\boldsymbol{v} \cdot \boldsymbol{n})\, \boldsymbol{n} + \sin\theta \boldsymbol{n} \times \boldsymbol{v}$$

$$\begin{aligned} \boldsymbol{v}' &= \cos\theta \boldsymbol{v} + (1 - \cos\theta)\, (\boldsymbol{v} \cdot \boldsymbol{n})\, \boldsymbol{n} + \sin\theta \boldsymbol{n} \times \boldsymbol{v} \\ &= \cos\theta \boldsymbol{v} + \left[1 - \left(1 - 2\sin^2\frac{\theta}{2}\right)\right] (\boldsymbol{v} \cdot \boldsymbol{n})\, \boldsymbol{n} + \left[2\sin\frac{\theta}{2}\cos\frac{\theta}{2}\right] \boldsymbol{n} \times \boldsymbol{v} \\ &= \cos\theta \boldsymbol{v} + \left[2\sin^2\frac{\theta}{2}\right] (\boldsymbol{v} \cdot \boldsymbol{n})\, \boldsymbol{n} + \left[2\sin\frac{\theta}{2}\cos\frac{\theta}{2}\right] \boldsymbol{n} \times \boldsymbol{v} \\ &= \cos\theta \boldsymbol{v} + 2\sin\frac{\theta}{2}\left[\left(\sin\frac{\theta}{2}\right)(\boldsymbol{v} \cdot \boldsymbol{n})\, \boldsymbol{n} + \left(\cos\frac{\theta}{2}\right) \boldsymbol{n} \times \boldsymbol{v}\right] \\ &= \cos\theta \boldsymbol{v} + 2\sin\frac{\theta}{2}\left[\left(\cos\frac{\theta}{2}\right) \boldsymbol{n} \times \boldsymbol{v} + \left(\sin\frac{\theta}{2}\right)(\boldsymbol{v} \cdot \boldsymbol{n})\, \boldsymbol{n}\right] \\ &= \left[2\cos^2\frac{\theta}{2} - 1\right] \boldsymbol{v} + 2\sin\frac{\theta}{2}\left[\left(\cos\frac{\theta}{2}\right) \boldsymbol{n} \times \boldsymbol{v} + \left(\sin\frac{\theta}{2}\right)(\boldsymbol{v} \cdot \boldsymbol{n})\, \boldsymbol{n}\right] \end{aligned}$$

34.2.2 quaternion operation

$$\begin{aligned} q =& a + b\mathrm{i} + c\mathrm{j} + d\mathrm{k} \\ \varrho =& \alpha + \beta\mathrm{i} + \gamma\mathrm{j} + \delta\mathrm{k} \end{aligned}$$

$$
\begin{aligned}
\begin{pmatrix} a \\ b \\ c \\ d \end{pmatrix} \begin{pmatrix} \alpha \\ \beta \\ \gamma \\ \delta \end{pmatrix} \leftarrow q\varrho = & (a + b\mathrm{i} + c\mathrm{j} + d\mathrm{k})(\alpha + \beta\mathrm{i} + \gamma\mathrm{j} + \delta\mathrm{k}) \\
& = a(\alpha + \beta\mathrm{i} + \gamma\mathrm{j} + \delta\mathrm{k}) \\
& + b\mathrm{i}(\alpha + \beta\mathrm{i} + \gamma\mathrm{j} + \delta\mathrm{k}) \\
& + c\mathrm{j}(\alpha + \beta\mathrm{i} + \gamma\mathrm{j} + \delta\mathrm{k}) \\
& + d\mathrm{k}(\alpha + \beta\mathrm{i} + \gamma\mathrm{j} + \delta\mathrm{k}) \\
& = a\alpha + a\beta\mathrm{i} + a\gamma\mathrm{j} + a\delta\mathrm{k} \\
& + b\alpha\mathrm{i} - b\beta + b\gamma\mathrm{k} - b\delta\mathrm{j} \\
& + c\alpha\mathrm{j} - c\beta\mathrm{k} - c\gamma + c\delta\mathrm{i} \\
& + d\alpha\mathrm{k} + d\beta\mathrm{j} - d\gamma\mathrm{i} - d\delta \\
& = (a\alpha - b\beta - c\gamma - d\delta) \\
& + (b\alpha + a\beta - d\gamma + c\delta)\mathrm{i} \\
& + (c\alpha + d\beta + a\gamma - b\delta)\mathrm{j} \\
& + (d\alpha - c\beta + b\gamma + a\delta)\mathrm{k} \\
\begin{pmatrix} a & -b & -c & -d \\ b & a & -d & c \\ c & d & a & -b \\ d & -c & b & a \end{pmatrix} \begin{pmatrix} \alpha \\ \beta \\ \gamma \\ \delta \end{pmatrix} \leftarrow & \\
& = (\alpha a - \beta b - \gamma c - \delta d) \\
& + (\beta a + \alpha b + \delta c - \gamma d)\mathrm{i} \\
& + (\gamma a - \delta b + \alpha c + \beta d)\mathrm{j} \\
& + (\delta a + \gamma b - \beta c + \alpha d)\mathrm{k} \\
\begin{pmatrix} \alpha & -\beta & -\gamma & -\delta \\ \beta & \alpha & \delta & -\gamma \\ \gamma & -\delta & \alpha & \beta \\ \delta & \gamma & -\beta & \alpha \end{pmatrix} \begin{pmatrix} a \\ b \\ c \\ d \end{pmatrix} \leftarrow &
\end{aligned}
$$

$$
qv \leftrightarrow \begin{pmatrix} a & -b & -c & -d \\ b & a & -d & c \\ c & d & a & -b \\ d & -c & b & a \end{pmatrix} \begin{pmatrix} v_0 \\ v_1 \\ v_2 \\ v_3 \end{pmatrix} = Q_\imath v
$$

concept like quaternion group representation[37.5]

$$
\begin{aligned}
L(q) = Q_\imath = \begin{pmatrix} a & -b & -c & -d \\ b & a & -d & c \\ c & d & a & -b \\ d & -c & b & a \end{pmatrix} = & \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} a + \begin{pmatrix} 0 & -1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & -1 \\ 0 & 0 & 1 & 0 \end{pmatrix} b \\
& + \begin{pmatrix} 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \end{pmatrix} c + \begin{pmatrix} 0 & 0 & 0 & -1 \\ 0 & 0 & -1 & 0 \\ 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{pmatrix} d \\
& \leftrightarrow 1a + \mathrm{i}b + \mathrm{j}c + \mathrm{k}d
\end{aligned}
$$

$$
\mathrm{ij} \leftrightarrow \begin{pmatrix} 0 & -1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & -1 \\ 0 & 0 & 1 & 0 \end{pmatrix} \begin{pmatrix} 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \end{pmatrix} = \begin{pmatrix} 0 & 0 & 0 & -1 \\ 0 & 0 & -1 & 0 \\ 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{pmatrix} \leftrightarrow \mathrm{k}
$$

$$
\mathrm{ji} \leftrightarrow \begin{pmatrix} 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \end{pmatrix} \begin{pmatrix} 0 & -1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & -1 \\ 0 & 0 & 1 & 0 \end{pmatrix} = - \begin{pmatrix} 0 & 0 & 0 & -1 \\ 0 & 0 & -1 & 0 \\ 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{pmatrix} \leftrightarrow -\mathrm{k}
$$

$$\mathrm{ki} \leftrightarrow \begin{pmatrix} 0 & 0 & 0 & -1 \\ 0 & 0 & -1 & 0 \\ 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} 0 & -1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & -1 \\ 0 & 0 & 1 & 0 \end{pmatrix} = \begin{pmatrix} 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \end{pmatrix} \leftrightarrow \mathrm{j}$$

$$Q_l = \begin{pmatrix} a & -b & -c & -d \\ b & a & -d & c \\ c & d & a & -b \\ d & -c & b & a \end{pmatrix} = \begin{pmatrix} a & -\begin{pmatrix} b & c & d \end{pmatrix} \\ \begin{pmatrix} b \\ c \\ d \end{pmatrix} & \begin{pmatrix} a & -d & c \\ d & a & -b \\ -c & b & a \end{pmatrix} \end{pmatrix} = \begin{pmatrix} q_0 & -\boldsymbol{q}^\mathsf{T} \\ \boldsymbol{q} & Q_i \end{pmatrix}, \begin{cases} a = q_0 \\ \boldsymbol{q} = \begin{pmatrix} b \\ c \\ d \end{pmatrix} \\ Q_i = \begin{pmatrix} a & -d & c \\ d & a & -b \\ -c & b & a \end{pmatrix} \end{cases}$$

Grassmann product

$$qv \leftrightarrow \begin{pmatrix} a & -b & -c & -d \\ b & a & -d & c \\ c & d & a & -b \\ d & -c & b & a \end{pmatrix} \begin{pmatrix} v_0 \\ v_1 \\ v_2 \\ v_3 \end{pmatrix} = Q_l v = \begin{pmatrix} q_0 & -\boldsymbol{q}^\mathsf{T} \\ \boldsymbol{q} & Q_i \end{pmatrix} \begin{pmatrix} v_0 \\ \boldsymbol{v} \end{pmatrix} = \begin{pmatrix} q_0 v_0 - \boldsymbol{q}^\mathsf{T} \boldsymbol{v} \\ \boldsymbol{q} v_0 + Q_i \boldsymbol{v} \end{pmatrix} = \begin{pmatrix} q_0 v_0 - \boldsymbol{q} \cdot \boldsymbol{v} \\ \boldsymbol{q} v_0 + q_0 \boldsymbol{v} + \boldsymbol{q} \times \boldsymbol{v} \end{pmatrix}$$

$$vq \leftrightarrow \begin{pmatrix} a & -b & -c & -d \\ b & a & d & -c \\ c & -d & a & b \\ d & c & -b & a \end{pmatrix} \begin{pmatrix} v_0 \\ v_1 \\ v_2 \\ v_3 \end{pmatrix} = Q_r v$$

$$\begin{aligned} R(q) = Q_r = \begin{pmatrix} a & -b & -c & -d \\ b & a & d & -c \\ c & -d & a & b \\ d & c & -b & a \end{pmatrix} =& a \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} + b \begin{pmatrix} 0 & -1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & -1 & 0 \end{pmatrix} \\ &+ c \begin{pmatrix} 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{pmatrix} + d \begin{pmatrix} 0 & 0 & 0 & -1 \\ 0 & 0 & 1 & 0 \\ 0 & -1 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{pmatrix} \\ \leftrightarrow& a1 + b\mathrm{i} + c\mathrm{j} + d\mathrm{k} \end{aligned}$$

$$\mathrm{ij} \leftrightarrow \begin{pmatrix} 0 & -1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & -1 & 0 \end{pmatrix} \begin{pmatrix} 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{pmatrix} = - \begin{pmatrix} 0 & 0 & 0 & -1 \\ 0 & 0 & 1 & 0 \\ 0 & -1 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{pmatrix} \leftrightarrow -\mathrm{k}$$

$$\mathrm{ji} \leftrightarrow \begin{pmatrix} 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{pmatrix} \begin{pmatrix} 0 & -1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & -1 & 0 \end{pmatrix} = \begin{pmatrix} 0 & 0 & 0 & -1 \\ 0 & 0 & 1 & 0 \\ 0 & -1 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{pmatrix} \leftrightarrow \mathrm{k}$$

$$\mathrm{ki} \leftrightarrow \begin{pmatrix} 0 & 0 & 0 & -1 \\ 0 & 0 & 1 & 0 \\ 0 & -1 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} 0 & -1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & -1 & 0 \end{pmatrix} = - \begin{pmatrix} 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{pmatrix} \leftrightarrow -\mathrm{j}$$

$$\begin{aligned} R(q) = Q_r = \begin{pmatrix} a & -b & -c & -d \\ b & a & d & -c \\ c & -d & a & b \\ d & c & -b & a \end{pmatrix} &\leftrightarrow \begin{pmatrix} a+bi & -c+di \\ c+di & a-bi \end{pmatrix} \\ &= \begin{pmatrix} \alpha & -\overline{\beta} \\ \beta & \overline{\alpha} \end{pmatrix}, \begin{cases} \alpha = a+bi & \overline{\alpha} = a-bi \\ \beta = c+di & \overline{\beta} = c-di \end{cases} \end{aligned}$$

$$Q_r=\begin{pmatrix}a&-b&-c&-d\\b&a&d&-c\\c&-d&a&b\\d&c&-b&a\end{pmatrix}=\begin{pmatrix}a&-\begin{pmatrix}b&c&d\end{pmatrix}\\\begin{pmatrix}b\\c\\d\end{pmatrix}&\begin{pmatrix}a&d&-c\\-d&a&b\\c&-b&a\end{pmatrix}\end{pmatrix}=\begin{pmatrix}q_0&-\boldsymbol{q}^{\mathsf{T}}\\\boldsymbol{q}&Q_i^{\mathsf{T}}\end{pmatrix},\begin{cases}a=q_0\\\boldsymbol{q}=\begin{pmatrix}b\\c\\d\end{pmatrix}\\Q_i=\begin{pmatrix}a&-d&c\\d&a&-b\\-c&b&a\end{pmatrix}\end{cases}$$

$$vq\leftrightarrow\begin{pmatrix}a&-b&-c&-d\\b&a&d&-c\\c&-d&a&b\\d&c&-b&a\end{pmatrix}\begin{pmatrix}v_0\\v_1\\v_2\\v_3\end{pmatrix}=Q_rv=\begin{pmatrix}q_0&-\boldsymbol{q}^{\mathsf{T}}\\\boldsymbol{q}&Q_i^{\mathsf{T}}\end{pmatrix}\begin{pmatrix}v_0\\\boldsymbol{v}\end{pmatrix}=\begin{pmatrix}v_0q_0-\boldsymbol{q}^{\mathsf{T}}\boldsymbol{v}\\v_0\boldsymbol{q}+Q_i^{\mathsf{T}}\boldsymbol{v}\end{pmatrix}=\begin{pmatrix}v_0q_0-\boldsymbol{v}\cdot\boldsymbol{q}\\v_0\boldsymbol{q}+q_0\boldsymbol{v}+\boldsymbol{v}\times\boldsymbol{q}\end{pmatrix}=\begin{pmatrix}q_0v_0-\boldsymbol{q}\cdot\boldsymbol{v}\\\boldsymbol{q}v_0+q_0\boldsymbol{v}-\boldsymbol{q}\times\boldsymbol{v}\end{pmatrix}$$

$$\begin{cases}qv\leftrightarrow Q_lv=\begin{pmatrix}q_0&-\boldsymbol{q}^{\mathsf{T}}\\\boldsymbol{q}&Q_i\end{pmatrix}\begin{pmatrix}v_0\\\boldsymbol{v}\end{pmatrix}=\begin{pmatrix}q_0v_0-\boldsymbol{q}^{\mathsf{T}}\boldsymbol{v}\\\boldsymbol{q}v_0+Q_i\boldsymbol{v}\end{pmatrix}=\begin{pmatrix}q_0v_0-\boldsymbol{q}\cdot\boldsymbol{v}\\\boldsymbol{q}v_0+q_0\boldsymbol{v}+\boldsymbol{q}\times\boldsymbol{v}\end{pmatrix}=\begin{pmatrix}q_0v_0-\boldsymbol{q}\cdot\boldsymbol{v}\\\boldsymbol{q}v_0+q_0\boldsymbol{v}+\boldsymbol{q}\times\boldsymbol{v}\end{pmatrix}\\vq\leftrightarrow Q_rv=\begin{pmatrix}q_0&-\boldsymbol{q}^{\mathsf{T}}\\\boldsymbol{q}&Q_i^{\mathsf{T}}\end{pmatrix}\begin{pmatrix}v_0\\\boldsymbol{v}\end{pmatrix}=\begin{pmatrix}v_0q_0-\boldsymbol{q}^{\mathsf{T}}\boldsymbol{v}\\v_0\boldsymbol{q}+Q_i^{\mathsf{T}}\boldsymbol{v}\end{pmatrix}=\begin{pmatrix}v_0q_0-\boldsymbol{v}\cdot\boldsymbol{q}\\v_0\boldsymbol{q}+\boldsymbol{v}q_0+\boldsymbol{v}\times\boldsymbol{q}\end{pmatrix}=\begin{pmatrix}q_0v_0-\boldsymbol{q}\cdot\boldsymbol{v}\\\boldsymbol{q}v_0+q_0\boldsymbol{v}-\boldsymbol{q}\times\boldsymbol{v}\end{pmatrix}\end{cases}$$

$$\begin{cases}qv=\begin{pmatrix}q_0\\\boldsymbol{q}\end{pmatrix}\begin{pmatrix}v_0\\\boldsymbol{v}\end{pmatrix}=L(q)\,v=Q_lv=\begin{pmatrix}q_0&-\boldsymbol{q}^{\mathsf{T}}\\\boldsymbol{q}&Q_i\end{pmatrix}\begin{pmatrix}v_0\\\boldsymbol{v}\end{pmatrix}\\vq=\begin{pmatrix}v_0\\\boldsymbol{v}\end{pmatrix}\begin{pmatrix}q_0\\\boldsymbol{q}\end{pmatrix}=R(q)\,v=Q_rv=\begin{pmatrix}q_0&-\boldsymbol{q}^{\mathsf{T}}\\\boldsymbol{q}&Q_i^{\mathsf{T}}\end{pmatrix}\begin{pmatrix}v_0\\\boldsymbol{v}\end{pmatrix}\end{cases}$$

$$\begin{cases}\boldsymbol{v}\xrightarrow{\text{rotate about }\boldsymbol{n}}\boldsymbol{v}'\\\begin{cases}\boldsymbol{v}=\boldsymbol{v}_{\parallel\boldsymbol{n}}+\boldsymbol{v}_{\perp\boldsymbol{n}}=\boldsymbol{v}_{\parallel}+\boldsymbol{v}_{\perp}=&\Rightarrow\boldsymbol{v}_{\perp}=\boldsymbol{v}-\boldsymbol{v}_{\parallel}\\\boldsymbol{v}'=\boldsymbol{v}'_{\parallel\boldsymbol{n}}+\boldsymbol{v}'_{\perp\boldsymbol{n}}=\boldsymbol{v}'_{\parallel}+\boldsymbol{v}'_{\perp}&\Rightarrow\boldsymbol{v}'_{\perp}=\boldsymbol{v}'-\boldsymbol{v}'_{\parallel}\end{cases}\\\begin{cases}\boldsymbol{u}=\boldsymbol{n}\times\boldsymbol{v}_{\perp}\\\|\boldsymbol{u}\|=\|\boldsymbol{n}\times\boldsymbol{v}_{\perp}\|=\|\boldsymbol{n}\|\,\|\boldsymbol{v}_{\perp}\|\sin\dfrac{\pi}{2}=\|\boldsymbol{v}_{\perp}\|\end{cases}\quad\begin{cases}\boldsymbol{u}\perp\boldsymbol{n}\\\boldsymbol{u}\perp\boldsymbol{v}_{\perp}\end{cases}\begin{cases}\boldsymbol{n}\perp\boldsymbol{v}_{\perp}\\\|\boldsymbol{n}\|=1\end{cases}\\\boldsymbol{v}'_{\parallel}=\boldsymbol{v}_{\parallel}\\\boldsymbol{v}'_{\perp}=\cos\theta\boldsymbol{v}_{\perp}+\sin\theta\,(\boldsymbol{n}\times\boldsymbol{v}_{\perp})\,,\theta=\angle\boldsymbol{v}'_{\perp}\boldsymbol{v}_{\perp}\end{cases}$$

$$\leftrightarrow\begin{cases}v\xrightarrow{\text{rotate about }n}v'\\\begin{cases}v=\begin{pmatrix}0\\\boldsymbol{v}\end{pmatrix}=\begin{pmatrix}0\\\boldsymbol{v}_{\parallel\boldsymbol{n}}\end{pmatrix}+\begin{pmatrix}0\\\boldsymbol{v}_{\perp\boldsymbol{n}}\end{pmatrix}=\begin{pmatrix}0\\\boldsymbol{v}_{\parallel}\end{pmatrix}+\begin{pmatrix}0\\\boldsymbol{v}_{\perp}\end{pmatrix}=v_{\parallel}+v_{\perp}\quad\Rightarrow v_{\perp}=v-v_{\parallel}\end{cases}\\\begin{cases}v'=\begin{pmatrix}0\\\boldsymbol{v}'\end{pmatrix}=\begin{pmatrix}0\\\boldsymbol{v}'_{\parallel\boldsymbol{n}}\end{pmatrix}+\begin{pmatrix}0\\\boldsymbol{v}'_{\perp\boldsymbol{n}}\end{pmatrix}=\begin{pmatrix}0\\\boldsymbol{v}'_{\parallel}\end{pmatrix}+\begin{pmatrix}0\\\boldsymbol{v}'_{\perp}\end{pmatrix}=v'_{\parallel}+v'_{\perp}\quad\Rightarrow v'_{\perp}=v'-v'_{\parallel}\end{cases}\\\begin{cases}u=\begin{pmatrix}0\\\boldsymbol{u}\end{pmatrix}=\begin{pmatrix}0\\\boldsymbol{n}\times\boldsymbol{v}_{\perp}\end{pmatrix}=\begin{pmatrix}00-0\\\boldsymbol{n}0+0\boldsymbol{v}_{\perp}+\boldsymbol{n}\times\boldsymbol{v}_{\perp}\end{pmatrix}=\begin{pmatrix}n_0v_{\perp0}-\boldsymbol{n}\cdot\boldsymbol{v}_{\perp}\\\boldsymbol{n}v_{\perp0}+n_0\boldsymbol{v}_{\perp}+\boldsymbol{n}\times\boldsymbol{v}_{\perp}\end{pmatrix}=nv_{\perp}\\\|\boldsymbol{u}\|=\|\boldsymbol{n}\times\boldsymbol{v}_{\perp}\|=\|\boldsymbol{n}\|\,\|\boldsymbol{v}_{\perp}\|\sin\dfrac{\pi}{2}=\|\boldsymbol{v}_{\perp}\|\end{cases}\quad\begin{cases}\boldsymbol{u}\perp\boldsymbol{n}\\\boldsymbol{u}\perp\boldsymbol{v}_{\perp}\end{cases}\begin{cases}\boldsymbol{n}\perp\boldsymbol{v}_{\perp}\\\|\boldsymbol{n}\|=1\end{cases}\\v'_{\parallel}=\begin{pmatrix}0\\\boldsymbol{v}'_{\parallel}\end{pmatrix}=\begin{pmatrix}0\\\boldsymbol{v}_{\parallel}\end{pmatrix}=v_{\parallel}\\v'_{\perp}=\begin{pmatrix}0\\\boldsymbol{v}'_{\perp}\end{pmatrix}=\cos\theta\begin{pmatrix}0\\\boldsymbol{v}_{\perp}\end{pmatrix}+\sin\theta\begin{pmatrix}0\\\boldsymbol{n}\times\boldsymbol{v}_{\perp}\end{pmatrix}=\cos\theta v_{\perp}+\sin\theta nv_{\perp}=(\cos\theta+\sin\theta n)\,v_{\perp}\end{cases}$$

$$v'_{\perp}=(\cos\theta+\sin\theta n)\,v_{\perp}$$

$$v'_{\perp}=pv_{\perp},p=\cos\theta+\sin\theta n=\cos\theta\begin{pmatrix}1\\\boldsymbol{0}\end{pmatrix}+\sin\theta\begin{pmatrix}0\\\boldsymbol{n}\end{pmatrix}=\begin{pmatrix}\cos\theta\\\sin\theta\boldsymbol{n}\end{pmatrix}$$

$$p = \begin{pmatrix} p_0 \\ \boldsymbol{p} \end{pmatrix} = \begin{pmatrix} \cos\theta \\ \sin\theta\boldsymbol{n} \end{pmatrix} = \cos\theta + \sin\theta n, \|\boldsymbol{n}\| = 1$$

$$|p| = \left\| \begin{pmatrix} \cos\theta \\ \sin\theta\boldsymbol{n} \end{pmatrix} \right\| = \sqrt{\begin{pmatrix} \cos\theta \\ \sin\theta\boldsymbol{n} \end{pmatrix}^{\mathsf{T}} \begin{pmatrix} \cos\theta \\ \sin\theta\boldsymbol{n} \end{pmatrix}} = \sqrt{\cos^2\theta + \sin^2\theta\,\|\boldsymbol{n}\|^2} = \sqrt{\cos^2\theta + \sin^2\theta} = 1$$

$$|p| = \left\| \begin{pmatrix} \cos\theta \\ \sin\theta\boldsymbol{n} \end{pmatrix} \right\| = 1$$

$$p^* = \overline{p} = \begin{pmatrix} p_0 \\ -\boldsymbol{p} \end{pmatrix} = \begin{pmatrix} \cos\theta \\ -\sin\theta\boldsymbol{n} \end{pmatrix} = \cos\theta - \sin\theta n$$

$$\begin{aligned}
p^*p = \overline{p}p &= \begin{pmatrix} p_0 \\ -\boldsymbol{p} \end{pmatrix} \begin{pmatrix} p_0 \\ \boldsymbol{p} \end{pmatrix} = \begin{pmatrix} p_0p_0 - (-\boldsymbol{p}) \cdot \boldsymbol{p} \\ -\boldsymbol{p}p_0 + p_0\boldsymbol{p} + (-\boldsymbol{p}) \times \boldsymbol{p} \end{pmatrix} \\
&= \begin{pmatrix} p_0p_0 + \boldsymbol{p} \cdot \boldsymbol{p} \\ \boldsymbol{0} + \boldsymbol{0} \end{pmatrix} = p_0p_0 + \boldsymbol{p} \cdot \boldsymbol{p} = p_0^2 + \|\boldsymbol{p}\|^2 = |p|^2 = 1^2 = 1 \\
&= \begin{pmatrix} p_0p_0 - \boldsymbol{p} \cdot (-\boldsymbol{p}) \\ \boldsymbol{p}p_0 + p_0(-\boldsymbol{p}) + \boldsymbol{p} \times (-\boldsymbol{p}) \end{pmatrix} = p\overline{p} = pp^*
\end{aligned}$$

$$p^*p = \overline{p}p = 1 = p\overline{p} = pp^* = |p|^2 = p_0^2 + \|\boldsymbol{p}\|^2 \Rightarrow p^* = \frac{1}{p} = p^{-1}$$

$$\begin{aligned}
p^2 = pp &= \begin{pmatrix} p_0p_0 - \boldsymbol{p} \cdot \boldsymbol{p} \\ \boldsymbol{p}p_0 + p_0\boldsymbol{p} + \boldsymbol{p} \times \boldsymbol{p} \end{pmatrix} = \begin{pmatrix} \cos^2\theta - \sin^2\theta\,\|\boldsymbol{n}\|^2 \\ 2(\sin\theta\boldsymbol{n})\cos\theta + \boldsymbol{0} \end{pmatrix} \\
&= \begin{pmatrix} \cos^2\theta - \sin^2\theta \\ 2\sin\theta\cos\theta\boldsymbol{n} \end{pmatrix} = \begin{pmatrix} \cos 2\theta \\ \sin 2\theta\boldsymbol{n} \end{pmatrix} \\
\begin{pmatrix} \cos\theta \\ \sin\theta\boldsymbol{n} \end{pmatrix}^2 &= \begin{pmatrix} \cos 2\theta \\ \sin 2\theta\boldsymbol{n} \end{pmatrix} \\
q^2 = \begin{pmatrix} \cos\frac{\theta}{2} \\ \sin\frac{\theta}{2}\boldsymbol{n} \end{pmatrix}^2 &= \begin{pmatrix} \cos\theta \\ \sin\theta\boldsymbol{n} \end{pmatrix} = p
\end{aligned}$$

$qv_\parallel = v_\parallel q$

$$\begin{aligned}
\begin{pmatrix} \alpha \\ \beta\boldsymbol{n} \end{pmatrix} v_{\parallel\boldsymbol{n}} = \begin{pmatrix} \alpha \\ \beta\boldsymbol{n} \end{pmatrix} v_\parallel &= \begin{pmatrix} \alpha \\ \beta\boldsymbol{n} \end{pmatrix} \begin{pmatrix} 0 \\ \boldsymbol{v}_\parallel \end{pmatrix} \\
&= \begin{pmatrix} \alpha 0 - \beta\boldsymbol{n} \cdot \boldsymbol{v}_\parallel \\ \beta\boldsymbol{n}0 + \alpha\boldsymbol{v}_\parallel + \beta\boldsymbol{n} \times \boldsymbol{v}_\parallel \end{pmatrix} = \begin{pmatrix} -\beta\boldsymbol{n} \cdot \boldsymbol{v}_\parallel \\ \alpha\boldsymbol{v}_\parallel + \boldsymbol{0} \end{pmatrix} = \begin{pmatrix} -\beta\boldsymbol{n} \cdot \boldsymbol{v}_\parallel \\ \alpha\boldsymbol{v}_\parallel \end{pmatrix} \\
&= \begin{pmatrix} 0\alpha - \boldsymbol{v}_\parallel \cdot \beta\boldsymbol{n} \\ \boldsymbol{v}_\parallel\alpha + 0\beta\boldsymbol{n} + \boldsymbol{v}_\parallel \times \beta\boldsymbol{n} \end{pmatrix} = \begin{pmatrix} -\boldsymbol{v}_\parallel \cdot \beta\boldsymbol{n} \\ \alpha\boldsymbol{v}_\parallel + \boldsymbol{0} \end{pmatrix} = \begin{pmatrix} -\beta\boldsymbol{n} \cdot \boldsymbol{v}_\parallel \\ \alpha\boldsymbol{v}_\parallel \end{pmatrix} \\
&= \begin{pmatrix} 0 \\ \boldsymbol{v}_\parallel \end{pmatrix} \begin{pmatrix} \alpha \\ \beta\boldsymbol{n} \end{pmatrix} = v_\parallel \begin{pmatrix} \alpha \\ \beta\boldsymbol{n} \end{pmatrix} = v_{\parallel\boldsymbol{n}} \begin{pmatrix} \alpha \\ \beta\boldsymbol{n} \end{pmatrix} \\
\begin{pmatrix} \alpha \\ \beta\boldsymbol{n} \end{pmatrix} v_\parallel &= v_\parallel \begin{pmatrix} \alpha \\ \beta\boldsymbol{n} \end{pmatrix}
\end{aligned}$$

$qv_\perp = v_\perp q^*$

$$\begin{aligned}
\begin{pmatrix} \alpha \\ \beta\boldsymbol{n} \end{pmatrix} v_{\perp\boldsymbol{n}} = \begin{pmatrix} \alpha \\ \beta\boldsymbol{n} \end{pmatrix} v_\perp &= \begin{pmatrix} \alpha \\ \beta\boldsymbol{n} \end{pmatrix} \begin{pmatrix} 0 \\ \boldsymbol{v}_\perp \end{pmatrix} \\
&= \begin{pmatrix} \alpha 0 - \beta\boldsymbol{n} \cdot \boldsymbol{v}_\perp \\ \beta\boldsymbol{n}0 + \alpha\boldsymbol{v}_\perp + \beta\boldsymbol{n} \times \boldsymbol{v}_\perp \end{pmatrix} = \begin{pmatrix} 0 - 0 \\ \alpha\boldsymbol{v}_\perp + \beta\boldsymbol{n} \times \boldsymbol{v}_\perp \end{pmatrix} = \begin{pmatrix} 0 \\ \alpha\boldsymbol{v}_\perp + \boldsymbol{v}_\perp \times (-\beta\boldsymbol{n}) \end{pmatrix} \\
&= \begin{pmatrix} 0\alpha - \boldsymbol{v}_\perp \cdot (-\beta\boldsymbol{n}) \\ \boldsymbol{v}_\perp\alpha + 0(-\beta\boldsymbol{n}) + \boldsymbol{v}_\perp \times (-\beta\boldsymbol{n}) \end{pmatrix} = \begin{pmatrix} 0 - 0 \\ \alpha\boldsymbol{v}_\perp + \boldsymbol{v}_\perp \times (-\beta\boldsymbol{n}) \end{pmatrix} = \begin{pmatrix} 0 \\ \alpha\boldsymbol{v}_\perp + \beta\boldsymbol{n} \times \boldsymbol{v}_\perp \end{pmatrix} \\
&= \begin{pmatrix} 0 \\ \boldsymbol{v}_\perp \end{pmatrix} \begin{pmatrix} \alpha \\ -\beta\boldsymbol{n} \end{pmatrix} = v_\perp \begin{pmatrix} \alpha \\ -\beta\boldsymbol{n} \end{pmatrix} = v_{\perp\boldsymbol{n}} \begin{pmatrix} \alpha \\ -\beta\boldsymbol{n} \end{pmatrix} \\
\begin{pmatrix} \alpha \\ \beta\boldsymbol{n} \end{pmatrix} v_\perp &= v_\perp \begin{pmatrix} \alpha \\ -\beta\boldsymbol{n} \end{pmatrix} = v_\perp \begin{pmatrix} \alpha \\ \beta\boldsymbol{n} \end{pmatrix}^* = v_\perp \overline{\begin{pmatrix} \alpha \\ \beta\boldsymbol{n} \end{pmatrix}}
\end{aligned}$$

$$q=\cos\frac{\theta}{2}+\sin\frac{\theta}{2}n\in\left\{p\middle|\begin{cases}p=\cos\theta+\sin\theta n\\ n=\begin{pmatrix}0\\ \boldsymbol{n}\end{pmatrix},\|\boldsymbol{n}\|=1\end{cases}\right\}\subset\left\{\begin{pmatrix}\alpha\\ \beta\boldsymbol{n}\end{pmatrix}\middle|\begin{cases}\alpha,\beta\in\mathbb{R}\\ \|\boldsymbol{n}\|=1\end{cases}\right\}$$

$$\Rightarrow\begin{cases}q^*q=\overline{q}q=1=q\overline{q}=qq^*=|q|^2=q_0^2+\|\boldsymbol{q}\|^2\Rightarrow q^*=\dfrac{1}{q}=q^{-1} & (u)\,\text{unit quaternion}\\ qv_\parallel=v_\parallel q & (c)\,\text{commutativity}\\ qv_\perp=v_\perp q^*=v_\perp\overline{q}\Leftrightarrow\begin{pmatrix}\cos\frac{\theta}{2}\\ \sin\frac{\theta}{2}\boldsymbol{n}\end{pmatrix}v_\perp=v_\perp\begin{pmatrix}\cos\frac{\theta}{2}\\ -\sin\frac{\theta}{2}\boldsymbol{n}\end{pmatrix} & (a)\,\text{anticommutativity}\end{cases}$$

$$\begin{aligned}v'=&v'_\parallel+v'_\perp,\begin{cases}v'_\parallel=v_\parallel & q=\cos\frac{\theta}{2}+\sin\frac{\theta}{2}n,|n|=1\\ v'_\perp=pv_\perp=q^2v_\perp & p=q^2=\cos\theta+\sin\theta n\end{cases}\\ =&v_\parallel+q^2v_\perp\overset{(u)}{=}1v_\parallel+qqv_\perp\\ \overset{(a)}{=}&qq^*v_\parallel+qv_\perp q^*\overset{(c)}{=}qv_\parallel q^*+qv_\perp q^*\\ =&q\left(v_\parallel+v_\perp\right)q^*=q\left(v\right)q^*=qvq^*\end{aligned}$$

$$\begin{aligned}v'=&qvq^{-1},q=\cos\frac{\theta}{2}+\sin\frac{\theta}{2}n,|n|=1\\ =&qvq^*=q\left(v_\parallel+v_\perp\right)q^*\\ =&qv_\parallel q^*+qv_\perp q^*=qq^*v_\parallel+qqv_\perp=1v_\parallel+q^2v_\perp\\ =&v_\parallel+pv_\perp,p=q^2=\cos\theta+\sin\theta n\end{aligned}$$

$$\begin{aligned}v'=qvq^*=q\left(vq^*\right)\leftrightarrow&L\left(q\right)R\left(q^*\right)v=Q_lQ_rv=\begin{pmatrix}q_0&-\boldsymbol{q}^\intercal\\ \boldsymbol{q}&Q_i\end{pmatrix}\begin{pmatrix}q_0&-\boldsymbol{q}^\intercal\\ \boldsymbol{q}&Q_i^\intercal\end{pmatrix}\begin{pmatrix}v_0\\ \boldsymbol{v}\end{pmatrix}\\ =\left(qv\right)q^*\leftrightarrow&R\left(q^*\right)L\left(q\right)v=Q_rQ_lv=\begin{pmatrix}q_0&-\boldsymbol{q}^\intercal\\ \boldsymbol{q}&Q_i^\intercal\end{pmatrix}\begin{pmatrix}q_0&-\boldsymbol{q}^\intercal\\ \boldsymbol{q}&Q_i\end{pmatrix}\begin{pmatrix}v_0\\ \boldsymbol{v}\end{pmatrix}\end{aligned}$$

$$\begin{pmatrix}q_0&-\boldsymbol{q}^\intercal\\ \boldsymbol{q}&Q_i\end{pmatrix}\begin{pmatrix}q_0&-\boldsymbol{q}^\intercal\\ \boldsymbol{q}&Q_i^\intercal\end{pmatrix}=\begin{pmatrix}q_0^2-\boldsymbol{q}^\intercal\boldsymbol{q}&-q_0\boldsymbol{q}^\intercal-\boldsymbol{q}^\intercal Q_i^\intercal\\ \boldsymbol{q}q_0+Q_i\boldsymbol{q}&-\boldsymbol{q}\boldsymbol{q}^\intercal+Q_iQ_i^\intercal\end{pmatrix}=L\left(q\right)R\left(q^*\right)$$

$$\begin{pmatrix}q_0&-\boldsymbol{q}^\intercal\\ \boldsymbol{q}&Q_i^\intercal\end{pmatrix}\begin{pmatrix}q_0&-\boldsymbol{q}^\intercal\\ \boldsymbol{q}&Q_i\end{pmatrix}=\begin{pmatrix}q_0^2-\boldsymbol{q}^\intercal\boldsymbol{q}&-q_0\boldsymbol{q}^\intercal-\boldsymbol{q}^\intercal Q_i\\ \boldsymbol{q}q_0+Q_i^\intercal\boldsymbol{q}&-\boldsymbol{q}\boldsymbol{q}^\intercal+Q_i^\intercal Q_i\end{pmatrix}=R\left(q^*\right)L\left(q\right)$$

$$q=a+b\mathrm{i}+c\mathrm{j}+d\mathrm{k}=\cos\frac{\theta}{2}+\sin\frac{\theta}{2}n=\cos\frac{\theta}{2}+b\mathrm{i}+c\mathrm{j}+d\mathrm{k}$$

$$\Rightarrow Q_iQ_i^\intercal=\begin{pmatrix}a&-d&c\\ d&a&-b\\ -c&b&a\end{pmatrix}\begin{pmatrix}a&d&-c\\ -d&a&b\\ c&-b&a\end{pmatrix}=\begin{pmatrix}a^2+d^2+c^2&-cb&-db\\ -cb&d^2+a^2+b^2&-dc\\ -db&-dc&c^2+b^2+a^2\end{pmatrix}$$

$$=Q_i^\intercal Q_i=\begin{pmatrix}a&d&-c\\ -d&a&b\\ c&-b&a\end{pmatrix}\begin{pmatrix}a&-d&c\\ d&a&-b\\ -c&b&a\end{pmatrix}=\begin{pmatrix}a^2+d^2+c^2&-cb&-db\\ -cb&d^2+a^2+b^2&-dc\\ -db&-dc&c^2+b^2+a^2\end{pmatrix}$$

$$\boldsymbol{q}^\intercal Q_i^\intercal=\begin{pmatrix}b&c&d\end{pmatrix}\begin{pmatrix}a&d&-c\\ -d&a&b\\ c&-b&a\end{pmatrix}=\begin{pmatrix}ba&ca&da\end{pmatrix}$$

$$=\boldsymbol{q}^\intercal Q_i=\begin{pmatrix}b&c&d\end{pmatrix}\begin{pmatrix}a&-d&c\\ d&a&-b\\ -c&b&a\end{pmatrix}=\begin{pmatrix}ba&ca&da\end{pmatrix}$$

$$Q_i\boldsymbol{q}=\begin{pmatrix}a&-d&c\\ d&a&-b\\ -c&b&a\end{pmatrix}\begin{pmatrix}b\\ c\\ d\end{pmatrix}=\begin{pmatrix}ab\\ ac\\ ad\end{pmatrix}$$

$$=Q_i^\intercal\boldsymbol{q}=\begin{pmatrix}a&d&-c\\ -d&a&b\\ c&-b&a\end{pmatrix}\begin{pmatrix}b\\ c\\ d\end{pmatrix}=\begin{pmatrix}ab\\ ac\\ ad\end{pmatrix}$$

$$
\begin{aligned}
&\begin{pmatrix} q_0 & -\boldsymbol{q}^\intercal \\ \boldsymbol{q} & Q_i \end{pmatrix}\begin{pmatrix} q_0 & -\boldsymbol{q}^\intercal \\ \boldsymbol{q} & Q_i^\intercal \end{pmatrix} = \begin{pmatrix} q_0^2 - \boldsymbol{q}^\intercal \boldsymbol{q} & -q_0\boldsymbol{q}^\intercal - \boldsymbol{q}^\intercal Q_i^\intercal \\ \boldsymbol{q} q_0 + Q_i \boldsymbol{q} & -\boldsymbol{q}\boldsymbol{q}^\intercal + Q_i Q_i^\intercal \end{pmatrix} = L(q)\, R(q^*) \\
=&\begin{pmatrix} q_0 & -\boldsymbol{q}^\intercal \\ \boldsymbol{q} & Q_i^\intercal \end{pmatrix}\begin{pmatrix} q_0 & -\boldsymbol{q}^\intercal \\ \boldsymbol{q} & Q_i \end{pmatrix} = \begin{pmatrix} q_0^2 - \boldsymbol{q}^\intercal \boldsymbol{q} & -q_0\boldsymbol{q}^\intercal - \boldsymbol{q}^\intercal Q_i \\ \boldsymbol{q} q_0 + Q_i^\intercal \boldsymbol{q} & -\boldsymbol{q}\boldsymbol{q}^\intercal + Q_i^\intercal Q_i \end{pmatrix} = R(q^*)\, L(q) \\
&\Downarrow \\
L(q)\, R(q^*) =& R(q^*)\, L(q) \\
&\updownarrow \\
v' = qvq^* = q\,(vq^*) =& (qv)\, q^*
\end{aligned}
$$

$$a = \cos\frac{\theta}{2} \Rightarrow \theta = 2\arccos a = 2\cos^{-1} a$$

$$\sin\frac{\theta}{2} n = \sin\frac{2\arccos a}{2} n = \sin\left(\cos^{-1} a\right) n$$

$$n = \begin{pmatrix} 0 \\ \boldsymbol{n} \end{pmatrix} = \begin{pmatrix} 0 \\ \hat{\boldsymbol{n}} \end{pmatrix}, \|\boldsymbol{n}\| = 1, \boldsymbol{n} = \begin{pmatrix} \sin\gamma\cos\alpha \\ \sin\gamma\sin\alpha \\ \cos\gamma \end{pmatrix}, \begin{cases} \alpha = \angle \hat{\boldsymbol{n}}_{xy}\hat{\boldsymbol{x}} \\ \gamma = \angle \hat{\boldsymbol{n}}\hat{\boldsymbol{z}} \end{cases}$$

$$
\begin{aligned}
q =& a + b\mathrm{i} + c\mathrm{j} + d\mathrm{k} \\
=& \cos\frac{\theta}{2} + \sin\frac{\theta}{2} n, n = \begin{pmatrix} 0 \\ \boldsymbol{n} \end{pmatrix} = \begin{pmatrix} 0 \\ \hat{\boldsymbol{n}} \end{pmatrix} = \begin{pmatrix} 0 \\ \sin\gamma\cos\alpha \\ \sin\gamma\sin\alpha \\ \cos\gamma \end{pmatrix}, \|\boldsymbol{n}\| = 1, \begin{cases} \alpha = \angle \hat{\boldsymbol{n}}_{xy}\hat{\boldsymbol{x}} \\ \gamma = \angle \hat{\boldsymbol{n}}\hat{\boldsymbol{z}} \end{cases} \\
=& \cos\frac{\theta}{2} + \sin\frac{\theta}{2}\boldsymbol{n}, \boldsymbol{n} = (\sin\gamma\cos\alpha)\,\mathrm{i} + (\sin\gamma\sin\alpha)\,\mathrm{j} + (\cos\gamma)\,\mathrm{k} \\
=& \cos\frac{\theta}{2} + \sin\frac{\theta}{2}\left[(\sin\gamma\cos\alpha)\,\mathrm{i} + (\sin\gamma\sin\alpha)\,\mathrm{j} + (\cos\gamma)\,\mathrm{k}\right] \\
=& \cos\frac{\theta}{2} + \boldsymbol{n}\sin\frac{\theta}{2} \\
q^{-1} = q^* =& \cos\frac{\theta}{2} - \boldsymbol{n}\sin\frac{\theta}{2} \\
=& \cos\frac{\theta}{2} - \sin\frac{\theta}{2}\left[(\sin\gamma\cos\alpha)\,\mathrm{i} + (\sin\gamma\sin\alpha)\,\mathrm{j} + (\cos\gamma)\,\mathrm{k}\right] \\
\boldsymbol{v}' = v' =& qvq^{-1} = q\boldsymbol{v}q^{-1}, \begin{cases} v = \boldsymbol{v} = v_1\mathrm{i} + v_2\mathrm{j} + v_3\mathrm{k} \\ v' = \boldsymbol{v}' = v_1'\mathrm{i} + v_2'\mathrm{j} + v_3'\mathrm{k} \end{cases} \\
=& \left\{\cos\frac{\theta}{2} + \sin\frac{\theta}{2}\left[(\sin\gamma\cos\alpha)\,\mathrm{i} + (\sin\gamma\sin\alpha)\,\mathrm{j} + (\cos\gamma)\,\mathrm{k}\right]\right\}\boldsymbol{v} \\
& \left\{\cos\frac{\theta}{2} - \sin\frac{\theta}{2}\left[(\sin\gamma\cos\alpha)\,\mathrm{i} + (\sin\gamma\sin\alpha)\,\mathrm{j} + (\cos\gamma)\,\mathrm{k}\right]\right\} \\
=& q(\theta, \gamma, \alpha)\,\boldsymbol{v} q^{-1}(\theta, \gamma, \alpha) \\
=& \left\{\cos\frac{\theta}{2} + \sin\frac{\theta}{2}\left[n_x\mathrm{i} + n_y\mathrm{j} + n_z\mathrm{k}\right]\right\}\boldsymbol{v}\left\{\cos\frac{\theta}{2} - \sin\frac{\theta}{2}\left[n_x\mathrm{i} + n_y\mathrm{j} + n_z\mathrm{k}\right]\right\} \\
=& \left\{\cos\frac{\theta}{2} + \sin\frac{\theta}{2}\left[n_x\mathrm{i} \pm \sqrt{1 - n_x^2 - n_z^2}\,\mathrm{j} + n_z\mathrm{k}\right]\right\}\boldsymbol{v} \\
& \left\{\cos\frac{\theta}{2} - \sin\frac{\theta}{2}\left[n_x\mathrm{i} \pm \sqrt{1 - n_x^2 - n_z^2}\,\mathrm{j} + n_z\mathrm{k}\right]\right\} \\
=& q(\theta, n_z, n_x)\,\boldsymbol{v} q^{-1}(\theta, n_z, n_x)
\end{aligned}
$$

$$\boldsymbol{v}' = v' = qvq^{-1} = q\boldsymbol{v}q^{-1}, \begin{cases} v = \boldsymbol{v} = v_1\mathrm{i} + v_2\mathrm{j} + v_3\mathrm{k} \\ v' = \boldsymbol{v}' = v_1'\mathrm{i} + v_2'\mathrm{j} + v_3'\mathrm{k} \\ q = \cos\frac{\theta}{2} + \sin\frac{\theta}{2}\left[(\sin\gamma\cos\alpha)\,\mathrm{i} + (\sin\gamma\sin\alpha)\,\mathrm{j} + (\cos\gamma)\,\mathrm{k}\right] \\ q^{-1} = \cos\frac{\theta}{2} - \sin\frac{\theta}{2}\left[(\sin\gamma\cos\alpha)\,\mathrm{i} + (\sin\gamma\sin\alpha)\,\mathrm{j} + (\cos\gamma)\,\mathrm{k}\right] \end{cases}$$

34.2.3 matrix form

$$q \overset{|q|=1}{=} a + b\mathrm{i} + c\mathrm{j} + d\mathrm{k}, \begin{cases} a = \cos\frac{\theta}{2} \\ b = \sin\frac{\theta}{2}\sin\gamma\cos\alpha \\ c = \sin\frac{\theta}{2}\sin\gamma\sin\alpha \\ d = \sin\frac{\theta}{2}\cos\gamma \end{cases}$$

$$L(q) = Q_l = \begin{pmatrix} a & -b & -c & -d \\ b & a & -d & c \\ c & d & a & -b \\ d & -c & b & a \end{pmatrix}$$

$$R(q) = Q_r = \begin{pmatrix} a & -b & -c & -d \\ b & a & d & -c \\ c & -d & a & b \\ d & c & -b & a \end{pmatrix}$$

$$R(q^*) = \begin{pmatrix} a & b & c & d \\ -b & a & -d & c \\ -c & d & a & -b \\ -d & -c & b & a \end{pmatrix}$$

$$\begin{aligned} v' = qvq^* = (qv)\,q^* &\leftrightarrow R(q^*)\,L(q)\,v \\ = q\,(vq^*) &\leftrightarrow L(q)\,R(q^*)\,v \\ &= \begin{pmatrix} a & -b & -c & -d \\ b & a & -d & c \\ c & d & a & -b \\ d & -c & b & a \end{pmatrix} \begin{pmatrix} a & b & c & d \\ -b & a & -d & c \\ -c & d & a & -b \\ -d & -c & b & a \end{pmatrix} v \\ &= \begin{pmatrix} a^2+b^2+c^2+d^2 & 0 & 0 & 0 \\ 0 & b^2+a^2-d^2-c^2 & -2ad+2bc & 2ac+2bd \\ 0 & 2ad+2bc & c^2-d^2+a^2-b^2 & 2cd-2ab \\ 0 & -2ac+2bd & 2cd+2ab & d^2-c^2-b^2+a^2 \end{pmatrix} v \\ &= \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1-2\left(d^2+c^2\right) & 2(bc-ad) & 2(bd+ac) \\ 0 & 2(bc+ad) & 1-2\left(d^2+b^2\right) & 2(cd-ab) \\ 0 & 2(bd-ac) & 2(cd+ab) & 1-2\left(c^2+b^2\right) \end{pmatrix} v \\ \begin{pmatrix} 0 \\ \boldsymbol{v}' \end{pmatrix} &= \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1-2\left(d^2+c^2\right) & 2(bc-ad) & 2(bd+ac) \\ 0 & 2(bc+ad) & 1-2\left(d^2+b^2\right) & 2(cd-ab) \\ 0 & 2(bd-ac) & 2(cd+ab) & 1-2\left(c^2+b^2\right) \end{pmatrix} \begin{pmatrix} 0 \\ \boldsymbol{v} \end{pmatrix} \end{aligned}$$

$$\boldsymbol{v}' = \begin{pmatrix} 1-2\left(d^2+c^2\right) & 2(bc-ad) & 2(bd+ac) \\ 2(bc+ad) & 1-2\left(d^2+b^2\right) & 2(cd-ab) \\ 2(bd-ac) & 2(cd+ab) & 1-2\left(c^2+b^2\right) \end{pmatrix} \boldsymbol{v}$$

$$q \overset{|q|=1}{=} a+b\mathrm{i}+c\mathrm{j}+d\mathrm{k}, \begin{cases} a=\cos\frac{\theta}{2} \\ b=\sin\frac{\theta}{2}\sin\gamma\cos\alpha \\ c=\sin\frac{\theta}{2}\sin\gamma\sin\alpha \\ d=\sin\frac{\theta}{2}\cos\gamma \end{cases}$$

34.2.4 exponential form

$$\begin{aligned} q =& a+b\mathrm{i}+c\mathrm{j}+d\mathrm{k} = a+\mathrm{i}b+\mathrm{j}c+\mathrm{k}d \\ =& a+r\frac{\mathrm{i}b+\mathrm{j}c+\mathrm{k}d}{r}, r^2=b^2+c^2+d^2 \\ \overset{|q|=1}{=}& \cos\frac{\theta}{2}+\sin\frac{\theta}{2}\boldsymbol{n}, \begin{cases} a=\cos\frac{\theta}{2} & \cos\frac{\theta}{2}=a \\ r=\sin\frac{\theta}{2} & \sin\frac{\theta}{2}=r \\ \frac{\mathrm{i}b+\mathrm{j}c+\mathrm{k}d}{r}=\boldsymbol{n} & \boldsymbol{n}=\frac{\mathrm{i}b+\mathrm{j}c+\mathrm{k}d}{r} \end{cases} \end{aligned}$$

$$\begin{aligned} \|\boldsymbol{n}\|^2 = \left(\frac{\mathrm{i}b+\mathrm{j}c+\mathrm{k}d}{r}\right)^2 =& \frac{-b^2-c^2-d^2+bc(\mathrm{ij}+\mathrm{ji})+cd(\mathrm{jk}+\mathrm{kj})+db(\mathrm{ki}+\mathrm{ik})}{r^2} \\ =& \frac{-b^2-c^2-d^2+bc(\mathrm{k}-\mathrm{k})+cd(\mathrm{i}-\mathrm{i})+db(\mathrm{j}-\mathrm{j})}{r^2} \\ =& \frac{-b^2-c^2-d^2+bc0+cd0+db0}{r^2} = \frac{-b^2-c^2-d^2}{r^2} \\ =& \frac{-b^2-c^2-d^2}{b^2+c^2+d^2} = -1 \\ \boldsymbol{n} = \frac{\mathrm{i}b+\mathrm{j}c+\mathrm{k}d}{r} =& \pm\sqrt{-1} \end{aligned}$$

$$\begin{aligned} q =& a+b\mathrm{i}+c\mathrm{j}+d\mathrm{k} = a+\mathrm{i}b+\mathrm{j}c+\mathrm{k}d \\ =& a+r\frac{\mathrm{i}b+\mathrm{j}c+\mathrm{k}d}{r}, r^2=b^2+c^2+d^2 \\ =& \sqrt{a^2+r^2}\left(\frac{a}{\sqrt{a^2+r^2}}+\frac{\mathrm{i}b+\mathrm{j}c+\mathrm{k}d}{r}\frac{r}{\sqrt{a^2+r^2}}\right) \\ =& \rho\left(\cos\phi+\sqrt{-1}\sin\phi\right), \begin{cases} \rho=\sqrt{a^2+r^2} \\ \tan\phi=\frac{r}{a} \Leftrightarrow \phi=\arctan\frac{r}{a}=\frac{\theta}{2} \end{cases} \\ =& \rho e^{\phi\sqrt{-1}} \overset{|q|=1}{=} |q|\, e^{\boldsymbol{n}\frac{\theta}{2}} = e^{\boldsymbol{n}\frac{\theta}{2}} \\ q =& \rho e^{\phi\sqrt{-1}} \overset{|q|=1}{=} e^{\boldsymbol{n}\frac{\theta}{2}}, \begin{cases} q=a+\mathrm{i}b+\mathrm{j}c+\mathrm{k}d \\ \rho=\sqrt{a^2+r^2}=|q|=\sqrt{a^2+b^2+c^2+d^2}=1 \\ \tan\frac{\theta}{2}=\frac{r}{a} \Leftrightarrow \frac{\theta}{2}=\arctan\frac{r}{a}=\arctan\frac{\pm\sqrt{b^2+c^2+d^2}}{a} \end{cases} \end{aligned}$$

$$
\begin{aligned}
\boldsymbol{v}' = v' =& qvq^{-1} = q\boldsymbol{v}q^{-1}, \begin{cases} v = \boldsymbol{v} = v_1\mathrm{i} + v_2\mathrm{j} + v_3\mathrm{k} \\ v' = \boldsymbol{v}' = v_1'\mathrm{i} + v_2'\mathrm{j} + v_3'\mathrm{k} \end{cases}, \boldsymbol{n} = \begin{pmatrix} \sin\gamma\cos\alpha \\ \sin\gamma\sin\alpha \\ \cos\gamma \end{pmatrix}, \begin{cases} \alpha = \angle\hat{\boldsymbol{n}}_{xy}\hat{\boldsymbol{x}} \\ \gamma = \angle\hat{\boldsymbol{n}}\hat{\boldsymbol{z}} \end{cases} \\
=& \left\{\cos\frac{\theta}{2} + \sin\frac{\theta}{2}\left[(\sin\gamma\cos\alpha)\,\mathrm{i} + (\sin\gamma\sin\alpha)\,\mathrm{j} + (\cos\gamma)\,\mathrm{k}\right]\right\}\boldsymbol{v} \\
& \left\{\cos\frac{\theta}{2} - \sin\frac{\theta}{2}\left[(\sin\gamma\cos\alpha)\,\mathrm{i} + (\sin\gamma\sin\alpha)\,\mathrm{j} + (\cos\gamma)\,\mathrm{k}\right]\right\} \\
=& q(\theta,\gamma,\alpha)\,\boldsymbol{v}q^{-1}(\theta,\gamma,\alpha) = q\left(\frac{\theta}{2},\gamma,\alpha\right)\boldsymbol{v}q^{-1}\left(\frac{\theta}{2},\gamma,\alpha\right) \\
=& \left\{\cos\frac{\theta}{2} + \sin\frac{\theta}{2}\left[n_x\mathrm{i} + n_y\mathrm{j} + n_z\mathrm{k}\right]\right\}\boldsymbol{v}\left\{\cos\frac{\theta}{2} - \sin\frac{\theta}{2}\left[n_x\mathrm{i} + n_y\mathrm{j} + n_z\mathrm{k}\right]\right\} \\
=& \left\{\cos\frac{\theta}{2} + \sin\frac{\theta}{2}\left[n_x\mathrm{i} \pm \sqrt{1 - n_x^2 - n_z^2}\mathrm{j} + n_z\mathrm{k}\right]\right\}\boldsymbol{v} \\
& \left\{\cos\frac{\theta}{2} - \sin\frac{\theta}{2}\left[n_x\mathrm{i} \pm \sqrt{1 - n_x^2 - n_z^2}\mathrm{j} + n_z\mathrm{k}\right]\right\} \\
=& q(\theta,n_z,n_x)\,\boldsymbol{v}q^{-1}(\theta,n_z,n_x) = q\left(\frac{\theta}{2},n_z,n_x\right)\boldsymbol{v}q^{-1}\left(\frac{\theta}{2},n_z,n_x\right) \\
=& \mathrm{e}^{\boldsymbol{n}\frac{\theta}{2}}\boldsymbol{v}\mathrm{e}^{-\boldsymbol{n}\frac{\theta}{2}}
\end{aligned}
$$

$$\boldsymbol{v}' = \mathrm{e}^{\boldsymbol{n}\frac{\theta}{2}}\boldsymbol{v}\mathrm{e}^{-\boldsymbol{n}\frac{\theta}{2}}$$

---

general quaternion exponential form

$$
\begin{aligned}
q =& \rho\mathrm{e}^{\phi\sqrt{-1}} = |q|\,\mathrm{e}^{\boldsymbol{n}\phi}, \begin{cases} q = a + \mathrm{i}b + \mathrm{j}c + \mathrm{k}d = a + r\dfrac{\mathrm{i}b + \mathrm{j}c + \mathrm{k}d}{r} = a + r\boldsymbol{n} \\ \rho = \sqrt{a^2 + r^2} = |q| = \sqrt{a^2 + b^2 + c^2 + d^2} \\ \tan\phi = \dfrac{r}{a} \Leftrightarrow \phi = \arctan\dfrac{r}{a} = \arctan\dfrac{\pm\sqrt{b^2 + c^2 + d^2}}{a} \end{cases} \\
=& \mathrm{e}^{\ln|q|}\mathrm{e}^{\boldsymbol{n}\phi} = \mathrm{e}^{\ln|q| + \boldsymbol{n}\phi}
\end{aligned}
$$

34.2.5 double cover

https://www.bilibili.com/video/BV1rj41117VW/?t=15m15s

$$
\begin{array}{cccccccccccccccccc}
\theta & 0 & \to & \frac{\pi}{2} & \to & \pi & \to & \frac{3\pi}{2} & \to & 2\pi & \to & \frac{5\pi}{2} & \to & 3\pi & \to & \frac{7\pi}{2} & \to & 4\pi \\
\frac{\theta}{2} & 0 & \to & \frac{\pi}{4} & \to & \frac{\pi}{2} & \to & \frac{3\pi}{4} & \to & \pi & \to & \frac{5\pi}{4} & \to & \frac{3\pi}{2} & \to & \frac{7\pi}{4} & \to & 2\pi \\
\mathrm{e}^{\boldsymbol{n}\frac{\theta}{2}} & 1 & \to & \mathrm{e}^{\boldsymbol{n}\frac{\pi}{4}} & \to & \boldsymbol{n} & \to & \mathrm{e}^{\boldsymbol{n}\frac{3\pi}{4}} & \to & -1 & \to & \mathrm{e}^{\boldsymbol{n}\frac{5\pi}{4}} & \to & -\boldsymbol{n} & \to & \mathrm{e}^{\boldsymbol{n}\frac{7\pi}{4}} & \to & 1 \\
\mathrm{e}^{-\boldsymbol{n}\frac{\theta}{2}} & 1 & \to & \mathrm{e}^{\boldsymbol{n}\frac{-\pi}{4}} & \to & -\boldsymbol{n} & \to & \mathrm{e}^{\boldsymbol{n}\frac{-3\pi}{4}} & \to & -1 & \to & \mathrm{e}^{\boldsymbol{n}\frac{-5\pi}{4}} & \to & \boldsymbol{n} & \to & \mathrm{e}^{\boldsymbol{n}\frac{-7\pi}{4}} & \to & 1 \\
\boldsymbol{v}' = \mathrm{e}^{\boldsymbol{n}\frac{\theta}{2}}\boldsymbol{v}\mathrm{e}^{-\boldsymbol{n}\frac{\theta}{2}} & \boldsymbol{v} & \to & & \to & -\boldsymbol{n}\boldsymbol{v}\boldsymbol{n} & \to & & \to & \boldsymbol{v} & \to & & \to & -\boldsymbol{n}\boldsymbol{v}\boldsymbol{n} & \to & & \to & \boldsymbol{v}
\end{array}
$$

$$SU(2) = \left\{1(\cdot)1, \mathrm{e}^{\boldsymbol{n}\frac{\theta}{2}}(\cdot)\,\mathrm{e}^{-\boldsymbol{n}\frac{\theta}{2}}\right\}$$

3D rotation about an arbitrary axis[34.1.4.2.1.3]

$$\boldsymbol{V}' = SR_z(\gamma)\,S^{-1}\boldsymbol{V}$$

$$\boldsymbol{v}' = SR_z(\theta)\,S^{-1}\boldsymbol{v}$$

$$SO(3) = \left\{1, SR_z(\theta)\,S^{-1}\right\} = \left\{\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}, SR_z(\theta)\,S^{-1}\right\}$$

$$SU(2) \cong 2 \cdot SO(3)$$

Special unitary group of degree 2 $SU(2)$ is the boss of special orthonormal group of degree 3 $SO(3)$, $SU(2)$ double covers $SO(3)$.

https://en.wikipedia.org/wiki/Special_unitary_group#The_group_SU(2)

https://en.wikipedia.org/wiki/Unitary_group

"Particle is group representation, spin is square root of 4-vector."?

「粒子是群表示，自旋是四維向量開根號。」？

34.2.6 right-hand vs. left-hand coordinates

Krasjet 四元數與三維旋轉 p.63 73

34.2.6.1 clock

$$\boldsymbol{v}' = \mathrm{e}^{\boldsymbol{n}\frac{\theta}{2}} \boldsymbol{v} \mathrm{e}^{-\boldsymbol{n}\frac{\theta}{2}}$$

$$\boldsymbol{v}' = \mathrm{e}^{-\boldsymbol{n}\frac{\theta}{2}} \boldsymbol{v} \mathrm{e}^{\boldsymbol{n}\frac{\theta}{2}}$$

right-hand = anticlockwise

$$[\boldsymbol{v}']_R = \left[\mathrm{e}^{\boldsymbol{n}\frac{\theta}{2}}\right]_R [\boldsymbol{v}]_R \left[\mathrm{e}^{-\boldsymbol{n}\frac{\theta}{2}}\right]_R$$

$$[\boldsymbol{v}']_L = \left[\mathrm{e}^{-\boldsymbol{n}\frac{\theta}{2}}\right]_R [\boldsymbol{v}]_L \left[\mathrm{e}^{\boldsymbol{n}\frac{\theta}{2}}\right]_R$$

$$[\boldsymbol{v}']_L = \left[\mathrm{e}^{\boldsymbol{n}\frac{\theta}{2}}\right]_L [\boldsymbol{v}]_L \left[\mathrm{e}^{-\boldsymbol{n}\frac{\theta}{2}}\right]_L$$

34.2.6.2 Rodrigues rotation in right-hand vs. left-hand coordinates

$$\boldsymbol{u} = \boldsymbol{n} \times \boldsymbol{v}_\perp$$

$$[\boldsymbol{u}]_R = [\boldsymbol{n}]_R \times_R [\boldsymbol{v}_\perp]_R = [\boldsymbol{n}]_L \times_L [\boldsymbol{v}_\perp]_L = [\boldsymbol{u}]_L$$

$$\begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \times_R \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \times_L \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}$$

$$[[\boldsymbol{u}]_L]_R = -[\boldsymbol{u}]_R$$

$$\left[\begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}_L\right]_R = -\begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 0 \\ -1 \\ 0 \end{pmatrix}_R$$

$$\det O = \pm 1$$

$\det O = -1$ will change handedness of the coordinate.

34.2.7 gimbal lock

https://krasjet.github.io/quaternion/bonus_gimbal_lock.pdf

34.2.8 Lie group

34.2.8.1 special unitary group

34.2.9 geometric algebra and Clifford algebra

https://en.wikipedia.org/wiki/Geometric_algebra

https://en.wikipedia.org/wiki/Clifford_algebra

John Vince_2009_Geometric Algebra_ An Algebraic System for Computer Games and Animation

https://www.amazon.com/Geometric-Algebra-Algebraic-Computer-Animation/dp/1848823789

34.2.10 dual quaternion

34.2.11 interpolation

34.2.11.1 Lerp

34.2.11.2 Nlerp

34.2.11.3 Slerp

34.3 3Blue1Brown

34.3.1 Ben Eater

https://eater.net/quaternions/video/intro

https://eater.net/quaternions

34.3.2 3B1B

34.3.3 Sutrabla

https://www.newscientist.com/article/mg20427391-600-alices-adventures-in-algebra-wonderland-solved/

https://threejs.org/

34.4 CCJou: LA Revelation

https://ccjou.wordpress.com/2014/04/21/四元數/

Chapter 35

DICOM

35.1 Innolitics: DICOM Standard Browser

https://dicom.innolitics.com/ciods

https://dicom.innolitics.com/ciods/parametric-map/parametric-map-multi-frame-functional-groups/52009229/0048021a/0040072a

35.2 David Clunie: Medical Image Format Site

https://www.dclunie.com/

35.3 Hsiao, Chia-Hung

https://www.youtube.com/@user-zp9yy1ln6u/videos

Chapter 36

tendon pathophysiology

Mark E. Schweitzer, MD

https://www.youtube.com/watch?v=uIo58pQgxhY

36.1 histology

Chapter 37

group theory

https://en.wikipedia.org/wiki/Group_(mathematics)

37.1 matrix group

subset of two-by-two matrices at least excluding zero matrix

$$\mathcal{M} = (\mathcal{M}, \cdot) \subset (\mathcal{M}_{2\times 2}(\mathbb{C}) - \{0\}, \cdot) = \mathcal{M}_{2\times 2}(\mathbb{C}) - \left\{ \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix} \right\}$$

matrix multiplication

$$\begin{aligned}
&\forall \langle M_1, M_2 \rangle \in \mathcal{M}^2, \exists M_1 M_2 \in \mathcal{M} [M_1 M_2 = M_1 \cdot M_2] \\
\Leftrightarrow &\cdot : \mathcal{M} \times \mathcal{M} = \mathcal{M}^2 \to \mathcal{M} \\
\Leftrightarrow &\cdot : \mathcal{M} \times \mathcal{M} \to \mathcal{M} \\
\Leftrightarrow &\mathcal{M} \times \mathcal{M} \xrightarrow{\cdot} \mathcal{M} \\
\Leftrightarrow &\mathcal{M}^2 \xrightarrow{\cdot} \mathcal{M}
\end{aligned}$$

matrix group

$$\begin{aligned}
&\begin{cases}
\forall \langle M_1, M_2, M_3 \rangle \in \mathcal{M}^3 [M_1 (M_2 M_3) = (M_1 M_2) M_3] & \text{associativity} \\
\exists I = I_2 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \in \mathcal{M}, \forall M \in \mathcal{M} [IM = M] & \text{left unit element} \\
\forall M \in \mathcal{M}, \exists M^{-1} \in \mathcal{M} \left[M^{-1} M = I \right] & \text{left inverse (element)}
\end{cases} \\
\Rightarrow &\mathcal{M} = (\mathcal{M}, \cdot) \text{ is a matrix group}
\end{aligned}$$

37.2 group definition and basic theorem

https://en.wikipedia.org/wiki/Group_(mathematics)#Elementary_consequences_of_the_group_axioms

Definition 37.1 (group)**.** group definition by a set and a binary operation on the set

$$\begin{aligned}
&\begin{cases}
\circ : G \times G = G^2 \to G & \text{binary operation} \\
\forall \langle g_1, g_2, g_3 \rangle \in G^3 [g_1 \circ (g_2 \circ g_3) = (g_1 \circ g_2) \circ g_3] & \text{associativity} \\
\exists e = \in G, \forall g \in G [e \circ g = g] & \text{left unit element} \\
\forall g \in G, \exists g^{-1} \in G \left[g^{-1} \circ g = e \right] & \text{left inverse (element)}
\end{cases} \\
\Leftrightarrow &G = (G, \circ) \text{ is a group}
\end{aligned}$$

Theorem 37.1. *group left inverses equal right inverses*

Proof:

to be proved

□

37.3 Polya enumeration theorem

37.4 complex group representation

37.4.1 complex basis group

$$\begin{aligned} G &= \{1, \mathrm{i}, -1, -\mathrm{i}\} \\ &= \left\{\mathrm{i}^0, \mathrm{i}^1, \mathrm{i}^2, \mathrm{i}^3\right\} \end{aligned}$$

$$\begin{aligned} &\forall \langle g_1, g_2 \rangle \in G^2, \exists g_1 g_2 \in G \left[g_1 g_2 = g_1 \cdot g_2\right] \\ \Leftrightarrow &\cdot : G \times G = G^2 \to G \end{aligned}$$

https://tex.stackexchange.com/questions/627708/tikz-how-to-put-tables-within-arbitrarily-placed-nodes

$\cdot$	1	i	-1	$-\mathrm{i}$
1	1	i	-1	$-\mathrm{i}$
i	i	-1	$-\mathrm{i}$	1
-1	-1	$-\mathrm{i}$	1	i
$-\mathrm{i}$	$-\mathrm{i}$	1	i	-1

$\cdot$	1	i	-1	$-\mathrm{i}$
1	1	i	-1	$-\mathrm{i}$
i	i	-1	$-\mathrm{i}$	1
-1	-1	$-\mathrm{i}$	1	i
$-\mathrm{i}$	$-\mathrm{i}$	1	i	-1

Figure 37.1: complex basis group table

37.4.2 $\mathbb{C} \to \mathcal{M}_{2\times 2}(\mathbb{R})$

$$1 \leftrightarrow \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = I_2 = I$$

$$\begin{aligned} c_1 \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} + c_2 \begin{pmatrix} a & b \\ c & d \end{pmatrix} =& x \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} + y \begin{pmatrix} a & b \\ c & d \end{pmatrix}, \langle a, b, c, d \rangle \in \mathbb{R}^4 \\ =& xI + yJ, J = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \mathcal{M}_{2\times 2}(\mathbb{R}), \langle x, y \rangle \in \mathbb{R}^2 \end{aligned}$$

$$J^2 = -I$$

$$J^2 = -I$$

$$\begin{pmatrix} a^2 + bc & ab + bd \\ ca + cd & cb + d^2 \end{pmatrix} = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \begin{pmatrix} a & b \\ c & d \end{pmatrix} = - \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix}$$

$$\begin{cases} a^2 + bc = -1 & b = 0 \Rightarrow a^2 = -1 \Rightarrow\Leftarrow a \in \mathbb{R} \Rightarrow b \neq 0 \\ ab + bd = 0 & (b = 0) \vee (a = -d) \overset{b \neq 0}{\Rightarrow} a = -d \\ ca + cd = 0 & \\ cb + d^2 = -1 & a^2 = d^2 \Rightarrow (a = d) \vee (a = -d) \end{cases} \quad \underset{\text{if } a=d}{\Rightarrow} \begin{cases} a = d \\ a = -d \end{cases}$$

$$\Rightarrow a = d = 0 \Rightarrow bc = -1$$

$$J = \begin{pmatrix} a & b \\ c & d \end{pmatrix} = \begin{pmatrix} a & b \\ \dfrac{-a^2 - 1}{b} & -a \end{pmatrix} = J(a, b), b \neq 0$$

$$J(a, b) = \begin{pmatrix} a & b \\ \dfrac{-a^2 - 1}{b} & -a \end{pmatrix}, b \neq 0$$

$$J(a = 1, b) = \begin{pmatrix} 1 & b \\ \dfrac{-2}{b} & -1 \end{pmatrix} \Rightarrow J^2(a = 1, b) = \begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix} = -I$$

$$xI + yJ(a = 1, b) = \begin{pmatrix} x + y & yb \\ y \cdot \dfrac{-2}{b} & x - y \end{pmatrix}$$

$$J(a = 0, b) = \begin{pmatrix} 0 & b \\ \dfrac{-1}{b} & 0 \end{pmatrix}$$

$$J(a = 0, b = 1) = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$$

$$\begin{aligned} J(a = 0, b = -1) &= \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} = -\begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} = -J(a = 0, b = 1) \\ \Rightarrow J^2(a = 0, b = -1) &= \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} = \begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix} = -I \end{aligned}$$

$$\begin{aligned} J = J(a = 0, b = -1) &= \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \\ &\Rightarrow \begin{cases} 1 \leftrightarrow I = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \\ \mathrm{i} \leftrightarrow J = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \end{cases} \\ \Rightarrow x + y\mathrm{i} \leftrightarrow & xI + yJ \\ &= x \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} + y \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} = \begin{pmatrix} x & -y \\ y & x \end{pmatrix} \end{aligned}$$

$$x + y\mathrm{i} \leftrightarrow \begin{pmatrix} x & -y \\ y & x \end{pmatrix} = xI + yJ$$

realizing

$$\mathbb{C} \to \mathcal{M}_{2\times 2}(\mathbb{R}) = \mathcal{M}_2(\mathbb{R})$$

37.4.3 (determinant of complex group representation) equivalent to (squared modulus of complex number)

$$\det(xI + yJ) = \det \begin{pmatrix} x & -y \\ y & x \end{pmatrix} = \begin{vmatrix} x & -y \\ y & x \end{vmatrix} = x^2 + y^2 = |x + y\mathrm{i}|^2$$

37.4.3.1 Lagrange identity

Lagrange identity

generalization of Brahmagupta–Fibonacci identity

specialization of Binet–Cauchy identity

cf. Euler identity[37.5.1.1]

$$\begin{aligned}
&\det\left[(aI+bJ)(cI+dJ)\right]\\
=&\det\left[\begin{pmatrix} a & -b \\ b & a \end{pmatrix}\begin{pmatrix} c & -d \\ d & c \end{pmatrix}\right]\\
=&\det\begin{pmatrix} ac-bd & -ad-bc \\ ad+bc & ac-bd \end{pmatrix} = \left|(ac-bd)+(ad+bc)\,\mathrm{i}\right|^2 = (ac-bd)^2+(ad+bc)^2\\
=&\left[\det\begin{pmatrix} a & -b \\ b & a \end{pmatrix}\right]\left[\det\begin{pmatrix} c & -d \\ d & c \end{pmatrix}\right] = |a+b\mathrm{i}|^2\,|c+d\mathrm{i}|^2 = \left(a^2+b^2\right)\left(c^2+d^2\right)
\end{aligned}$$

$$|a+b\mathrm{i}|^2\,|c+d\mathrm{i}|^2 = \left(a^2+b^2\right)\left(c^2+d^2\right) = (ac-bd)^2+(ad+bc)^2$$

$$\begin{aligned}
&\det\left[(x_1I+y_1J)(x_2I+y_2J)\right]\\
=&\det\left[\begin{pmatrix} x_1 & -y_1 \\ y_1 & x_1 \end{pmatrix}\begin{pmatrix} x_2 & -y_2 \\ y_2 & x_2 \end{pmatrix}\right]\\
=&\det\begin{pmatrix} x_1x_2-y_1y_2 & -x_1y_2-y_1x_2 \\ x_1y_2+y_1x_2 & x_1x_2-y_1y_2 \end{pmatrix} = \left|(x_1x_2-y_1y_2)+(x_1y_2+y_1x_2)\,\mathrm{i}\right|^2 = (x_1x_2-y_1y_2)^2+(x_1y_2+y_1x_2)^2\\
=&\left[\det\begin{pmatrix} x_1 & -y_1 \\ y_1 & x_1 \end{pmatrix}\right]\left[\det\begin{pmatrix} x_2 & -y_2 \\ y_2 & x_2 \end{pmatrix}\right] = |x_1+y_1\mathrm{i}|^2\,|x_2+y_2\mathrm{i}|^2 = \left(x_1^2+y_1^2\right)\left(x_2^2+y_2^2\right)
\end{aligned}$$

$$|x_1+y_1\mathrm{i}|^2\,|x_2+y_2\mathrm{i}|^2 = \left(x_1^2+y_1^2\right)\left(x_2^2+y_2^2\right) = (x_1x_2-y_1y_2)^2+(x_1y_2+y_1x_2)^2$$

37.4.4 Euler formula proved by complex group representation

$$\begin{aligned}
\begin{pmatrix} x' \\ y' \end{pmatrix} &= \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix} = R_\theta\begin{pmatrix} x \\ y \end{pmatrix}\\
&= \begin{pmatrix} \cos n\frac{\theta}{n} & -\sin n\frac{\theta}{n} \\ \sin n\frac{\theta}{n} & \cos n\frac{\theta}{n} \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} \cos\frac{\theta}{n} & -\sin\frac{\theta}{n} \\ \sin\frac{\theta}{n} & \cos\frac{\theta}{n} \end{pmatrix}^n\begin{pmatrix} x \\ y \end{pmatrix} = R^n_{\frac{\theta}{n}}\begin{pmatrix} x \\ y \end{pmatrix}
\end{aligned}$$

$$\begin{aligned}
\lim_{n\to\infty} R_{\frac{\theta}{n}} = \lim_{n\to\infty}\begin{pmatrix} \cos\frac{\theta}{n} & -\sin\frac{\theta}{n} \\ \sin\frac{\theta}{n} & \cos\frac{\theta}{n} \end{pmatrix} &= \begin{pmatrix} 1 & -\frac{\theta}{n} \\ \frac{\theta}{n} & 1 \end{pmatrix}\\
&= \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} + \frac{\theta}{n}\begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}\\
&= \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} + \frac{\theta}{n}\begin{pmatrix} \cos\frac{\pi}{2} & -\sin\frac{\pi}{2} \\ \sin\frac{\pi}{2} & \cos\frac{\pi}{2} \end{pmatrix}\\
&= I + \frac{\theta}{n}R_{\frac{\pi}{2}}
\end{aligned}$$

$$\lim_{n\to\infty} R_{\frac{\theta}{n}} = \lim_{n\to\infty}\begin{pmatrix} \cos\frac{\theta}{n} & -\sin\frac{\theta}{n} \\ \sin\frac{\theta}{n} & \cos\frac{\theta}{n} \end{pmatrix} = \begin{pmatrix} 1 & -\frac{\theta}{n} \\ \frac{\theta}{n} & 1 \end{pmatrix} = I + \frac{\theta}{n}R_{\frac{\pi}{2}}$$

$$\begin{aligned}
\begin{pmatrix} x' \\ y' \end{pmatrix} = \lim_{n\to\infty}\begin{pmatrix} x' \\ y' \end{pmatrix} &= \lim_{n\to\infty} R^n_{\frac{\theta}{n}}\begin{pmatrix} x \\ y \end{pmatrix} = \left[\lim_{n\to\infty} R^n_{\frac{\theta}{n}}\right]\left[\lim_{n\to\infty}\begin{pmatrix} x \\ y \end{pmatrix}\right] = \left[\lim_{n\to\infty} R^n_{\frac{\theta}{n}}\right]\begin{pmatrix} x \\ y \end{pmatrix}\\
&= \lim_{n\to\infty}\left[\lim_{n\to\infty} R_{\frac{\theta}{n}}\right]^n\begin{pmatrix} x \\ y \end{pmatrix} = \lim_{n\to\infty}\left[I + \frac{\theta}{n}R_{\frac{\pi}{2}}\right]^n\begin{pmatrix} x \\ y \end{pmatrix}\\
&= \lim_{n\to\infty}\left[I + \frac{\theta}{n}\begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}\right]^n\begin{pmatrix} x \\ y \end{pmatrix} = \lim_{n\to\infty}\left[I + \frac{\theta J}{n}\right]^n\begin{pmatrix} x \\ y \end{pmatrix}, J = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}\\
&= \mathrm{e}^{J\theta}\begin{pmatrix} x \\ y \end{pmatrix}
\end{aligned}$$

$$\begin{cases} \begin{pmatrix} x' \\ y' \end{pmatrix} = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} \\ \begin{pmatrix} x' \\ y' \end{pmatrix} = \mathrm{e}^{J\theta} \begin{pmatrix} x \\ y \end{pmatrix} \end{cases}$$

$$\Rightarrow \mathrm{e}^{J\theta} = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix} \overset{x+y\mathrm{i}\leftrightarrow\begin{pmatrix} x & -y \\ y & x \end{pmatrix}=xI+yJ}{\Rightarrow} \mathrm{e}^{\mathrm{i}\theta} = \cos\theta + \mathrm{i}\sin\theta$$

□

37.5 quaternion group representation

https://groupprops.subwiki.org/wiki/Linear_representation_theory_of_quaternion_group#Two-dimensional_irreducible_representation_over_a_splitting_field

$$\begin{aligned} q &= a + b\mathrm{i} + c\mathrm{j} + d\mathrm{k} = a + \mathrm{i}b + \mathrm{j}c + \mathrm{k}d, \begin{cases} q \in \mathbb{H} \\ a,b,c,d \in \mathbb{R} \quad \Leftrightarrow \langle a,b,c,d\rangle \in \mathbb{R}^4 \end{cases} \\ &= w = t + x\mathrm{i} + y\mathrm{j} + z\mathrm{k} = t + \mathrm{i}x + \mathrm{j}y + \mathrm{k}z, \begin{cases} w \in \mathbb{H} \\ t,x,y,z \in \mathbb{R} \quad \Leftrightarrow \langle t,x,y,z\rangle \in \mathbb{R}^4 \end{cases} \\ &= a1 + b\mathrm{i} + c\mathrm{j} + d\mathrm{k} = t1 + x\mathrm{i} + y\mathrm{j} + z\mathrm{k} = x_0 1 + \boldsymbol{e}_i x_i \end{aligned}$$

37.5.1 $\mathbb{H} \to \mathcal{M}_{2\times 2}(\mathbb{C})$

$$1 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = 1_2 = 1$$

$$\begin{aligned} \boldsymbol{e} &= \begin{pmatrix} a & b \\ c & d \end{pmatrix} = \begin{pmatrix} a & b \\ \frac{-a^2-1}{b} & -a \end{pmatrix} \\ &= \begin{cases} \begin{pmatrix} a & b \\ \frac{-a^2-1}{b} & -a \end{pmatrix} = \begin{pmatrix} 0 & b \\ \frac{-1}{b} & 0 \end{pmatrix} = \begin{pmatrix} 0 & \beta \\ -\beta & 0 \end{pmatrix} \Rightarrow \boldsymbol{e}_2 = J = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} = \mathrm{j} & a = 0 \\ \begin{pmatrix} a & b \\ \frac{-a^2-1}{b} & -a \end{pmatrix} = \begin{pmatrix} \alpha & \beta \\ \beta & -\alpha \end{pmatrix} & a \neq 0 \end{cases} \end{aligned}$$

$$\begin{pmatrix} \alpha^2+\beta^2 & 0 \\ 0 & \beta^2+\alpha^2 \end{pmatrix} = \begin{pmatrix} \alpha & \beta \\ \beta & -\alpha \end{pmatrix} \begin{pmatrix} \alpha & \beta \\ \beta & -\alpha \end{pmatrix} = \boldsymbol{e}^2 = -1 = -\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix}$$
$$\Downarrow$$
$$\alpha^2 + \beta^2 = -1 \Leftrightarrow \beta^2 + \alpha^2 = -1$$

$$\alpha^2 + \beta^2 = -1 \Rightarrow \langle\alpha,\beta\rangle \notin \mathbb{R}^2 \Rightarrow\Leftarrow \langle\alpha,\beta\rangle \in \mathbb{R}^2 \Rightarrow \alpha^2 + \beta^2 \geq 0$$

quaternion group has no irreducible two-dimensional representation over the reals [1]

$$\langle\alpha,\beta\rangle \in \mathbb{C}^2 - \mathbb{R}^2$$

$$\alpha^2 + \beta^2 = -1 = \beta^2 + \alpha^2$$

$$\begin{pmatrix} \beta & \alpha \\ \alpha & -\beta \end{pmatrix} \begin{pmatrix} \beta & \alpha \\ \alpha & -\beta \end{pmatrix} = \begin{pmatrix} \beta^2+\alpha^2 & 0 \\ 0 & \alpha^2+\beta^2 \end{pmatrix} = \begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix} = -\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = -1$$

[1] https://groupprops.subwiki.org/wiki/Linear_representation_theory_of_quaternion_group#Four-dimensional_irreducible_representation_over_a_non-splitting_field

$$1=\begin{pmatrix}1&0\\0&1\end{pmatrix},\boldsymbol{e}_1=\begin{pmatrix}\alpha&\beta\\\beta&-\alpha\end{pmatrix},\boldsymbol{e}_2=\begin{pmatrix}0&-1\\1&0\end{pmatrix}$$

$$1=\begin{pmatrix}1&0\\0&1\end{pmatrix},\mathrm{i}=\begin{pmatrix}\alpha&\beta\\\beta&-\alpha\end{pmatrix},\mathrm{j}=\begin{pmatrix}0&-1\\1&0\end{pmatrix}$$

$$\mathrm{ij}=\begin{pmatrix}\alpha&\beta\\\beta&-\alpha\end{pmatrix}\begin{pmatrix}0&-1\\1&0\end{pmatrix}=\begin{pmatrix}\beta&-\alpha\\-\alpha&-\beta\end{pmatrix}=\mathrm{k}$$

$$1=\begin{pmatrix}1&0\\0&1\end{pmatrix},\boldsymbol{e}_1=\begin{pmatrix}\alpha&\beta\\\beta&-\alpha\end{pmatrix},\boldsymbol{e}_2=\begin{pmatrix}0&-1\\1&0\end{pmatrix},\boldsymbol{e}_3=\begin{pmatrix}\beta&-\alpha\\-\alpha&-\beta\end{pmatrix}$$

$$1=\begin{pmatrix}1&0\\0&1\end{pmatrix},\mathrm{i}=\begin{pmatrix}\alpha&\beta\\\beta&-\alpha\end{pmatrix},\mathrm{j}=\begin{pmatrix}0&-1\\1&0\end{pmatrix},\mathrm{k}=\begin{pmatrix}\beta&-\alpha\\-\alpha&-\beta\end{pmatrix}$$

$$\mathrm{jk}=\begin{pmatrix}0&-1\\1&0\end{pmatrix}\begin{pmatrix}\beta&-\alpha\\-\alpha&-\beta\end{pmatrix}=\begin{pmatrix}\alpha&\beta\\\beta&-\alpha\end{pmatrix}=\mathrm{i}$$

$$\begin{array}{ccccc}
 & \alpha^2+\beta^2=-1 & \begin{cases}\alpha=\sqrt{-1}\\\beta=0\end{cases} & \begin{cases}\alpha=\sqrt{-2}\\\beta=1\end{cases} & \beta=\alpha^2,n\in\{1,2,4,5\}\\
1 & \begin{pmatrix}1&0\\0&1\end{pmatrix} & \begin{pmatrix}1&0\\0&1\end{pmatrix} & \begin{pmatrix}1&0\\0&1\end{pmatrix} & \begin{pmatrix}1&0\\0&1\end{pmatrix}\\
-1 & \begin{pmatrix}-1&0\\0&-1\end{pmatrix} & \begin{pmatrix}-1&0\\0&-1\end{pmatrix} & \begin{pmatrix}-1&0\\0&-1\end{pmatrix} & \begin{pmatrix}-1&0\\0&-1\end{pmatrix}\\
\mathrm{i} & \begin{pmatrix}\alpha&\beta\\\beta&-\alpha\end{pmatrix} & \begin{pmatrix}\sqrt{-1}&0\\0&-\sqrt{-1}\end{pmatrix} & \begin{pmatrix}\sqrt{-2}&1\\1&-\sqrt{-2}\end{pmatrix} & \begin{pmatrix}\mathrm{e}^{\pi\frac{n}{3}\sqrt{-1}}&\mathrm{e}^{\pi\frac{2n}{3}\sqrt{-1}}\\\mathrm{e}^{\pi\frac{2n}{3}\sqrt{-1}}&-\mathrm{e}^{\pi\frac{n}{3}\sqrt{-1}}\end{pmatrix}\\
-\mathrm{i} & \begin{pmatrix}-\alpha&-\beta\\-\beta&\alpha\end{pmatrix} & \begin{pmatrix}-\sqrt{-1}&0\\0&\sqrt{-1}\end{pmatrix} & \begin{pmatrix}-\sqrt{-2}&-1\\-1&\sqrt{-2}\end{pmatrix} & \begin{pmatrix}-\mathrm{e}^{\pi\frac{n}{3}\sqrt{-1}}&-\mathrm{e}^{\pi\frac{2n}{3}\sqrt{-1}}\\-\mathrm{e}^{\pi\frac{2n}{3}\sqrt{-1}}&\mathrm{e}^{\pi\frac{n}{3}\sqrt{-1}}\end{pmatrix}\\
\mathrm{j} & \begin{pmatrix}0&-1\\1&0\end{pmatrix} & \begin{pmatrix}0&-1\\1&0\end{pmatrix} & \begin{pmatrix}0&-1\\1&0\end{pmatrix} & \begin{pmatrix}0&-1\\1&0\end{pmatrix}\\
-\mathrm{j} & \begin{pmatrix}0&1\\-1&0\end{pmatrix} & \begin{pmatrix}0&1\\-1&0\end{pmatrix} & \begin{pmatrix}0&1\\-1&0\end{pmatrix} & \begin{pmatrix}0&1\\-1&0\end{pmatrix}\\
\mathrm{k} & \begin{pmatrix}\beta&-\alpha\\-\alpha&-\beta\end{pmatrix} & \begin{pmatrix}0&-\sqrt{-1}\\-\sqrt{-1}&0\end{pmatrix} & \begin{pmatrix}1&-\sqrt{-2}\\-\sqrt{-2}&-1\end{pmatrix} & \begin{pmatrix}\mathrm{e}^{\pi\frac{2n}{3}\sqrt{-1}}&-\mathrm{e}^{\pi\frac{n}{3}\sqrt{-1}}\\-\mathrm{e}^{\pi\frac{n}{3}\sqrt{-1}}&-\mathrm{e}^{\pi\frac{2n}{3}\sqrt{-1}}\end{pmatrix}\\
-\mathrm{k} & \begin{pmatrix}-\beta&\alpha\\\alpha&\beta\end{pmatrix} & \begin{pmatrix}0&\sqrt{-1}\\\sqrt{-1}&0\end{pmatrix} & \begin{pmatrix}-1&\sqrt{-2}\\\sqrt{-2}&1\end{pmatrix} & \begin{pmatrix}-\mathrm{e}^{\pi\frac{2n}{3}\sqrt{-1}}&\mathrm{e}^{\pi\frac{n}{3}\sqrt{-1}}\\\mathrm{e}^{\pi\frac{n}{3}\sqrt{-1}}&\mathrm{e}^{\pi\frac{2n}{3}\sqrt{-1}}\end{pmatrix}
\end{array}$$

$$\begin{aligned}
-1 =&\alpha^2+\beta^2\\
\overset{\beta=\alpha^2}{=}&\alpha^2+\alpha^4\\
\alpha^4+\alpha^2+1 =&0,\alpha^4+\alpha^2+1=\left(\alpha^2+\alpha+1\right)\left(\alpha^2-\alpha+1\right)\\
\left(\alpha^2-1\right)\left(\alpha^4+\alpha^2+1\right) =&0\\
\alpha^6-1 =&0\\
\alpha^6 =&1=\mathrm{e}^{2\pi k\sqrt{-1}},k\in\mathbb{Z}\\
\alpha =&\mathrm{e}^{2\pi\frac{n}{6}\sqrt{-1}},n\in\{0,1,2,3,4,5\}-\{0,3\}\\
=&\mathrm{e}^{\pi\frac{n}{3}\sqrt{-1}},n\in\{1,2,4,5\}
\end{aligned}$$

$$
\begin{array}{ccccc}
 & \alpha^2+\beta^2=-1 & \begin{cases}\alpha=i\\ \beta=0\end{cases} & \begin{cases}\alpha=\sqrt{2}i\\ \beta=1\end{cases} & \omega=\mathrm{e}^{i\pi\frac{n}{3}}, n\in\{1,2,4,5\} \\
1 & \begin{pmatrix}1&0\\0&1\end{pmatrix} & \begin{pmatrix}1&0\\0&1\end{pmatrix} & \begin{pmatrix}1&0\\0&1\end{pmatrix} & \begin{pmatrix}1&0\\0&1\end{pmatrix} \\
-1 & \begin{pmatrix}-1&0\\0&-1\end{pmatrix} & \begin{pmatrix}-1&0\\0&-1\end{pmatrix} & \begin{pmatrix}-1&0\\0&-1\end{pmatrix} & \begin{pmatrix}-1&0\\0&-1\end{pmatrix} \\
\mathrm{i} & \begin{pmatrix}\alpha&\beta\\\beta&-\alpha\end{pmatrix} & \begin{pmatrix}i&0\\0&-i\end{pmatrix} & \begin{pmatrix}\sqrt{2}i&1\\1&-\sqrt{2}i\end{pmatrix} & \begin{pmatrix}\omega&\omega^2\\\omega^2&-\omega\end{pmatrix} \\
-\mathrm{i} & \begin{pmatrix}-\alpha&-\beta\\-\beta&\alpha\end{pmatrix} & \begin{pmatrix}-i&0\\0&i\end{pmatrix} & \begin{pmatrix}-\sqrt{2}i&-1\\-1&\sqrt{2}i\end{pmatrix} & \begin{pmatrix}-\omega&-\omega^2\\-\omega^2&\omega\end{pmatrix} \\
\mathrm{j} & \begin{pmatrix}0&-1\\1&0\end{pmatrix} & \begin{pmatrix}0&-1\\1&0\end{pmatrix} & \begin{pmatrix}0&-1\\1&0\end{pmatrix} & \begin{pmatrix}0&-1\\1&0\end{pmatrix} \\
-\mathrm{j} & \begin{pmatrix}0&1\\-1&0\end{pmatrix} & \begin{pmatrix}0&1\\-1&0\end{pmatrix} & \begin{pmatrix}0&1\\-1&0\end{pmatrix} & \begin{pmatrix}0&1\\-1&0\end{pmatrix} \\
\mathrm{k} & \begin{pmatrix}\beta&-\alpha\\-\alpha&-\beta\end{pmatrix} & \begin{pmatrix}0&-i\\-i&0\end{pmatrix} & \begin{pmatrix}1&-\sqrt{2}i\\-\sqrt{2}i&-1\end{pmatrix} & \begin{pmatrix}\omega^2&-\omega\\-\omega&-\omega^2\end{pmatrix} \\
-\mathrm{k} & \begin{pmatrix}-\beta&\alpha\\\alpha&\beta\end{pmatrix} & \begin{pmatrix}0&i\\i&0\end{pmatrix} & \begin{pmatrix}-1&\sqrt{2}i\\\sqrt{2}i&1\end{pmatrix} & \begin{pmatrix}-\omega^2&\omega\\\omega&\omega^2\end{pmatrix}
\end{array}
$$

realizing

$$\mathbb{H}\to\mathcal{M}_{2\times 2}(\mathbb{C})=\mathcal{M}_2(\mathbb{C})$$

37.5.1.1 Euler identity

cf. Lagrange identity[37.4.3.1]

$$
\begin{aligned}
&\det(a+b\mathrm{i}+c\mathrm{j}+d\mathrm{k})\\
=&\det\left[a\begin{pmatrix}1&0\\0&1\end{pmatrix}+b\begin{pmatrix}i&0\\0&-i\end{pmatrix}+c\begin{pmatrix}0&-1\\1&0\end{pmatrix}+d\begin{pmatrix}0&-i\\-i&0\end{pmatrix}\right]\\
=&\det\begin{pmatrix}a+bi&-c-di\\c-di&a-bi\end{pmatrix}=\begin{vmatrix}a+bi&-c-di\\c-di&a-bi\end{vmatrix}\\
=&\left(a^2+b^2\right)+\left(c^2+d^2\right)=a^2+b^2+c^2+d^2
\end{aligned}
$$

$$\det(a+b\mathrm{i}+c\mathrm{j}+d\mathrm{k})=\det\begin{pmatrix}a+bi&-c-di\\c-di&a-bi\end{pmatrix}=a^2+b^2+c^2+d^2$$

$$\det\left[\left(q_{10}+q_{11}\mathrm{i}+q_{12}\mathrm{j}+q_{13}\mathrm{k}\right)\left(q_{20}+q_{21}\mathrm{i}+q_{22}\mathrm{j}+q_{23}\mathrm{k}\right)\right]=\det\left[\left(a+b\mathrm{i}+c\mathrm{j}+d\mathrm{k}\right)\left(\alpha+\beta\mathrm{i}+\gamma\mathrm{j}+\delta\mathrm{k}\right)\right]$$
$$=\det\left\{\left[q_{10}\begin{pmatrix}1&0\\0&1\end{pmatrix}+q_{11}\begin{pmatrix}i&0\\0&-i\end{pmatrix}+q_{12}\begin{pmatrix}0&-1\\1&0\end{pmatrix}+q_{13}\begin{pmatrix}0&-i\\-i&0\end{pmatrix}\right]\right.$$
$$\left.\left[q_{20}\begin{pmatrix}1&0\\0&1\end{pmatrix}+q_{21}\begin{pmatrix}i&0\\0&-i\end{pmatrix}+q_{22}\begin{pmatrix}0&-1\\1&0\end{pmatrix}+q_{23}\begin{pmatrix}0&-i\\-i&0\end{pmatrix}\right]\right\}$$
$$=\det\left\{\left[a\begin{pmatrix}1&0\\0&1\end{pmatrix}+b\begin{pmatrix}i&0\\0&-i\end{pmatrix}+c\begin{pmatrix}0&-1\\1&0\end{pmatrix}+d\begin{pmatrix}0&-i\\-i&0\end{pmatrix}\right]\right.$$
$$\left.\left[\alpha\begin{pmatrix}1&0\\0&1\end{pmatrix}+\beta\begin{pmatrix}i&0\\0&-i\end{pmatrix}+\gamma\begin{pmatrix}0&-1\\1&0\end{pmatrix}+\delta\begin{pmatrix}0&-i\\-i&0\end{pmatrix}\right]\right\}$$
$$=\det\left\{\begin{pmatrix}a+bi&-c-di\\c-di&a-bi\end{pmatrix}\begin{pmatrix}\alpha+\beta i&-\gamma-\delta i\\\gamma-\delta i&\alpha-\beta i\end{pmatrix}\right\}$$
$$=\det\begin{pmatrix}a+bi&-c-di\\c-di&a-bi\end{pmatrix}\det\begin{pmatrix}\alpha+\beta i&-\gamma-\delta i\\\gamma-\delta i&\alpha-\beta i\end{pmatrix}=\left(a^2+b^2+c^2+d^2\right)\left(\alpha^2+\beta^2+\gamma^2+\delta^2\right)$$
$$=\det\left\{\begin{pmatrix}[a+bi]\,[\alpha+\beta i]-[c+di]\,[\gamma-\delta i]&[a+bi]\,[-\gamma-\delta i]-[c+di]\,[\alpha-\beta i]\\ [c-di]\,[\alpha+\beta i]+[a-bi]\,[\gamma-\delta i]&[c-di]\,[-\gamma-\delta i]+[a-bi]\,[\alpha-\beta i]\end{pmatrix}\right\}$$
$$=\det\left\{\begin{pmatrix}(a\alpha-b\beta-c\gamma-d\delta)+i\,(a\beta+b\alpha+c\delta-d\gamma)&-(a\gamma-b\delta+c\alpha+d\beta)-i\,(a\delta+b\gamma-c\beta+d\alpha)\\(a\gamma-b\delta+c\alpha+d\beta)-i\,(a\delta+b\gamma-c\beta+d\alpha)&(a\alpha-b\beta-c\gamma-d\delta)-i\,(a\beta+b\alpha+c\delta-d\gamma)\end{pmatrix}\right\}$$
$$=\det\left\{(a\alpha-b\beta-c\gamma-d\delta)+(a\beta+b\alpha+c\delta-d\gamma)\,\mathrm{i}+(a\gamma-b\delta+c\alpha+d\beta)\,\mathrm{j}+(a\delta+b\gamma-c\beta+d\alpha)\,\mathrm{k}\right\}$$
$$=(a\alpha-b\beta-c\gamma-d\delta)^2+(a\beta+b\alpha+c\delta-d\gamma)^2+(a\gamma-b\delta+c\alpha+d\beta)^2+(a\delta+b\gamma-c\beta+d\alpha)^2$$
$$=(q_{10}q_{20}-q_{11}q_{21}-q_{12}q_{22}-q_{13}q_{23})^2+(q_{10}q_{21}+q_{11}q_{20}+q_{12}q_{23}-q_{13}q_{22})^2$$
$$+(q_{10}q_{22}-q_{11}q_{23}+q_{12}q_{20}+q_{13}q_{21})^2+(q_{10}q_{23}+q_{11}q_{22}-q_{12}q_{21}+q_{13}q_{20})^2$$

$$\begin{aligned}&\left(a^2+b^2+c^2+d^2\right)\left(\alpha^2+\beta^2+\gamma^2+\delta^2\right)\\&=(a\alpha-b\beta-c\gamma-d\delta)^2+(a\beta+b\alpha+c\delta-d\gamma)^2\\&\quad+(a\gamma-b\delta+c\alpha+d\beta)^2+(a\delta+b\gamma-c\beta+d\alpha)^2\end{aligned}$$

Theorem 37.2. *For any two integers greater than zero, their multiplication can be the summation of squared four integers greater than zero.*

$$\forall\left\langle m_1,m_2\right\rangle\in\left(\mathbb{N}\cup\{0\}\right)^2,\exists\left\langle k_1,k_2,k_3,k_4\right\rangle\in\left(\mathbb{N}\cup\{0\}\right)^4\left[m_1m_2=k_1^2+k_2^2+k_3^2+k_4^2\right]$$

Proof:

Let $\begin{cases}m_1=a^2+b^2+c^2+d^2&\langle a,b,c,d\rangle\in\left(\mathbb{N}\cup\{0\}\right)^4\\m_2=\alpha^2+\beta^2+\gamma^2+\delta^2&\langle\alpha,\beta,\gamma,\delta\rangle\in\left(\mathbb{N}\cup\{0\}\right)^4\end{cases}\overset{\text{closure property}}{\Rightarrow}\begin{pmatrix}m_1\\m_2\end{pmatrix}=\begin{pmatrix}a^2+b^2+c^2+d^2\\\alpha^2+\beta^2+\gamma^2+\delta^2\end{pmatrix}\in\left(\mathbb{N}\cup\{0\}\right)^2$,

$$\begin{aligned}m_1m_2&=\left(a^2+b^2+c^2+d^2\right)\left(\alpha^2+\beta^2+\gamma^2+\delta^2\right)\\&\overset{\text{Euler identity}}{=}(a\alpha-b\beta-c\gamma-d\delta)^2+(a\beta+b\alpha+c\delta-d\gamma)^2\\&+(a\gamma-b\delta+c\alpha+d\beta)^2+(a\delta+b\gamma-c\beta+d\alpha)^2\\&=\left|a\alpha-b\beta-c\gamma-d\delta\right|^2\\&+\left|a\beta+b\alpha+c\delta-d\gamma\right|^2\\&+\left|a\gamma-b\delta+c\alpha+d\beta\right|^2\\&+\left|a\delta+b\gamma-c\beta+d\alpha\right|^2\\&=k_1^2+k_2^2+k_3^2+k_4^2,\begin{pmatrix}k_1\\k_2\\k_3\\k_4\end{pmatrix}=\begin{pmatrix}\left|a\alpha-b\beta-c\gamma-d\delta\right|\\\left|a\beta+b\alpha+c\delta-d\gamma\right|\\\left|a\gamma-b\delta+c\alpha+d\beta\right|\\\left|a\delta+b\gamma-c\beta+d\alpha\right|\end{pmatrix}\in\left(\mathbb{N}\cup\{0\}\right)^4\end{aligned}$$

$$\because \begin{cases} \langle a,b,c,d\rangle \in (\mathbb{N}\cup\{0\})^4 \\ \langle \alpha,\beta,\gamma,\delta\rangle \in (\mathbb{N}\cup\{0\})^4 \end{cases}$$
$$\overset{\text{closure property}}{\Rightarrow} \begin{pmatrix} |a\alpha - b\beta - c\gamma - d\delta| \\ |a\beta + b\alpha + c\delta - d\gamma| \\ |a\gamma - b\delta + c\alpha + d\beta| \\ |a\delta + b\gamma - c\beta + d\alpha| \end{pmatrix} \in (\mathbb{N}\cup\{0\})^4$$

□

37.5.2 $\mathbb{H} \to \mathcal{M}_{4\times 4}(\mathbb{R})$

https://groupprops.subwiki.org/wiki/Linear_representation_theory_of_quaternion_group#Four-dimensional_irreducible_representation_over_a_non-splitting_field

$$\mathbb{H} \to \mathcal{M}_2(\mathbb{C}) \overset{\mathbb{C}\to\mathcal{M}_2(\mathbb{R})}{\to} \mathcal{M}_{4\times 4}(\mathbb{R}) = \mathcal{M}_4(\mathbb{R})$$

$$\mathbb{C} \to \mathcal{M}_2(\mathbb{R}) \Leftarrow \begin{cases} 1 \to \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \\ i \to \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \end{cases}$$

$$1 \to \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \to \begin{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} & \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix} \\ \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix} & \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \end{pmatrix} \to \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

$$-1 \to \begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix} \to \begin{pmatrix} \begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix} & \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix} \\ \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix} & \begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix} \end{pmatrix} \to \begin{pmatrix} -1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix}$$

$$\mathrm{i} \to \begin{pmatrix} i & 0 \\ 0 & -i \end{pmatrix} \to \begin{pmatrix} \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} & \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix} \\ \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix} & \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \end{pmatrix} \to \begin{pmatrix} 0 & -1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & -1 & 0 \end{pmatrix}$$

$$-\mathrm{i} \to \begin{pmatrix} -i & 0 \\ 0 & i \end{pmatrix} \to \begin{pmatrix} \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} & \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix} \\ \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix} & \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \end{pmatrix} \to \begin{pmatrix} 0 & 1 & 0 & 0 \\ -1 & 0 & 0 & 0 \\ 0 & 0 & 0 & -1 \\ 0 & 0 & 1 & 0 \end{pmatrix}$$

$$\mathrm{j} \to \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \to \begin{pmatrix} \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix} & \begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix} \\ \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} & \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix} \end{pmatrix} \to \begin{pmatrix} 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{pmatrix}$$

$$-\mathrm{j} \to \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \to \begin{pmatrix} \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix} & \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \\ \begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix} & \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix} \end{pmatrix} \to \begin{pmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ -1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \end{pmatrix}$$

$$\mathrm{k} \to \begin{pmatrix} 0 & -i \\ -i & 0 \end{pmatrix} \to \begin{pmatrix} \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix} & \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \\ \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} & \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix} \end{pmatrix} \to \begin{pmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & -1 & 0 \\ 0 & 1 & 0 & 0 \\ -1 & 0 & 0 & 0 \end{pmatrix}$$

$$-\mathrm{k} \to \begin{pmatrix} 0 & i \\ i & 0 \end{pmatrix} \to \begin{pmatrix} \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix} & \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \\ \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} & \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix} \end{pmatrix} \to \begin{pmatrix} 0 & 0 & 0 & -1 \\ 0 & 0 & 1 & 0 \\ 0 & -1 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{pmatrix}$$

some examinations

$$\mathrm{ij} \to \begin{pmatrix} 0 & -1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & -1 & 0 \end{pmatrix} \begin{pmatrix} 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{pmatrix} = \begin{pmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & -1 & 0 \\ 0 & 1 & 0 & 0 \\ -1 & 0 & 0 & 0 \end{pmatrix} \leftarrow \mathrm{k}$$

$$\mathrm{ji} \to \begin{pmatrix} 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{pmatrix} \begin{pmatrix} 0 & -1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & -1 & 0 \end{pmatrix} = \begin{pmatrix} 0 & 0 & 0 & -1 \\ 0 & 0 & 1 & 0 \\ 0 & -1 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{pmatrix} \leftarrow -\mathrm{k}$$

$$\mathrm{i}^2 = \mathrm{ii} \to \begin{pmatrix} 0 & -1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & -1 & 0 \end{pmatrix} \begin{pmatrix} 0 & -1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & -1 & 0 \end{pmatrix} = \begin{pmatrix} -1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix} \leftarrow -1$$

	$\begin{cases}\alpha = i \\ \beta = 0\end{cases}$	$\begin{cases}1 \to \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \\ i \to \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}\end{cases}$	$\begin{cases}\alpha = \sqrt{2}i \\ \beta = 1\end{cases}$	$\begin{cases}1 \to \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \\ i \to \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}\end{cases}$
1	$\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$	$\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$	$\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$	$\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$
−1	$\begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix}$	$\begin{pmatrix} -1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix}$	$\begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix}$	$\begin{pmatrix} -1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix}$
i	$\begin{pmatrix} i & 0 \\ 0 & -i \end{pmatrix}$	$\begin{pmatrix} 0 & -1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & -1 & 0 \end{pmatrix}$	$\begin{pmatrix} \sqrt{2}i & 1 \\ 1 & -\sqrt{2}i \end{pmatrix}$	$\begin{pmatrix} 0 & -\sqrt{2} & 1 & 0 \\ \sqrt{2} & 0 & 0 & 1 \\ 1 & 0 & 0 & \sqrt{2} \\ 0 & 1 & -\sqrt{2} & 0 \end{pmatrix}$
−i	$\begin{pmatrix} -i & 0 \\ 0 & i \end{pmatrix}$	$\begin{pmatrix} 0 & 1 & 0 & 0 \\ -1 & 0 & 0 & 0 \\ 0 & 0 & 0 & -1 \\ 0 & 0 & 1 & 0 \end{pmatrix}$	$\begin{pmatrix} -\sqrt{2}i & -1 \\ -1 & \sqrt{2}i \end{pmatrix}$	$\begin{pmatrix} 0 & -1 & -1 & 0 \\ 1 & 0 & 0 & -1 \\ -1 & 0 & 0 & 1 \\ 0 & -1 & -1 & 0 \end{pmatrix}$
j	$\begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$	$\begin{pmatrix} 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{pmatrix}$	$\begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$	$\begin{pmatrix} 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{pmatrix}$
−j	$\begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$	$\begin{pmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ -1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \end{pmatrix}$	$\begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$	$\begin{pmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ -1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \end{pmatrix}$
k	$\begin{pmatrix} 0 & -i \\ -i & 0 \end{pmatrix}$	$\begin{pmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & -1 & 0 \\ 0 & 1 & 0 & 0 \\ -1 & 0 & 0 & 0 \end{pmatrix}$	$\begin{pmatrix} 1 & -\sqrt{2}i \\ -\sqrt{2}i & -1 \end{pmatrix}$	$\begin{pmatrix} 1 & 0 & 0 & \sqrt{2} \\ 0 & 1 & -\sqrt{2} & 0 \\ 0 & \sqrt{2} & -1 & 0 \\ -\sqrt{2} & 0 & 0 & -1 \end{pmatrix}$
−k	$\begin{pmatrix} 0 & i \\ i & 0 \end{pmatrix}$	$\begin{pmatrix} 0 & 0 & 0 & -1 \\ 0 & 0 & 1 & 0 \\ 0 & -1 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{pmatrix}$	$\begin{pmatrix} -1 & \sqrt{2}i \\ \sqrt{2}i & 1 \end{pmatrix}$	$\begin{pmatrix} -1 & 0 & 0 & -\sqrt{2} \\ 0 & -1 & \sqrt{2} & 0 \\ 0 & -\sqrt{2} & 1 & 0 \\ \sqrt{2} & 0 & 0 & 1 \end{pmatrix}$

some examinations

$$\text{ij} \rightarrow \begin{pmatrix} 0 & -\sqrt{2} & 1 & 0 \\ \sqrt{2} & 0 & 0 & 1 \\ 1 & 0 & 0 & \sqrt{2} \\ 0 & 1 & -\sqrt{2} & 0 \end{pmatrix} \begin{pmatrix} 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 & \sqrt{2} \\ 0 & 1 & -\sqrt{2} & 0 \\ 0 & \sqrt{2} & -1 & 0 \\ -\sqrt{2} & 0 & 0 & -1 \end{pmatrix} \leftarrow \text{k}$$

Chapter 38

tensor

38.1 Einstein summation convention

38.1.1 dummy index

38.1.2 free index

38.2 Elliot Schneider

- Elliot Schneider: Physics with Elliot

38.2.1 Fundamentals of Cartesian Tensors

https://courses.physicswithelliot.com/products/part-i-fundamentals-of-cartesian-tensors/categories/2154478208/posts/2176662131

https://courses.physicswithelliot.com/products/part-i-fundamentals-of-cartesian-tensors/categories/2154478208/posts/2174076341

https://courses.physicswithelliot.com/products/part-i-fundamentals-of-cartesian-tensors/categories/2154478208/posts/2174076344

38.2.2 Fundamentals of Curvilinear Tensors

38.2.3 Fundamentals of Spacetime Tensors

Chapter 39

dual space

dual space and linear functional

https://ccjou.wordpress.com/2011/06/13/線性泛函與對偶空間/

39.1 linear equations

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n =& b_1 \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n =& b_2 \\ & \vdots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n =& b_m \end{aligned}$$

$$\begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{pmatrix}$$

$$A\boldsymbol{x} = \boldsymbol{b}$$

$$\begin{pmatrix} a_{i1} & a_{i2} & \cdots & a_{in} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} b_i \end{pmatrix} = b_i$$

$$\begin{pmatrix} a_{i1} & \cdots & a_{ij} & \cdots & a_{in} \end{pmatrix} \begin{pmatrix} x_1 \\ \vdots \\ x_j \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} b_i \end{pmatrix} = b_i$$

$$\begin{pmatrix} \cdots & a_{ij} & \cdots \end{pmatrix} \begin{pmatrix} \vdots \\ x_j \\ \vdots \end{pmatrix} = b_i$$

$$\boldsymbol{a}_i^\intercal \boldsymbol{x} = \boldsymbol{a}_i \cdot \boldsymbol{x} = b_i$$

$$a_{ij}x_j = \boldsymbol{a}_i^\intercal \boldsymbol{x} = \boldsymbol{a}_i \cdot \boldsymbol{x} = b_i$$

$$\boldsymbol{a}_i \cdot \boldsymbol{x} = \boldsymbol{a}_i^\intercal \boldsymbol{x} = a_{ij}x_j = \cdots + a_{ij}x_j + \cdots$$

if finite,

$$\boldsymbol{a}_i \cdot \boldsymbol{x} = \boldsymbol{a}_i^\intercal \boldsymbol{x} = a_{ij}x_j = a_{i1}x_1 + \cdots + a_{ij}x_j + \cdots + a_{in}x_n$$

39.2 matrix multiplication

$$\begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix} \begin{pmatrix} x_{11} & x_{12} & \cdots & x_{1p} \\ x_{21} & x_{22} & \cdots & x_{2p} \\ \vdots & \vdots & \ddots & \vdots \\ x_{n1} & x_{n2} & \cdots & x_{np} \end{pmatrix} = \begin{pmatrix} b_{11} & b_{12} & \cdots & b_{1p} \\ b_{21} & b_{22} & \cdots & b_{2p} \\ \vdots & \vdots & \ddots & \vdots \\ b_{m1} & b_{m2} & \cdots & b_{mp} \end{pmatrix}$$

$$AX = B$$

$$\begin{pmatrix} a_{i1} & a_{i2} & \cdots & a_{in} \end{pmatrix} \begin{pmatrix} x_{1k} \\ x_{2k} \\ \vdots \\ x_{nk} \end{pmatrix} = \begin{pmatrix} b_{ik} \end{pmatrix} = b_{ik}$$

$$\begin{pmatrix} a_{i1} & \cdots & a_{ij} & \cdots & a_{in} \end{pmatrix} \begin{pmatrix} x_{1k} \\ \vdots \\ x_{jk} \\ \vdots \\ x_{nk} \end{pmatrix} = \begin{pmatrix} b_{ik} \end{pmatrix} = b_{ik}$$

$$\begin{pmatrix} \cdots & a_{ij} & \cdots \end{pmatrix} \begin{pmatrix} \vdots \\ x_{jk} \\ \vdots \end{pmatrix} = b_{ik}$$

$$a_{ij}x_j = \boldsymbol{a}_i^\intercal \boldsymbol{x} = \boldsymbol{a}_i \cdot \boldsymbol{x} = b_i$$

$$\boldsymbol{a}_i \cdot \boldsymbol{x} = \boldsymbol{a}_i^\intercal \boldsymbol{x} = a_{ij}x_j$$

`\iddots` in MathJax

https://math.meta.stackexchange.com/questions/23273/mathjax-and-iddots-udots-or-reflectbox

$$b_{ik} = a_{ij}x_{jk} = \boldsymbol{a}_i \cdot \boldsymbol{x}_k = \boldsymbol{a}_i^\intercal \boldsymbol{x}_k$$

$$\text{row}\,(A)\,\text{col}\,(X) = b_{\text{row,col}}$$

$$\boldsymbol{a}_i \cdot \boldsymbol{x}_k = \boldsymbol{a}_i^\intercal \boldsymbol{x}_k = a_{ij}x_{jk}$$

$$\boldsymbol{a}_i \cdot \boldsymbol{x}_j = \boldsymbol{a}_i^\intercal \boldsymbol{x}_j = a_{ik}x_{kj}$$

$$\boldsymbol{a}_i \cdot \boldsymbol{x}_j = \boldsymbol{a}_i^\intercal \boldsymbol{x}_j = a_{ik}x_{kj} = \cdots + a_{ik}x_{kj} + \cdots$$

if finite,

$$\boldsymbol{a}_i \cdot \boldsymbol{x}_j = \boldsymbol{a}_i^\intercal \boldsymbol{x}_j = a_{ik}x_{kj} = a_{i1}x_{1j} + \cdots + a_{ik}x_{kj} + \cdots + a_{in}x_{nj}$$

39.3 functional

(innder product or dot product) or linear equations

$$\boldsymbol{a}_i \cdot \boldsymbol{x} = \boldsymbol{a}_i^\intercal \boldsymbol{x} = a_{ij} x_j = \cdots + a_{ij} x_j + \cdots$$

if finite,

$$\boldsymbol{a}_i \cdot \boldsymbol{x} = \boldsymbol{a}_i^\intercal \boldsymbol{x} = a_{ij} x_j = a_{i1} x_1 + \cdots + a_{ij} x_j + \cdots + a_{in} x_n$$

---

actually, several rows

$$\begin{array}{ccccccc} \vdots & & \vdots & & \vdots & & \vdots \\ \boldsymbol{a}_i \cdot \boldsymbol{x} & = & \boldsymbol{a}_i^\intercal \boldsymbol{x} & = & a_{ij} x_j & = & \cdots + a_{ij} x_j + \cdots \\ \vdots & & \vdots & & \vdots & & \vdots \end{array}$$

if finite,

$$\begin{array}{ccccccc} \vdots & & \vdots & & \vdots & & \vdots \\ \boldsymbol{a}_i \cdot \boldsymbol{x} & = & \boldsymbol{a}_i^\intercal \boldsymbol{x} & = & a_{ij} x_j & = & a_{i1} x_1 + \cdots + a_{ij} x_j + \cdots + a_{in} x_n \\ \vdots & & \vdots & & \vdots & & \vdots \end{array}$$

---

in functional aspect,

$$\begin{array}{ccccccccc} \vdots & & \vdots & & \vdots & & \vdots & & \vdots \\ f_i(\boldsymbol{x}) & = & \boldsymbol{a}_i \cdot \boldsymbol{x} & = & \boldsymbol{a}_i^\intercal \boldsymbol{x} & = & a_{ij} x_j & = & \cdots + a_{ij} x_j + \cdots \\ \vdots & & \vdots & & \vdots & & \vdots & & \vdots \end{array}$$

if finite,

$$\begin{array}{ccccccccc} \vdots & & \vdots & & \vdots & & \vdots & & \vdots \\ f_i(\boldsymbol{x}) & = & \boldsymbol{a}_i \cdot \boldsymbol{x} & = & \boldsymbol{a}_i^\intercal \boldsymbol{x} & = & a_{ij} x_j & = & a_{i1} x_1 + \cdots + a_{ij} x_j + \cdots + a_{in} x_n \\ \vdots & & \vdots & & \vdots & & \vdots & & \vdots \end{array}$$

$$\begin{array}{ccccccccccccc} \vdots & & \vdots & & \vdots & & \vdots & & \vdots & & \vdots & & \vdots \\ f_i(\cdots, x_j, \cdots) & = & f_i(x_j) & = & f_i(\boldsymbol{x}) & = & \boldsymbol{a}_i \cdot \boldsymbol{x} & = & \boldsymbol{a}_i^\intercal \boldsymbol{x} & = & a_{ij} x_j & = & \cdots + a_{ij} x_j + \cdots \\ \vdots & & \vdots & & \vdots & & \vdots & & \vdots & & \vdots & & \vdots \end{array}$$

if finite,

$$\begin{array}{ccccccccccccc} \vdots & & \vdots & & \vdots & & \vdots & & \vdots & & \vdots & & \vdots \\ f_i(x_1, \cdots, x_j, \cdots, x_n) & = & f_i(x_j) & = & f_i(\boldsymbol{x}) & = & \boldsymbol{a}_i \cdot \boldsymbol{x} & = & \boldsymbol{a}_i^\intercal \boldsymbol{x} & = & a_{ij} x_j & = & a_{i1} x_1 + \cdots + a_{ij} x_j + \cdots + a_{in} x_n \\ \vdots & & \vdots & & \vdots & & \vdots & & \vdots & & \vdots & & \vdots \end{array}$$

---

or simply

$$f_i(\cdots, x_j, \cdots) = f_i(x_j) = f_i(\boldsymbol{x}) = \boldsymbol{a}_i \cdot \boldsymbol{x}_j = \boldsymbol{a}_i^\intercal \boldsymbol{x}_j = a_{ik} x_{kj} = \cdots + a_{ij} x_j + \cdots$$

if scalar with complex as the field,

$$f_i : \mathbb{C}^\infty \to \mathbb{C}$$

if scalar with a field,

$$f_i : \mathbb{F}^\infty \to \mathbb{F}$$

or more abstract notation,

$$f_i : F^\infty \to F$$

if scalar with real as the field,

$$f_i : \mathbb{R}^\infty \to \mathbb{R}$$

---

if finite,

$$f_i\left(x_1, \cdots, x_j, \cdots, x_n\right) = f_i\left(x_j\right) = f_i\left(\boldsymbol{x}\right) = \boldsymbol{a}_i \cdot \boldsymbol{x}_j = \boldsymbol{a}_i^\intercal \boldsymbol{x}_j = a_{ij}x_j = a_{i1}x_1 + \cdots + a_{ij}x_j + \cdots + a_{in}x_n$$

$$\begin{aligned} f_i\left(x_1, x_2, \cdots, x_n\right) =& f_i\left(x_1, \cdots, x_j, \cdots, x_n\right) = f_i\left(x_j\right) = f_i\left(\boldsymbol{x}\right) \\ =& \boldsymbol{a}_i \cdot \boldsymbol{x}_j = \boldsymbol{a}_i^\intercal \boldsymbol{x}_j = a_{ij}x_j = a_{i1}x_1 + \cdots + a_{ij}x_j + \cdots + a_{in}x_n \end{aligned}$$

if scalar with complex as the field,

$$f_i : \mathbb{C}^n \to \mathbb{C}$$

if scalar with a field,

$$f_i : \mathbb{F}^n \to \mathbb{F}$$

or more abstract notation,

$$f_i : F^n \to F$$

if scalar with real as the field,

$$f_i : \mathbb{R}^n \to \mathbb{R}$$

---

functionals are a set of fuctions mapping n-dimensional vectors to scalars

$$f_i : F^n \to F$$

$$\begin{matrix} \vdots & & \vdots & & \vdots & & \vdots & & \vdots & & \vdots & & \vdots \\ f_i\left(\cdots, x_j, \cdots\right) & = & f_i\left(x_j\right) & = & f_i\left(\boldsymbol{x}\right) & = & \boldsymbol{a}_i \cdot \boldsymbol{x} & = & \boldsymbol{a}_i^\intercal \boldsymbol{x} & = & a_{ij}x_j & = & \cdots + a_{ij}x_j + \cdots \\ \vdots & & \vdots & & \vdots & & \vdots & & \vdots & & \vdots & & \vdots \end{matrix}$$

$$f = \left\{f_i | f_i : F^\infty \to F\right\} = \left\{ \begin{matrix} \vdots \\ f_i\left(\cdots, x_j, \cdots\right) = f_i\left(x_j\right) = f_i\left(\boldsymbol{x}\right) = \boldsymbol{a}_i \cdot \boldsymbol{x} = \boldsymbol{a}_i^\intercal \boldsymbol{x} = a_{ij}x_j, \\ \vdots \end{matrix} \right\}$$

if finite,

$$\begin{matrix} \vdots & & \vdots & & \vdots & & \vdots & & \vdots & & \vdots & & \vdots \\ f_i\left(x_1, \cdots, x_j, \cdots, x_n\right) & = & f_i\left(x_j\right) & = & f_i\left(\boldsymbol{x}\right) & = & \boldsymbol{a}_i \cdot \boldsymbol{x} & = & \boldsymbol{a}_i^\intercal \boldsymbol{x} & = & a_{ij}x_j & = & a_{i1}x_1 + \cdots + a_{ij}x_j + \cdots + a_{in}x_n \\ \vdots & & \vdots & & \vdots & & \vdots & & \vdots & & \vdots & & \vdots \end{matrix}$$

$$f = \{f_i | f_i : F^n \to F\} = \left\{ \begin{array}{c} \vdots \\ f_i(x_1, x_2, \cdots, x_n) = f_i(\boldsymbol{x}) = a_{i1}x_1 + \cdots + a_{ij}x_j + \cdots + a_{in}x_n, \\ \vdots \end{array} \right\}$$

---

linear functionals are a set of fuctions mapping vectors to scalars linearly

$$f = \{f_i | f_i : F^n \to F\} = \left\{ \begin{array}{c} \vdots \\ f_i(x_1, x_2, \cdots, x_n) = f_i(\boldsymbol{x}) = a_{i1}x_1 + \cdots + a_{ij}x_j + \cdots + a_{in}x_n, \\ \vdots \end{array} \right\}$$

39.4 definition of linear functional

functionals generalized to general vector space

$$f = \{f_i | f_i : V \to F\}$$

linear functionals generalized to general vector space is a linear transformation

$$f = \left\{ f_i \middle| \begin{cases} f_i : V \to F \\ \forall \langle \boldsymbol{x}, \boldsymbol{y} \rangle \in V^2 \left[f_i(\boldsymbol{x} + \boldsymbol{y}) = f_i(\boldsymbol{x}) + f_i(\boldsymbol{y}) \right] \\ \forall \boldsymbol{x} \in V, c \in F \left[f_i(c\boldsymbol{x}) = c f_i(\boldsymbol{x}) \right] \end{cases} \right\}$$

$$f = \left\{ f_i \middle| \begin{cases} f_i : V \to F \\ \forall \boldsymbol{x}, \boldsymbol{y} \in V \left[f_i(\boldsymbol{x} + \boldsymbol{y}) = f_i(\boldsymbol{x}) + f_i(\boldsymbol{y}) \right] \\ \forall \boldsymbol{x} \in V, c \in F \left[f_i(c\boldsymbol{x}) = c f_i(\boldsymbol{x}) \right] \end{cases} \right\}$$

if $\begin{cases} F = \mathbb{C} \\ V = \mathbb{C}^n \end{cases}$,

$$f = \left\{ f_i \middle| \begin{cases} f_i : \mathbb{C}^n \to \mathbb{C} \\ \forall \langle \boldsymbol{x}, \boldsymbol{y} \rangle \in (\mathbb{C}^n)^2 \left[f_i(\boldsymbol{x} + \boldsymbol{y}) = f_i(\boldsymbol{x}) + f_i(\boldsymbol{y}) \right] \\ \forall \boldsymbol{x} \in \mathbb{C}^n, c \in \mathbb{C} \left[f_i(c\boldsymbol{x}) = c f_i(\boldsymbol{x}) \right] \end{cases} \right\}$$

$$f = \left\{ f_i \middle| \begin{cases} f_i : \mathbb{C}^n \to \mathbb{C} \\ \forall \boldsymbol{x}, \boldsymbol{y} \in \mathbb{C}^n \left[f_i(\boldsymbol{x} + \boldsymbol{y}) = f_i(\boldsymbol{x}) + f_i(\boldsymbol{y}) \right] \\ \forall \boldsymbol{x} \in \mathbb{C}^n, c \in \mathbb{C} \left[f_i(c\boldsymbol{x}) = c f_i(\boldsymbol{x}) \right] \end{cases} \right\}$$

then

$$f_i(\boldsymbol{x}) = a_{i1}x_1 + \cdots + a_{ij}x_j + \cdots + a_{in}x_n$$

satisfying

$$\begin{aligned} f_i(\boldsymbol{x} + \boldsymbol{y}) =& a_{i1}(x + y)_1 + \cdots + a_{ij}(x + y)_j + \cdots + a_{in}(x + y)_n \\ =& a_{i1}(x_1 + y_1) + \cdots + a_{ij}(x_j + y_j) + \cdots + a_{in}(x_n + y_n) \\ =& (a_{i1}x_1 + a_{i1}y_1) + \cdots + (a_{ij}x_j + a_{ij}y_j) + \cdots + (a_{in}x_n + a_{in}y_n) \\ =& (a_{i1}x_1 + \cdots + a_{ij}x_j + \cdots + a_{in}x_n) + (a_{i1}y_1 + \cdots + a_{ij}y_j + \cdots + a_{in}y_n) \\ =& f_i(\boldsymbol{x}) + f_i(\boldsymbol{y}) \end{aligned}$$

$$\begin{aligned} f_i(c\boldsymbol{x}) =& a_{i1}(cx)_1 + \cdots + a_{ij}(cx)_j + \cdots + a_{in}(cx)_n \\ =& a_{i1}(cx_1) + \cdots + a_{ij}(cx_j) + \cdots + a_{in}(cx_n) \\ =& c(a_{i1}x_1) + \cdots + c(a_{ij}x_j) + \cdots + c(a_{in}x_n) \\ =& c(a_{i1}x_1 + \cdots + a_{ij}x_j + \cdots + a_{in}x_n) \\ =& c f_i(\boldsymbol{x}) \end{aligned}$$

---

different functional has different a_{ij}

let

$$a_{ij} = f_i(\boldsymbol{e}_j), \boldsymbol{e}_j = \left\langle \overbrace{0, \cdots, 0}^{j-1}, 1, 0, \cdots, 0 \right\rangle = \begin{pmatrix} 0 & \cdots & 0 & 1 & 0 & \cdots & 0 \end{pmatrix}^{\mathsf{T}} = \begin{pmatrix} 0 \\ \vdots \\ 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix} = \begin{pmatrix} 0 \\ \vdots \\ 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}_{n\times 1}$$

if f_i is a linear functional, then

$$\begin{aligned} f_i(\boldsymbol{x}) =& f_i(x_1\boldsymbol{e}_1 + \cdots + x_j\boldsymbol{e}_j + \cdots + x_n\boldsymbol{e}_n) \\ =& f_i(x_1\boldsymbol{e}_1) + \cdots + f_i(x_j\boldsymbol{e}_j) + \cdots + f_i(x_n\boldsymbol{e}_n) \\ =& x_1 f_i(\boldsymbol{e}_1) + \cdots + x_j f_i(\boldsymbol{e}_j) + \cdots + x_n f_i(\boldsymbol{e}_n) \\ =& x_1 a_{i1} + \cdots + x_j a_{ij} + \cdots + x_n a_{in} \\ =& a_{i1}x_1 + \cdots + a_{ij}x_j + \cdots + a_{in}x_n \\ =& f_i(\boldsymbol{x}) \end{aligned}$$

39.5 set of all linear transformations is a vector space

https://ccjou.wordpress.com/2011/04/08/線性變換集合構成向量空間/

vector space[40]

https://math.stackexchange.com/questions/2381942/the-set-of-all-linear-maps-tv-w-is-a-vector-space

$$T : V \to W \Leftrightarrow \forall \boldsymbol{v} \in V, \exists! \boldsymbol{w} \in W \left[\boldsymbol{w} = T(\boldsymbol{v})\right]$$

$$\begin{cases} V, W \text{ are vector spaces} \\ T : V \to W \\ \begin{cases} \forall \boldsymbol{u}, \boldsymbol{v} \in V \left[T(\boldsymbol{u} + \boldsymbol{v}) = T(\boldsymbol{u}) + T(\boldsymbol{v})\right] \\ \forall \boldsymbol{v} \in V, c \in F \left[T(c\boldsymbol{v}) = cT(\boldsymbol{v})\right] \end{cases} & \text{linearity} \end{cases}$$
$$\Leftrightarrow T \text{ is a linear tranformation}$$

$$\begin{cases} V, W \text{ are vector spaces, both over } F \\ T, U : V \to W \begin{cases} T : V \to W \\ U : V \to W \end{cases} \\ T, U \text{ are both linear tranformations} \\ \boldsymbol{v} \in V \\ c \in F \end{cases}$$

There is still linearity over linear transformations

(va)

$$\begin{aligned} (T + U)(\boldsymbol{u} + \boldsymbol{v}) =& T(\boldsymbol{u} + \boldsymbol{v}) + U(\boldsymbol{u} + \boldsymbol{v}) \\ &= \left[T(\boldsymbol{u}) + T(\boldsymbol{v})\right] + \left[U(\boldsymbol{u}) + U(\boldsymbol{v})\right] \\ &= \left[T(\boldsymbol{u}) + U(\boldsymbol{u})\right] + \left[T(\boldsymbol{v}) + U(\boldsymbol{v})\right] \\ &= (T + U)(\boldsymbol{u}) + (T + U)(\boldsymbol{v}) \end{aligned}$$

(sm)

$$\begin{aligned}(T+U)(c\boldsymbol{v}) =& T(c\boldsymbol{v})+U(c\boldsymbol{v})\\ =& cT(\boldsymbol{v})+cU(\boldsymbol{v})\\ =& c[T(\boldsymbol{v})+U(\boldsymbol{v})]\\ =& c(T+U)(\boldsymbol{v})\end{aligned}$$

so we can define

$$\begin{cases}(T+U)(\boldsymbol{v})=T(\boldsymbol{v})+U(\boldsymbol{v}) & \text{linear transformation addition}\\ (cT)(\boldsymbol{v})=cT(\boldsymbol{v}) & \text{scalar linear transformation multiplication}\end{cases}$$

---

the set of all linear tranformations is a vector space

$\mathcal{T}$ is the set of all linear tranformations

$$\begin{cases}F & (f)\, F \text{ is a field}\\ \mathcal{T}\neq\emptyset & (ne)\,\text{nonempty set}\\ +:\mathcal{T}\times\mathcal{T}=\mathcal{T}^2\xrightarrow{+}\mathcal{T}\Leftrightarrow\forall T,U\in\mathcal{T},\exists S\in\mathcal{T}[S=T+U] & (va)\,\text{vector addition}\\ \cdot:F\times\mathcal{T}\xrightarrow{\cdot}\mathcal{T}\Leftrightarrow\forall c\in F,\forall T\in\mathcal{T},\exists U\in\mathcal{T}[U=cT=c\cdot T] & (sm)\,\text{scalar multiplication}\\ \begin{cases}\forall S,T,U\in\mathcal{T}[S+(T+U)=(S+T)+U] & (a)\\ \forall T,U\in\mathcal{T}[T+U=U+T] & (c)\\ \exists! O\in\mathcal{T},\forall T\in\mathcal{T}[O+T=T] & (e)\\ \forall T\in\mathcal{T},\exists!-T\in\mathcal{T}[(-T)+T=O] & (i)\end{cases} & (va)\,\text{vector addition axioms}\\ \begin{cases}\forall b,c\in F,T\in\mathcal{T}[b(cT)=(bc)T] & (a)\\ \exists!1\in F,\forall T\in\mathcal{T}[1T=T] & (e)\\ \forall c\in F,T,U\in\mathcal{T}[c(T+U)=cT+cU] & (dv)\\ \forall b,c\in F,T\in\mathcal{T}[(b+c)T=bT+cT] & (ds)\end{cases} & (sm)\,\text{scalar multiplication axioms}\end{cases}$$
$$\Leftrightarrow\mathcal{T}=\mathcal{T}(F,+,\cdot)=(\mathcal{T},F,+,\cdot)\text{ is a vector space over the field } F$$
$$\Leftrightarrow\mathcal{T}\text{ is a vector space}$$

Selected proofs of 8 vector space axioms due to some trivial field and vector space properties:

$(va)\,(a)$

$$\begin{aligned}(S+(T+U))(\boldsymbol{v}) =& S(\boldsymbol{v})+(T+U)(\boldsymbol{v})\\ =& S(\boldsymbol{v})+T(\boldsymbol{v})+U(\boldsymbol{v})\\ =& (S+T)(\boldsymbol{v})+U(\boldsymbol{v})\\ =& ((S+T)+U)(\boldsymbol{v})\end{aligned}$$

$(va)\,(c)$

$$\begin{aligned}(T+U)(\boldsymbol{v}) =& T(\boldsymbol{v})+U(\boldsymbol{v})\\ =& U(\boldsymbol{v})+T(\boldsymbol{v})\\ =& (U+T)(\boldsymbol{v})\end{aligned}$$

$(va)\,(e)$

$$O(\boldsymbol{v})=0\boldsymbol{w}\in W$$

$$\begin{aligned}(O+T)(\boldsymbol{v}) =& O(\boldsymbol{v})+T(\boldsymbol{v})\\ =& 0\boldsymbol{w}+T(\boldsymbol{v})\\ =& T(\boldsymbol{v})\end{aligned}$$

$$O_1(\boldsymbol{v})-O_2(\boldsymbol{v})=0\boldsymbol{w}-0\boldsymbol{w}=0\boldsymbol{w}\Rightarrow O_1(\boldsymbol{v})=O_2(\boldsymbol{v})$$

$(sm)\,(dv)$

$$\begin{aligned}(c(T+U))(\boldsymbol{v}) =& c(T+U)(\boldsymbol{v})\\ =& c[T(\boldsymbol{v})+U(\boldsymbol{v})]\\ =& cT(\boldsymbol{v})+cU(\boldsymbol{v})\\ =& (cT+cU)(\boldsymbol{v})\end{aligned}$$

The set of all linear tranformations $\mathcal{T}$ is a vector space.

□

39.6 definition of dual space

$$\begin{aligned}&V^* = L(V,F)\\ =&f = \left\{ f_i \middle| \begin{cases} f_i : V \to F & \text{functional mapping vector to field scalar}\\ \begin{cases}\forall \boldsymbol{x},\boldsymbol{y} \in V\,[f_i(\boldsymbol{x}+\boldsymbol{y}) = f_i(\boldsymbol{x}) + f_i(\boldsymbol{y})]\\ \forall \boldsymbol{x} \in V, c \in F\,[f_i(c\boldsymbol{x}) = cf_i(\boldsymbol{x})]\end{cases} & (L)\,\text{linearity}\end{cases}\right\}\\ \Leftrightarrow& V^* \text{ is a dual space, a set of linear functionals } f_i \text{ mapping vectors in the vector space } V \text{ to scalars in the field } F\end{aligned}$$

vector space[40]

https://web.math.sinica.edu.tw/mathmedia/HTMLarticle18.jsp?mID=31304

https://web.math.sinica.edu.tw/mathmedia/author18.jsp?query_filter=%E9%BE%94%E6%98%87

39.7 double dual

double dual = second dual

https://ccjou.wordpress.com/2014/04/10/雙重對偶/

Chapter 40

vector space

https://ccjou.wordpress.com/2010/04/15/同構的向量空間/

40.1 What is a vector?

What is a vector? or **What is an element in a vector space?**

Binary operations defined on a vector space satisfying some properties is more important than **what is a vector**.

ultimate answer: double dual concept[40.4.1.2]

40.2 vector space definition

https://tex.stackexchange.com/a/141489 multiline node

group

1. vector addition clousre
2. identity $\mathbf{0}$
3. inverse
4. associativity

$\rightarrow$ abelian group

4 + 1. commutativity

$\rightarrow$ linear space

1. scalar multiplicaiton clousre
2. indentity 1
3. associativity
4. scalar distributivity
5. vector distributivity

Figure 40.1: vector space construction

$$\begin{cases} F \text{ is a field} & (f)\text{ field} \\ V \neq \emptyset & (ne)\text{ nonempty set} \\ + : V \times V = V^2 \xrightarrow{+} V \Leftrightarrow \forall \boldsymbol{u}, \boldsymbol{v} \in V, \exists! \boldsymbol{w} \in V \left[\boldsymbol{w} = \boldsymbol{u} + \boldsymbol{v}\right] & (va)\text{ vector addition} \\ \cdot : F \times V \xrightarrow{\cdot} V \Leftrightarrow \forall s \in F, \forall \boldsymbol{v} \in V, \exists! \boldsymbol{u} \in V \left[\boldsymbol{u} = s\boldsymbol{v} = s \cdot \boldsymbol{v}\right] & (sm)\text{ scalar multiplication} \\ \begin{cases} \exists! \mathbf{0} \in V, \forall \boldsymbol{v} \in V \left[\mathbf{0} + \boldsymbol{v} = \boldsymbol{v}\right] & (e)\text{ identity} \\ \forall \boldsymbol{v} \in V, \exists! -\boldsymbol{v} \in V \left[(-\boldsymbol{v}) + \boldsymbol{v} = \mathbf{0}\right] & (i)\text{ inverse} \\ \forall \boldsymbol{u}, \boldsymbol{v}, \boldsymbol{w} \in V \left[\boldsymbol{u} + (\boldsymbol{v} + \boldsymbol{w}) = (\boldsymbol{u} + \boldsymbol{v}) + \boldsymbol{w}\right] & (a)\text{ associativity} \\ \forall \boldsymbol{u}, \boldsymbol{v} \in V \left[\boldsymbol{u} + \boldsymbol{v} = \boldsymbol{v} + \boldsymbol{u}\right] & (c)\text{ commutativity} \end{cases} & (va)\text{ axioms} \\ \begin{cases} \exists! 1 \in F, \forall \boldsymbol{v} \in V \left[1\boldsymbol{v} = \boldsymbol{v}\right] & (e)\text{ identity} \\ \forall s, t \in F, \boldsymbol{v} \in V \left[s\left(t\boldsymbol{v}\right) = (st)\,\boldsymbol{v}\right] & (a)\text{ associativity} \\ \forall s, t \in F, \boldsymbol{v} \in V \left[(s + t)\,\boldsymbol{v} = s\boldsymbol{v} + t\boldsymbol{v}\right] & (ds)\text{ scalar distributivity} \\ \forall s \in F, \boldsymbol{u}, \boldsymbol{v} \in V \left[s\left(\boldsymbol{u} + \boldsymbol{v}\right) = s\boldsymbol{u} + s\boldsymbol{v}\right] & (dv)\text{ vector distributivity} \end{cases} & (sm)\text{ axioms} \end{cases}$$

$$\Leftrightarrow V = V\left(F, +, \cdot\right) = \left(V, F, +, \cdot\right) \text{ is a vector space over the field } F$$
$$\Leftrightarrow V \text{ is a vector space}$$

40.2.1 commutative group structure of vector space

(va) axioms = vector addition axioms

$$
\begin{aligned}
&V=(V,+)\ \text{ is a commutative group} \Leftrightarrow V=(V,+)\ \text{ is an abelian group}\\
\Leftrightarrow &\begin{cases} V=(V,+)=(V,+_V)\ \text{ is a group} & (g)\,\text{group}\\ \forall \boldsymbol{u},\boldsymbol{v}\in V\,[\boldsymbol{u}+\boldsymbol{v}=\boldsymbol{v}+\boldsymbol{u}] & (c)\,\text{commutativity}\end{cases}\\
\Leftrightarrow &\begin{cases} \begin{cases} +:V\times V=V^2\overset{+}{\to}V\Leftrightarrow\forall\boldsymbol{u},\boldsymbol{v}\in V,\exists!\boldsymbol{w}\in V\,[\boldsymbol{w}=\boldsymbol{u}+\boldsymbol{v}] & (cl)\,\text{closure}\\ \exists!\boldsymbol{0}\in V,\forall\boldsymbol{v}\in V\,[\boldsymbol{0}+\boldsymbol{v}=\boldsymbol{v}] & (e)\,\text{identity}\\ \forall\boldsymbol{v}\in V,\exists!-\boldsymbol{v}\in V\,[(-\boldsymbol{v})+\boldsymbol{v}=\boldsymbol{0}] & (i)\,\text{inverse}\\ \forall\boldsymbol{u},\boldsymbol{v},\boldsymbol{w}\in V\,[\boldsymbol{u}+(\boldsymbol{v}+\boldsymbol{w})=(\boldsymbol{u}+\boldsymbol{v})+\boldsymbol{w}] & (a)\,\text{associativity}\end{cases} & (g)\\ \forall\boldsymbol{u},\boldsymbol{v}\in V\,[\boldsymbol{u}+\boldsymbol{v}=\boldsymbol{v}+\boldsymbol{u}] & (c)\end{cases}
\end{aligned}
$$

$$
\begin{aligned}
&V\ \text{ is a vector space}\\
\Leftrightarrow &V=V(F,+,\cdot)=(V,F,+,\cdot)\ \text{ is a vector space over the field } F\\
\Leftrightarrow &\begin{cases} F\ \text{ is a field} & (f)\,\text{field}\\ V\neq\emptyset & (ne)\,\text{nonempty set}\\ V=(V,+)\ \text{ is a commutative group}\Leftrightarrow V=(V,+)\ \text{ is an abelian group} & (va)\,\text{vector addition}\\ \cdot:F\times V\overset{\cdot}{\to}V\Leftrightarrow\forall s\in F,\forall\boldsymbol{v}\in V,\exists!\boldsymbol{u}\in V\,[\boldsymbol{u}=s\boldsymbol{v}=s\cdot\boldsymbol{v}] & (sm)\,\text{scalar multiplication}\\ \begin{cases} \exists!1\in F,\forall\boldsymbol{v}\in V\,[1\boldsymbol{v}=\boldsymbol{v}] & (e)\,\text{identity}\\ \forall s,t\in F,\boldsymbol{v}\in V\,[s(t\boldsymbol{v})=(st)\boldsymbol{v}] & (a)\,\text{associativity}\\ \forall s,t\in F,\boldsymbol{v}\in V\,[(s+t)\boldsymbol{v}=s\boldsymbol{v}+t\boldsymbol{v}] & (ds)\,\text{scalar distributivity}\\ \forall s\in F,\boldsymbol{u},\boldsymbol{v}\in V\,[s(\boldsymbol{u}+\boldsymbol{v})=s\boldsymbol{u}+s\boldsymbol{v}] & (dv)\,\text{vector distributivity}\end{cases} & (sm)\,\text{axioms}\end{cases}\\
\Leftrightarrow &\begin{cases} F\ \text{ is a field} & (f)\,\text{field}\\ V\neq\emptyset & (ne)\,\text{nonempty set}\\ \begin{cases} V=(V,+)=(V,+_V)\ \text{ is a group} & (g)\,\text{group}\\ \forall\boldsymbol{u},\boldsymbol{v}\in V\,[\boldsymbol{u}+\boldsymbol{v}=\boldsymbol{v}+\boldsymbol{u}] & (c)\,\text{commutativity}\end{cases} & (va)\,\text{vector addition}\\ \cdot=\cdot_{F\times V}:F\times V\overset{\cdot}{\to}V\Leftrightarrow\forall s\in F,\forall\boldsymbol{v}\in V,\exists!\boldsymbol{u}\in V\,[\boldsymbol{u}=s\boldsymbol{v}=s\cdot\boldsymbol{v}] & (sm)\,\text{scalar multiplication}\\ \begin{cases} \exists!1\in F,\forall\boldsymbol{v}\in V\,[1\boldsymbol{v}=\boldsymbol{v}] & (e)\,\text{identity}\\ \forall s,t\in F,\boldsymbol{v}\in V\,[s(t\boldsymbol{v})=(st)\boldsymbol{v}] & (a)\,\text{associativity}\\ \forall s,t\in F,\boldsymbol{v}\in V\,[(s+t)\boldsymbol{v}=s\boldsymbol{v}+t\boldsymbol{v}] & (ds)\,\text{scalar distributivity}\\ \forall s\in F,\boldsymbol{u},\boldsymbol{v}\in V\,[s(\boldsymbol{u}+\boldsymbol{v})=s\boldsymbol{u}+s\boldsymbol{v}] & (dv)\,\text{vector distributivity}\end{cases} & (sm)\,\text{axioms}\end{cases}\\
\Leftrightarrow &\begin{cases} F=F(+_F,\cdot_F)=(F,+_F,\cdot_F)=(F,+,\cdot)\ \text{ is a field} & (f)\\ V\neq\emptyset & (ne)\\ \begin{cases} \begin{cases} +:V\times V=V^2\overset{+}{\to}V\Leftrightarrow\forall\boldsymbol{u},\boldsymbol{v}\in V,\exists!\boldsymbol{w}\in V\,[\boldsymbol{w}=\boldsymbol{u}+\boldsymbol{v}] & (cl)\,\text{closure}\\ \exists!\boldsymbol{0}\in V,\forall\boldsymbol{v}\in V\,[\boldsymbol{0}+\boldsymbol{v}=\boldsymbol{v}] & (e)\,\text{identity}\\ \forall\boldsymbol{v}\in V,\exists!-\boldsymbol{v}\in V\,[(-\boldsymbol{v})+\boldsymbol{v}=\boldsymbol{0}] & (i)\,\text{inverse}\\ \forall\boldsymbol{u},\boldsymbol{v},\boldsymbol{w}\in V\,[\boldsymbol{u}+(\boldsymbol{v}+\boldsymbol{w})=(\boldsymbol{u}+\boldsymbol{v})+\boldsymbol{w}] & (a)\,\text{associativity}\end{cases} & (g)\\ \forall\boldsymbol{u},\boldsymbol{v}\in V\,[\boldsymbol{u}+\boldsymbol{v}=\boldsymbol{v}+\boldsymbol{u}] & (c)\end{cases} & (va)\\ \begin{cases} \cdot:F\times V\overset{\cdot}{\to}V\Leftrightarrow\forall s\in F,\forall\boldsymbol{v}\in V,\exists!\boldsymbol{u}\in V\,[\boldsymbol{u}=s\boldsymbol{v}=s\cdot\boldsymbol{v}] & (cl)\,\text{closure}\\ \exists!1\in F,\forall\boldsymbol{v}\in V\,[1\boldsymbol{v}=\boldsymbol{v}] & (e)\,\text{identity}\\ \forall s,t\in F,\boldsymbol{v}\in V\,[s(t\boldsymbol{v})=s\cdot_{F\times V}(t\cdot_{F\times V}\boldsymbol{v})=(s\cdot_F t)\cdot_{F\times V}\boldsymbol{v}=(st)\boldsymbol{v}] & (a)\,\text{associativity}\\ \forall s,t\in F,\boldsymbol{v}\in V\,[(s+t)\boldsymbol{v}=(s+_F t)\boldsymbol{v}=s\boldsymbol{v}+_V t\boldsymbol{v}=s\boldsymbol{v}+t\boldsymbol{v}] & (ds)\,\text{scalar distributivity}\\ \forall s\in F,\boldsymbol{u},\boldsymbol{v}\in V\,[s(\boldsymbol{u}+\boldsymbol{v})=s\boldsymbol{u}+s\boldsymbol{v}] & (dv)\,\text{vector distributivity}\end{cases} & (sm)\end{cases}
\end{aligned}
$$

40.2.2 scalar distributivity

$(sm)\,(ds)$

$$\forall s,t\in F,\boldsymbol{v}\in V\,[(s+t)\boldsymbol{v}=s\boldsymbol{v}+t\boldsymbol{v}]$$

$$\forall s,t\in F,\boldsymbol{v}\in V\,[(s+_F t)\boldsymbol{v}=s\boldsymbol{v}+t\boldsymbol{v}]$$

$$\forall s,t\in F,\boldsymbol{v}\in V\,[(s+_F t)\boldsymbol{v}=s\boldsymbol{v}+_V t\boldsymbol{v}]$$

40.3 linearity

$$\begin{cases} f\left(x+y\right)=f\left(x\right)+f\left(y\right) & \text{additivity} \\ f\left(\lambda x\right)=\lambda f\left(x\right) & \text{homogeneity} \end{cases}$$
$$\Leftrightarrow f\left(\lambda x+y\right)=\lambda f\left(x\right)+f\left(y\right)$$
$$\Leftrightarrow f \text{ is linear}$$

40.3.1 linear structure of vector space

$$\forall s\in F,\boldsymbol{u},\boldsymbol{v}\in V\left[\boldsymbol{u}+s\boldsymbol{v}\in V\right]$$

$$\forall s\in F,\forall\boldsymbol{u},\boldsymbol{v}\in V\left[\boldsymbol{u}+s\boldsymbol{v}\in V\right]$$

$$\forall s\in F,\left\langle\boldsymbol{u},\boldsymbol{v}\right\rangle\in V^{2}\left[\boldsymbol{u}+s\boldsymbol{v}\in V\right]$$

$$\begin{cases} \forall\boldsymbol{u},\boldsymbol{v}\in V\left[\boldsymbol{u}+\boldsymbol{v}\in V\right] & \text{vector addition closure} \\ \forall s\in F,\boldsymbol{v}\in V\left[s\boldsymbol{v}\in V\right] & \text{scalar multiplication closure} \end{cases}$$
$$\Leftrightarrow\begin{cases} \forall\boldsymbol{u},\boldsymbol{v}\in V\left[\boldsymbol{u}+\boldsymbol{v}\in V\right] & (a)\,\text{additivity} \\ \forall s\in F,\boldsymbol{v}\in V\left[s\boldsymbol{v}\in V\right] & (h)\,\text{homogeneity} \end{cases}$$
$$\Leftrightarrow\forall s\in F,\boldsymbol{u},\boldsymbol{v}\in V\left[\boldsymbol{u}+s\boldsymbol{v}\in V\right]\quad (l)\,\text{linearity}$$

40.3.2 linear transformation or linear map

$$\begin{cases} V,W\text{ are vector spaces} \\ T:V\to W \\ \begin{cases} \forall\boldsymbol{u},\boldsymbol{v}\in V\left[T\left(\boldsymbol{u}+\boldsymbol{v}\right)=T\left(\boldsymbol{u}\right)+T\left(\boldsymbol{v}\right)\right] & (a)\,\text{additivity} \\ \forall\boldsymbol{v}\in V,c\in F\left[T\left(c\boldsymbol{v}\right)=cT\left(\boldsymbol{v}\right)\right] & (h)\,\text{homogeneity} \end{cases} \end{cases}\quad (L)$$
$$\Leftrightarrow\begin{cases} V,W\text{ are vector spaces} \\ T:V\to W \\ \forall\boldsymbol{u},\boldsymbol{v}\in V,c\in F\left[T\left(\boldsymbol{u}+c\boldsymbol{v}\right)=T\left(\boldsymbol{u}\right)+cT\left(\boldsymbol{v}\right)\right] \quad (l)\,\text{linearity} \end{cases}$$
$$\Leftrightarrow T\text{ is a linear map from }V\text{ to }W$$
$$\Leftrightarrow T\text{ is a linear tranformation}$$

40.4 vector space example

- arrow vector
- number
  - integer
  - real
  - complex
  - quaternion
- function
  - polynomial function
  - continuous function
- matrix
  - real matrix
  - complex matrix
- reciprocal space

applications in different disciplines

- math
  - recursive number series
  - Fourier series
- physics
  - electrical circuit: linear response / superposition theorem in linear circuit / linear network
- chemistry

– balancing chemical equation

40.4.1 reciprocal space

reciprocal space = 倒易空間

$$\begin{cases} \boldsymbol{e}_1 = \boldsymbol{a} & \boldsymbol{a} \times \boldsymbol{b} \neq \boldsymbol{0} \\ \boldsymbol{e}_2 = \boldsymbol{b} & \boldsymbol{b} \times \boldsymbol{c} \neq \boldsymbol{0} \\ \boldsymbol{e}_3 = \boldsymbol{c} & \boldsymbol{c} \times \boldsymbol{a} \neq \boldsymbol{0} \end{cases} \Rightarrow \begin{cases} \boldsymbol{e}_1' = \dfrac{\boldsymbol{b} \times \boldsymbol{c}}{\Omega} \\ \boldsymbol{e}_2' = \dfrac{\boldsymbol{c} \times \boldsymbol{a}}{\Omega} \\ \boldsymbol{e}_3' = \dfrac{\boldsymbol{a} \times \boldsymbol{b}}{\Omega} \end{cases},$$

$$\Omega = \boldsymbol{a} \cdot (\boldsymbol{b} \times \boldsymbol{c}) = \boldsymbol{b} \cdot (\boldsymbol{c} \times \boldsymbol{a}) = \boldsymbol{c} \cdot (\boldsymbol{a} \times \boldsymbol{b})$$

reciprocal space as dual space and contravariant vector

$$\begin{aligned} &\operatorname{span}\{\boldsymbol{e}_1, \boldsymbol{e}_2, \boldsymbol{e}_3\} = \operatorname{span}\{\boldsymbol{a}, \boldsymbol{b}, \boldsymbol{c}\} = V \\ = \mathbb{R}^3 =&\operatorname{span}\{\boldsymbol{e}_1', \boldsymbol{e}_2', \boldsymbol{e}_3'\} = \operatorname{span}\left\{\frac{\boldsymbol{b} \times \boldsymbol{c}}{\Omega}, \frac{\boldsymbol{c} \times \boldsymbol{a}}{\Omega}, \frac{\boldsymbol{a} \times \boldsymbol{b}}{\Omega}\right\} \\ =&\operatorname{span}\{\boldsymbol{e}^1, \boldsymbol{e}^2, \boldsymbol{e}^3\} = \operatorname{span}\{\boldsymbol{e}^*\}_{*\in\{1,2,3\}} = V^* \end{aligned}$$

40.4.1.1 Kronecker delta

$$\begin{pmatrix} \boldsymbol{e}_1 \cdot \boldsymbol{e}_1' & \boldsymbol{e}_1 \cdot \boldsymbol{e}_2' & \boldsymbol{e}_1 \cdot \boldsymbol{e}_3' \\ \boldsymbol{e}_2 \cdot \boldsymbol{e}_1' & \boldsymbol{e}_2 \cdot \boldsymbol{e}_2' & \boldsymbol{e}_2 \cdot \boldsymbol{e}_3' \\ \boldsymbol{e}_3 \cdot \boldsymbol{e}_1' & \boldsymbol{e}_3 \cdot \boldsymbol{e}_2' & \boldsymbol{e}_3 \cdot \boldsymbol{e}_3' \end{pmatrix} = [\delta_{ij}] = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} = \begin{pmatrix} \boldsymbol{e}^1 \cdot \boldsymbol{e}_1 & \boldsymbol{e}^1 \cdot \boldsymbol{e}_2 & \boldsymbol{e}^1 \cdot \boldsymbol{e}_3 \\ \boldsymbol{e}^2 \cdot \boldsymbol{e}_1 & \boldsymbol{e}^2 \cdot \boldsymbol{e}_2 & \boldsymbol{e}^2 \cdot \boldsymbol{e}_3 \\ \boldsymbol{e}^3 \cdot \boldsymbol{e}_1 & \boldsymbol{e}^3 \cdot \boldsymbol{e}_2 & \boldsymbol{e}^3 \cdot \boldsymbol{e}_3 \end{pmatrix}$$

Kronecker delta

$$\boldsymbol{e}_i \cdot \boldsymbol{e}_j' = \delta_{ij} = \begin{cases} 1 & i = j \\ 0 & i \neq j \end{cases}$$

Kronecker delta tensor = Kronecker tensor

$$\boldsymbol{e}^i (\boldsymbol{e}_j) = \boldsymbol{e}^i \cdot \boldsymbol{e}_j = \delta^i_j = \delta^i{}_j = \begin{cases} 1 & i = j \\ 0 & i \neq j \end{cases}$$

$$\boldsymbol{v} = v_a \boldsymbol{a} + v_b \boldsymbol{b} + v_c \boldsymbol{c} = v_1 \boldsymbol{e}_1 + v_2 \boldsymbol{e}_2 + v_3 \boldsymbol{e}_3$$

$$\boldsymbol{e}^1 \cdot \boldsymbol{v} = \boldsymbol{v} \cdot \boldsymbol{e}_1' = v_1 \boldsymbol{e}_1 \cdot \boldsymbol{e}_1' + v_2 \boldsymbol{e}_2 \cdot \boldsymbol{e}_1' + v_3 \boldsymbol{e}_3 \cdot \boldsymbol{e}_1' = v_1$$

$$\boldsymbol{e}^2 \cdot \boldsymbol{v} = \boldsymbol{v} \cdot \boldsymbol{e}_2' = v_1 \boldsymbol{e}_1 \cdot \boldsymbol{e}_2' + v_2 \boldsymbol{e}_2 \cdot \boldsymbol{e}_2' + v_3 \boldsymbol{e}_3 \cdot \boldsymbol{e}_2' = v_2$$

$$\boldsymbol{e}^3 \cdot \boldsymbol{v} = \boldsymbol{v} \cdot \boldsymbol{e}_3' = v_1 \boldsymbol{e}_1 \cdot \boldsymbol{e}_3' + v_2 \boldsymbol{e}_2 \cdot \boldsymbol{e}_3' + v_3 \boldsymbol{e}_3 \cdot \boldsymbol{e}_3' = v_3$$

$$\begin{aligned} \boldsymbol{v} =& v_1 \boldsymbol{e}_1 + v_2 \boldsymbol{e}_2 + v_3 \boldsymbol{e}_3 \\ =& (\boldsymbol{v} \cdot \boldsymbol{e}_1') \boldsymbol{e}_1 + (\boldsymbol{v} \cdot \boldsymbol{e}_2') \boldsymbol{e}_2 + (\boldsymbol{v} \cdot \boldsymbol{e}_3') \boldsymbol{e}_3 \\ =& (\boldsymbol{e}^1 \cdot \boldsymbol{v}) \boldsymbol{e}_1 + (\boldsymbol{e}^2 \cdot \boldsymbol{v}) \boldsymbol{e}_2 + (\boldsymbol{e}^3 \cdot \boldsymbol{v}) \boldsymbol{e}_3 \\ =& \boldsymbol{e}^1 (\boldsymbol{v}) \boldsymbol{e}_1 + \boldsymbol{e}^2 (\boldsymbol{v}) \boldsymbol{e}_2 + \boldsymbol{e}^3 (\boldsymbol{v}) \boldsymbol{e}_3 \end{aligned}$$

$$\begin{cases} \boldsymbol{e}^1 (\boldsymbol{v}) = \boldsymbol{e}^1 \cdot \boldsymbol{v} = \boldsymbol{v} \cdot \boldsymbol{e}_1' = v_1 \\ \boldsymbol{e}^2 (\boldsymbol{v}) = \boldsymbol{e}^2 \cdot \boldsymbol{v} = \boldsymbol{v} \cdot \boldsymbol{e}_2' = v_2 \\ \boldsymbol{e}^3 (\boldsymbol{v}) = \boldsymbol{e}^3 \cdot \boldsymbol{v} = \boldsymbol{v} \cdot \boldsymbol{e}_3' = v_3 \end{cases}$$

$$\boldsymbol{e}^i(\boldsymbol{e}_j) = \boldsymbol{e}^i \cdot \boldsymbol{e}_j = \delta^i_j = \delta^i{}_j = \begin{cases} 1 & i = j \\ 0 & i \neq j \end{cases}$$

reciprocal space is a dual space of its original vector space

$$\begin{aligned} V =& \text{span}\{\boldsymbol{e}_1, \boldsymbol{e}_2, \boldsymbol{e}_3\} = \{v_1\boldsymbol{e}_1 + v_2\boldsymbol{e}_2 + v_3\boldsymbol{e}_3\} \\ =& \left\{\sum_{j=1}^{3} v_j\boldsymbol{e}_j\right\} = \left\{v_j\boldsymbol{e}_j \middle| \begin{cases} v_j \in F \\ \boldsymbol{e}_j \in F^3 \end{cases}\right\} = \{\boldsymbol{v}|\boldsymbol{v} \in V\} \\ V^* =& \text{span}\{\boldsymbol{e}^1, \boldsymbol{e}^2, \boldsymbol{e}^3\} = \{v^{*1}\boldsymbol{e}^1 + v^{*2}\boldsymbol{e}^2 + v^{*3}\boldsymbol{e}^3\} \\ =& \left\{\sum_{i=1}^{3} v^{*i}\boldsymbol{e}^i\right\} = \left\{v^{*i}\boldsymbol{e}^i \middle| \begin{cases} v^{*i} \in F \\ \boldsymbol{e}^i \in F^3 \end{cases}\right\} = \{\boldsymbol{v}^*|\boldsymbol{v}^* \in V^*\} \end{aligned}$$

$$\begin{aligned} \boldsymbol{v}^*(\boldsymbol{v}) &= (v^{*i}\boldsymbol{e}^i)(\boldsymbol{v}), \boldsymbol{v} \in V \\ &= (v^{*1}\boldsymbol{e}^1 + v^{*2}\boldsymbol{e}^2 + v^{*3}\boldsymbol{e}^3)(\boldsymbol{v}) \\ &= v^{*1}\boldsymbol{e}^1(\boldsymbol{v}) + v^{*2}\boldsymbol{e}^2(\boldsymbol{v}) + v^{*3}\boldsymbol{e}^3(\boldsymbol{v}) \\ &= v^{*1}v_1 + v^{*2}v_2 + v^{*3}v_3 \in F \end{aligned}$$

element in dual space is a functional or mapping from its original vector space to the field

$$\boldsymbol{v}^* : V \to F$$

$$V \xrightarrow{\boldsymbol{v}^*} F$$

$V^* = \{\boldsymbol{v}^*|\boldsymbol{v}^* : V \to F\}$

$$\begin{array}{cccccccccccc} & & & V & = \{ & \boldsymbol{e}_1 & \boldsymbol{e}_2 & \boldsymbol{e}_3 & \boldsymbol{v} & \cdots\} \\ & \boldsymbol{e}^1 & : & \downarrow & & \downarrow & \downarrow & \downarrow & \downarrow & \\ V^* = \{ & \boldsymbol{e}^2 & \} & F & \supseteq \{ & 1 & 0 & 0 & v_1 & \cdots\} \\ & \boldsymbol{e}^3 & & V & = \{ & \boldsymbol{e}_1 & \boldsymbol{e}_2 & \boldsymbol{e}_3 & \boldsymbol{v} & \cdots\} \\ & \boldsymbol{v}^* & : & \downarrow & & \downarrow & \downarrow & \downarrow & \downarrow & \\ & \vdots & & F & \supseteq \{ & v^{*1} & v^{*2} & v^{*3} & v^{*i}v_i & \cdots\} \end{array}$$

$$\begin{aligned} V^* = \{\boldsymbol{v}^*|\boldsymbol{v}^* \in V^*\} &= \{\boldsymbol{v}^*|\boldsymbol{v}^* : V \to F\} \\ &= \left\{\boldsymbol{v}^* \middle| V \xrightarrow{\boldsymbol{v}^*} F\right\} \\ &= \{\boldsymbol{\omega}|\boldsymbol{\omega} : V \to F\} \\ &= \left\{\omega^i\boldsymbol{e}^i \middle| \begin{cases} \omega^i \in F \\ \boldsymbol{e}^i \in F^3 \end{cases}\right\} \end{aligned}$$

By defining vector addition and scalar multiplication on the dual space

$$\begin{cases} + : V^* \times V^* \to V^* \Leftrightarrow \forall \boldsymbol{\omega}_1, \boldsymbol{\omega}_2 \in V^*, \exists! (\boldsymbol{\omega}_1 + \boldsymbol{\omega}_2) \in V^* [(\boldsymbol{\omega}_1 + \boldsymbol{\omega}_2)(\boldsymbol{v}) = \boldsymbol{\omega}_1(\boldsymbol{v}) + \boldsymbol{\omega}_2(\boldsymbol{v})] \\ \cdot : F \times V^* \to V^* \Leftrightarrow \forall k \in F, \forall \boldsymbol{\omega} \in V^*, \exists! (k\boldsymbol{\omega}) \in V^* [(k\boldsymbol{\omega})(\boldsymbol{v}) = k \cdot \boldsymbol{\omega}(\boldsymbol{v})] \\ \forall \boldsymbol{\omega} \in V^*, \exists! \boldsymbol{0} \in V^* [(\boldsymbol{\omega} + \boldsymbol{0})(\boldsymbol{v}) = \boldsymbol{\omega}(\boldsymbol{v}) + \boldsymbol{0}(\boldsymbol{v}) = \boldsymbol{\omega}(\boldsymbol{v})] \end{cases}$$

the dual space also becomes a vector space.

40.4.1.2 double dual concept

double dual space = second dual space

$$
\begin{aligned}
V^{**} &= (V^*)^* \\
&= \{\boldsymbol{\omega}^* | \boldsymbol{\omega}^* : V^* \to F\} \\
&= \{\boldsymbol{\omega}^* | \boldsymbol{\omega}^* \in V^{**}\}
\end{aligned}
$$

$$
\begin{aligned}
V^{**} = (V^*)^* &= \operatorname{span}\{\boldsymbol{e}^\mu\}^*_{\mu\in\{1,2,3\}} \\
&= \operatorname{span}\{\boldsymbol{e}^1, \boldsymbol{e}^2, \boldsymbol{e}^3\}^* \\
&= \operatorname{span}\{\boldsymbol{e}^{1*}, \boldsymbol{e}^{2*}, \boldsymbol{e}^{3*}\} \\
&= \operatorname{span}\{\boldsymbol{e}^{\nu*}\}_{\nu\in\{1,2,3\}}
\end{aligned}
$$

$$
\begin{aligned}
\boldsymbol{\omega}^*(\boldsymbol{\omega}) &= (\omega^{*\nu}\boldsymbol{e}^{\nu*})(\boldsymbol{\omega}), \boldsymbol{\omega} \in V^* \\
&= (\omega^{*1}\boldsymbol{e}^{1*} + \omega^{*2}\boldsymbol{e}^{2*} + \omega^{*3}\boldsymbol{e}^{3*})(\boldsymbol{\omega}) \\
&= \omega^{*1}\boldsymbol{e}^{1*}(\boldsymbol{\omega}) + \omega^{*2}\boldsymbol{e}^{2*}(\boldsymbol{\omega}) + \omega^{*3}\boldsymbol{e}^{3*}(\boldsymbol{\omega}) \\
&= \omega^{*1}\omega_1 + \omega^{*2}\omega_2 + \omega^{*3}\omega_3 \in F
\end{aligned}
$$

$V^{**} = \{\boldsymbol{\omega}^* | \boldsymbol{\omega}^* : V^* \to F\}$

$$
\begin{array}{cccccccccc}
 & & & V^* & =\{ & \boldsymbol{e}^1 & \boldsymbol{e}^2 & \boldsymbol{e}^3 & \boldsymbol{\omega} & \cdots\} \\
 & \boldsymbol{e}^{1*} & : & \downarrow & & \downarrow & \downarrow & \downarrow & \downarrow & \\
V^{**} = \{ & \boldsymbol{e}^{2*} & \} & F & \supseteq\{ & 1 & 0 & 0 & \omega^1 & \cdots\} \\
 & \boldsymbol{e}^{3*} & & V^* & =\{ & \boldsymbol{e}_1 & \boldsymbol{e}_2 & \boldsymbol{e}_3 & \boldsymbol{v} & \cdots\} \\
 & \boldsymbol{\omega}^* & : & \downarrow & & \downarrow & \downarrow & \downarrow & \downarrow & \\
 & \vdots & & F & \supseteq\{ & \omega^{1*} & \omega^{2*} & \omega^{3*} & \omega^{\mu*}\omega^\mu & \cdots\}
\end{array}
$$

$$
\begin{cases}
\boldsymbol{e}^{1*}(\boldsymbol{\omega}) = \boldsymbol{e}^{1*} \cdot \boldsymbol{\omega} = \boldsymbol{\omega} \cdot \boldsymbol{e}^{1*} = \boldsymbol{\omega}(\boldsymbol{e}^{1*}) \\
\boldsymbol{e}^{2*}(\boldsymbol{\omega}) = \boldsymbol{e}^{2*} \cdot \boldsymbol{\omega} = \boldsymbol{\omega} \cdot \boldsymbol{e}^{2*} = \boldsymbol{\omega}(\boldsymbol{e}^{2*}) \\
\boldsymbol{e}^{3*}(\boldsymbol{\omega}) = \boldsymbol{e}^{3*} \cdot \boldsymbol{\omega} = \boldsymbol{\omega} \cdot \boldsymbol{e}^{3*} = \boldsymbol{\omega}(\boldsymbol{e}^{3*})
\end{cases}
$$

$$
\boldsymbol{\omega}^*(\boldsymbol{\omega}) = \boldsymbol{\omega}^* \cdot \boldsymbol{\omega} = \boldsymbol{\omega} \cdot \boldsymbol{\omega}^* = \boldsymbol{\omega}(\boldsymbol{\omega}^*)
$$

i.e. $\boldsymbol{f}$ acts on $\boldsymbol{x}$ equivalent to $\boldsymbol{x}$ acts on $\boldsymbol{f}$

$$
\boldsymbol{x}(\boldsymbol{f}) = \boldsymbol{x} \cdot \boldsymbol{f} = \boldsymbol{f} \cdot \boldsymbol{x} = \boldsymbol{f}(\boldsymbol{x})
$$

$$
\boldsymbol{e}^\mu(\boldsymbol{e}_\nu) = \boldsymbol{e}^\mu \cdot \boldsymbol{e}_\nu = \delta^\mu_\nu = \delta^\mu{}_\nu = \begin{cases} 1 & \mu = \nu \\ 0 & \mu \neq \nu \end{cases}
$$

$$
\begin{aligned}
\boldsymbol{e}^{\nu*}(\boldsymbol{e}^\mu) \overset{\text{def.}}{=} \boldsymbol{e}_\nu \cdot \boldsymbol{e}^\mu &= \boldsymbol{e}^\mu \cdot \boldsymbol{e}_\nu = \boldsymbol{e}^\mu(\boldsymbol{e}_\nu) \\
&\Downarrow \\
V^{**} = \operatorname{span}\{\boldsymbol{e}^{\nu*}\}_{\nu\in\{1,2,3\}} &\cong \operatorname{span}\{\boldsymbol{e}_\nu\}_{\nu\in\{1,2,3\}} = V \\
V^{**} &\cong V \\
&\Downarrow \begin{cases} V^{**} \cong V & V, V^{**} \text{ are isomorphism} \\ & \text{independent of choice of bases} \end{cases} \\
V, V^{**} &\text{ are naturally isomorphism}
\end{aligned}
$$

$V^{**} = \{\boldsymbol{\omega}^* | \boldsymbol{\omega}^* : V^* \to F\} \cong V = \{v | v : V^* \to F\}$

$$\begin{array}{cccccccccc}
 & & & V^* & =\{ & \boldsymbol{e}^1 & \boldsymbol{e}^2 & \boldsymbol{e}^3 & \boldsymbol{\omega} & \cdots\} \\
 & \boldsymbol{e}^{1*} & : & \downarrow & & \downarrow & \downarrow & \downarrow & \downarrow & \\
V^{**}=\{ & \boldsymbol{e}^{2*} & \} & F & \supseteq\{ & 1 & 0 & 0 & \omega^1 & \cdots\} \\
 & \boldsymbol{e}^{3*} & & V^* & =\{ & \boldsymbol{e}^1 & \boldsymbol{e}^2 & \boldsymbol{e}^3 & \boldsymbol{\omega} & \cdots\} \\
 & \boldsymbol{\omega}^* & : & \downarrow & & \downarrow & \downarrow & \downarrow & \downarrow & \\
 & \vdots & & F & \supseteq\{ & \omega^{1*} & \omega^{2*} & \omega^{3*} & \omega^{\mu*}\omega^{\mu} & \cdots\} \\
 & & & V^* & =\{ & \boldsymbol{e}^1 & \boldsymbol{e}^2 & \boldsymbol{e}^3 & \boldsymbol{v}^* & \cdots\} \\
 & \boldsymbol{e}_1 & : & \downarrow & & \downarrow & \downarrow & \downarrow & \downarrow & \\
\cong V=\{ & \boldsymbol{e}_2 & \} & F & \supseteq\{ & 1 & 0 & 0 & v^{*1} & \cdots\} \\
 & \boldsymbol{e}_3 & & V^* & =\{ & \boldsymbol{e}^1 & \boldsymbol{e}^2 & \boldsymbol{e}^3 & \boldsymbol{v}^* & \cdots\} \\
 & \boldsymbol{v} & : & \downarrow & & \downarrow & \downarrow & \downarrow & \downarrow & \\
 & \vdots & & F & \supseteq\{ & v_1 & v_2 & v_3 & v_\mu v^{*\mu} & \cdots\}
\end{array}$$

$$V \cong V^{**}$$

$$V \cong V^{**} = \{\boldsymbol{\omega}^* | \boldsymbol{\omega}^* : V^* \to F\}$$

$$V = \{v | v : V^* \to F\}$$

i.e. **vector space is a set of functionals or mappings from its dual space to the field**, answering What is a vector?[40.1], and satifying Fig: 40.1.

40.5 field

https://web.math.sinica.edu.tw/math_media/d312/31202.pdf

40.6 module

https://web.math.sinica.edu.tw/mathmedia/HTMLarticle18.jsp?mID=31304

40.7 subspace

Chapter 41

$\mathrm{d}f$

41.1 $\mathrm{d}f$ decomposed with partials as a set of basis in vector space

$$f = \{f_i\} = \{f_1, f_2, \cdots\} = \{f, g, \cdots\}$$

$$\boldsymbol{v} : f \to F$$

$$\boldsymbol{v}(af + bg) = a\boldsymbol{v}(f) + b\boldsymbol{v}(g)$$

$$\boldsymbol{v}(fg) = f|_P \boldsymbol{v}(g) + \boldsymbol{v}(f)\, g|_P$$

$$\frac{\mathrm{d}}{\mathrm{d}x}\left[f(x)\, g(x)\right]|_{x=x_0} = f(x_0)\frac{\mathrm{d}}{\mathrm{d}x}g(x)\,|_{x=x_0} + \frac{\mathrm{d}}{\mathrm{d}x}f(x)\,|_{x=x_0}\, g(x_0)$$

$$V = \{\boldsymbol{v} | \boldsymbol{v} : f \to F\}$$

$$\begin{aligned} f =& f(\boldsymbol{x}) \\ =& f(x_1, \cdots, x_j, \cdots, x_n) \\ =& f(x^1, \cdots, x^j, \cdots, x^n) \end{aligned}$$

$$\boldsymbol{x} = \langle x^1, \cdots, x^j, \cdots, x^n \rangle$$

$$\boldsymbol{x}(t) = \langle x^1(t), \cdots, x^j(t), \cdots, x^n(t) \rangle$$

$$\begin{aligned} \frac{\mathrm{d}f}{\mathrm{d}t} =& \frac{\mathrm{d}x^1}{\mathrm{d}t}\frac{\partial f}{\partial x^1} + \cdots + \frac{\mathrm{d}x^j}{\mathrm{d}t}\frac{\partial f}{\partial x^j} + \cdots + \frac{\mathrm{d}x^n}{\mathrm{d}t}\frac{\partial f}{\partial x^n} \\ =& \cdots + \frac{\mathrm{d}x^j}{\mathrm{d}t}\frac{\partial f}{\partial x^j} + \cdots = \frac{\mathrm{d}x^j}{\mathrm{d}t}\frac{\partial f}{\partial x^j} = \frac{\mathrm{d}x^j}{\mathrm{d}t}\partial_j f \end{aligned}$$

$$\begin{aligned} V =& \mathrm{span}\{\boldsymbol{e}_1, \cdots, \boldsymbol{e}_j, \cdots, \boldsymbol{e}_n\} \\ =& \mathrm{span}\left\{\frac{\partial}{\partial x^1}|_P, \cdots, \frac{\partial}{\partial x^j}|_P, \cdots, \frac{\partial}{\partial x^n}|_P\right\} \\ =& \mathrm{span}\{\boldsymbol{\partial}_1, \cdots, \boldsymbol{\partial}_j, \cdots, \boldsymbol{\partial}_n\} \\ =& \left\{\boldsymbol{\partial}_t \middle| \boldsymbol{\partial}_t = a_j\boldsymbol{e}_j = a_j\boldsymbol{\partial}_j = a_j\frac{\partial}{\partial x^j}|_P\right\} \\ =& \left\{\frac{\partial}{\partial t}|_P \middle| \frac{\partial}{\partial t}|_P = a_1\frac{\partial}{\partial x^1}|_P + \cdots + a_j\frac{\partial}{\partial x^j}|_P + \cdots + a_n\frac{\partial}{\partial x^n}|_P\right\} \end{aligned}$$

41.2 dual space of span of partials

$$V^* = \{\boldsymbol{\omega}_f | \boldsymbol{\omega}_f : V \to F\}$$

$$\boldsymbol{\omega}_f(\boldsymbol{e}_j) = \boldsymbol{\omega}_f(\boldsymbol{\partial}_j) = \boldsymbol{\omega}_f\left(\frac{\partial}{\partial x^j}|_P\right) = \frac{\partial f}{\partial x^j}|_P \in F$$

$$\begin{aligned}\boldsymbol{\omega}_{fg}(\boldsymbol{\partial}_j) = \frac{\partial fg}{\partial x^j}|_P =& f|_P \frac{\partial g}{\partial x^j}|_P + \frac{\partial f}{\partial x^j}|_P g|_P \\ =& f|_P \boldsymbol{\omega}_g(\boldsymbol{\partial}_j) + \boldsymbol{\omega}_f(\boldsymbol{\partial}_j) g|_P\end{aligned}$$

$$\boldsymbol{\omega}_{x^i}(\boldsymbol{\partial}_j) = \boldsymbol{\omega}_{x^i}\left(\frac{\partial}{\partial x^j}|_P\right) = \frac{\partial x^i}{\partial x^j}|_P = \delta_{ij} = \begin{cases}1 & i = j \\ 0 & i \neq j\end{cases}$$

$$\begin{aligned}V^* =& \{\boldsymbol{\omega}_f | \boldsymbol{\omega}_f : V \to F\} = \left\{\boldsymbol{\omega}_f \middle| \begin{cases}\boldsymbol{\omega}_f(\boldsymbol{e}_j) = \boldsymbol{\omega}_f(\boldsymbol{\partial}_j) = \boldsymbol{\omega}_f\left(\frac{\partial}{\partial x^j}|_P\right) = \frac{\partial f}{\partial x^j}|_P \in F \\ \boldsymbol{\omega}_{fg}(\boldsymbol{\partial}_j) = f|_P \boldsymbol{\omega}_g(\boldsymbol{\partial}_j) + \boldsymbol{\omega}_f(\boldsymbol{\partial}_j) g|_P \\ \boldsymbol{\omega}_{x^i}(\boldsymbol{\partial}_j) = \boldsymbol{\omega}_{x^i}\left(\frac{\partial}{\partial x^j}|_P\right) = \frac{\partial x^i}{\partial x^j}|_P = \delta_{ij} = \begin{cases}1 & i = j \\ 0 & i \neq j\end{cases}\end{cases}\right\} \\ =& \{\mathrm{d}f | \mathrm{d}f : V \to F\} = \left\{\mathrm{d}f \middle| \begin{cases}\mathrm{d}f(\boldsymbol{e}_j) = \mathrm{d}f(\boldsymbol{\partial}_j) = \mathrm{d}f\left(\frac{\partial}{\partial x^j}|_P\right) = \frac{\partial f}{\partial x^j}|_P \in F \\ \mathrm{d}fg(\boldsymbol{\partial}_j) = f|_P(\mathrm{d}g) + (\mathrm{d}f) g|_P \\ \mathrm{d}x^i(\boldsymbol{\partial}_j) = \mathrm{d}x^i\left(\frac{\partial}{\partial x^j}|_P\right) = \frac{\partial x^i}{\partial x^j}|_P = \delta_{ij} = \begin{cases}1 & i = j \\ 0 & i \neq j\end{cases}\end{cases}\right\}\end{aligned}$$

$$\mathrm{d}x^i\left(\frac{\partial}{\partial x^j}|_P\right) = \delta_{ij} = \boldsymbol{e}^i \cdot \boldsymbol{e}_j \Rightarrow \begin{cases}\boldsymbol{e}^i = \mathrm{d}x^i \\ \boldsymbol{e}_j = \frac{\partial}{\partial x^j}|_P\end{cases}$$

$$\begin{aligned}V^* =& \{\mathrm{d}f | \mathrm{d}f : V \to F\} = \left\{\mathrm{d}f \middle| \begin{cases}\mathrm{d}f(\boldsymbol{e}_j) = \mathrm{d}f(\boldsymbol{\partial}_j) = \mathrm{d}f\left(\frac{\partial}{\partial x^j}|_P\right) = \frac{\partial f}{\partial x^j}|_P \in F \\ \mathrm{d}fg(\boldsymbol{\partial}_j) = f|_P(\mathrm{d}g) + (\mathrm{d}f) g|_P \\ \mathrm{d}x^i(\boldsymbol{\partial}_j) = \mathrm{d}x^i\left(\frac{\partial}{\partial x^j}|_P\right) = \frac{\partial x^i}{\partial x^j}|_P = \delta_{ij} = \begin{cases}1 & i = j \\ 0 & i \neq j\end{cases}\end{cases}\right\} \\ =& \mathrm{span}\{\mathrm{d}x^1, \cdots, \mathrm{d}x^i, \cdots, \mathrm{d}x^n\} = \mathrm{span}\{\boldsymbol{e}^1, \cdots, \boldsymbol{e}^j, \cdots, \boldsymbol{e}^n\}\end{aligned}$$

41.3 directional derivative

$$\begin{aligned}\mathrm{d}f(\boldsymbol{v}) =& \mathrm{d}f(v_j \boldsymbol{e}_j) = v_j \mathrm{d}f(\boldsymbol{e}_j) \\ =& v_j \mathrm{d}f(\boldsymbol{\partial}_j) = v_j \frac{\partial f}{\partial x^j}|_P \\ =& v_1 \frac{\partial f}{\partial x^1}|_P + + v_j \frac{\partial f}{\partial x^j}|_P + \cdots + v_n \frac{\partial f}{\partial x^n}|_P \\ =& \begin{pmatrix}v_1 & \cdots & v_j & \cdots & v_n\end{pmatrix} \boldsymbol{\nabla} f\end{aligned}$$

$$\overset{\frown}{PQ} = C(t) - C(0) = Q - P$$

$$\boldsymbol{v} = \frac{\partial}{\partial t}|_P$$

$$\begin{aligned}\mathrm{d}f(s\boldsymbol{v}) =& \mathrm{d}f\left(s\frac{\partial}{\partial t}|_P\right) = s\frac{\partial f}{\partial t}|_P \\ =& s\boldsymbol{v}(f) = s \cdot \lim_{t \to 0} \frac{f(C(t)) - f(C(0))}{t} \\ \approx& s \cdot \frac{f(Q) - f(P)}{s} = f(Q) - f(P) = \Delta f\end{aligned}$$

41.4 coefficient of linear combination for vector space and dual space

$$
\begin{aligned}
V =& \{\boldsymbol{v} | \boldsymbol{v} : f \to F\} \\
=& \text{span}\{\boldsymbol{e}_1, \cdots, \boldsymbol{e}_j, \cdots, \boldsymbol{e}_n\} \\
=& \text{span}\left\{\frac{\partial}{\partial x^1}|_P, \cdots, \frac{\partial}{\partial x^j}|_P, \cdots, \frac{\partial}{\partial x^n}|_P\right\} = \text{span}\{\boldsymbol{\partial}_1, \cdots, \boldsymbol{\partial}_j, \cdots, \boldsymbol{\partial}_n\} \\
=& \left\{\boldsymbol{\partial}_t \middle| \boldsymbol{\partial}_t = a_j \boldsymbol{e}_j = a_j \boldsymbol{\partial}_j = a_j \frac{\partial}{\partial x^j}|_P\right\} \\
=& \left\{\frac{\partial}{\partial t}|_P \middle| \frac{\partial}{\partial t}|_P = a_1 \frac{\partial}{\partial x^1}|_P + \cdots + a_j \frac{\partial}{\partial x^j}|_P + \cdots + a_n \frac{\partial}{\partial x^n}|_P\right\} \\
V^* =& \{\mathrm{d}f | \mathrm{d}f : V \to F\} \\
=& \text{span}\{\boldsymbol{e}^1, \cdots, \boldsymbol{e}^i, \cdots, \boldsymbol{e}^n\} \\
=& \text{span}\{\mathrm{d}x^1, \cdots, \mathrm{d}x^i, \cdots, \mathrm{d}x^n\} \\
=& \{\mathrm{d}f | \mathrm{d}f = b^i \boldsymbol{e}^i = b^i \mathrm{d}x^i\} \\
=& \{\mathrm{d}f | \mathrm{d}f = b^1 \mathrm{d}x^1 + \cdots + b^i \mathrm{d}x^i + \cdots + b^n \mathrm{d}x^n\}
\end{aligned}
$$

---

or more simply to be comparison

$$
\begin{array}{llllllllll}
V & = \text{span}\{ & \boldsymbol{e}_j = & \partial_j & \} = \{ & \boldsymbol{v} = & \partial_t|_P & = a_j \, \boldsymbol{e}_j = & a_j \, \partial_j|_P & : f \to F \quad \} \\
V^* & = \text{span}\{ & \boldsymbol{e}^i = & \mathrm{d}x^i & \} = \{ & \boldsymbol{\omega} = & \mathrm{d}f & = b^i \, \boldsymbol{e}^i = & b^i \, \mathrm{d}x^i & : V \to F \quad \}
\end{array}
$$

$$
\begin{cases}
\mathrm{d}x^i(\boldsymbol{\partial}_j) = \mathrm{d}x^i\left(\frac{\partial}{\partial x^j}|_P\right) = \frac{\partial x^i}{\partial x^j}|_P = \delta_{ij} = \begin{cases} 1 & i = j \\ 0 & i \neq j \end{cases} \\
\boldsymbol{\partial}_t = a_j \boldsymbol{e}_j = a_j \boldsymbol{\partial}_j \Leftrightarrow \frac{\partial}{\partial t}|_P = a_1 \frac{\partial}{\partial x^1}|_P + \cdots + a_j \frac{\partial}{\partial x^j}|_P + \cdots + a_n \frac{\partial}{\partial x^n}|_P
\end{cases}
$$

$$
\begin{aligned}
&\Rightarrow \begin{cases}
\mathrm{d}x^i(\boldsymbol{\partial}_t) = \mathrm{d}x^i\left(\frac{\partial}{\partial t}|_P\right) = \frac{\partial x^i}{\partial t}|_P \\
\mathrm{d}x^i(\boldsymbol{\partial}_t) = \mathrm{d}x^i(a_j \boldsymbol{\partial}_j) = a_j \mathrm{d}x^i(\boldsymbol{\partial}_j) = a_j \delta_{ij} = a_i
\end{cases} \qquad \Rightarrow a_i = \mathrm{d}x^i(\boldsymbol{\partial}_t) = \frac{\partial x^i}{\partial t}|_P \\
&\Rightarrow a_i = \frac{\partial x^i}{\partial t}|_P \Rightarrow a_j = \frac{\partial x^j}{\partial t}|_P = \partial_t x^j|_P \\
&\Rightarrow \frac{\partial}{\partial t}|_P = a_i \frac{\partial}{\partial x^i}|_P = \frac{\partial x^i}{\partial t}|_P \frac{\partial}{\partial x^i}|_P = \frac{\partial x^i}{\partial t} \frac{\partial}{\partial x^i}|_P \Rightarrow \frac{\partial}{\partial t} = \frac{\partial x^i}{\partial t} \frac{\partial}{\partial x^i} \\
&\Rightarrow \partial_t|_P = \frac{\partial x^j}{\partial t} \partial_j|_P \Leftrightarrow \partial_t|_P = \partial_t x^j \partial_j|_P
\end{aligned}
$$

$$
\begin{aligned}
\mathrm{d}f =& b^i \boldsymbol{e}^i = b^i \mathrm{d}x^i \\
\frac{\partial f}{\partial x^j} = \mathrm{d}f(\boldsymbol{\partial}_j) = \mathrm{d}f(\boldsymbol{e}_j) =& b^i \boldsymbol{e}^i \cdot \boldsymbol{e}_j = b^i \delta_{ij} = b^j \\
b^j =& \frac{\partial f}{\partial x^j} \\
b^i =& \frac{\partial f}{\partial x^i} = \partial_i f \\
\mathrm{d}f = b^i \boldsymbol{e}^i = b^i \mathrm{d}x^i =& \frac{\partial f}{\partial x^i} \mathrm{d}x^i \\
\mathrm{d}f =& \frac{\partial f}{\partial x^i} \mathrm{d}x^i \\
\mathrm{d}f =& \partial_i f \mathrm{d}x^i
\end{aligned}
$$

$$
\begin{array}{llllllllll}
V & = \text{span}\{ & \boldsymbol{e}_j = & \partial_j & \} = \{ & \boldsymbol{v} = & \partial_t|_P & = a_j \, \boldsymbol{e}_j = & \partial_t x^j \, \partial_j|_P & : f \to F \quad \} \\
V^* & = \text{span}\{ & \boldsymbol{e}^i = & \mathrm{d}x^i & \} = \{ & \boldsymbol{\omega} = & \mathrm{d}f & = b^i \, \boldsymbol{e}^i = & \partial_i f \, \mathrm{d}x^i & : V \to F \quad \}
\end{array}
$$

41.5 change of basis / change of coordinate

$$
\frac{\partial}{\partial t} = \frac{\partial x^i}{\partial t} \frac{\partial}{\partial x^i} \overset{t = x'^j}{\Rightarrow} \frac{\partial}{\partial x'^j} = \frac{\partial x^i}{\partial x'^j} \frac{\partial}{\partial x^i} = \frac{\partial x^1}{\partial x'^j} \frac{\partial}{\partial x^1} + \frac{\partial x^2}{\partial x'^j} \frac{\partial}{\partial x^2} + \frac{\partial x^3}{\partial x'^j} \frac{\partial}{\partial x^3}
$$

$$\begin{aligned} \mathrm{d}f =& \frac{\partial f}{\partial x^i}\mathrm{d}x^i \\ f = x'^j \Downarrow & \\ \mathrm{d}x'^j =& \frac{\partial x'^j}{\partial x^i}\mathrm{d}x^i \end{aligned}$$

$$\begin{cases} \dfrac{\partial}{\partial x'^j} = \dfrac{\partial x^i}{\partial x'^j}\dfrac{\partial}{\partial x^i} = \sum_i \dfrac{\partial x^i}{\partial x'^j}\dfrac{\partial}{\partial x^i} \\ \mathrm{d}x'^j = \dfrac{\partial x'^j}{\partial x^i}\mathrm{d}x^i = \sum_i \dfrac{\partial x'^j}{\partial x^i}\mathrm{d}x^i \end{cases}$$

Chapter 42

determinant

42.1 induction of determinant axioms

42.2 determinant axioms

determinant axioms

$$\begin{cases}\begin{cases}\det(\boldsymbol{u},\boldsymbol{v})=\det(\boldsymbol{u}+s\boldsymbol{v},\boldsymbol{v})\\ \det(\boldsymbol{u},\boldsymbol{v})=\det(\boldsymbol{u},\boldsymbol{v}+s\boldsymbol{u})\end{cases} & \text{translation invariance}\\ \begin{cases}\det(s\boldsymbol{u},\boldsymbol{v})=s\det(\boldsymbol{u},\boldsymbol{v})\\ \det(\boldsymbol{u},s\boldsymbol{v})=s\det(\boldsymbol{u},\boldsymbol{v})\end{cases} & \text{scaling}\\ \begin{cases}\det(\boldsymbol{u}_1+\boldsymbol{u}_2,\boldsymbol{v})=\det(\boldsymbol{u}_1,\boldsymbol{v})+\det(\boldsymbol{u}_2,\boldsymbol{v})\\ \det(\boldsymbol{u},\boldsymbol{v}_1+\boldsymbol{v}_2)=\det(\boldsymbol{u},\boldsymbol{v}_1)+\det(\boldsymbol{u},\boldsymbol{v}_2)\end{cases} & \text{decomposition}\end{cases}$$

$$\Leftrightarrow\begin{cases}\begin{cases}\det(\boldsymbol{u},\boldsymbol{v})=\det(\boldsymbol{u}+s\boldsymbol{v},\boldsymbol{v})\\ \det(\boldsymbol{u},\boldsymbol{v})=\det(\boldsymbol{u},\boldsymbol{v}+s\boldsymbol{u})\end{cases} & \text{translation invariance}\\ \begin{cases}\det(\boldsymbol{u}_1+s\boldsymbol{u}_2,\boldsymbol{v})=\det(\boldsymbol{u}_1,\boldsymbol{v})+s\det(\boldsymbol{u}_2,\boldsymbol{v})\\ \det(\boldsymbol{u},\boldsymbol{v}_1+s\boldsymbol{v}_2)=\det(\boldsymbol{u},\boldsymbol{v}_1)+s\det(\boldsymbol{u},\boldsymbol{v}_2)\end{cases} & \text{linearity}\end{cases}$$

42.3 determinant theorems or properties

42.4 ellipse area

42.5 Cramer rule geometry perspective

Chapter 43

hypergeometric function

43.1 linear space of function

Taylor vs. Fourier[@ref(taylor-vs.-fourier)]

$$f(x) = a_0 x^0 + a_1 x^1 + a_2 x^2 + \cdots = \sum_{k=0}^{\infty} a_k x^k$$

$$f(x) = x_0 x^0 + x_1 x^1 + x_2 x^2 + \cdots = \sum_{k=0}^{\infty} x_k x^k$$

Def: 45.3

$$\langle f|g\rangle = \int_a^b \overline{f(x)} g(x)\,\mathrm{d}x \overset{f,g:\mathbb{R}\to\mathbb{R}}{=} \int_a^b f(x)\,g(x)\,\mathrm{d}x$$

Dirac bracket[45.5]

$$\left\langle x^2 \middle| x\right\rangle \overset{x^2,x:\mathbb{R}\to\mathbb{R}}{=} \int_a^b x^2 x \mathrm{d}x = \int_a^b x^3 \mathrm{d}x = \left[\frac{x^4}{4}\right]_a^b \not\equiv 0$$

$$x^0 \not\perp x^1, x^1 \not\perp x^2, \cdots$$

$$\langle x_m | x^n\rangle = \int_a^b x_m x^n \mathrm{d}x = \delta_{mn}$$

$$\langle 1 | x^n\rangle = \int_a^b x_0 x^n \mathrm{d}x = \delta_{0n} \Rightarrow x_0 = \delta(x) = \delta(x-0)$$

$$\langle x_m | x^n\rangle = \int_a^b x_m x^n \mathrm{d}x = \delta_{mn} \Rightarrow x_m = \frac{(-1)^m}{m!}\delta^{(m)}(x)$$

$$|f\rangle = 1\,|f\rangle = \left(\sum_i \left|\hat{f}_i\right\rangle \left\langle \hat{f}_i\right|\right)|f\rangle = \sum_i \left|\hat{f}_i\right\rangle \left\langle \hat{f}_i \middle| f\right\rangle$$

$$\begin{aligned}
|f\rangle =& 1\,|f\rangle = \left(\sum_i \left|\hat{f}_i\right\rangle \left\langle \hat{f}_i\right|\right)|f\rangle = \sum_i \left|\hat{f}_i\right\rangle \left\langle \hat{f}_i\middle|f\right\rangle \\
=& 1\,|f\rangle = \left(\sum_n |x^n\rangle\langle x^n|\right)|f\rangle = \sum_n |x^n\rangle\langle x^n|f\rangle = \sum_n \langle x^n|f\rangle\,|x^n\rangle \\
\langle x^n|\,|f\rangle =& \langle x^n|f\rangle = \int_a^b x_n f(x)\,\mathrm{d}x = \int_a^b \frac{(-1)^n}{n!}\delta^{(n)}(x)\,f(x)\,\mathrm{d}x = \frac{f^{(n)}(0)}{n!} \\
|f\rangle =& \sum_n \langle x^n|f\rangle\,|x^n\rangle = \sum_n \frac{f^{(n)}(0)}{n!}\,|x^n\rangle \\
|f\rangle =& \sum_n \frac{f^{(n)}(0)}{n!}\,|x^n\rangle \\
& \Downarrow \\
f(x) =& \sum_n \frac{f^{(n)}(0)}{n!}x^n
\end{aligned}$$

43.2 beta function

$$\begin{aligned}
\binom{n}{k} = \mathrm{C}_k^n =& \frac{n!}{(n-k)!k!} \\
=& \frac{n(n-1)\cdots(n-k+1)}{k!}, \begin{cases} n\in\mathbb{N} \\ k\in(\{0\}\cup\mathbb{N}) \end{cases}
\end{aligned}$$

$$\binom{r}{k} = \begin{cases} \dfrac{r(r-1)\cdots(r-k+1)}{k!} & k\geq 0, k\in\mathbb{Z} \\ 0 & k<0, k\in\mathbb{Z} \end{cases}$$

$$\sum_{k=0}^{n}\binom{r}{k}(\cdot)$$

$$\sum_{k=-\infty}^{n}\binom{r}{k}(\cdot) = (0+0+\cdots) + \sum_{k=0}^{n}\binom{r}{k}(\cdot)$$

$$\sum_{k=-\infty}^{\infty}\binom{r}{k}(\cdot)$$

$$n! = \Gamma(n+1) = \int_0^\infty x^{(n+1)-1}\mathrm{e}^{-x}\mathrm{d}x$$

$$\Gamma(z) = \int_0^\infty x^{z-1}\mathrm{e}^{-x}\mathrm{d}x$$

$$\Gamma(z+1) = z\Gamma(z)$$

$$\Gamma(z)\Gamma(1-z) = \frac{\pi}{\sin(\pi z)}$$

$$\begin{aligned}
\Gamma(z)\Gamma(1-z) =& \frac{\pi}{\sin(\pi z)} \\
[\Gamma(z)\Gamma(1-z)]_{z=-n} =& \left[\frac{\pi}{\sin(\pi z)}\right]_{z=-n}, n\in\mathbb{N} \\
\Gamma(-n)\,n! = \Gamma(n+1) = \Gamma(-n)\Gamma(1-(-n)) =& \frac{\pi}{\sin(\pi(-n))} = \frac{\pi}{-\sin(n\pi)} \\
\Gamma(-n) =& \frac{-\pi}{n!\sin(n\pi)} = \frac{-\pi}{n!0} \to -\infty, n\in\mathbb{N}
\end{aligned}$$

$$\begin{aligned}\binom{n}{k} = \mathrm{C}_k^n =& \frac{n!}{(n-k)!k!}\\ =& \frac{\Gamma(n+1)}{\Gamma(n-k+1)\Gamma(k+1)}\end{aligned}$$

$$\binom{n}{k} = \frac{\Gamma(n+1)}{\Gamma(n-k+1)\Gamma(k+1)}$$

$$\binom{n}{k} = \frac{\Gamma(n+1)}{\Gamma(n-k+1)\Gamma(k+1)} \overset{k\le 0}{\equiv} \begin{cases} \dfrac{\Gamma(n+1)}{\Gamma(n+1)\Gamma(1)} = \dfrac{\Gamma(n+1)}{\Gamma(n+1)1} = 1 & k=0 \\ \dfrac{\Gamma(n+1)}{\Gamma(n-k+1)(-\infty)} = 0 & k \le -1, k \in \mathbb{Z} \end{cases}$$

beta function $=\beta$ function

Definition 43.1. beta function $=\beta$ function

$$B(p,q) = \int_0^1 x^{p-1}(1-x)^{q-1}\,\mathrm{d}x = \frac{\Gamma(p)\Gamma(q)}{\Gamma(p+q)}$$

$$\begin{aligned}\binom{n}{k} =& \frac{\Gamma(n+1)}{\Gamma(n-k+1)\Gamma(k+1)}\\ \left[\binom{n}{k}\right]\begin{cases}n=a+b\\k=a\end{cases} =& \left[\frac{\Gamma(n+1)}{\Gamma(n-k+1)\Gamma(k+1)}\right]\begin{cases}n=a+b\\k=a\end{cases}\\ \binom{a+b}{a} =& \frac{\Gamma(a+b+1)}{\Gamma(a+b-a+1)\Gamma(a+1)}\\ =& \frac{\Gamma(a+b+1)}{\Gamma(b+1)\Gamma(a+1)}\end{aligned}$$

$$\begin{aligned}B(p,q) =& \frac{\Gamma(p)\Gamma(q)}{\Gamma(p+q)}\\ [B(p,q)]\begin{cases}p=a+1\\q=b+1\end{cases} =& \left[\frac{\Gamma(p)\Gamma(q)}{\Gamma(p+q)}\right]\begin{cases}p=a+1\\q=b+1\end{cases}\\ B(a+1,b+1) =& \frac{\Gamma(a+1)\Gamma(b+1)}{\Gamma(a+1+b+1)}\\ =& \frac{\Gamma(a+1)\Gamma(b+1)}{\Gamma([a+b+1]+1)} = \frac{\Gamma(a+1)\Gamma(b+1)}{[a+b+1]\Gamma(a+b+1)}\end{aligned}$$

$$\begin{aligned}\binom{a+b}{a} =& \frac{\Gamma(a+b+1)}{\Gamma(b+1)\Gamma(a+1)} = \frac{1}{\dfrac{\Gamma(b+1)\Gamma(a+1)}{\Gamma(a+b+1)}}\\ =& \frac{1}{[a+b+1]\dfrac{\Gamma(b+1)\Gamma(a+1)}{[a+b+1]\Gamma(a+b+1)}}\\ =& \frac{1}{[a+b+1]B(a+1,b+1)}\end{aligned}$$

---

https://en.wikipedia.org/wiki/Beta_function

https://en.wikipedia.org/wiki/Beta_function#Other_identities_and_formulas

https://en.wikipedia.org/wiki/Beta_function#Multivariate_beta_function

https://www.bilibili.com/video/BV1pa4y1P7Da/?t=4m

43.2.1 Wallis product formula

https://en.wikipedia.org/wiki/Wallis_product

https://www.bilibili.com/video/BV1pa4y1P7Da/?t=4m10s

https://www.math.sinica.edu.tw/mathmedia/journals/4739?keywords%5B%5D=Paul+Dirac

43.3 gamma function

43.4 recursion

43.5 mean and variance of discrete probability distributions

Chapter 44

Feynman method

44.1 Feynman method of differentiation / derivative

分式微分不是難而是煩

44.1.1 principle

Theorem 44.1. *Feynman method of differentiation / derivative*

$$f(t) = k[u(t)]^a [v(t)]^b [w(t)]^c \cdots$$
$$\Downarrow$$
$$f'(t) = f(t)\left[a\frac{u'(t)}{u(t)} + b\frac{v'(t)}{v(t)} + c\frac{w'(t)}{w(t)} + \cdots\right]$$

Proof:

$$f(t) = k[u(t)]^a [v(t)]^b [w(t)]^c \cdots$$

$$f = ku^a v^b w^c \cdots = k \cdot u^a \cdot v^b \cdot w^c \cdot \cdots$$

$$\begin{aligned}
f =& ku^a v^b w^c \cdots = k \cdot u^a \cdot v^b \cdot w^c \cdot \cdots \\
\ln f =& \ln\left(ku^a v^b w^c \cdots\right) = \ln k + \ln u^a + \ln v^b + \ln w^c + \cdots \\
=& \ln k + a\ln u + b\ln v + c\ln w + \cdots \\
\frac{\mathrm{d}}{\mathrm{d}t}\ln f =& \frac{\mathrm{d}}{\mathrm{d}t}\left(\ln k + a\ln u + b\ln v + c\ln w + \cdots\right) \\
\frac{\frac{\mathrm{d}}{\mathrm{d}t}f}{f} =& 0 + \frac{\mathrm{d}}{\mathrm{d}t}(a\ln u) + \frac{\mathrm{d}}{\mathrm{d}t}(b\ln v) + \frac{\mathrm{d}}{\mathrm{d}t}(c\ln w) + \cdots \\
=& a\frac{\mathrm{d}}{\mathrm{d}t}\ln u + b\frac{\mathrm{d}}{\mathrm{d}t}\ln v + c\frac{\mathrm{d}}{\mathrm{d}t}\ln w + \cdots \\
=& a\frac{\frac{\mathrm{d}}{\mathrm{d}t}u}{u} + b\frac{\frac{\mathrm{d}}{\mathrm{d}t}v}{v} + c\frac{\frac{\mathrm{d}}{\mathrm{d}t}w}{w} + \cdots \\
\frac{f'}{f} =& a\frac{u'}{u} + b\frac{v'}{v} + c\frac{w'}{w} + \cdots \\
f' =& f\left(a\frac{u'}{u} + b\frac{v'}{v} + c\frac{w'}{w} + \cdots\right) \\
f'(t) =& f(t)\left[a\frac{u'(t)}{u(t)} + b\frac{v'(t)}{v(t)} + c\frac{w'(t)}{w(t)} + \cdots\right]
\end{aligned}$$

□

44.1.2 examples

$f(x) = x^x$

$$(x^x)' = x^x + x^x \ln x$$

$$\begin{aligned}
f(x) =& x^x \\
\ln f(x) =& x \ln x \\
\frac{\mathrm{d}}{\mathrm{d}x} \ln f(x) =& \frac{\mathrm{d}}{\mathrm{d}x} [x \ln x] \\
\frac{f'(x)}{f(x)} =& [x \ln x]' \\
f'(x) =& f(x) [x \ln x]' = x^x \left[(x)' \ln x + x (\ln x)' \right] \\
=& x^x \left[1 \ln x + x \frac{1}{x} \right] = x^x [\ln x + 1] = x^x [1 + \ln x] \\
=& x^x + x^x \ln x
\end{aligned}$$

$\square$

3D delta function[45.2.2]

$$\Delta \left(\frac{1}{r} \right) = \boldsymbol{\nabla}^2 \left(\frac{1}{\sqrt{\boldsymbol{r}^2}} \right) = \boldsymbol{\nabla} \cdot \boldsymbol{\nabla} \left(\frac{1}{\sqrt{\boldsymbol{r} \cdot \boldsymbol{r}}} \right) = \boldsymbol{\nabla} \cdot \left(\frac{\partial}{\partial x}, \frac{\partial}{\partial y}, \frac{\partial}{\partial z} \right) \left(\frac{1}{\sqrt{x^2 + y^2 + z^2}} \right)$$

$$\frac{\partial}{\partial x} \frac{1}{\sqrt{x^2 + y^2 + z^2}}$$

$$\begin{aligned}
&\frac{\partial}{\partial x} \frac{1}{\sqrt{x^2 + y^2 + z^2}} = \frac{\partial}{\partial x} \left[\left(x^2 + y^2 + z^2 \right)^{\frac{-1}{2}} \right] \\
&= \left(x^2 + y^2 + z^2 \right)^{\frac{-1}{2}} \left[\frac{-1}{2} \frac{2x}{x^2 + y^2 + z^2} \right] \\
&= \frac{-x}{\left(x^2 + y^2 + z^2 \right)^{\frac{-3}{2}}}
\end{aligned}$$

$$\boldsymbol{\nabla} \left(\frac{1}{r} \right) = \boldsymbol{\nabla} \left(\frac{1}{\sqrt{\boldsymbol{r} \cdot \boldsymbol{r}}} \right) = \left(\frac{\partial}{\partial x}, \frac{\partial}{\partial y}, \frac{\partial}{\partial z} \right) \left(\frac{1}{\sqrt{x^2 + y^2 + z^2}} \right) = \left(\frac{-x}{(x^2 + y^2 + z^2)^{\frac{3}{2}}}, \frac{-y}{(x^2 + y^2 + z^2)^{\frac{3}{2}}}, \frac{-z}{(x^2 + y^2 + z^2)^{\frac{3}{2}}} \right)$$

$$\begin{aligned}
\boldsymbol{\nabla} \cdot \boldsymbol{\nabla} \left(\frac{1}{r} \right) =& \boldsymbol{\nabla} \cdot \left(\frac{-x}{(x^2 + y^2 + z^2)^{\frac{3}{2}}}, \frac{-y}{(x^2 + y^2 + z^2)^{\frac{3}{2}}}, \frac{-z}{(x^2 + y^2 + z^2)^{\frac{3}{2}}} \right) \\
=& \frac{\partial}{\partial x} \frac{-x}{(x^2 + y^2 + z^2 + \epsilon^2)^{\frac{3}{2}}} + \frac{\partial}{\partial y} \frac{-y}{(x^2 + y^2 + z^2 + \epsilon^2)^{\frac{3}{2}}} + \frac{\partial}{\partial z} \frac{-z}{(x^2 + y^2 + z^2 + \epsilon^2)^{\frac{3}{2}}}
\end{aligned}$$

$$\frac{\partial}{\partial x} \frac{-x}{(x^2 + y^2 + z^2 + \epsilon^2)^{\frac{3}{2}}}$$

$$\begin{aligned}
&\frac{\partial}{\partial x} \frac{-x}{(x^2 + y^2 + z^2)^{\frac{3}{2}}} \\
=& \frac{-x}{(x^2 + y^2 + z^2)^{\frac{3}{2}}} \left[1 \cdot \frac{1}{x} + \frac{-3}{2} \frac{2x}{x^2 + y^2 + z^2} \right] \\
=& \frac{-1}{(x^2 + y^2 + z^2)^{\frac{3}{2}}} + \frac{3x^2}{(x^2 + y^2 + z^2)^{\frac{5}{2}}} \\
=& \frac{-\left(x^2 + y^2 + z^2 \right) + 3x^2}{(x^2 + y^2 + z^2)^{\frac{5}{2}}} = \frac{2x^2 - y^2 - z^2}{(x^2 + y^2 + z^2)^{\frac{5}{2}}}
\end{aligned}$$

$$\boldsymbol{\nabla}\cdot\boldsymbol{\nabla}\left(\frac{1}{r}\right)=\frac{\partial}{\partial x}\frac{-x}{\left(x^2+y^2+z^2+\epsilon^2\right)^{\frac{3}{2}}}+\frac{\partial}{\partial y}\frac{-y}{\left(x^2+y^2+z^2+\epsilon^2\right)^{\frac{3}{2}}}+\frac{\partial}{\partial z}\frac{-z}{\left(x^2+y^2+z^2+\epsilon^2\right)^{\frac{3}{2}}}$$
$$=\frac{+2x^2-y^2-z^2}{\left(x^2+y^2+z^2\right)^{\frac{5}{2}}}+\frac{-x^2+2y^2-z^2}{\left(x^2+y^2+z^2\right)^{\frac{5}{2}}}+\frac{-x^2-y^2+2z^2}{\left(x^2+y^2+z^2\right)^{\frac{5}{2}}}=0$$

$$\Delta\left(\frac{1}{r}\right)=\boldsymbol{\nabla}\cdot\boldsymbol{\nabla}\left(\frac{1}{r}\right)=0$$

$\square$

$$\frac{6\left(1+2t^2\right)\left(t^3-t\right)^2}{\sqrt{t+5t^2}\,(4t)^{\frac{3}{2}}}+\frac{\sqrt{1+2t}}{t+\sqrt{1+t^2}}$$

$$\frac{\mathrm{d}}{\mathrm{d}t}\left[\frac{6\left(1+2t^2\right)\left(t^3-t\right)^2}{\sqrt{t+5t^2}\,(4t)^{\frac{3}{2}}}+\frac{\sqrt{1+2t}}{t+\sqrt{1+t^2}}\right]$$

$$\frac{\mathrm{d}}{\mathrm{d}t}\left[\frac{6\left(1+2t^2\right)\left(t^3-t\right)^2}{\sqrt{t+5t^2}\,(4t)^{\frac{3}{2}}}\right]$$

$$\frac{\mathrm{d}}{\mathrm{d}t}\left[\frac{6\left(1+2t^2\right)\left(t^3-t\right)^2}{\sqrt{t+5t^2}\,(4t)^{\frac{3}{2}}}\right]$$
$$=\frac{6\left(1+2t^2\right)\left(t^3-t\right)^2}{\sqrt{t+5t^2}\,(4t)^{\frac{3}{2}}}\cdot[$$
$$=\frac{6\left[1+2t^2\right]\left(t^3-t\right)^2}{\sqrt{t+5t^2}\,(4t)^{\frac{3}{2}}}\cdot[1\cdot\frac{1}{1+2t^2}\cdot 4t,\left[1+2t^2\right]\rightarrow\begin{cases}\text{exponential :} & 1\\ \text{linear to denominator :} & \frac{1}{1+2t^2}\\ \text{differentiation} & 4t\end{cases}$$
$$=\frac{6\left(1+2t^2\right)\left[\left(t^3-t\right)^2\right]}{\sqrt{t+5t^2}\,(4t)^{\frac{3}{2}}}\cdot[\frac{4t}{1+2t^2}+2\cdot\frac{1}{t^3-t}\cdot\left(3t^2-1\right)$$
$$,\left[\left(t^3-t\right)^2\right]\rightarrow\begin{cases}\text{exponential :} & 2\\ \text{linear to denominator :} & \frac{1}{t^3-t}\\ \text{differentiation} & 3t^2-1\end{cases}$$
$$=\frac{6\left(1+2t^2\right)\left(t^3-t\right)^2}{\left[\sqrt{t+5t^2}\right](4t)^{\frac{3}{2}}}\cdot[\frac{4t}{1+2t^2}+\frac{6t^2-2}{t^3-t}+\frac{-1}{2}\cdot\frac{1}{t+5t^2}\cdot(1+10t)$$
$$,\left[\sqrt{t+5t^2}\right]\rightarrow\begin{cases}\text{exponential :} & \frac{-1}{2}\\ \text{linear to denominator :} & \frac{1}{t+5t^2}\\ \text{differentiation} & 1+10t\end{cases}$$
$$=\frac{6\left(1+2t^2\right)\left(t^3-t\right)^2}{\sqrt{t+5t^2}\left[(4t)^{\frac{3}{2}}\right]}\cdot\left[\frac{4t}{1+2t^2}+\frac{6t^2-2}{t^3-t}-\frac{1+10t}{2t+10t^2}+\frac{-3}{2}\cdot\frac{1}{4t}\cdot 4\right]$$
$$,\left[(4t)^{\frac{3}{2}}\right]\rightarrow\begin{cases}\text{exponential :} & \frac{-3}{2}\\ \text{linear to denominator :} & \frac{1}{4t}\\ \text{differentiation} & 4\end{cases}$$
$$=\frac{6\left(1+2t^2\right)\left(t^3-t\right)^2}{\sqrt{t+5t^2}\,(4t)^{\frac{3}{2}}}\left[\frac{4t}{1+2t^2}+\frac{6t^2-2}{t^3-t}-\frac{1+10t}{2t+10t^2}-\frac{3}{2t}\right]$$

$$\frac{\mathrm{d}}{\mathrm{d}t}\left[\frac{\sqrt{1+2t}}{t+\sqrt{1+t^2}}\right]$$

$$\begin{aligned}&\frac{\mathrm{d}}{\mathrm{d}t}\left[\frac{\sqrt{1+2t}}{t+\sqrt{1+t^2}}\right]\\=&\frac{\sqrt{1+2t}}{t+\left[\sqrt{1+t^2}\right]}\left[\frac{1}{2}\frac{2}{1+2t}+(-1)\,\frac{1+\left[\frac{1}{2}\frac{2t}{\sqrt{1+t^2}}\right]}{t+\sqrt{1+t^2}}\right]\end{aligned}$$

□

44.2 Feynman method of integration / integral

44.3 path integral

https://www.youtube.com/watch?v=Sp5SvdDh2u8

https://en.wikipedia.org/wiki/Path_integral_formulation

https://www.youtube.com/watch?v=7yjv-gLHFqg

Chapter 45

Hilbert space

```r
knitr::opts_chunk$set(fig.pos = "H", out.extra = "")
```

45.1 Taylor expansion or Taylor series

Lemma 45.1. *Newton-Leibniz formula = N-LF, equivalent to first mean value theorem for derivatives / Lagrange mean value theorem = MVTdL / MVTd1*

牛頓-萊布尼茨公式 Newton-Leibniz formula = N-LF, equivalent to 第一微分均值定理 / 拉格朗日均值定理 first mean value theorem for derivatives / Lagrange mean value theorem = MVTdL / MVTd1 **??**

$$\int_a^b f'(x)\,\mathrm{d}x = f(b) - f(a) \tag{45.1}$$

$$\left\{\begin{array}{l}\left\{\begin{array}{ll} f:[a,b]\to\mathbb{R} & (0)\\ f \text{ continuous on } [a,b] & (1)\\ f \text{ differentiable on } (a,b) & (2)\end{array}\right.\\ \left\{\begin{array}{ll} f':[a,b]\to\mathbb{R}\ (4) & a<b\ (3)\\ f' \text{ continuous on } [a,b] & (5)\end{array}\right.\end{array}\right. \Rightarrow \exists c\in(a,b)\left(f'(c)=\frac{\int_a^b f'(x)\,\mathrm{d}x}{b-a}\right) \quad \text{MVTi1 } \mathbf{??}$$

$$\Rightarrow \begin{array}{c} \exists c\in(a,b)\left\{f'(c)(b-a)=f(b)-f(a)\right\}\ 45.1\\ \Updownarrow\\ \int_a^b f'(x)\,\mathrm{d}x = f(b)-f(a)\ 45.1\end{array}$$

($\Downarrow$):

$$\left\{\begin{array}{l}\left\{\begin{array}{ll} f:[a,b]\to\mathbb{R} & (0)\\ f \text{ continuous on } [a,b] & (1)\\ f \text{ differentiable on } (a,b) & (2)\end{array}\right. \overset{??}{\Rightarrow} \exists c\in(a,b)\left\{f'(c)(b-a)=f(b)-f(a)\right\} \quad \text{MVTd1 } \mathbf{??}\\ [a,b]=[x_0,x_n]=[x_0,x_1]\cup[x_0,x_2]\cup\cdots\cup[x_{n-1},x_n] \quad 45.2\end{array}\right.$$

$$\overset{??}{\underset{}{\overset{??}{\Rightarrow}}} \int_a^b f'(x)\,\mathrm{d}x = f(b)-f(a)$$

($\Uparrow$):

$$\left\{\begin{array}{l}\left\{\begin{array}{ll} f:[a,b]\to\mathbb{R} & (0)\\ f \text{ continuous on } [a,b] & (1)\\ f \text{ differentiable on } (a,b) & (2)\end{array}\right. \overset{??}{\overset{??}{\Rightarrow}} \int_a^b f'(x)\,\mathrm{d}x = f(b)-f(a) \quad \text{N-LF } 45.1\\ \left\{\begin{array}{ll} f':[a,b]\to\mathbb{R}\ (4) & a<b\ (3)\\ f' \text{ continuous on } [a,b] & (5)\end{array}\right. \Rightarrow \exists c\in(a,b)\left(f'(c)=\frac{\int_a^b f'(x)\,\mathrm{d}x}{b-a}\right) \quad \text{MVTi1 } \mathbf{??}\end{array}\right.$$

$$\Rightarrow \exists c\in(a,b)\left\{f'(c)(b-a)=f(b)-f(a)\right\}$$

Proof: ($\Downarrow$)

重疊端點分割 $[a,b]$ 成 n 部分聯集 ($n \in \mathbb{N}$)

$$[a,b]=[x_0,x_n]=[x_0,x_1]\cup[x_0,x_2]\cup\cdots\cup[x_{n-1},x_n] \tag{45.2}$$

$$\begin{cases}\begin{cases}f:[a,b]\to\mathbb{R} & (0)\\ f \text{ continuous on } [a,b] & (1)\\ f \text{ differentiable on } (a,b) & (2)\end{cases} \overset{??}{\Rightarrow} \exists c\in(a,b)\left\{f'(c)(b-a)=f(b)-f(a)\right\} \quad \text{MVTd1 } \textbf{??}\\ [a,b]=[x_0,x_n]=[x_0,x_1]\cup[x_0,x_2]\cup\cdots\cup[x_{n-1},x_n] \qquad 45.2\end{cases}$$

$$\Rightarrow\begin{cases}(f:[x_0,x_n]\to\mathbb{R})=(f:[x_0,x_1]\cup[x_0,x_2]\cup\cdots\cup[x_{n-1},x_n]\to\mathbb{R}) & \Leftarrow(0)\\ f \text{ continuous on } [x_0,x_1]\cup[x_0,x_2]\cup\cdots\cup[x_{n-1},x_n] & \Leftarrow(1)\\ f \text{ differentiable on } (x_0,x_1]\cup[x_0,x_2]\cup\cdots\cup[x_{n-1},x_n) & \Leftarrow(2)\end{cases}$$

$$\Rightarrow\begin{cases}f:[x_0,x_1]\to\mathbb{R}, f:[x_0,x_2]\to\mathbb{R},\cdots,f:[x_{n-1},x_n]\to\mathbb{R} & \Leftarrow(0)\\ f \text{ continuous on } [x_0,x_1],[x_0,x_2],\cdots,[x_{n-1},x_n] & \Leftarrow(1)\\ f \text{ differentiable on } (x_0,x_1],[x_0,x_2],\cdots,[x_{n-1},x_n)\Rightarrow f \text{ differentiable on } (x_0,x_1),(x_0,x_2),\cdots,(x_{n-1},x_n) & \Leftarrow(2)\end{cases}$$

$$\overset{??}{\underset{}{\overset{??}{\Rightarrow}}}\begin{cases}\exists c_1\in(x_0,x_1)\left\{f'(c_1)(x_1-x_0)=f(x_1)-f(x_0)\right\}\\ \exists c_2\in(x_1,x_2)\left\{f'(c_2)(x_2-x_1)=f(x_2)-f(x_1)\right\}\\ \vdots\\ \exists c_k\in(x_{k-1},x_k)\left\{f'(c_k)(x_k-x_{k-1})=f(x_k)-f(x_{k-1})\right\} \quad \forall k\in\mathbb{N}\cap[1,n]\\ \vdots\\ \exists c_n\in(x_{n-1},x_n)\left\{f'(c_n)(x_n-x_{n-1})=f(x_n)-f(x_{n-1})\right\}\end{cases}$$

$$\begin{aligned}
&\Rightarrow f'(c_k)(x_k-x_{k-1})=f(x_k)-f(x_{k-1})\\
&\Rightarrow \sum_k f'(c_k)(x_k-x_{k-1})=\sum_k f(x_k)-f(x_{k-1})\\
&\Rightarrow \sum_{k=1}^n f'(c_k)(x_k-x_{k-1})=\sum_{k=1}^n f(x_k)-f(x_{k-1})\\
&\quad =[f(x_1)-f(x_0)]+[f(x_2)-f(x_1)]+\cdots+[f(x_n)-f(x_{n-1})]\\
&\quad =f(x_n)-f(x_0)=f(b)-f(a)\\
&\quad \Rightarrow \sum_{k=1}^n f'(c_k)(x_k-x_{k-1})=f(b)-f(a) \overset{\Delta x_k=x_k-x_{k-1}}{\Leftrightarrow} \sum_{k=1}^n f'(c_k)\Delta x_k=f(b)-f(a)\\
&\Rightarrow \lim_{n\to\infty}\lim_{\Delta x_k=\frac{x_n-x_0}{n}}\sum_{k=1}^n f'(c_k)\Delta x_k=\lim_{n\to\infty}\lim_{\Delta x_k=\frac{b-a}{n}\to 0} f(b)-f(a)=f(b)-f(a)\\
&\Rightarrow \int_a^b f'(x)\,\mathrm{d}x=\int_{x_0}^{x_n} f'(x)\,\mathrm{d}x=f(b)-f(a)\\
&\Rightarrow \int_a^b f'(x)\,\mathrm{d}x=f(b)-f(a)
\end{aligned}$$

$$\int_a^b f'(x)\,\mathrm{d}x=f(b)-f(a)$$

得到 牛頓-萊布尼茨公式 Newton-Leibniz formula = N-LF [eq:N-LF]

□

Proof: ($\Uparrow$)

$$\left\{\begin{array}{ll}\left\{\begin{array}{ll}f:[a,b]\to\mathbb{R} & (0)\\ f\ \text{continuous on}\ [a,b] & (1)\\ f\ \text{differentiable on}\ (a,b) & (2)\end{array}\right. \overset{??}{\underset{??}{\Rightarrow}} \int_a^b f'(x)\,\mathrm{d}x = f(b)-f(a) & \text{N-LF 45.1}\\ \left\{\begin{array}{ll}f:[a,b]\to\mathbb{R} & a<b\ (3)\\ f\ \text{continuous on}\ [a,b] & (1)\end{array}\right. \Rightarrow \exists c\in(a,b)\left(f(c)=\frac{\int_a^b f(x)\,\mathrm{d}x}{b-a}\right) & \text{MVTi1 ??}\end{array}\right.$$

$$\Rightarrow\left\{\begin{array}{ll}\left\{\begin{array}{ll}f:[a,b]\to\mathbb{R} & (0)\\ f\ \text{continuous on}\ [a,b] & (1)\\ f\ \text{differentiable on}\ (a,b) & (2)\end{array}\right. \overset{??}{\underset{??}{\Rightarrow}} \int_a^b f'(x)\,\mathrm{d}x = f(b)-f(a) & \text{N-LF 45.1}\\ \left\{\begin{array}{ll}f:[a,b]\to\mathbb{R} & a<b\ (3)\\ f\ \text{continuous on}\ [a,b] & (1)\end{array}\right. \Rightarrow \exists c\in(a,b)\left\{f(c)(b-a)=\int_a^b f(x)\,\mathrm{d}x\right\} & \text{MVTi1 ??}\end{array}\right.$$

$$\Rightarrow\left\{\begin{array}{ll}\left\{\begin{array}{ll}f:[a,b]\to\mathbb{R} & (0)\\ f\ \text{continuous on}\ [a,b] & (1)\\ f\ \text{differentiable on}\ (a,b) & (2)\end{array}\right. \overset{??}{\underset{??}{\Rightarrow}} \int_a^b f'(x)\,\mathrm{d}x = f(b)-f(a) & \text{N-LF 45.1}\\ \left\{\begin{array}{ll}\text{if}\ f':[a,b]\to\mathbb{R} & (4)\\ \text{if}\ f'\ \text{continuous on}\ [a,b] & (5)\end{array}\right. \Rightarrow \exists c\in(a,b)\left\{f'(c)(b-a)=\int_a^b f'(x)\,\mathrm{d}x\right\} & \text{MVTi1 ??}\end{array}\right.$$

$$\Rightarrow\exists c\in(a,b)\left\{f'(c)(b-a)=\int_a^b f'(x)\,\mathrm{d}x = f(b)-f(a)\right\}$$

$$\Rightarrow\exists c\in(a,b)\left\{f'(c)(b-a)=f(b)-f(a)\right\}$$

$$\exists c\in(a,b)\left[f'(c)(b-a)=f(b)-f(a)\right]$$

得到 第一微分均值定理 / 拉格朗日均值定理 first mean value theorem for derivatives / Lagrange mean value theorem = MVTdL / MVTd1 [thm:MVTd1]

□

$$\begin{aligned}\int_a^b f'(x)\,\mathrm{d}x =& f(b)-f(a)\\ \int_{x_0}^x f'(t)\,\mathrm{d}t =& f(x)-f(x_0)\\ f(x) =& f(x_0)+\int_{x_0}^x f'(t)\,\mathrm{d}t\end{aligned}$$

$$f(x)=f(x_0)+\int_{x_0}^x f'(t)\,\mathrm{d}t$$

Lemma 45.2. *univariable product rule*

$$\mathrm{d}(uv)=(\mathrm{d}u)v+u\mathrm{d}v=v\mathrm{d}u+u\mathrm{d}v$$

$$\frac{\mathrm{d}(uv)}{\mathrm{d}t}=v\frac{\mathrm{d}u}{\mathrm{d}t}+u\frac{\mathrm{d}v}{\mathrm{d}t}$$

Lemma 45.3. *integration by parts*

$$\begin{aligned}
\int \frac{\mathrm{d}\left(uv\right)}{\mathrm{d}t}\,\mathrm{d}t &= \int v\frac{\mathrm{d}u}{\mathrm{d}t} + u\frac{\mathrm{d}v}{\mathrm{d}t}\mathrm{d}t \\
\int \mathrm{d}\left(uv\right) = & = \int v\frac{\mathrm{d}u}{\mathrm{d}t}\mathrm{d}t + \int u\frac{\mathrm{d}v}{\mathrm{d}t}\mathrm{d}t \\
uv = & = \int v\mathrm{d}u + \int u\mathrm{d}v \\
\int u\frac{\mathrm{d}v}{\mathrm{d}t}\mathrm{d}t &= uv - \int v\frac{\mathrm{d}u}{\mathrm{d}t}\mathrm{d}t \text{ "switching derivatives"} \qquad (45.3) \\
\int u\mathrm{d}v &= uv - \int v\mathrm{d}u \text{ "switching differentials"} \qquad (45.4)
\end{aligned}$$

Theorem 45.1. *univariable Taylor theorem with the remainder in integral form*

$$f\left(x\right) = \sum_{k=0}^{n} \frac{f^{(k)}\left(a\right)}{k!}\left(x-a\right)^{k} + \int_{a}^{x} f^{(n+1)}\left(t\right)\frac{\left(t-x\right)^{n}}{n!}\mathrm{d}t$$

$$f\left(x\right) = \sum_{k=0}^{n} \frac{f^{(k)}\left(a\right)}{k!}\left(x-a\right)^{k} + \int_{a}^{x} f^{(n+1)}\left(t\right)\frac{\left(x-t\right)^{n}}{n!}\mathrm{d}t$$

Proof:

$$\begin{aligned}
& f(x) - f(a) && (45.5)\\
\overset{45.1}{=} & \int_a^x f'(t)\,\mathrm{d}t = \int_a^x f'(t)\frac{\mathrm{d}(t-x)}{\mathrm{d}t}\mathrm{d}t \overset{45.3}{=} \left[f'(t)(t-x)\right]_{t=a}^x - \int_a^x (t-x)\frac{\mathrm{d}f'(t)}{\mathrm{d}t}\mathrm{d}t && (45.6)\\
= & \left[-f'(a)(a-x)\right] - \int_a^x f''(t)(t-x)\,\mathrm{d}t = f'(a)(x-a) - \int_a^x f''(t)\frac{\mathrm{d}\frac{(t-x)^2}{2}}{\mathrm{d}t}\mathrm{d}t && (45.7)\\
= & f'(a)(x-a) - \left(\left[f''(t)\frac{(t-x)^2}{2}\right]_{t=a}^x - \int_a^x \frac{(t-x)^2}{2}\frac{\mathrm{d}f''(t)}{\mathrm{d}t}\mathrm{d}t\right) && (45.8)\\
= & f'(a)(x-a) - \left(\left[-f''(a)\frac{(x-a)^2}{2}\right] - \int_a^x f'''(t)\frac{(t-x)^2}{2}\mathrm{d}t\right) && (45.9)\\
= & f'(a)(x-a) + f''(a)\frac{(x-a)^2}{2} + \int_a^x f'''(t)\frac{(t-x)^2}{2}\mathrm{d}t && (45.10)\\
= & f'(a)(x-a) + f''(a)\frac{(x-a)^2}{2} + \int_a^x f'''(t)\frac{\mathrm{d}\frac{(t-x)^3}{2\cdot 3}}{\mathrm{d}t}\mathrm{d}t && (45.11)\\
= & f'(a)(x-a) + f''(a)\frac{(x-a)^2}{2} + \left(\left[f'''(t)\frac{(t-x)^3}{2\cdot 3}\right]_{t=a}^x - \int_a^x \frac{(t-x)^3}{2\cdot 3}\frac{\mathrm{d}f'''(t)}{\mathrm{d}t}\mathrm{d}t\right) && (45.12)\\
= & f'(a)(x-a) + f''(a)\frac{(x-a)^2}{2} + \left(\left[f'''(a)\frac{(x-a)^3}{2\cdot 3}\right] - \int_a^x f^{(4)}(t)\frac{(t-x)^3}{2\cdot 3}\mathrm{d}t\right) && (45.13)\\
= & f'(a)(x-a) + f''(a)\frac{(x-a)^2}{2} + f'''(t)\frac{(x-a)^3}{2\cdot 3} - \int_a^x f^{(4)}(t)\frac{(t-x)^3}{2\cdot 3}\mathrm{d}t && (45.14)\\
= & f'(a)(x-a) + f''(a)\frac{(x-a)^2}{2} + f'''(t)\frac{(x-a)^3}{2\cdot 3} - \int_a^x f^{(4)}(t)\frac{\mathrm{d}\frac{(t-x)^4}{2\cdot 3\cdot 4}}{\mathrm{d}t}\mathrm{d}t && (45.15)\\
= & f'(a)(x-a) + f''(a)\frac{(x-a)^2}{2} + f'''(t)\frac{(x-a)^3}{2\cdot 3} && (45.16)\\
& - \left(\left[f^{(4)}(t)\frac{(t-x)^4}{2\cdot 3\cdot 4}\right]_{t=a}^x - \int_a^x \frac{(t-x)^4}{2\cdot 3\cdot 4}\frac{\mathrm{d}f^{(4)}(t)}{\mathrm{d}t}\mathrm{d}t\right) && (45.17)\\
= & f'(t)(x-a) + f''(a)\frac{(x-a)^2}{2} + f'''(t)\frac{(x-a)^3}{2\cdot 3} && (45.18)\\
& - \left(\left[-f^{(4)}(a)\frac{(x-a)^4}{2\cdot 3\cdot 4}\right] - \int_a^x f^{(5)}(t)\frac{(t-x)^4}{2\cdot 3\cdot 4}\mathrm{d}t\right) && (45.19)\\
= & f'(t)(x-a) + f''(a)\frac{(x-a)^2}{2} + f'''(a)\frac{(x-a)^3}{2\cdot 3} + f^{(4)}(a)\frac{(x-a)^4}{2\cdot 3\cdot 4} && (45.20)\\
\vdots & + \int_a^x f^{(5)}(t)\frac{(t-x)^4}{2\cdot 3\cdot 4}\mathrm{d}t && (45.21)\\
= & \sum_{k=1}^n \frac{f^{(k)}(a)}{k!}(x-a)^k + \int_a^x f^{(n+1)}(t)\frac{(t-x)^n}{n!}\mathrm{d}t \text{ remainder in integral form} && (45.22)
\end{aligned}$$

□

Proof:

$$
\begin{aligned}
&f(x)-f(a)\\
&\overset{45.1}{=}\int_a^x f'(t)\,\mathrm{d}t=\int_a^x f'(t)\frac{\mathrm{d}(t-x)}{\mathrm{d}t}\mathrm{d}t\overset{45.3}{=}\left[f'(t)(t-x)\right]_{t=a}^x-\int_a^x (t-x)\frac{\mathrm{d}f'(t)}{\mathrm{d}t}\mathrm{d}t\\
&=\left[-f'(a)(a-x)\right]-\int_a^x f''(t)(t-x)\,\mathrm{d}t=f'(a)(x-a)+\int_a^x f''(t)(x-t)\,\mathrm{d}t\\
&=f'(a)(x-a)-\int_a^x f''(t)(t-x)\,\mathrm{d}t=\sum_{k=1}^{1}\frac{f^{(k)}(a)}{k!}(x-a)^k+\int_a^x f''(t)(x-t)\,\mathrm{d}t\\
&=f'(a)(x-a)-\int_a^x f''(t)\frac{\mathrm{d}\frac{(t-x)^2}{2}}{\mathrm{d}t}\mathrm{d}t\\
&=f'(a)(x-a)-\left(\left[f''(t)\frac{(t-x)^2}{2}\right]_{t=a}^x-\int_a^x\frac{(t-x)^2}{2}\frac{\mathrm{d}f''(t)}{\mathrm{d}t}\mathrm{d}t\right)\\
&=f'(a)(x-a)-\left(\left[-f''(a)\frac{(x-a)^2}{2}\right]-\int_a^x f'''(t)\frac{(t-x)^2}{2}\mathrm{d}t\right)\\
&=f'(a)(x-a)+f''(a)\frac{(x-a)^2}{2}+\int_a^x f'''(t)\frac{(t-x)^2}{2}\mathrm{d}t\\
&=\sum_{k=1}^{2}\frac{f^{(k)}(a)}{k!}(x-a)^k+\int_a^x f'''(t)\frac{(t-x)^2}{2}\mathrm{d}t=\sum_{k=1}^{2}\frac{f^{(k)}(a)}{k!}(x-a)^k+\int_a^x f'''(t)\frac{(x-t)^2}{2!}\mathrm{d}t\\
&=f'(a)(x-a)+f''(a)\frac{(x-a)^2}{2}+\int_a^x f'''(t)\frac{\mathrm{d}\frac{(t-x)^3}{2\cdot 3}}{\mathrm{d}t}\mathrm{d}t\\
&=f'(a)(x-a)+f''(a)\frac{(x-a)^2}{2}+\left(\left[f'''(t)\frac{(t-x)^3}{2\cdot 3}\right]_{t=a}^x-\int_a^x\frac{(t-x)^3}{2\cdot 3}\frac{\mathrm{d}f'''(t)}{\mathrm{d}t}\mathrm{d}t\right)\\
&=f'(a)(x-a)+f''(a)\frac{(x-a)^2}{2}+\left(\left[f'''(a)\frac{(x-a)^3}{2\cdot 3}\right]-\int_a^x f^{(4)}(t)\frac{(t-x)^3}{2\cdot 3}\mathrm{d}t\right)\\
&=f'(a)(x-a)+f''(a)\frac{(x-a)^2}{2}+f'''(t)\frac{(x-a)^3}{2\cdot 3}-\int_a^x f^{(4)}(t)\frac{(t-x)^3}{2\cdot 3}\mathrm{d}t\\
&=\sum_{k=1}^{3}\frac{f^{(k)}(a)}{k!}(x-a)^k-\int_a^x f^{(4)}(t)\frac{(t-x)^3}{2\cdot 3}\mathrm{d}t=\sum_{k=1}^{3}\frac{f^{(k)}(a)}{k!}(x-a)^k+\int_a^x f^{(4)}(t)\frac{(x-t)^3}{3!}\mathrm{d}t\\
&\overset{\vdots}{=}\sum_{k=1}^{n}\frac{f^{(k)}(a)}{k!}(x-a)^k+\int_a^x f^{(n+1)}(t)\frac{(x-t)^n}{n!}\mathrm{d}t
\end{aligned}
$$

$$
\begin{aligned}
f(x)-f(a)=&\sum_{k=1}^{n}\frac{f^{(k)}(a)}{k!}(x-a)^k+\int_a^x f^{(n+1)}(t)\frac{(x-t)^n}{n!}\mathrm{d}t\\
f(x)=&f(a)+\sum_{k=1}^{n}\frac{f^{(k)}(a)}{k!}(x-a)^k+\int_a^x f^{(n+1)}(t)\frac{(x-t)^n}{n!}\mathrm{d}t\\
=&\frac{f^{(0)}(a)}{0!}(x-a)^0+\sum_{k=1}^{n}\frac{f^{(k)}(a)}{k!}(x-a)^k+\int_a^x f^{(n+1)}(t)\frac{(x-t)^n}{n!}\mathrm{d}t\\
=&\sum_{k=0}^{n}\frac{f^{(k)}(a)}{k!}(x-a)^k+\int_a^x f^{(n+1)}(t)\frac{(x-t)^n}{n!}\mathrm{d}t
\end{aligned}
$$

□

Theorem 45.2. *univariable Taylor theorem with the remainder in Lagrange form*

$$
f(x)=\sum_{k=0}^{n}\frac{f^{(k)}(a)}{k!}(x-a)^k+\frac{f^{(n+1)}(\xi)}{(n+1)!}(x-a)^{n+1}
$$

Proof:

by 連續函數 極值定理 / 最大最小值定理 / 最小最大值定理 continuous function extreme value theorem = CFEVT / extreme value theorem = EVT **??** and 連續函數 介值定理 / 中間值定理 continuous function intermediate value theorem = CFIVT / intermediate value theorem = IVT **??**

$$\begin{aligned}
\text{let } f(x) &= \sum_{k=0}^{n} \frac{f^{(k)}(a)}{k!}(x-a)^k + r_n(x) && (45.23)\\
r_n(x) &= \int_a^x f^{(n+1)}(t)\frac{(t-x)^n}{n!}\mathrm{d}t && (45.24)\\
\text{let } n &\in \{2k-1 | k \in \mathbb{N}\} && (45.25)\\
\text{let } f(t) &\overset{??}{\in} [m, M] \subseteq f((a,x)) \text{ when } a < x && (45.26)\\
\int_a^x m\frac{(t-x)^n}{n!}\mathrm{d}t \leq \quad r_n(x) &\leq \int_a^x M\frac{(t-x)^n}{n!}\mathrm{d}t && (45.27)\\
m\int_a^x \frac{(t-x)^n}{n!}\mathrm{d}t = &= M\int_a^x \frac{(t-x)^n}{n!}\mathrm{d}t && (45.28)\\
m\left[\frac{(t-x)^{n+1}}{(n+1)!}\right]_{t=a}^x = &= M\left[\frac{(t-x)^{n+1}}{(n+1)!}\right]_{t=a}^x && (45.29)\\
m\frac{(x-a)^{n+1}}{(n+1)!} = &= M\frac{(x-a)^{n+1}}{(n+1)!} && (45.30)\\
&\Downarrow \quad ?? && (45.31)\\
r_n(x) &\overset{\exists\xi\in(a,x)}{=} f^{(n+1)}(\xi)\frac{(x-a)^{n+1}}{(n+1)!} && (45.32)\\
f(x) &= \sum_{k=0}^{n} \frac{f^{(k)}(a)}{k!}(x-a)^k + f^{(n+1)}(\xi)\frac{(x-a)^{n+1}}{(n+1)!} && (45.33)\\
&= \sum_{k=0}^{n} \frac{f^{(k)}(a)}{k!}(x-a)^k + \frac{f^{(n+1)}(\xi)}{(n+1)!}(x-a)^{n+1} && (45.34)
\end{aligned}$$

□

Theorem 45.3. *univariable Taylor theorem with the remainder in big O form*

$$f(a+h) = \sum_{k=0}^{n} \frac{f^{(k)}(a)}{k!}h^k + \mathrm{O}\left(h^{n+1}\right)$$

Proof:

$$\begin{aligned}
\text{if } \left|f^{(n+1)}(t)\right| &\leq K\,\forall t \in (a,x) && (45.35)\\
|r_n(x)| = \left|\frac{f^{(n+1)}(\xi)}{(n+1)!}(x-a)^{n+1}\right| &\leq \frac{K}{(n+1)!}|x-a|^{n+1} && (45.36)\\
r_n(x) &\in \mathrm{O}\left((x-a)^{n+1}\right) && (45.37)\\
\text{let } R_n(h) &= r_n(a+h) && (45.38)\\
R_n(h) &\in \mathrm{O}\left(h^{n+1}\right) && (45.39)\\
f(a+h) &= \sum_{k=0}^{n} \frac{f^{(k)}(a)}{k!}h^k + \frac{f^{(n+1)}(\xi)}{(n+1)!}h^{n+1} && (45.40)\\
&= \sum_{k=0}^{n} \frac{f^{(k)}(a)}{k!}h^k + \mathrm{O}\left(h^{n+1}\right) && (45.41)
\end{aligned}$$

□

$$f(x)=\sum_{k=0}^{n}\frac{f^{(k)}(a)}{k!}(x-a)^k+\int_a^x f^{(n+1)}(t)\frac{(x-t)^n}{n!}\mathrm{d}t$$

$f(x)=\sin(x)$

$$\sin(x)=\sum_{k=0}^{n}\frac{\sin^{(k)}(a)}{k!}(x-a)^k+\int_a^x \sin^{(n+1)}(t)\frac{(x-t)^n}{n!}\mathrm{d}t$$

$a=0$

$$\sin(x)=\sum_{k=0}^{n}\frac{\sin^{(k)}(0)}{k!}(x-0)^k+\int_0^x \sin^{(n+1)}(t)\frac{(x-t)^n}{n!}\mathrm{d}t$$

$$\begin{aligned}
\sin(x)=&\sum_{k=0}^{n}\frac{\sin^{(k)}(0)}{k!}(x-0)^k+\int_0^x \sin^{(n+1)}(t)\frac{(x-t)^n}{n!}\mathrm{d}t\\
=&\sum_{k=0}^{n}\frac{\sin^{(k)}(0)}{k!}x^k+\int_0^x \sin^{(n+1)}(t)\frac{(x-t)^n}{n!}\mathrm{d}t\\
=&\frac{\sin^{(0)}(0)}{0!}x^0+\frac{\sin^{(1)}(0)}{1!}x^1+\frac{\sin^{(2)}(0)}{2!}x^2+\frac{\sin^{(3)}(0)}{3!}x^3\\
&+\int_0^x \sin^{(3+1)}(t)\frac{(x-t)^3}{3!}\mathrm{d}t\\
=&\frac{0}{0!}x^0+\frac{\cos(0)}{1!}x^1+\frac{-\sin(0)}{2!}x^2+\frac{-\cos(0)}{3!}x^3\\
&+\int_0^x \sin^{(4)}(t)\frac{(x-t)^3}{3!}\mathrm{d}t\\
=&0+\frac{1}{1}x+0-\frac{1}{6}x^3+\int_0^x \sin(t)\frac{(x-t)^3}{6}\mathrm{d}t\\
=&x-\frac{1}{6}x^3+\frac{1}{6}\int_0^x (x-t)^3\sin(t)\,\mathrm{d}t\\
\sin(x)=&x-\frac{1}{6}x^3+\frac{1}{6}\int_0^x (x-t)^3\sin(t)\,\mathrm{d}t
\end{aligned}$$

$$\begin{aligned}
\sin(x)=&\sum_{k=0}^{n}\frac{\sin^{(k)}(0)}{k!}(x-0)^k+\int_0^x \sin^{(n+1)}(t)\frac{(x-t)^n}{n!}\mathrm{d}t\\
=&\sum_{k=0}^{n}\frac{\sin^{(k)}(0)}{k!}x^k+\int_0^x \sin^{(n+1)}(t)\frac{(x-t)^n}{n!}\mathrm{d}t\\
=&\frac{\sin^{(0)}(0)}{0!}x^0+\frac{\sin^{(1)}(0)}{1!}x^1+\frac{\sin^{(2)}(0)}{2!}x^2+\frac{\sin^{(3)}(0)}{3!}x^3+\frac{\sin^{(4)}(0)}{4!}x^4\\
&+\int_0^x \sin^{(4+1)}(t)\frac{(x-t)^4}{4!}\mathrm{d}t\\
=&\frac{0}{0!}x^0+\frac{\cos(0)}{1!}x^1+\frac{-\sin(0)}{2!}x^2+\frac{-\cos(0)}{3!}x^3+\frac{\sin(0)}{4!}x^4\\
&+\int_0^x \sin^{(5)}(t)\frac{(x-t)^4}{4!}\mathrm{d}t\\
=&0+\frac{1}{1}x+0-\frac{1}{6}x^3+0+\int_0^x \cos(t)\frac{(x-t)^4}{24}\mathrm{d}t\\
=&x-\frac{1}{6}x^3+\frac{1}{24}\int_0^x (x-t)^4\cos(t)\,\mathrm{d}t\\
\sin(x)=&x-\frac{1}{6}x^3+\frac{1}{24}\int_0^x (x-t)^4\cos(t)\,\mathrm{d}t
\end{aligned}$$

$$\begin{aligned}\sin(x) =& x-\frac{1}{6}x^3+\frac{1}{6}\int_0^x (x-t)^3\sin(t)\,\mathrm{d}t\\ =& x-\frac{1}{6}x^3+\frac{1}{24}\int_0^x (x-t)^4\cos(t)\,\mathrm{d}t\\ &\Downarrow\\ \frac{1}{6}\int_0^x (x-t)^3\sin(t)\,\mathrm{d}t =& \frac{1}{24}\int_0^x (x-t)^4\cos(t)\,\mathrm{d}t\\ &\Downarrow\\ \int_0^x (x-t)^4\cos(t)\,\mathrm{d}t =& 4\int_0^x (x-t)^3\sin(t)\,\mathrm{d}t\end{aligned}$$

$$f(x)=\sum_{k=0}^{n}\frac{f^{(k)}(a)}{k!}(x-a)^k+\int_a^x f^{(n+1)}(t)\frac{(x-t)^n}{n!}\mathrm{d}t$$

$f(x)=\mathrm{e}^x=\exp(x)$

$$\mathrm{e}^x=\sum_{k=0}^{n}\frac{\exp^{(k)}(a)}{k!}(x-a)^k+\int_a^x \exp^{(n+1)}(t)\frac{(x-t)^n}{n!}\mathrm{d}t$$

$a=0$

$$\mathrm{e}^x=\sum_{k=0}^{n}\frac{\exp^{(k)}(0)}{k!}(x-0)^k+\int_0^x \exp^{(n+1)}(t)\frac{(x-t)^n}{n!}\mathrm{d}t$$

$$\begin{aligned}\mathrm{e}^x =& \sum_{k=0}^{n}\frac{\exp^{(k)}(0)}{k!}(x-0)^k+\int_0^x \exp^{(n+1)}(t)\frac{(x-t)^n}{n!}\mathrm{d}t\\ =& \sum_{k=0}^{n}\frac{\exp^{(k)}(0)}{k!}x^k+\int_0^x \exp^{(n+1)}(t)\frac{(x-t)^n}{n!}\mathrm{d}t\\ =& \sum_{k=0}^{n}\frac{\exp(0)}{k!}x^k+\int_0^x \exp(t)\frac{(x-t)^n}{n!}\mathrm{d}t\\ =& \sum_{k=0}^{n}\frac{\mathrm{e}^0}{k!}x^k+\int_0^x \mathrm{e}^t\frac{(x-t)^n}{n!}\mathrm{d}t=\sum_{k=0}^{n}\frac{1}{k!}x^k+\int_0^x \mathrm{e}^t\frac{(x-t)^n}{n!}\mathrm{d}t\\ =& \sum_{k=0}^{n}\frac{x^k}{k!}+\frac{1}{n!}\int_0^x (x-t)^n\,\mathrm{e}^t\mathrm{d}t\\ \mathrm{e}^x =& \sum_{k=0}^{n}\frac{x^k}{k!}+\frac{1}{n!}\int_0^x (x-t)^n\,\mathrm{e}^t\mathrm{d}t\end{aligned}$$

$$\mathrm{e}^x=\sum_{k=0}^{n}\frac{x^k}{k!}+\frac{1}{n!}\int_0^x (x-t)^n\,\mathrm{e}^t\mathrm{d}t$$

$$\lim_{n\to\infty}\frac{1+n+\frac{n^2}{2}+\cdots+\frac{n^n}{n!}}{\mathrm{e}^n}$$

$$
\begin{aligned}
&\lim_{n\to\infty}\frac{1+n+\dfrac{n^2}{2}+\cdots+\dfrac{n^n}{n!}}{\mathrm{e}^n}, \wedge \mathrm{e}^x=\sum_{k=0}^{n}\frac{x^k}{k!}+\frac{1}{n!}\int_0^x (x-t)^n\,\mathrm{e}^t\mathrm{d}t\\
&=\lim_{n\to\infty}\frac{1+n+\dfrac{n^2}{2}+\cdots+\dfrac{n^n}{n!}}{\left[\sum_{k=0}^{n}\frac{x^k}{k!}+\frac{1}{n!}\int_0^x (x-t)^n\,\mathrm{e}^t\mathrm{d}t\right]_{x=n}}\\
&=\lim_{n\to\infty}\frac{\sum_{k=0}^{n}\frac{n^k}{k!}}{\sum_{k=0}^{n}\frac{n^k}{k!}+\frac{1}{n!}\int_0^n (n-t)^n\,\mathrm{e}^t\mathrm{d}t}=\lim_{n\to\infty}\frac{1}{1+\dfrac{\frac{1}{n!}\int_0^n (n-t)^n\,\mathrm{e}^t\mathrm{d}t}{\sum_{k=0}^{n}\frac{n^k}{k!}}}
\end{aligned}
$$

$$
\begin{aligned}
&\lim_{n\to\infty}\frac{1+n+\dfrac{n^2}{2}+\cdots+\dfrac{n^n}{n!}}{\mathrm{e}^n}, \wedge \mathrm{e}^x=\sum_{k=0}^{n}\frac{x^k}{k!}+\frac{1}{n!}\int_0^x (x-t)^n\,\mathrm{e}^t\mathrm{d}t\\
&=\lim_{n\to\infty}\frac{\sum_{k=0}^{n}\frac{n^k}{k!}}{\mathrm{e}^n}, \wedge \mathrm{e}^n=\sum_{k=0}^{n}\frac{n^k}{k!}+\frac{1}{n!}\int_0^n (n-t)^n\,\mathrm{e}^t\mathrm{d}t\Rightarrow\sum_{k=0}^{n}\frac{n^k}{k!}=\mathrm{e}^n-\frac{1}{n!}\int_0^n (n-t)^n\,\mathrm{e}^t\mathrm{d}t\\
&=\lim_{n\to\infty}\frac{\mathrm{e}^n-\frac{1}{n!}\int_0^n (n-t)^n\,\mathrm{e}^t\mathrm{d}t}{\mathrm{e}^n}=\lim_{n\to\infty}\left[1-\frac{1}{n!}\int_0^n (n-t)^n\,\mathrm{e}^{t-n}\mathrm{d}t\right],\ \text{if }\lim_{n\to\infty}\frac{1}{n!}\int_0^n (t-n)^n\,\mathrm{e}^{t-n}\mathrm{d}t\in\mathbb{R}\\
&=\lim_{n\to\infty}1-\lim_{n\to\infty}\frac{1}{n!}\int_0^n (n-t)^n\,\mathrm{e}^{t-n}\mathrm{d}t=1-\lim_{n\to\infty}\frac{1}{n!}\int_0^n (n-t)^n\,\mathrm{e}^{t-n}\mathrm{d}t
\end{aligned}
$$

$\lim_{n\to\infty}\frac{1}{n!}\int_0^n (n-t)^n\,\mathrm{e}^{t-n}\mathrm{d}t$

$$
\begin{aligned}
&\lim_{n\to\infty}\frac{1}{n!}\int_0^n (n-t)^n\,\mathrm{e}^{t-n}\mathrm{d}t, x=n-t\\
&=\lim_{n\to\infty}\frac{1}{n!}\int_{t=0}^{t=n} (n-t)^n\,\mathrm{e}^{t-n}\mathrm{d}t, t=n-x\\
&=\lim_{n\to\infty}\frac{1}{n!}\int_{n-x=0}^{n-x=n} (n-(n-x))^n\,\mathrm{e}^{(n-x)-n}\mathrm{d}\,(n-x)\\
&=\lim_{n\to\infty}\frac{1}{n!}\left(-\int_{x=n}^{x=0} x^n\mathrm{e}^{-x}\mathrm{d}x\right)=\lim_{n\to\infty}\frac{1}{n!}\left(\int_{x=0}^{x=n} x^n\mathrm{e}^{-x}\mathrm{d}x\right)\\
&=\lim_{n\to\infty}\frac{1}{n!}\int_0^n x^n\mathrm{e}^{-x}\mathrm{d}x
\end{aligned}
$$

$\int x^n\mathrm{e}^{-x}\mathrm{d}x$

difficult process

$$
\begin{aligned}
\int x^n\mathrm{e}^{-x}\mathrm{d}x&=-\int x^n\mathrm{d}\mathrm{e}^{-x}\\
&=-\left[x^n\mathrm{e}^{-x}-\int\mathrm{e}^{-x}\mathrm{d}x^n\right]\\
&=-\left[x^n\mathrm{e}^{-x}-\int\mathrm{e}^{-x}nx^{n-1}\mathrm{d}x\right]\\
&=-x^n\mathrm{e}^{-x}+n\int x^{n-1}\mathrm{e}^{-x}\mathrm{d}x
\end{aligned}
$$

$$
\int x^n\mathrm{e}^{-x}\mathrm{d}x=-x^n\mathrm{e}^{-x}+n\int x^{n-1}\mathrm{e}^{-x}\mathrm{d}x
$$

$$\begin{aligned}\int_0^n x^n \mathrm{e}^{-x}\mathrm{d}x &= \left[-x^n\mathrm{e}^{-x}\right]_{x=0}^n + n\int_0^n x^{n-1}\mathrm{e}^{-x}\mathrm{d}x \\ &= \left[-n^n\mathrm{e}^{-n} - \left(-0^n\mathrm{e}^{-0}\right)\right] + n\int_0^n x^{n-1}\mathrm{e}^{-x}\mathrm{d}x \\ &= -\,n^n\mathrm{e}^{-n} + n\int_0^n x^{n-1}\mathrm{e}^{-x}\mathrm{d}x\end{aligned}$$

$$\int_0^n x^n \mathrm{e}^{-x}\mathrm{d}x = -n^n\mathrm{e}^{-n} + n\int_0^n x^{n-1}\mathrm{e}^{-x}\mathrm{d}x$$

more formal process with gamma function

https://www.wolframalpha.com/input?i2d=true&i=Divide%5BIntegrate%5BPower%5Bx%2Cn%5DPower%5Be%2C-x%5D%2C%7Bx%2C0%2Cn%7D%5D%2Cn%21%5D

https://www.wolframalpha.com/input?i2d=true&i=Limit%5BDivide%5BIntegrate%5BPower%5Bx%2Cn%5DPower%5Be%2C-x%5D%2C%7Bx%2C0%2Cn%7D%5D%2Cn%21%5D%2Cn-%3E%E2%88%9E%5D

to be proved

45.1.1 Elliot Schneider

https://www.youtube.com/watch?v=HQsZG8Yxb7w

https://www.youtube.com/watch?v=HQsZG8Yxb7w&t=12m25s

$$f(x) = \sum_{k=0}^{n} \frac{f^{(k)}(x_0)}{k!}(x-x_0)^k + \frac{f^{(n+1)}(\xi)}{(n+1)!}(x-x_0)^{n+1}$$

$$f(x_0+\epsilon) = \sum_{k=0}^{n} \frac{f^{(k)}(x_0)}{k!}\epsilon^k + \mathrm{O}\left(\epsilon^{n+1}\right), \epsilon = x - x_0$$

$$f(x+\epsilon) = \sum_{n=0}^{\infty} \frac{f^{(n)}(x)}{n!}\epsilon^n$$

$$\begin{aligned}f(x+\epsilon) &= \sum_{n=0}^{\infty} \frac{f^{(n)}(x)}{n!}\epsilon^n = \sum_{n=0}^{\infty} \frac{\frac{\mathrm{d}^n}{\mathrm{d}x^n}f(x)}{n!}\epsilon^n = \sum_{n=0}^{\infty} \frac{\left(\frac{\mathrm{d}}{\mathrm{d}x}\right)^n f(x)}{n!}\epsilon^n \\ &= \sum_{n=0}^{\infty} \frac{\left(\epsilon\frac{\mathrm{d}}{\mathrm{d}x}\right)^n f(x)}{n!} = \left(\sum_{n=0}^{\infty} \frac{\left(\epsilon\frac{\mathrm{d}}{\mathrm{d}x}\right)^n}{n!}\right) f(x) = \mathrm{e}^{\epsilon\frac{\mathrm{d}}{\mathrm{d}x}} f(x)\end{aligned}$$

---

univariable Taylor operator

$$f(x+\epsilon) = \mathrm{e}^{\epsilon\frac{\mathrm{d}}{\mathrm{d}x}} f(x)$$

$$\mathrm{e}^{\epsilon\frac{\mathrm{d}}{\mathrm{d}x}} = \sum_{n=0}^{\infty} \frac{\left(\epsilon\frac{\mathrm{d}}{\mathrm{d}x}\right)^n}{n!} = 1 + \epsilon\frac{\mathrm{d}}{\mathrm{d}x} + \frac{1}{2}\epsilon^2\frac{\mathrm{d}^2}{\mathrm{d}x^2} + \cdots$$

$f(x) = mx + b,$

$$\begin{aligned}
\mathrm{e}^{\epsilon\frac{\mathrm{d}}{\mathrm{d}x}} f(x) = & \left(1 + \epsilon\frac{\mathrm{d}}{\mathrm{d}x} + \frac{1}{2}\epsilon^2\frac{\mathrm{d}^2}{\mathrm{d}x^2} + \cdots\right)(mx+b) \\
= & 1(mx+b) + \epsilon\frac{\mathrm{d}}{\mathrm{d}x}(mx+b) + \frac{1}{2}\epsilon^2\frac{\mathrm{d}^2}{\mathrm{d}x^2}(mx+b) + \cdots \\
= & (mx+b) + (\epsilon m) + \left(\frac{1}{2}\epsilon^2 0\right) + (0 + 0 + \cdots) \\
= & (mx+b) + (\epsilon m) + (0 + 0 + \cdots) \\
= & (mx+b) + (\epsilon m) + 0 \\
= & mx + b + m\epsilon = m(x+\epsilon) + b = f(x+\epsilon)
\end{aligned}$$

---

https://www.youtube.com/watch?v=HQsZG8Yxb7w&t=15m34s

multivariable ...

$$\begin{aligned}
f(x+\epsilon) = & \mathrm{e}^{\epsilon\frac{\mathrm{d}}{\mathrm{d}x}} f(x) \\
f(x+\epsilon_x, y+\epsilon_y, z+\epsilon_z) = & f(x, y+\epsilon_y, z+\epsilon_z) \\
& + \epsilon\frac{\partial}{\partial x} f(x, y+\epsilon_y, z+\epsilon_z) \\
& + \frac{1}{2}\epsilon^2\frac{\partial^2}{\partial x^2} f(x, y+\epsilon_y, z+\epsilon_z) \\
& + \cdots
\end{aligned}$$

$$f(x+\epsilon) = \mathrm{e}^{\epsilon\frac{\mathrm{d}}{\mathrm{d}x}} f(x)$$

$$f(x+\epsilon_x, y+\epsilon_y, z+\epsilon_z) = f(\boldsymbol{r}+\boldsymbol{\epsilon}) = \mathrm{e}^{\boldsymbol{\epsilon}\cdot\boldsymbol{\nabla}} f(\boldsymbol{r}), \begin{cases} \boldsymbol{r} = (x, y, z) \\ \boldsymbol{\epsilon} = (\epsilon_x, \epsilon_y, \epsilon_z) \\ \boldsymbol{\nabla} = \left(\frac{\partial}{\partial x}, \frac{\partial}{\partial y}, \frac{\partial}{\partial z}\right) \end{cases}$$

$$\boldsymbol{\epsilon}\cdot\boldsymbol{\nabla} = (\epsilon_x, \epsilon_y, \epsilon_z)\cdot\left(\frac{\partial}{\partial x}, \frac{\partial}{\partial y}, \frac{\partial}{\partial z}\right) = \epsilon_x\frac{\partial}{\partial x} + \epsilon_y\frac{\partial}{\partial y} + \epsilon_z\frac{\partial}{\partial z}$$

---

multivariable Taylor operator

$$f(\boldsymbol{r}+\boldsymbol{\epsilon}) = \mathrm{e}^{\boldsymbol{\epsilon}\cdot\boldsymbol{\nabla}} f(\boldsymbol{r})$$

---

https://www.youtube.com/watch?v=HQsZG8Yxb7w&t=17m36s

physics

45.1.1.1 making complicated equastion simple

linearize

simple pendulum

sin(x) to x linearization

potential energy

$$f(x_0+\epsilon) = \sum_{k=0}^{n}\frac{f^{(k)}(x_0)}{k!}\epsilon^k + \mathrm{O}\left(\epsilon^{n+1}\right)$$

$$U(x_0+\epsilon) = \sum_{k=0}^{n}\frac{U^{(k)}(x_0)}{k!}\epsilon^k + \mathrm{O}\left(\epsilon^{n+1}\right) = U(x_0) + U'(x_0)x + \frac{1}{2}U''(x_0)x^2 + \cdots$$

$$F = \frac{-\mathrm{d}U}{\mathrm{d}x} = -U'(x_0) - U''(x_0)x - \cdots \overset{\text{if } U'(x_0)=0}{=} -U''(x_0)x = -kx, k = U''(x_0)$$

45.1.1.2 non-relativistic limit

https://www.youtube.com/watch?v=HQsZG8Yxb7w&t=22m32s

$$
\begin{aligned}
E =&\sqrt{(mc^2)^2+(pc)^2} = \sqrt{m^2c^4+p^2c^2}\\
=&mc^2\sqrt{1+\left(\frac{pc}{mc^2}\right)^2} = mc^2\sqrt{1+\left(\frac{p}{mc}\right)^2} = mc^2\left(1+\left(\frac{p}{mc}\right)^2\right)^{\frac{1}{2}}\\
=&mc^2\left[\frac{1}{0!}+\frac{1}{1!}\left(\frac{1}{2}\right)\left(\left(\frac{p}{mc}\right)^2\right)^1+\frac{1}{2!}\left(\frac{1}{2}\right)\left(\frac{1}{2}-1\right)\left(\left(\frac{p}{mc}\right)^2\right)^2+\cdots\right]\\
=&mc^2\left[1+\frac{1}{2}\left(\frac{p}{mc}\right)^2-\frac{1}{8}\left(\frac{p}{mc}\right)^4+\cdots\right] = mc^2\left[1+\frac{1}{2}\frac{p^2}{m^2c^2}-\frac{1}{8}\frac{p^4}{m^4c^4}+\cdots\right]\\
=&mc^2+\frac{1}{2}\frac{p^2}{m}-\frac{1}{8}\frac{p^4}{m^3c^2}+\cdots = mc^2+\frac{p^2}{2m}-\frac{p^4}{8m^3c^2}+\cdots\\
=&E_0+\frac{p^2}{2m}-\frac{p^4}{8m^3c^2}+\cdots, \begin{cases} E_0 = mc^2 \\ \frac{p^2}{2m} = \frac{(mv)^2}{2m} = \frac{mv^2}{2} = \frac{1}{2}mv^2 \end{cases}
\end{aligned}
$$

sluppy here $p = \gamma mv$

https://www.youtube.com/watch?v=HQsZG8Yxb7w&t=24m50s

leading relativistic correction

$$
-\frac{p^4}{8m^3c^2}
$$

binding energy of hydrogen atom

https://www.youtube.com/watch?v=HQsZG8Yxb7w&t=25m40s

fine-structure constant

45.1.1.3 quantum momentum operator

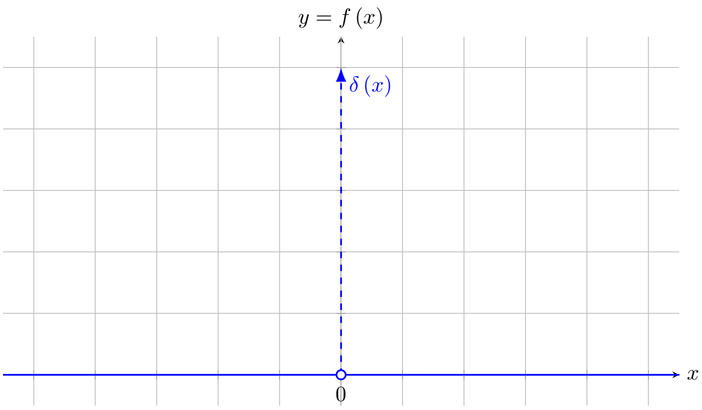
https://www.youtube.com/watch?v=HQsZG8Yxb7w&t=29m13s

45.2 Dirac delta function

Dirac function = Dirac delta function

https://tikz.net/delta_function/

https://tikz.dev/tikz-arrows

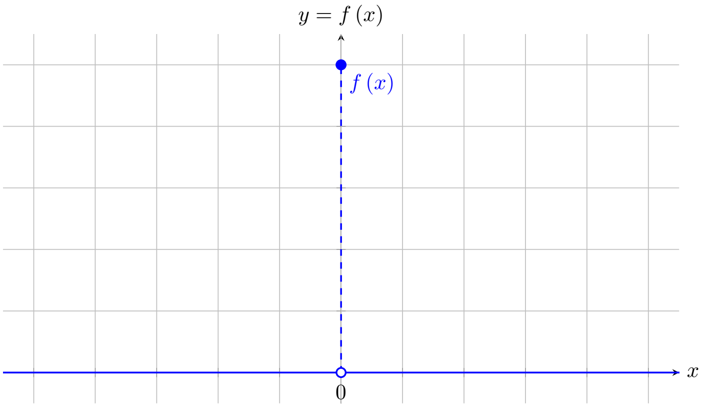


$$
\delta(x) = \begin{cases} 0 & x \neq 0 \\ \infty & x = 0 \end{cases}, \int_{-\infty}^{\infty} \delta(x)\,\mathrm{d}x = 1
$$

Figure 45.1: Dirac function = Dirac delta function

$$
\operatorname{supp}(f) = \left\{ x \,\middle|\, \begin{cases} x \in \mathcal{D} \\ f(x) \neq 0 \end{cases} \right\}
$$

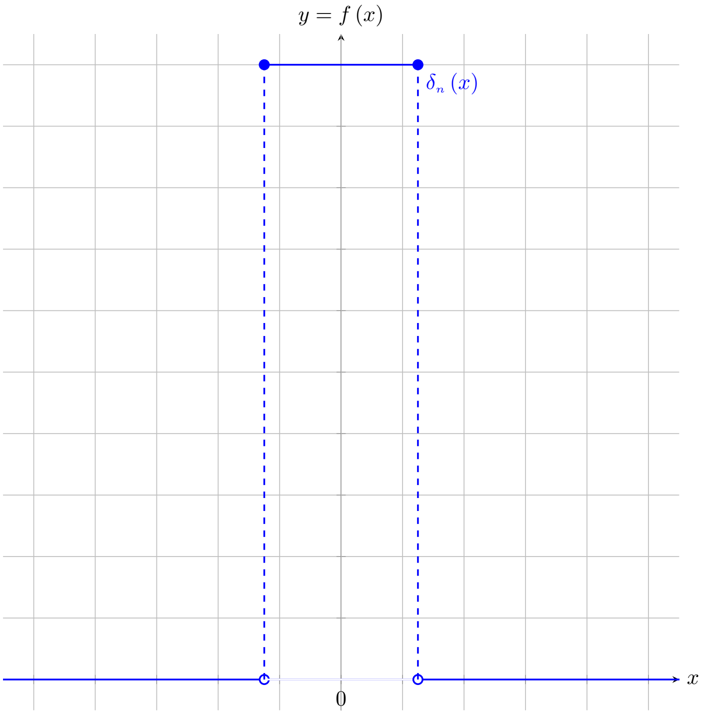
$$\operatorname{supp}(\delta) = \left\{ x \middle| \begin{cases} x \in \mathcal{D} = \mathbb{R} \\ \delta(x) \neq 0 \end{cases} \right\} = \{0\}$$



$$f(x) = \begin{cases} 0 & x \neq 0 \\ 1 & x = 0 \end{cases}$$

Figure 45.2: $f(x) = 1$ if $x = 0$ else 0

https://tex.stackexchange.com/questions/45275/tikz-get-values-for-predefined-dash-patterns



$$\delta_n(x) = \begin{cases} 0 & |x| > \frac{1}{2n} \\ n & |x| \leq \frac{1}{2n} \end{cases}, \forall n \in \mathbb{N}$$

$$\delta(x) = \{\delta_n(x) | n \in \mathbb{N}\} = \{\delta_1(x), \delta_2(x), \cdots\} = \lim_{n \to \infty} \delta_n(x)$$

Figure 45.3: $\delta_n(x)$

$$\delta_n(x) = \begin{cases} 0 & |x| > \frac{1}{2n} \Leftrightarrow \begin{cases} x > \frac{1}{2n} \\ x < \frac{-1}{2n} \end{cases} \\ n & |x| \leq \frac{1}{2n} \Leftrightarrow \frac{-1}{2n} \leq x \leq \frac{1}{2n} \Leftrightarrow x \in \left[\frac{-1}{2n}, \frac{1}{2n}\right] \end{cases}, \forall n \in \mathbb{N}$$

$$\int_{-\infty}^{\infty} \delta_n(x)\,\mathrm{d}x = \int_{-\frac{1}{2n}}^{\frac{1}{2n}} n\mathrm{d}x = n \int_{-\frac{1}{2n}}^{\frac{1}{2n}} \mathrm{d}x = n\,[x]_{\frac{-1}{2n}}^{\frac{1}{2n}} = n\left[\frac{1}{2n} - \frac{-1}{2n}\right] = n \cdot \frac{1}{n} = 1$$

$$
\begin{aligned}
\int_{-\infty}^{\infty}\delta(x)f(x)\,\mathrm{d}x &= \lim_{n\to\infty}\int_{-\infty}^{\infty}\delta_n(x)f(x)\,\mathrm{d}x \\
&= \lim_{n\to\infty}\int_{-\frac{1}{2n}}^{\frac{1}{2n}}\delta_n(x)f(x)\,\mathrm{d}x \\
&= \lim_{n\to\infty}\int_{-\frac{1}{2n}}^{\frac{1}{2n}} nf(x)\,\mathrm{d}x, \begin{cases} f(x)\in[m,M]\subseteq f\left(\left[\frac{-1}{2n},\frac{1}{2n}\right]\right) \\ \Downarrow \\ m\le f(x)\le M \end{cases}
\end{aligned}
$$

$$\int_{-\frac{1}{2n}}^{\frac{1}{2n}} nm\mathrm{d}x \le \int_{-\frac{1}{2n}}^{\frac{1}{2n}} nf(x)\,\mathrm{d}x \le \int_{-\frac{1}{2n}}^{\frac{1}{2n}} nM\mathrm{d}x$$

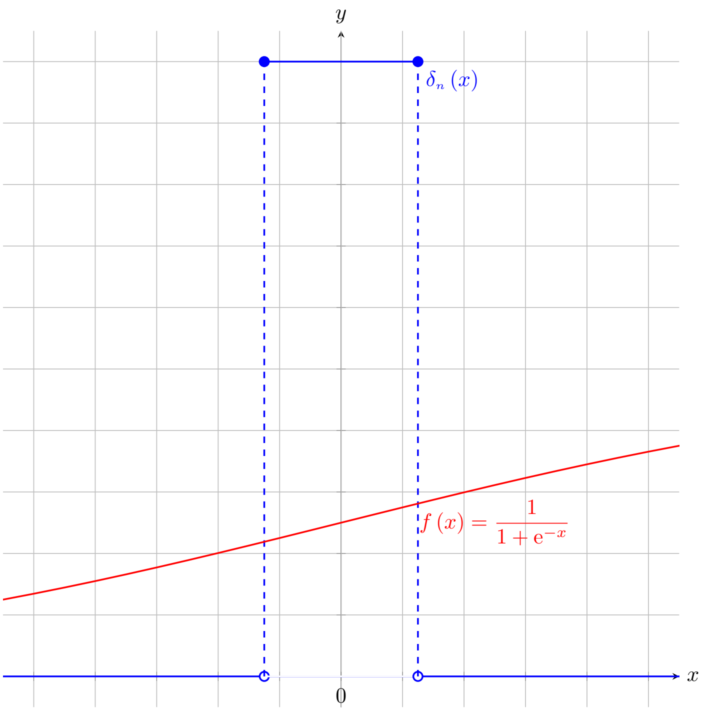
$$m = m\cdot 1 = m\int_{-\frac{1}{2n}}^{\frac{1}{2n}} n\mathrm{d}x \le \int_{-\frac{1}{2n}}^{\frac{1}{2n}} nf(x)\,\mathrm{d}x \le M\int_{-\frac{1}{2n}}^{\frac{1}{2n}} n\mathrm{d}x = M\cdot 1 = M$$

$$\Downarrow\mathbf{??}$$

$$\int_{-\frac{1}{2n}}^{\frac{1}{2n}} nf(x)\,\mathrm{d}x \overset{\exists\xi_n\in\left(\frac{-1}{2n},\frac{1}{2n}\right)}{=} f(\xi_n)\int_{-\frac{1}{2n}}^{\frac{1}{2n}} n\mathrm{d}x = f(\xi_n)\cdot 1 = f(\xi_n)$$

$$
\begin{aligned}
\int_{-\infty}^{\infty}\delta(x)f(x)\,\mathrm{d}x = \lim_{n\to\infty}\int_{-\frac{1}{2n}}^{\frac{1}{2n}} nf(x)\,\mathrm{d}x &= \lim_{n\to\infty} f(\xi_n), \xi_n\in\left(\frac{-1}{2n},\frac{1}{2n}\right) \\
&= f(0) \\
\int_{-\infty}^{\infty}\delta(x)f(x)\,\mathrm{d}x &= f(0)
\end{aligned}
$$

$$\int_{-\infty}^{\infty}\delta(x)f(x)\,\mathrm{d}x = f(0)$$



$$
\begin{aligned}
\int_{-\infty}^{\infty}\delta(x)f(x)\,\mathrm{d}x &= \lim_{n\to\infty}\int_{-\frac{1}{2n}}^{\frac{1}{2n}}\delta_n(x)f(x)\,\mathrm{d}x \\
&= \lim_{n\to\infty}\int_{-\frac{1}{2n}}^{\frac{1}{2n}} nf(x)\,\mathrm{d}x \\
&= \lim_{n\to\infty} f(\xi_n), \xi_n\in\left(\frac{-1}{2n},\frac{1}{2n}\right)
\end{aligned}
$$

$$\int_{-\infty}^{\infty}\delta(x)f(x)\,\mathrm{d}x = f(0)$$

$$\int_{-\infty}^{\infty}\delta(x-0)f(x)\,\mathrm{d}x = f(0)$$

Figure 45.4: $\int_{-\infty}^{\infty}\delta(x)f(x)\,\mathrm{d}x = f(0)$

$$\int_{-\infty}^{\infty}\delta(x)f(x)\,\mathrm{d}x = f(0)$$

$$\int_{-\infty}^{\infty}\delta(x-0)f(x)\,\mathrm{d}x = f(0)$$

$$\int_{-\infty}^{\infty} \delta(x - x_0) f(x) \, \mathrm{d}x = f(x_0)$$

$$\int_{-\infty}^{\infty} \delta(x - x') f(x) \, \mathrm{d}x = f(x')$$

---

https://www.youtube.com/watch?v=nDa3cqFk80o

$$\begin{aligned}
&\left\{ \{f_n(x) | n \in \mathbb{N}\} \middle| \begin{cases} f_n : \mathbb{R} \to \mathbb{R} \\ \int_{-\infty}^{\infty} f_n(x) \, \mathrm{d}x = 1 \end{cases} \right\} \\
=& \left\{ \left\{ \frac{n}{2} \mathrm{e}^{-n|x|} \middle| n \in \mathbb{N} \right\}, \right. \\
& \left\{ \frac{1}{\pi} \frac{n}{n^2 x^2 + 1} \middle| n \in \mathbb{N} \right\}, \\
& \left\{ \frac{n}{\sqrt{\pi}} \mathrm{e}^{-n^2 x^2} \middle| n \in \mathbb{N} \right\}, \\
& \left. \left\{ \frac{1}{\pi} \frac{\sin(nx)}{x} \middle| n \in \mathbb{N} \right\}, \cdots \right\}
\end{aligned}$$

$$\left\{ \delta(x) \middle| \begin{cases} \delta : \mathbb{R} \to \mathbb{R} \\ \int_{-\infty}^{\infty} \delta(x) \cdot 1 \mathrm{d}x = 1 \end{cases} \right\}$$

$$\left\{ \delta(x) \middle| \begin{cases} \delta : \mathbb{R} \to \mathbb{R} \\ \int_{-\infty}^{\infty} \delta(x - 0) \cdot f(x) \, \mathrm{d}x = f(0) \end{cases} \right\}$$

In measure theory, we can define the distance of two functions by

$$d(f, g) = \sqrt{\int_{-\infty}^{\infty} [f(x) - g(x)]^2 \, \mathrm{d}x}$$

for real distance of two square delta function approximations,

$$\begin{aligned}
d(\delta_m, \delta_n) =& \sqrt{\int_{-\infty}^{\infty} [\delta_m(x) - \delta_n(x)]^2 \, \mathrm{d}x} \\
=& \sqrt{\int_{-\infty}^{\infty} [\delta_m(x)]^2 - 2\delta_m(x) \delta_n(x) + [\delta_n(x)]^2 \, \mathrm{d}x} \\
=& \sqrt{\int_{-\infty}^{\infty} [\delta_m(x)]^2 \, \mathrm{d}x - 2 \int_{-\infty}^{\infty} \delta_m(x) \delta_n(x) \, \mathrm{d}x + \int_{-\infty}^{\infty} [\delta_n(x)]^2 \, \mathrm{d}x} \\
=& \sqrt{\int_{-\infty}^{\infty} \delta_m(x) \delta_m(x) \, \mathrm{d}x - 2 \int_{-\infty}^{\infty} \delta_m(x) \delta_n(x) \, \mathrm{d}x + \int_{-\infty}^{\infty} \delta_n(x) \delta_n(x) \, \mathrm{d}x} \\
=& \sqrt{\delta_m(\xi_m) - 2 \int_{-\infty}^{\infty} \delta_m(x) \delta_n(x) \, \mathrm{d}x + \delta_n(\xi_n)}, \begin{cases} \xi_m \in \left(\frac{-1}{2m}, \frac{1}{2m} \right) \\ \xi_n \in \left(\frac{-1}{2n}, \frac{1}{2n} \right) \end{cases} \\
=& \sqrt{m - 2 \int_{-\infty}^{\infty} \delta_m(x) \delta_n(x) \, \mathrm{d}x + n}, \text{ if } m > n \\
=& \sqrt{m - 2\delta_n(\xi_n) + n}, \xi_n \in \left(\frac{-1}{2n}, \frac{1}{2n} \right) \\
=& \sqrt{m - 2n + n} = \sqrt{m - n} \in \mathbb{R} \\
d(\delta_m, \delta_n) \overset{m \geq n}{=}& \sqrt{m - n} \in \mathbb{R}
\end{aligned}$$

$\langle d(\delta_m, \delta_n) \rangle_{n \in \mathbb{N}} = \left\langle \sqrt{m - n} \right\rangle_{n \in \mathbb{N}}$ is not a Cauchy series, not even mentioned convergence

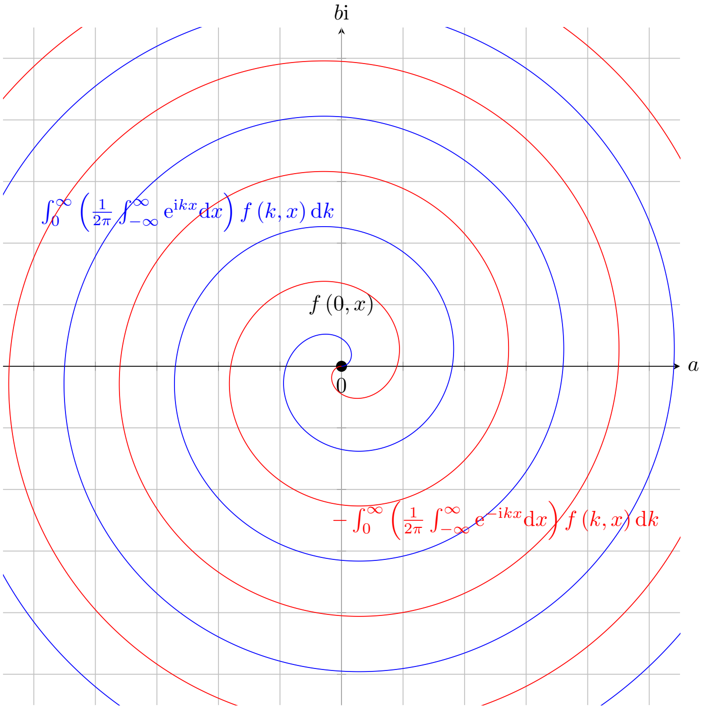
Def: 45.2

45.2.1 complex delta function

$$\delta\left(k\right)=\frac{1}{2\pi}\int_{-\infty}^{\infty}\mathrm{e}^{\mathrm{i}kx}\mathrm{d}x$$

$$\begin{aligned}
&\int_{-\infty}^{\infty}\left(\frac{1}{2\pi}\int_{-\infty}^{\infty}\mathrm{e}^{\mathrm{i}kx}\mathrm{d}x\right)f\left(k,x\right)\mathrm{d}k\\
&=\lim_{k'\to 0}\left[\int_{-\infty}^{k'-}\left(\frac{1}{2\pi}\int_{-\infty}^{\infty}\mathrm{e}^{\mathrm{i}kx}\mathrm{d}x\right)f\left(k,x\right)\mathrm{d}k+\int_{k'-}^{k'+}\left(\frac{1}{2\pi}\int_{-\infty}^{\infty}\mathrm{e}^{\mathrm{i}kx}\mathrm{d}x\right)f\left(k,x\right)\mathrm{d}k+\int_{k'+}^{\infty}\left(\frac{1}{2\pi}\int_{-\infty}^{\infty}\mathrm{e}^{\mathrm{i}kx}\mathrm{d}x\right)f\left(k,x\right)\mathrm{d}k\right]\\
&\approx\lim_{k'\to 0}\left[\int_{-\infty}^{0}\left(\frac{1}{2\pi}\int_{-\infty}^{\infty}\mathrm{e}^{\mathrm{i}kx}\mathrm{d}x\right)f\left(k,x\right)\mathrm{d}k+\int_{k'-}^{k'+}\left(\frac{1}{2\pi}\int_{-\infty}^{\infty}\mathrm{e}^{\mathrm{i}0x}\mathrm{d}x\right)f\left(0,x\right)\mathrm{d}k+\int_{0}^{\infty}\left(\frac{1}{2\pi}\int_{-\infty}^{\infty}\mathrm{e}^{\mathrm{i}kx}\mathrm{d}x\right)f\left(k,x\right)\mathrm{d}k\right]\\
&\approx\lim_{k'\to 0}\left[\int_{-\infty}^{0}\left(\frac{1}{2\pi}\int_{-\infty}^{\infty}\mathrm{e}^{\mathrm{i}kx}\mathrm{d}x\right)f\left(k,x\right)\mathrm{d}k+\int_{k'-}^{k'+}\left(\frac{1}{2\pi}\int_{-\infty}^{\infty}1\mathrm{d}x\right)f\left(0,x\right)\mathrm{d}k+\int_{0}^{\infty}\left(\frac{1}{2\pi}\int_{-\infty}^{\infty}\mathrm{e}^{\mathrm{i}kx}\mathrm{d}x\right)f\left(k,x\right)\mathrm{d}k\right]\\
&\approx\lim_{k'\to 0}\left[-\int_{0}^{\infty}\left(\frac{1}{2\pi}\int_{-\infty}^{\infty}\mathrm{e}^{-\mathrm{i}kx}\mathrm{d}x\right)f\left(k,x\right)\mathrm{d}k+\int_{k'-}^{k'+}\left(\frac{1}{2\pi}\int_{-\infty}^{\infty}1\mathrm{d}x\right)f\left(0,x\right)\mathrm{d}k+\int_{0}^{\infty}\left(\frac{1}{2\pi}\int_{-\infty}^{\infty}\mathrm{e}^{\mathrm{i}kx}\mathrm{d}x\right)f\left(k,x\right)\mathrm{d}k\right]\\
&\approx\lim_{k'\to 0}\left[-\int_{0}^{\infty}\left(\frac{1}{2\pi}\int_{-\infty}^{\infty}\mathrm{e}^{-\mathrm{i}kx}\mathrm{d}x\right)f\left(x\right)\mathrm{d}k+\int_{k'-}^{k'+}\left(\frac{1}{2\pi}\infty\right)f\left(0,x\right)\mathrm{d}k+\int_{0}^{\infty}\left(\frac{1}{2\pi}\int_{-\infty}^{\infty}\mathrm{e}^{\mathrm{i}kx}\mathrm{d}x\right)f\left(k,x\right)\mathrm{d}k\right]\\
&\approx\lim_{k'\to 0}\left[\int_{k'-}^{k'+}\frac{1}{\mathrm{d}k}f\left(0,x\right)\mathrm{d}k+\int_{0}^{\infty}\left(\frac{1}{2\pi}\int_{-\infty}^{\infty}\mathrm{e}^{\mathrm{i}kx}-\mathrm{e}^{-\mathrm{i}kx}\mathrm{d}x\right)f\left(k,x\right)\mathrm{d}k\right]\\
&\approx\lim_{k'\to 0}\left[f\left(0,x\right)+\int_{0}^{\infty}\left(\frac{1}{2\pi}0\right)f\left(k,x\right)\mathrm{d}k\right]\approx f\left(0,x\right)\\
&f\left(0,x\right)=\int_{-\infty}^{\infty}\left(\frac{1}{2\pi}\int_{-\infty}^{\infty}\mathrm{e}^{\mathrm{i}kx}\mathrm{d}x\right)f\left(k,x\right)\mathrm{d}k=\int_{-\infty}^{\infty}\delta\left(k\right)f\left(k,x\right)\mathrm{d}k=\int_{-\infty}^{\infty}\delta\left(k-0\right)f\left(k,x\right)\mathrm{d}k
\end{aligned}$$

https://tex.stackexchange.com/questions/150138/how-can-i-create-a-polar-plot-on-a-cartesian-grid



$$\begin{aligned}
f\left(0,x\right)&=\int_{-\infty}^{\infty}\left(\frac{1}{2\pi}\int_{-\infty}^{\infty}\mathrm{e}^{\mathrm{i}kx}\mathrm{d}x\right)f\left(k,x\right)\mathrm{d}k\\
&=\int_{-\infty}^{\infty}\delta\left(k-0\right)f\left(k,x\right)\mathrm{d}k
\end{aligned}$$

Figure 45.5: $\int_{-\infty}^{\infty}\left(\frac{1}{2\pi}\int_{-\infty}^{\infty}\mathrm{e}^{\mathrm{i}kx}\mathrm{d}x\right)f\left(k,x\right)\mathrm{d}k$

According to Fourier transform and inverse transform, Fourier analysis[50]

$$f(t)=\frac{1}{2\pi}\int_{-\infty}^{\infty}\mathrm{e}^{\mathrm{i}t\cdot\omega}\left(\int_{-\infty}^{\infty}\mathrm{e}^{-\mathrm{i}\omega\cdot s}f(s)\,\mathrm{d}s\right)\mathrm{d}\omega$$

$$f(x)=\frac{1}{2\pi}\int_{-\infty}^{\infty}\mathrm{e}^{\mathrm{i}xk}\left(\int_{-\infty}^{\infty}\mathrm{e}^{-\mathrm{i}kx'}f(x')\,\mathrm{d}x'\right)\mathrm{d}k$$

or

$$f(t)=\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty}\mathrm{e}^{\mathrm{i}t\cdot\omega}\left(\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty}\mathrm{e}^{-\imath\omega\cdot s}f(s)\,\mathrm{d}s\right)\mathrm{d}\omega$$

$$f(t)=\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty}\mathrm{e}^{\mathrm{i}xk}\left(\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty}\mathrm{e}^{-\mathrm{i}kx'}f(x')\,\mathrm{d}x'\right)\mathrm{d}k$$

---

complex function if analytic always with better properties than real function

$$\begin{aligned}&\left\{\{\delta(k)|k\in\mathbb{R}\}\middle|\begin{cases}\delta:\mathbb{R}\to\mathbb{C}\\ \int_{-\infty}^{\infty}\delta(k-0)\,f(k,x)\,\mathrm{d}k=f(0,x)\end{cases}\right\}\\ =&\left\{\left\{K\int_{-\infty}^{\infty}\mathrm{e}^{\mathrm{i}kx}(\cdot)\,\mathrm{d}k\middle|k\in\mathbb{R}\right\},\cdots\right\}\end{aligned}$$

$$\delta(k)=\frac{1}{2\pi}\int_{-\infty}^{\infty}\mathrm{e}^{\mathrm{i}kx}(\cdot)\,\mathrm{d}x$$

$$\delta(x)=\frac{1}{2\pi}\int_{-\infty}^{\infty}\mathrm{e}^{\mathrm{i}kx}(\cdot)\,\mathrm{d}k$$

$$\delta(x')=\frac{1}{2\pi}\int_{-\infty}^{\infty}\mathrm{e}^{\mathrm{i}kx'}(\cdot)\,\mathrm{d}k$$

$\psi(x)=\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty}\mathrm{e}^{\mathrm{i}kx}\psi(k)\,\mathrm{d}k$

$$\psi(k)=\frac{1}{\sqrt{2\pi}}\int\mathrm{e}^{-\mathrm{i}kx}\psi(x)\,\mathrm{d}x$$

$$\begin{aligned}&\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty}\mathrm{e}^{\mathrm{i}kx}\psi(k)\,\mathrm{d}k\\ =&\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty}\mathrm{e}^{\mathrm{i}kx}\left(\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty}\mathrm{e}^{-\mathrm{i}kx'}\psi(x')\,\mathrm{d}x'\right)\mathrm{d}k\\ =&\frac{1}{2\pi}\int_{-\infty}^{\infty}\left(\int_{-\infty}^{\infty}\mathrm{e}^{\mathrm{i}kx}\mathrm{e}^{-\mathrm{i}kx'}\psi(x')\,\mathrm{d}x'\right)\mathrm{d}k\\ \overset{\text{Fubini}}{=}&\frac{1}{2\pi}\int_{-\infty}^{\infty}\left(\int_{-\infty}^{\infty}\mathrm{e}^{\mathrm{i}k\left(x-x'\right)}\psi(x')\,\mathrm{d}k\right)\mathrm{d}x'\\ =&\int_{-\infty}^{\infty}\left(\frac{1}{2\pi}\int_{-\infty}^{\infty}\mathrm{e}^{\mathrm{i}k\left(x-x'\right)}\psi(x')\,\mathrm{d}k\right)\mathrm{d}x'\\ =&\int_{-\infty}^{\infty}\psi(x')\left(\frac{1}{2\pi}\int_{-\infty}^{\infty}\mathrm{e}^{\mathrm{i}k\left(x-x'\right)}\mathrm{d}k\right)\mathrm{d}x'\\ =&\int_{-\infty}^{\infty}\psi(x')\,\delta(x-x')\,\mathrm{d}x'=\int_{-\infty}^{\infty}\delta(x-x')\,\psi(x')\,\mathrm{d}x'=\psi(x)\end{aligned}$$

$$\psi(x)=\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty}\mathrm{e}^{\mathrm{i}kx}\psi(k)\,\mathrm{d}k$$

$\int_{-\infty}^{\infty} |\psi(k)|^2 \, \mathrm{d}k = \int_{-\infty}^{\infty} |\psi(x)|^2 \, \mathrm{d}x$

$$\begin{cases} \psi(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \mathrm{e}^{\mathrm{i}kx} \psi(k) \, \mathrm{d}k \\ \psi(k) = \dfrac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \mathrm{e}^{-\mathrm{i}kx} \psi(x) \, \mathrm{d}x \end{cases}$$

$$\begin{aligned} \psi^*(k) = \overline{\psi}(k) = & \overline{\frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \mathrm{e}^{-\mathrm{i}kx} \psi(x) \, \mathrm{d}x} \\ = & \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \overline{\mathrm{e}^{-\mathrm{i}kx} \psi(x)} \mathrm{d}x \\ = & \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \overline{\mathrm{e}^{-\mathrm{i}kx}} \, \overline{\psi(x)} \mathrm{d}x \\ = & \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \mathrm{e}^{\mathrm{i}kx} \psi^*(x) \, \mathrm{d}x \end{aligned}$$

$$\begin{aligned} & \int_{-\infty}^{\infty} |\psi(k)|^2 \, \mathrm{d}k = \int_{-\infty}^{\infty} \psi^*(k) \, \psi(k) \, \mathrm{d}k \\ = & \int_{-\infty}^{\infty} \left(\frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \mathrm{e}^{\mathrm{i}kx} \psi^*(x) \, \mathrm{d}x \right) \left(\frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \mathrm{e}^{-\mathrm{i}kx'} \psi(x') \, \mathrm{d}x' \right) \mathrm{d}k \\ = & \frac{1}{2\pi} \int_{-\infty}^{\infty} \left(\int_{-\infty}^{\infty} \mathrm{e}^{\mathrm{i}kx} \psi^*(x) \, \mathrm{d}x \right) \left(\int_{-\infty}^{\infty} \mathrm{e}^{-\mathrm{i}kx'} \psi(x') \, \mathrm{d}x' \right) \mathrm{d}k \\ = & \frac{1}{2\pi} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \psi^*(x) \int_{-\infty}^{\infty} \mathrm{e}^{\mathrm{i}kx} \mathrm{e}^{-\mathrm{i}kx'} \psi(x') \, \mathrm{d}x' \mathrm{d}x \mathrm{d}k \\ = & \frac{1}{2\pi} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \psi^*(x) \int_{-\infty}^{\infty} \mathrm{e}^{\mathrm{i}k\left(x-x'\right)} \psi(x') \, \mathrm{d}x' \mathrm{d}x \mathrm{d}k \\ = & \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \psi^*(x) \frac{1}{2\pi} \int_{-\infty}^{\infty} \mathrm{e}^{\mathrm{i}k\left(x-x'\right)} \psi(x') \, \mathrm{d}x' \mathrm{d}x \mathrm{d}k \\ \overset{\text{Fubini}}{=} & \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \psi^*(x) \frac{1}{2\pi} \int_{-\infty}^{\infty} \mathrm{e}^{\mathrm{i}k\left(x-x'\right)} \psi(x') \, \mathrm{d}k \mathrm{d}x' \mathrm{d}x \\ = & \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \psi^*(x) \, \psi(x') \frac{1}{2\pi} \int_{-\infty}^{\infty} \mathrm{e}^{\mathrm{i}k\left(x-x'\right)} \mathrm{d}k \mathrm{d}x' \mathrm{d}x \\ = & \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \psi^*(x) \, \psi(x') \, \delta(x-x') \, \mathrm{d}x' \mathrm{d}x \\ = & \int_{-\infty}^{\infty} \psi^*(x) \int_{-\infty}^{\infty} \psi(x') \, \delta(x-x') \, \mathrm{d}x' \mathrm{d}x \\ = & \int_{-\infty}^{\infty} \psi^*(x) \, \psi(x) \, \mathrm{d}x = \int_{-\infty}^{\infty} |\psi(x)|^2 \, \mathrm{d}x \end{aligned}$$

$$\int_{-\infty}^{\infty} |\psi(k)|^2 \, \mathrm{d}k = \int_{-\infty}^{\infty} |\psi(x)|^2 \, \mathrm{d}x$$

---

convolution

$\psi(x) = \int_{-\infty}^{\infty} \psi_1(y) \, \psi_2(x-y) \, \mathrm{d}y \Rightarrow \psi(k) = \sqrt{2\pi} \psi_1(k) \, \psi_2(k)$

$$\psi(x) = \int_{-\infty}^{\infty} \psi_1(y) \, \psi_2(x-y) \, \mathrm{d}y$$

$$\frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \mathrm{e}^{\mathrm{i}k_1 y} \psi_1(k_1) \, \mathrm{d}k_1$$

$$\psi_2(x-y) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \mathrm{e}^{\mathrm{i}k_2(x-y)} \psi_2(k_2) \, \mathrm{d}k_2$$

$$\begin{cases} \psi(x) = & \int_{-\infty}^{\infty} \psi_1(y) \, \psi_2(x-y) \, \mathrm{d}y \\ \psi_1(y) = & \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \mathrm{e}^{\mathrm{i}k_1 y} \psi_1(k_1) \, \mathrm{d}k_1 \\ \psi_2(x-y) = & \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \mathrm{e}^{\mathrm{i}k_2(x-y)} \psi_2(k_2) \, \mathrm{d}k_2 \end{cases}$$

$$\begin{aligned}
\psi(k) &= \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty} \mathrm{e}^{-\mathrm{i}kx}\psi(x)\,\mathrm{d}x = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty} \mathrm{e}^{-\mathrm{i}kx}\int_{-\infty}^{\infty}\psi_1(y)\,\psi_2(x-y)\,\mathrm{d}y\mathrm{d}x \\
&= \frac{1}{\left(\sqrt{2\pi}\right)^3}\int_{-\infty}^{\infty}\mathrm{e}^{-\mathrm{i}kx}\int_{-\infty}^{\infty}\left(\int_{-\infty}^{\infty}\mathrm{e}^{\mathrm{i}k_1 y}\psi_1(k_1)\,\mathrm{d}k_1\right)\left(\int_{-\infty}^{\infty}\mathrm{e}^{\mathrm{i}k_2(x-y)}\psi_2(k_2)\,\mathrm{d}k_2\right)\mathrm{d}y\mathrm{d}x \\
&= \frac{1}{\left(\sqrt{2\pi}\right)^3}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\mathrm{e}^{\mathrm{i}k_1 y}\mathrm{e}^{-\mathrm{i}k_2 y}\psi_1(k_1)\,\mathrm{e}^{-\mathrm{i}kx}\mathrm{e}^{\mathrm{i}k_2 x}\psi_2(k_2)\,\mathrm{d}k_1\mathrm{d}k_2\mathrm{d}y\mathrm{d}x \\
&= \frac{1}{\left(\sqrt{2\pi}\right)^3}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\mathrm{e}^{\mathrm{i}(k_1-k_2)y}\psi_1(k_1)\,\mathrm{e}^{\mathrm{i}(k_2-k)x}\psi_2(k_2)\,\mathrm{d}k_1\mathrm{d}k_2\mathrm{d}y\mathrm{d}x \\
&\overset{\text{Fubini}}{=} \frac{1}{\left(\sqrt{2\pi}\right)^3}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\mathrm{e}^{\mathrm{i}(k_1-k_2)y}\psi_1(k_1)\,\mathrm{e}^{\mathrm{i}(k_2-k)x}\psi_2(k_2)\,\mathrm{d}y\mathrm{d}x\mathrm{d}k_1\mathrm{d}k_2 \\
&= \frac{1}{\left(\sqrt{2\pi}\right)^3}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\left(\int_{-\infty}^{\infty}\mathrm{e}^{\mathrm{i}(k_1-k_2)y}\psi_1(k_1)\,\mathrm{d}y\right)\left(\int_{-\infty}^{\infty}\mathrm{e}^{\mathrm{i}(k_2-k)x}\psi_2(k_2)\,\mathrm{d}x\right)\mathrm{d}k_1\mathrm{d}k_2
\end{aligned}$$

$$\begin{aligned}
&\frac{1}{\left(\sqrt{2\pi}\right)^3}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\left(\int_{-\infty}^{\infty}\mathrm{e}^{\mathrm{i}(k_1-k_2)y}\psi_1(k_1)\,\mathrm{d}y\right)\left(\int_{-\infty}^{\infty}\mathrm{e}^{\mathrm{i}(k_2-k)x}\psi_2(k_2)\,\mathrm{d}x\right)\mathrm{d}k_1\mathrm{d}k_2 \\
=&\sqrt{2\pi}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\left[\frac{1}{2\pi}\int_{-\infty}^{\infty}\mathrm{e}^{\mathrm{i}(k_1-k_2)y}\psi_1(k_1)\,\mathrm{d}y\right]\left[\frac{1}{2\pi}\int_{-\infty}^{\infty}\mathrm{e}^{\mathrm{i}(k_2-k)x}\psi_2(k_2)\,\mathrm{d}x\right]\mathrm{d}k_1\mathrm{d}k_2 \\
=&\sqrt{2\pi}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\left[\psi_1(k_1)\,\frac{1}{2\pi}\int_{-\infty}^{\infty}\mathrm{e}^{\mathrm{i}(k_1-k_2)y}\mathrm{d}y\right]\left[\psi_2(k_2)\,\frac{1}{2\pi}\int_{-\infty}^{\infty}\mathrm{e}^{\mathrm{i}(k_2-k)x}\mathrm{d}x\right]\mathrm{d}k_1\mathrm{d}k_2 \\
=&\sqrt{2\pi}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\left[\psi_1(k_1)\,\delta(k_1-k_2)\right]\left[\psi_2(k_2)\,\delta(k_2-k)\right]\mathrm{d}k_1\mathrm{d}k_2 \\
\overset{\text{Fubini}}{=}&\sqrt{2\pi}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\left[\psi_1(k_1)\,\delta(k_1-k_2)\right]\left[\psi_2(k_2)\,\delta(k_2-k)\right]\mathrm{d}k_2\mathrm{d}k_1 \\
=&\sqrt{2\pi}\int_{-\infty}^{\infty}\left[\psi_1(k_1)\right]\left[\int_{-\infty}^{\infty}\delta(k_1-k_2)\,\psi_2(k_2)\,\delta(k_2-k)\,\mathrm{d}k_2\right]\mathrm{d}k_1 \\
=&\sqrt{2\pi}\int_{-\infty}^{\infty}\left[\psi_1(k_1)\right]\left[\delta(k_1-k)\,\psi_2(k)\right]\mathrm{d}k_1 = \sqrt{2\pi}\psi_2(k)\int_{-\infty}^{\infty}\psi_1(k_1)\,\delta(k_1-k)\,\mathrm{d}k_1 \\
=&\sqrt{2\pi}\psi_2(k)\,\psi_1(k) = \sqrt{2\pi}\psi_1(k)\,\psi_2(k)
\end{aligned}$$

$$\begin{gathered}
\begin{cases}
\psi(x) = & \int_{-\infty}^{\infty}\psi_1(y)\,\psi_2(x-y)\,\mathrm{d}y \\
\psi_1(y) = & \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty}\mathrm{e}^{\mathrm{i}k_1 y}\psi_1(k_1)\,\mathrm{d}k_1 \\
\psi_2(x-y) = & \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty}\mathrm{e}^{\mathrm{i}k_2(x-y)}\psi_2(k_2)\,\mathrm{d}k_2
\end{cases} \\
\Downarrow \\
\psi(k) = \sqrt{2\pi}\psi_1(k)\,\psi_2(k)
\end{gathered}$$

45.2.2 3D delta function

https://www.youtube.com/watch?v=Y8y965ZAmQE

$$\delta(\boldsymbol{r}-\boldsymbol{r}') = \delta(x-x')\,\delta(y-y')\,\delta(z-z')$$

Poisson equation

Feynman method of differentiation / derivative[44.1]

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

$\Delta\left(\frac{1}{r}\right)=-4\pi\delta\left(\boldsymbol{r}\right)$

https://math.stackexchange.com/questions/3774483/derivatives-of-frac1r-and-dirac-delta-function

$$\begin{aligned}
\Delta\frac{1}{r\left(\epsilon\right)} =&\boldsymbol{\nabla}^2\frac{1}{\sqrt{\boldsymbol{r}^2+\epsilon^2}}=\boldsymbol{\nabla}\cdot\boldsymbol{\nabla}\frac{1}{\sqrt{x^2+y^2+z^2+\epsilon^2}}\\
=&\boldsymbol{\nabla}\cdot\left(\frac{\partial}{\partial x},\frac{\partial}{\partial y},\frac{\partial}{\partial z}\right)\left(\frac{1}{\sqrt{x^2+y^2+z^2+\epsilon^2}}\right)\\
=&\boldsymbol{\nabla}\cdot\left(\frac{-x}{(x^2+y^2+z^2+\epsilon^2)^{\frac{3}{2}}},\frac{-y}{(x^2+y^2+z^2+\epsilon^2)^{\frac{3}{2}}},\frac{-z}{(x^2+y^2+z^2+\epsilon^2)^{\frac{3}{2}}}\right)\\
=&\frac{\partial}{\partial x}\frac{-x}{(x^2+y^2+z^2+\epsilon^2)^{\frac{3}{2}}}+\frac{\partial}{\partial y}\frac{-y}{(x^2+y^2+z^2+\epsilon^2)^{\frac{3}{2}}}+\frac{\partial}{\partial z}\frac{-z}{(x^2+y^2+z^2+\epsilon^2)^{\frac{3}{2}}}\\
=&\frac{+2x^2-y^2-z^2-\epsilon^2}{(x^2+y^2+z^2+\epsilon^2)^{\frac{5}{2}}}+\frac{-x^2+2y^2-z^2-\epsilon^2}{(x^2+y^2+z^2+\epsilon^2)^{\frac{5}{2}}}+\frac{-x^2-y^2+2z^2-\epsilon^2}{(x^2+y^2+z^2+\epsilon^2)^{\frac{5}{2}}}\\
=&\frac{-3\epsilon^2}{(x^2+y^2+z^2+\epsilon^2)^{\frac{5}{2}}}=\frac{-3\epsilon^2}{(\boldsymbol{r}^2+\epsilon^2)^{\frac{5}{2}}}
\end{aligned}$$

$$\Delta\left(\frac{1}{r}\right)=\boldsymbol{\nabla}^2\left(\frac{1}{r}\right)=\boldsymbol{\nabla}^2\frac{1}{|\boldsymbol{r}|}=\boldsymbol{\nabla}^2\frac{1}{|\boldsymbol{r}-\boldsymbol{0}|}=\Delta\frac{1}{|\boldsymbol{r}-\boldsymbol{0}|}=\Delta\left(\boldsymbol{r}-\boldsymbol{0}\right)=\Delta\left(\boldsymbol{r}\right)$$

$$\Delta\frac{1}{r\left(\epsilon\right)}=\Delta_\epsilon\left(r\right)=\Delta_\epsilon\left(\boldsymbol{r}\right)=\Delta_\epsilon\left(\boldsymbol{r}-\boldsymbol{0}\right)$$

$$\lim_{\epsilon\to 0}\Delta_\epsilon\left(\boldsymbol{r}\right)=\lim_{\epsilon\to 0}\Delta\frac{1}{r\left(\epsilon\right)}=\Delta\left(\frac{1}{r}\right)=\Delta\left(\boldsymbol{r}\right)$$

$$K_3\int_{\mathbb{R}^3}f\left(\boldsymbol{r}\right)\Delta\left(\boldsymbol{r}\right)\mathrm{d}\boldsymbol{r}=\lim_{\epsilon\to 0}K_3\int_{\mathbb{R}^3}f\left(\boldsymbol{r}\right)\Delta_\epsilon\left(\boldsymbol{r}\right)\mathrm{d}\boldsymbol{r}=\lim_{\epsilon\to 0}K_3\int_{\mathbb{R}^3}f\left(\boldsymbol{r}\right)\Delta_\epsilon\left(\boldsymbol{r}-\boldsymbol{0}\right)\mathrm{d}\boldsymbol{r}=f\left(\boldsymbol{0}\right)$$

$$K_1\int_{\mathbb{R}^1}f\left(x\right)\delta\left(x\right)\mathrm{d}x=\lim_{n\to\infty}K_1\int_{\mathbb{R}^1}f\left(x\right)\delta_n\left(x\right)\mathrm{d}x=K_1\int_{\mathbb{R}^1}f\left(x\right)\delta_n\left(x-0\right)\mathrm{d}x=f\left(0\right)$$

$$\int_{\mathbb{R}^3} f(\boldsymbol{r})\,\Delta_\epsilon(\boldsymbol{r})\,\mathrm{d}\boldsymbol{r} = \int_{\mathbb{R}^3} f(\boldsymbol{r})\,\frac{-3\epsilon^2}{(\boldsymbol{r}^2+\epsilon^2)^{\frac{5}{2}}}\mathrm{d}\boldsymbol{r}, \begin{cases} \boldsymbol{r} = r\boldsymbol{n}_{\mathbb{S}} \\ \boldsymbol{n}_{\mathbb{S}} = \begin{pmatrix} \sin\phi\cos\theta \\ \sin\phi\sin\theta \\ \cos\phi \end{pmatrix} \end{cases}$$

$$= \int_0^\infty \int_{\mathbb{S}^2} f(r\boldsymbol{n}_{\mathbb{S}})\,\frac{-3\epsilon^2}{(r^2+\epsilon^2)^{\frac{5}{2}}} r^2\sigma(\mathrm{d}\boldsymbol{n}_S)\,\mathrm{d}r, \begin{cases} \mathbb{S}^2 = \text{unit sphere centered at the origin} \\ \sigma(\mathrm{d}\boldsymbol{n}_S) \;\text{ is the surface measure of } \mathbb{S}^2 \\ r^2\sigma(\mathrm{d}\boldsymbol{n}_S) = r^2\sin\phi\mathrm{d}\phi\mathrm{d}\theta \\ r = \epsilon s \end{cases}$$

$$= \int_0^\infty \int_{\mathbb{S}^2} f(\epsilon s\boldsymbol{n}_{\mathbb{S}})\,\frac{-3\epsilon^2}{\left((\epsilon s)^2+\epsilon^2\right)^{\frac{5}{2}}}(\epsilon s)^2\,\sigma(\mathrm{d}\boldsymbol{n}_S)\,\mathrm{d}\epsilon s$$

$$= \int_0^\infty \int_{\mathbb{S}^2} f(\epsilon s\boldsymbol{n}_{\mathbb{S}})\,\frac{-3\epsilon^5}{\epsilon^5(s^2+1)^{\frac{5}{2}}} s^2\sigma(\mathrm{d}\boldsymbol{n}_S)\,\mathrm{d}s = -\int_0^\infty \int_{\mathbb{S}^2} f(\epsilon s\boldsymbol{n}_{\mathbb{S}})\,\frac{3s^2}{(s^2+1)^{\frac{5}{2}}}\sigma(\mathrm{d}\boldsymbol{n}_{\mathbb{S}})\,\mathrm{d}s$$

$$= -\int_0^\infty \int_0^{2\pi}\int_0^{2\pi} f(\epsilon s\boldsymbol{n}_{\mathbb{S}})\,\frac{3s^2}{(s^2+1)^{\frac{5}{2}}}\sin\phi\mathrm{d}\phi\mathrm{d}\theta\mathrm{d}s = -\int_0^\infty 4\pi f(\epsilon s\boldsymbol{n}_{\mathbb{S}})\,\frac{3s^2}{(s^2+1)^{\frac{5}{2}}}\mathrm{d}s$$

$$\begin{aligned} &\lim_{\epsilon\to 0}\int_{\mathbb{R}^3} f(\boldsymbol{r})\,\Delta_\epsilon(\boldsymbol{r})\,\mathrm{d}\boldsymbol{r} \\ &= \lim_{\epsilon\to 0}\left[-\int_0^\infty 4\pi f(\epsilon s\boldsymbol{n}_{\mathbb{S}})\,\frac{3s^2}{(s^2+1)^{\frac{5}{2}}}\mathrm{d}s\right] \\ &= -4\pi f(\boldsymbol{0})\int_0^\infty \frac{3s^2}{(s^2+1)^{\frac{5}{2}}}\mathrm{d}s \end{aligned}$$

https://www.integral-calculator.com/

$$\begin{aligned} &\int \frac{3s^2}{(s^2+1)^{\frac{5}{2}}}\mathrm{d}s,\, s = \tan u \\ &= \int \frac{3\tan^2 u}{(\tan^2 u+1)^{\frac{5}{2}}}\mathrm{d}\tan u = \int \frac{3\tan^2 u}{(\tan^2 u+1)^{\frac{5}{2}}}\sec^2 u\mathrm{d}u \\ &= \int \frac{3\tan^2 u}{(\sec^2 u)^{\frac{5}{2}}}\sec^2 u\mathrm{d}u = \int \frac{3\tan^2 u}{\sec^5 u}\sec^2 u\mathrm{d}u = \int \frac{3\tan^2 u}{\sec^3 u}\mathrm{d}u \\ &= \int \frac{3\left(\frac{\sin u}{\cos u}\right)^2}{\left(\frac{1}{\cos u}\right)^3}\mathrm{d}u = 3\int \sin^2 u\cos u\mathrm{d}u = 3\int \sin^2 u\mathrm{d}\sin u = 3\frac{\sin^3 u}{3} + C \\ &= \sin^3 u + C = \left(\frac{s}{\sqrt{s^2+1}}\right)^3 + C = \frac{s^3}{(s^2+1)^{\frac{3}{2}}} + C \overset{s\geq 0}{=} \frac{1}{(1+s^{-2})^{\frac{3}{2}}} + C \end{aligned}$$

$$\int_0^\infty \frac{3s^2}{(s^2+1)^{\frac{5}{2}}}\mathrm{d}s = \left[\frac{s^3}{(s^2+1)^{\frac{3}{2}}}\right]_0^\infty = \left[\frac{1}{(1+s^{-2})^{\frac{3}{2}}}\right]_{s=\infty} - \left[\frac{s^3}{(s^2+1)^{\frac{3}{2}}}\right]_{s=0} = 1 - 0 = 1$$

$$\begin{aligned} &\lim_{\epsilon\to 0}\int_{\mathbb{R}^3} f(\boldsymbol{r})\,\Delta_\epsilon(\boldsymbol{r})\,\mathrm{d}\boldsymbol{r} \\ &= \lim_{\epsilon\to 0}\left[-\int_0^\infty 4\pi f(\epsilon s\boldsymbol{n}_{\mathbb{S}})\,\frac{3s^2}{(s^2+1)^{\frac{5}{2}}}\mathrm{d}s\right] \\ &= -4\pi f(\boldsymbol{0})\int_0^\infty \frac{3s^2}{(s^2+1)^{\frac{5}{2}}}\mathrm{d}s = -4\pi f(\boldsymbol{0})\,1 = -4\pi f(\boldsymbol{0}) \end{aligned}$$

$$\int_{\mathbb{R}^3} f(\boldsymbol{r})\,\Delta(\boldsymbol{r})\,\mathrm{d}\boldsymbol{r} = \lim_{\epsilon\to 0}\int_{\mathbb{R}^3} f(\boldsymbol{r})\,\Delta_\epsilon(\boldsymbol{r})\,\mathrm{d}\boldsymbol{r} = -4\pi f(\boldsymbol{0}) = -4\pi\int_{\mathbb{R}^3} f(\boldsymbol{r})\,\delta(\boldsymbol{r})\,\mathrm{d}\boldsymbol{r}$$

$$\Delta\left(\frac{1}{r}\right) = \Delta(\boldsymbol{r}) = -4\pi\delta(\boldsymbol{r})$$

Big Delta reciprocal r is minus 4 pi delta r.

$\square$

$$\Delta\left(\frac{1}{r}\right)=-4\pi\delta\left(\boldsymbol{r}\right)$$

---

https://math.stackexchange.com/questions/368155/where-does-the-relation-nabla21-r-4-pi-delta3-bf-r-between-laplacian

---

$$\begin{aligned}&\left\{\{\delta\left(k\right)|k\in\mathbb{R}\}\middle|\begin{cases}\delta:\mathbb{R}\to\mathbb{C}\\ \int_{-\infty}^{\infty}\delta\left(k-0\right)f\left(k,x\right)\mathrm{d}k=f\left(0,x\right)\end{cases}\right\}\\ &=\left\{\left\{K\int_{-\infty}^{\infty}\mathrm{e}^{\mathrm{i}kx}\left(\cdot\right)\mathrm{d}k\middle|k\in\mathbb{R}\right\},\cdots\right\}\end{aligned}$$

[34.1.3]

$$\begin{aligned}&\left\{\{\delta\left(\boldsymbol{k}\right)|\boldsymbol{k}\in\mathbb{R}^3\}\middle|\begin{cases}\delta:\mathbb{R}^3\to\mathbb{H}\\ \int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\delta\left(\boldsymbol{k}-\boldsymbol{0}\right)f\left(\boldsymbol{k},\boldsymbol{r}\right)\mathrm{d}^3\boldsymbol{r}=f\left(\boldsymbol{0},\boldsymbol{r}\right)\end{cases}\right\}\\ &=\left\{\left\{K\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\mathrm{e}^{\mathrm{i}\boldsymbol{k}\cdot\boldsymbol{r}}\left(\cdot\right)\mathrm{d}^3\boldsymbol{k}\middle|\boldsymbol{k}\in\mathbb{R}^3\right\},\cdots\right\}\end{aligned}$$

Columb law

$$F_{\mathrm{e}}=k\frac{Qq}{r^2}$$

$$\boldsymbol{F}_{\mathrm{e}}=k\frac{Qq}{r^2}\hat{\boldsymbol{r}}$$

electric field from point electric charge

$$\begin{aligned}\boldsymbol{E}\left(\boldsymbol{r}\right)=\boldsymbol{E}\left(r\right)=\boldsymbol{E}=&k\frac{q}{r^2}\hat{\boldsymbol{r}}=k_{\mathrm{e}}\frac{q}{|\boldsymbol{r}|^2}\frac{\boldsymbol{r}}{|\boldsymbol{r}|}=k_{\mathrm{e}}q\frac{\boldsymbol{r}}{|\boldsymbol{r}|^3}=k_{\mathrm{e}}q\frac{\boldsymbol{r}}{\left(\boldsymbol{r}^2\right)^{\frac{3}{2}}}\\ =&k_{\mathrm{e}}q\frac{\boldsymbol{r}}{\left(\boldsymbol{r}^2\right)^{\frac{3}{2}}}=k_{\mathrm{e}}q\left(-\boldsymbol{\nabla}\frac{1}{\sqrt{\boldsymbol{r}\cdot\boldsymbol{r}}}\right)=-k_{\mathrm{e}}q\boldsymbol{\nabla}\frac{1}{\|\boldsymbol{r}\|}\end{aligned}$$

$$\boldsymbol{E}\left(\boldsymbol{r}\right)=-k_{\mathrm{e}}q\boldsymbol{\nabla}\frac{1}{\|\boldsymbol{r}\|}$$

electric field from tiny point electric charge

$$\begin{aligned}\mathrm{d}\boldsymbol{E}\left(\boldsymbol{r}\right)=\mathrm{d}\boldsymbol{E}\left(r\right)=\mathrm{d}\boldsymbol{E}=&k\frac{\mathrm{d}q}{r^2}\hat{\boldsymbol{r}}=k\frac{\mathrm{d}q\left(r\right)}{r^2}\hat{\boldsymbol{r}}=k_{\mathrm{e}}\mathrm{d}q\left(\boldsymbol{r}\right)\frac{\boldsymbol{r}}{\left(\boldsymbol{r}^2\right)^{\frac{3}{2}}}\\ =&k_{\mathrm{e}}\mathrm{d}q\left(\boldsymbol{r}\right)\frac{\boldsymbol{r}}{\left(\boldsymbol{r}^2\right)^{\frac{3}{2}}}=k_{\mathrm{e}}\mathrm{d}q\left(\boldsymbol{r}\right)\boldsymbol{\nabla}\frac{1}{\|\boldsymbol{r}\|}\end{aligned}$$

$$\mathrm{d}\boldsymbol{E}\left(\boldsymbol{r}\right)=-k_{\mathrm{e}}\mathrm{d}q\left(\boldsymbol{r}\right)\boldsymbol{\nabla}\frac{1}{\|\boldsymbol{r}\|}$$

$$\mathrm{d}q=\rho\mathrm{d}V\Leftrightarrow\rho=\frac{\mathrm{d}q}{\mathrm{d}V}$$

$$\mathrm{d}q\left(\boldsymbol{r}\right)=\rho\left(\boldsymbol{r}\right)\mathrm{d}V\Leftrightarrow\rho\left(\boldsymbol{r}\right)=\frac{\mathrm{d}q\left(\boldsymbol{r}\right)}{\mathrm{d}V}$$

$$q=\int_q \mathrm{d}q\left(\boldsymbol{r}\right)=\int_V \rho\left(\boldsymbol{r}\right)\mathrm{d}V=\iiint_V \rho\left(\boldsymbol{r}\right)\mathrm{d}^3\boldsymbol{r}$$

$$\begin{aligned}\boldsymbol{E}=&\iiint_V \boldsymbol{E}\left(\boldsymbol{r}\right)\mathrm{d}^3\boldsymbol{r}=\iiint_V\left(-k_\mathrm{e}\rho\left(\boldsymbol{r}\right)\boldsymbol{\nabla}\frac{1}{\|\boldsymbol{r}\|}\right)\mathrm{d}^3\boldsymbol{r}\\=&-k_\mathrm{e}\iiint_V\left(\rho\left(\boldsymbol{r}\right)\boldsymbol{\nabla}\frac{1}{\|\boldsymbol{r}\|}\right)\mathrm{d}^3\boldsymbol{r}\\=&-k_\mathrm{e}\iiint_V\left(\rho\left(\boldsymbol{r}\right)\boldsymbol{\nabla}\frac{1}{\|\boldsymbol{r}-\boldsymbol{0}\|}\right)\mathrm{d}^3\boldsymbol{r}=\boldsymbol{E}\left(\boldsymbol{0}\right)\end{aligned}$$

$$\begin{aligned}&\Downarrow\begin{cases}Q\text{ at }\boldsymbol{r}\\\mathrm{d}q\text{ or }\rho\text{ at }\boldsymbol{r}',\text{ totally }V\end{cases}\\\boldsymbol{E}\left(\boldsymbol{r}\right)=&\iiint_V\left(-k_\mathrm{e}\rho\left(\boldsymbol{r}'\right)\boldsymbol{\nabla}\frac{1}{\|\boldsymbol{r}-\boldsymbol{r}'\|}\right)\mathrm{d}^3\boldsymbol{r}'=\iiint_V\left(-k_\mathrm{e}\rho\left(\boldsymbol{r}'\right)\boldsymbol{\nabla}_{\boldsymbol{r}}\frac{1}{\|\boldsymbol{r}-\boldsymbol{r}'\|}\right)\mathrm{d}^3\boldsymbol{r}'\\\boldsymbol{\nabla}_{\boldsymbol{r}}\cdot\boldsymbol{E}\left(\boldsymbol{r}\right)=&\boldsymbol{\nabla}_{\boldsymbol{r}}\cdot\iiint_V\left(-k_\mathrm{e}\rho\left(\boldsymbol{r}'\right)\boldsymbol{\nabla}_{\boldsymbol{r}}\frac{1}{\|\boldsymbol{r}-\boldsymbol{r}'\|}\right)\mathrm{d}^3\boldsymbol{r}'=\iiint_V\left(-k_\mathrm{e}\rho\left(\boldsymbol{r}'\right)\boldsymbol{\nabla}_{\boldsymbol{r}}\cdot\boldsymbol{\nabla}_{\boldsymbol{r}}\frac{1}{\|\boldsymbol{r}-\boldsymbol{r}'\|}\right)\mathrm{d}^3\boldsymbol{r}'\\=&\iiint_V\left(-k_\mathrm{e}\rho\left(\boldsymbol{r}'\right)\boldsymbol{\nabla}_{\boldsymbol{r}}^2\frac{1}{\|\boldsymbol{r}-\boldsymbol{r}'\|}\right)\mathrm{d}^3\boldsymbol{r}'=\iiint_V\left(-k_\mathrm{e}\rho\left(\boldsymbol{r}'\right)\Delta_{\boldsymbol{r}}\frac{1}{\|\boldsymbol{r}-\boldsymbol{r}'\|}\right)\mathrm{d}^3\boldsymbol{r}'\\=&\iiint_V f\left(\boldsymbol{r}'\right)\Delta\left(\boldsymbol{r}-\boldsymbol{r}'\right)\mathrm{d}^3\boldsymbol{r}',f\left(\boldsymbol{r}'\right)=-k_\mathrm{e}\rho\left(\boldsymbol{r}'\right)\\=&\iiint_V f\left(\boldsymbol{r}'\right)\delta\left(\boldsymbol{r}-\boldsymbol{r}'\right)\mathrm{d}^3\boldsymbol{r}',f\left(\boldsymbol{r}'\right)=-k_\mathrm{e}\rho\left(\boldsymbol{r}'\right)\\=&-4\pi f\left(\boldsymbol{r}'=\boldsymbol{r}\right)=-4\pi\left[-k_\mathrm{e}\rho\left(\boldsymbol{r}'=\boldsymbol{r}\right)\right]\\=&-4\pi f\left(\boldsymbol{r}\right)=-4\pi\left[-k_\mathrm{e}\rho\left(\boldsymbol{r}\right)\right]=4\pi k_\mathrm{e}\rho\left(\boldsymbol{r}\right)\\\boldsymbol{\nabla}\cdot\boldsymbol{E}\left(\boldsymbol{r}\right)=&4\pi k_\mathrm{e}\rho\left(\boldsymbol{r}\right)\end{aligned}$$

$$\boldsymbol{\nabla}\cdot\boldsymbol{E}\left(\boldsymbol{r}\right)=4\pi k_\mathrm{e}\rho\left(\boldsymbol{r}\right)$$

---

https://www.bilibili.com/video/BV1qu411578c/?t=7m26s

https://en.wikipedia.org/wiki/Helmholtz_equation

Helmholtz equation

$$\left(\Delta+k^2\right)\frac{\mathrm{e}^{\pm\mathrm{i}\boldsymbol{k}\cdot\boldsymbol{r}}}{r}=-4\pi\delta\left(\boldsymbol{r}\right)$$

$$\left(\boldsymbol{\nabla}_{\boldsymbol{r}}^2+\boldsymbol{k}^2\right)\frac{\mathrm{e}^{\pm\mathrm{i}\boldsymbol{k}\cdot\boldsymbol{r}}}{\|\boldsymbol{r}\|}=-4\pi\delta\left(\boldsymbol{r}\right)$$

45.2.3 quantum state

https://www.bilibili.com/video/BV1qu411578c/?t=7m45s

45.2.4 rigorous definition of delta function

$$\int_{-\infty}^{\infty}\delta\left(x-0\right)\cdot f\left(x\right)\mathrm{d}x=f\left(0\right)$$

$$\int\delta\left(x\right)f\left(x\right)\mathrm{d}x=f\left(0\right)$$

$$\int\delta\left(x\right)\left(\cdot\right)\mathrm{d}x:f\left(x\right)\to f\left(0\right)$$

$$f\left(x\right)\overset{\int\delta(x)(\cdot)\mathrm{d}x}{\to}f\left(0\right)$$

$$f \text{ must be with good properties}$$

$$\text{continuous linear functional} \begin{cases} \text{function-like} & \eta_{\varphi}(f) = \int \varphi(x) f(x) \, \mathrm{d}x \\ \text{non-function-like, e.g.} & \delta_{x'}(f) = f(x') \end{cases}$$

general function with "formally" integral notation

$$\delta_{x'}(f) = \int \delta(x - x') f(x) \, \mathrm{d}x = f(x')$$

https://www.bilibili.com/video/BV1qu411578c/?t=11m36s

Quantum mechanics is founded on Hilbert space; Hoever, Dirac delta function is not a linear functional on Hilbert space.?

One of the reason is that we cannot find 1-1 of Dirac delta function, but many-to-one or one-to-many functions.

量子力學本不需 Dirac delta function, 其 卻仍大行其道, 實 廣義函數 為其負重前行.

45.3 Fourier expansion or Fourier series

vector space definition[40.2]

Fig: 40.1

linear space of function[43.1]

https://www.bilibili.com/video/BV1PX4y167RS

https://www.bilibili.com/video/BV1m24y1A74K/?=3m2s

45.3.1 Taylor vs. Fourier

$$f(x) = x_0 x^0 + x_1 x^1 + x_2 x^2 + \cdots = \sum_{k=0}^{\infty} x_k x^k, x_k = \frac{f^{(k)}(0)}{k!}$$

$$f(r) = r_0 r^0 + r_1 r^1 + r_2 r^2 + \cdots = \sum_{k=0}^{\infty} r_k r^k, r_k = \frac{f^{(k)}(0)}{k!}$$

$$f(\theta) = a_0 \cos(0\theta) + a_1 \cos(1\theta) + a_2 \cos(2\theta) + \cdots = \sum_{k=0}^{\infty} a_k \cos(k\theta)$$

$$a_k = \int_{-\infty}^{\infty} \cos(k\theta) f(\theta) \, \mathrm{d}\theta$$

$$\begin{aligned} f(\theta) &= a_0 \cos(0\theta) + a_1 \cos(1\theta) + a_2 \cos(2\theta) + \cdots = \sum_{k=0}^{\infty} a_k \cos(k\theta) \\ &= a_0 \Re\left[\mathrm{e}^{\mathrm{i}0\theta}\right] + a_1 \Re\left[\mathrm{e}^{\mathrm{i}1\theta}\right] + a_2 \Re\left[\mathrm{e}^{\mathrm{i}2\theta}\right] + \cdots = \sum_{k=0}^{\infty} a_k \Re\left[\mathrm{e}^{\mathrm{i}k\theta}\right] = \sum_{k=0}^{\infty} a_k \Re\left[\left(\mathrm{e}^{\mathrm{i}\theta}\right)^k\right] \\ &= \Re\left[a_0 \mathrm{e}^{\mathrm{i}0\theta} + a_1 \mathrm{e}^{\mathrm{i}1\theta} + a_2 \mathrm{e}^{\mathrm{i}2\theta} + \cdots\right] = \Re\left[\sum_{k=0}^{\infty} a_k \mathrm{e}^{\mathrm{i}k\theta}\right] = \Re\left[\sum_{k=0}^{\infty} a_k \left(\mathrm{e}^{\mathrm{i}\theta}\right)^k\right] \end{aligned}$$

$$f(x) = x_0 x^0 + x_1 x^1 + x_2 x^2 + \cdots = \sum_{k=0}^{\infty} x_k x^k$$

$$f(z) = z_0 z^0 + z_1 z^1 + z_2 z^2 + \cdots = \sum_{k=0}^{\infty} z_k z^k$$

$$z = r\mathrm{e}^{\mathrm{i}\theta}$$

$$
\begin{aligned}
f(z) =& z_0 z^0 + z_1 z^1 + z_2 z^2 + \cdots = \sum_{k=0}^{\infty} z_k z^k \\
f\left(re^{i\theta}\right) =& z_0 \left(re^{i\theta}\right)^0 + z_1 \left(re^{i\theta}\right)^1 + z_2 \left(re^{i\theta}\right)^2 + \cdots = \sum_{k=0}^{\infty} z_k \left(re^{i\theta}\right)^k \\
= f(r) =& z_0 \left(e^{i\theta}\right)^0 r^0 + z_1 \left(e^{i\theta}\right)^1 r^1 + z_2 \left(e^{i\theta}\right)^2 r^2 + \cdots = \sum_{k=0}^{\infty} z_k \left(e^{i\theta}\right)^k r^k \\
=& r_0 r^0 + r_1 r^1 + r_2 r^2 + \cdots = \sum_{k=0}^{\infty} r_k r^k, r_k = z_k \left(e^{i\theta}\right)^k \\
= f\left(e^{i\theta}\right) =& z_0 r^0 \left(e^{i\theta}\right)^0 + z_1 r^1 \left(e^{i\theta}\right)^1 + z_2 r^2 \left(e^{i\theta}\right)^2 + \cdots = \sum_{k=0}^{\infty} z_k r^k \left(e^{i\theta}\right)^k \\
=& \left(e^{i\theta}\right)_0 \left(e^{i\theta}\right)^0 + \left(e^{i\theta}\right)_1 \left(e^{i\theta}\right)^1 + \left(e^{i\theta}\right)_2 \left(e^{i\theta}\right)^2 + \cdots = \sum_{k=0}^{\infty} \left(e^{i\theta}\right)_k \left(e^{i\theta}\right)^k \\
& , \left(e^{i\theta}\right)_k = z_k r^k
\end{aligned}
$$

45.4 Hilbert space

Definition 45.1. sequence limit

$$\lim_{n\to\infty} a_n = a \Leftrightarrow \forall \epsilon > 0, \exists N \in \mathbb{N}, \forall n \in \mathbb{N} \left(n > N \Rightarrow |a_n - a| < \epsilon\right)$$

$$x_1, x_2, \cdots, x_n, x_{n+1}, \cdots$$

- to find convergence
  - monotone convergence theorem
  - split into known sequence
  - fraction: Stolz theorem
  - sequence itself
    * ratio $\frac{x_m}{x_n}$, but not good for some terms being zero
    * difference $x_m - x_n$

Theorem 45.4. *Cauchy convergence theorem*

$$
\begin{gathered}
\{x_n\}_{n\in\mathbb{N}} \text{ is convergent} \\
\Updownarrow \\
\forall \epsilon > 0, \exists N \in \mathbb{N}, \forall m, n \in \mathbb{N} \left(m, n > N \Rightarrow |x_m - x_n| < \epsilon\right) \Leftrightarrow \{x_n\}_{n\in\mathbb{N}} \text{ is a Cauchy sequence}
\end{gathered}
$$

Definition 45.2. Cauchy sequence

$$
\begin{gathered}
\{x_n\}_{n\in\mathbb{N}} \text{ is a Cauchy sequence} \\
\Updownarrow \\
\forall \epsilon > 0, \exists N \in \mathbb{N}, \forall m, n \in \mathbb{N} \left(m, n > N \Rightarrow |x_m - x_n| < \epsilon\right)
\end{gathered}
$$

$$\text{convergence} \begin{cases} |x_n - x| < \epsilon & \text{convergent sequence} \\ |x_m - x_n| < \epsilon & \text{Cauchy sequence} \end{cases}$$

$$\text{convergence} \begin{cases} |x_n - x| < \epsilon & \text{convergent sequence} \\ \Downarrow \text{easy proof} & \Uparrow \text{hard proof, weak condition needs all real number as supporting base} \\ |x_m - x_n| < \epsilon & \text{Cauchy sequence} \end{cases}$$

$$\text{convergence} \begin{cases} |x_n - x| < \epsilon & \text{convergent sequence} \\ \Downarrow & \Uparrow x_n, x_m, x \in \mathbb{R} \text{ with real completeness} \\ |x_m - x_n| < \epsilon & \text{Cauchy sequence} \end{cases}$$

45.4.1 inner product space

vector space definition[40.2]

Fig: 40.1

group

1. vector addition clousre
2. identity **0**
3. inverse
4. associativity

→

abelian group

4 + 1. commutativity

→

linear space

1. scalar multiplicaiton clousre
2. indentity 1
3. associativity
4. scalar distributivity
5. vector distributivity

→

inner product space

inner product $\langle \cdot | \cdot \rangle$

1. $\langle f | g+h \rangle = \langle f|g \rangle + \langle f|h \rangle$
2. $\langle f | \lambda g \rangle = \lambda \langle f|g \rangle$
3. $\langle f|g \rangle = \overline{\langle g|f \rangle} = \langle g|f \rangle^*$
4. $\langle f|f \rangle \underset{f=0 \Leftrightarrow \langle f|f \rangle = 0}{\geq} 0$

Figure 45.6: inner product space construction

$$\langle \cdot | \cdot \rangle : \begin{cases} \langle f|g+h \rangle = \langle f|g \rangle + \langle f|h \rangle \\ \langle f|cg \rangle = c \langle f|g \rangle \\ \langle f|g \rangle = \overline{\langle g|f \rangle} = \langle g|f \rangle^* \\ \langle f|f \rangle \underset{f=0 \Leftrightarrow \langle f|f \rangle = 0}{\geq} 0 \end{cases}$$

definition of distance

Definition 45.3. distance

$$d(f,g) = d\langle f|g \rangle = \sqrt{\langle f-g|f-g \rangle}$$

For real number,

$$\begin{cases} d(f,g) = d\langle f|g \rangle = \sqrt{\langle f-g|f-g \rangle} \\ f,g \in \mathbb{R} \\ \langle f|g \rangle = fg = f \cdot g \end{cases}$$
$$\Downarrow$$
$$d\langle f|g \rangle = \sqrt{(f-g)^2} = |f-g|$$

For n-dimensional vector

$$\begin{cases} d(f,g) = d\langle f|g \rangle = \sqrt{\langle f-g|f-g \rangle} \\ f,g \in F^n \\ \langle f|g \rangle = \sum\limits_{k=1}^{n} f_k g_k \end{cases}$$
$$\Downarrow$$
$$d\langle f|g \rangle = \sqrt{\sum_{k=1}^{n} (f_k - g_k)^2}$$

For continuous complex function on $[a,b]$,

$$\begin{cases} d(f,g) = d\langle f|g \rangle = \sqrt{\langle f-g|f-g \rangle} \\ f,g \in C_{\mathbb{C}}([a,b]) & \text{continuous complex function} \\ \langle f|g \rangle = \int_a^b \overline{f(x)} g(x)\,\mathrm{d}x \end{cases}$$
$$\Downarrow$$
$$\underline{d\langle f|g \rangle = \sqrt{\int_a^b \overline{[f-g](x)}\,[f-g](x)\,\mathrm{d}x} = \sqrt{\int_a^b |[f-g](x)|^2\,\mathrm{d}x}}$$

For real number sequence,

$$\begin{cases} d\left(f,g\right)=d\left\langle f|g\right\rangle =\sqrt{\left\langle f-g|f-g\right\rangle} \\ f,g\in\mathbb{R} \\ \left\langle f|g\right\rangle = fg = f\cdot g \end{cases}$$
$$\Downarrow$$
$$d\left\langle f|g\right\rangle = \sqrt{\left(f-g\right)^2} = \left|f-g\right|$$

$$\text{convergence}\begin{cases} \left|x_n-x\right|<\epsilon & \text{convergent sequence} \\ \left|x_m-x_n\right|<\epsilon & \text{Cauchy sequence} \end{cases}$$

$$\text{convergence}\begin{cases} \left|x_n-x\right|<\epsilon & \text{convergent sequence} \\ \Downarrow & \Uparrow x_n, x_m, x\in\mathbb{R} \text{ with real completeness} \\ \left|x_m-x_n\right|<\epsilon & \text{Cauchy sequence} \end{cases}$$

For complex function sequence,

$$\begin{cases} d\left(f,g\right)=d\left\langle f|g\right\rangle =\sqrt{\left\langle f-g|f-g\right\rangle} & \\ f,g\in C_{\mathbb{C}}\left(\left[a,b\right]\right) & \text{continuous complex function} \\ \left\langle f|g\right\rangle = \int_a^b \overline{f\left(x\right)} g\left(x\right) \mathrm{d}x & \end{cases}$$
$$\Downarrow$$
$$d\left\langle f|g\right\rangle = \sqrt{\int_a^b \overline{\left[f-g\right]\left(x\right)}\left[f-g\right]\left(x\right)\mathrm{d}x} = \sqrt{\int_a^b \left|\left[f-g\right]\left(x\right)\right|^2 \mathrm{d}x}$$

$$\text{function convergence}\begin{cases} d\left\langle f_n|f\right\rangle<\epsilon & \text{convergent function sequence} \\ d\left\langle f_m|f_n\right\rangle<\epsilon & \text{Cauchy function sequence} \end{cases}$$

$$\text{function convergence}\begin{cases} d\left\langle f_n|f\right\rangle<\epsilon & \text{convergent function sequence} \\ \Downarrow & \not\Uparrow \text{ not always continuous} \\ d\left\langle f_m|f_n\right\rangle<\epsilon & \text{Cauchy function sequence} \end{cases}$$

group

1. vector addition clousre
2. identity **0**
3. inverse
4. associativity

→ abelian group

4 + 1. commutativity

→ linear space

1. scalar multiplicaiton clousre
2. indentity 1
3. associativity
4. scalar distributivity
5. vector distributivity

→ inner product space

inner product $\langle\cdot|\cdot\rangle$
1. $\langle f|g+h\rangle = \langle f|g\rangle + \langle f|h\rangle$
2. $\langle f|\lambda g\rangle = \lambda\langle f|g\rangle$
3. $\langle f|g\rangle = \overline{\langle g|f\rangle} = \langle g|f\rangle^*$
4. $\langle f|f\rangle \underset{f=0\Leftrightarrow\langle f|f\rangle=0}{\geq} 0$

↓

completeness:
complete if
all Cauchy function sequence convergent

Hilbert space

Figure 45.7: Hilbert space construction

45.5 Dirac bracket

reciprocal space[40.4.1]

double dual concept[40.4.1.2]

$$\begin{array}{ccccccccccc} & & & V & =\{ & \boldsymbol{e}_1 & \boldsymbol{e}_2 & \boldsymbol{e}_3 & \boldsymbol{v} & \cdots\} \\ & \boldsymbol{e}^1 & : & \downarrow & & \downarrow & \downarrow & \downarrow & \downarrow & \\ V^*=\{ & \boldsymbol{e}^2 & \} & F & \supseteq\{ & 1 & 0 & 0 & v_1 & \cdots\} \\ & \boldsymbol{e}^3 & & V & =\{ & \boldsymbol{e}_1 & \boldsymbol{e}_2 & \boldsymbol{e}_3 & \boldsymbol{v} & \cdots\} \\ & \boldsymbol{v}^* & : & \downarrow & & \downarrow & \downarrow & \downarrow & \downarrow & \\ & \vdots & & F & \supseteq\{ & v^{*1} & v^{*2} & v^{*3} & v^{*i}v_i & \cdots\} \end{array}$$

dual space of span of partials[41.2]

coefficient of linear combination for vector space and dual space[41.4]

$$\begin{array}{ccccccccccc} & & & V & = \{ & \boldsymbol{e}_1 & \boldsymbol{e}_2 & \boldsymbol{e}_3 & \boldsymbol{v} & \cdots \} \\ & \boldsymbol{e}^1 & : & \downarrow & & \downarrow & \downarrow & \downarrow & \downarrow & \\ V^* = \{ & \boldsymbol{e}^2 & \} & F & \supseteq \{ & 1 & 0 & 0 & v_1 & \cdots \} \\ & \boldsymbol{e}^3 & & V & = \{ & \boldsymbol{e}_1 & \boldsymbol{e}_2 & \boldsymbol{e}_3 & \boldsymbol{v} & \cdots \} \\ & \boldsymbol{v}^* & : & \downarrow & & \downarrow & \downarrow & \downarrow & \downarrow & \\ & \vdots & & F & \supseteq \{ & v^{*1} & v^{*2} & v^{*3} & v^{*i} v_i & \cdots \} \end{array}$$

inner product space[45.4.1]

$$\begin{cases} d\left(f,g\right) = d\left\langle f|g\right\rangle = \sqrt{\left\langle f-g|f-g\right\rangle} \\ f,g \in C_{\mathbb{C}}\left(\left[a,b\right]\right) & \text{continuous complex function} \\ \left(f,g\right) = \left\langle f|g\right\rangle = \int_a^b \overline{f\left(x\right)} g\left(x\right) \mathrm{d}x = \int_a^b f^*\left(x\right) g\left(x\right) \mathrm{d}x \end{cases}$$

$$\Downarrow$$

$$d\left\langle f|g\right\rangle = \sqrt{\int_a^b \overline{\left[f-g\right]\left(x\right)} \left[f-g\right]\left(x\right) \mathrm{d}x} = \sqrt{\int_a^b \left|\left[f-g\right]\left(x\right)\right|^2 \mathrm{d}x}$$

$$v^*\left(f\right) = \eta_f\left(\cdot\right) = \left(f,\cdot\right) = \left\langle f|\cdot\right\rangle = \int_a^b \overline{f\left(x\right)} \left(\cdot\right) \mathrm{d}x = \int_a^b f^*\left(x\right)\left(\cdot\right) \mathrm{d}x$$

$$\eta_f\left(g\right) = \left(f,g\right) = \left\langle f|g\right\rangle = \int_a^b \overline{f\left(x\right)} g\left(x\right) \mathrm{d}x = \int_a^b f^*\left(x\right) g\left(x\right) \mathrm{d}x$$

https://en.wikipedia.org/wiki/Antilinear_map

$$\begin{cases} v^*\left(f+g\right) = v^*\left(f\right) + v^*\left(g\right) & \text{addivity} = \text{superposition} \\ v^*\left(\lambda f\right) = \lambda^* v^*\left(f\right) = \overline{\lambda} v^*\left(f\right) & \text{conjugate homogeneity} = \text{complex conjugate} \end{cases}$$

$$\begin{aligned} & \begin{cases} v^*\left(f+g\right) = v^*\left(f\right) + v^*\left(g\right) & \text{addivity} = \text{superposition} \\ v^*\left(\lambda f\right) = \lambda^* v^*\left(f\right) = \overline{\lambda} v^*\left(f\right) & \text{conjugate homogeneity} = \text{complex conjugate} \end{cases} \\ \Leftrightarrow & \left\{ v^*\left(f+g\right) = \lambda^* v^*\left(f\right) + v^*\left(g\right) = \overline{\lambda} v^*\left(f\right) + v^*\left(g\right) \quad \text{antilinearity} \right. \end{aligned}$$

$$\begin{array}{ccccccccccc} & & & V & = \{ & \hat{f}_1 & \hat{f}_2 & \hat{f}_3 & f & \cdots \} \\ & \hat{f}^1 = \left(\hat{f}_1,\cdot\right) & : & \downarrow & & \downarrow & \downarrow & \downarrow & \downarrow & \\ V^* = \{ & \hat{f}^2 = \left(\hat{f}_2,\cdot\right) & \} & F & \supseteq \{ & \left(\hat{f}_1,\hat{f}_1\right) = 1 & \left(\hat{f}_1,\hat{f}_2\right) = 0 & \left(\hat{f}_1,\hat{f}_3\right) = 0 & \left(\hat{f}_1,f\right) & \cdots \} \\ & \hat{f}^3 = \left(\hat{f}_3,\cdot\right) & & V & = \{ & \hat{f}_1 & \hat{f}_2 & \hat{f}_3 & f & \cdots \} \\ & f^* = \left(f,\cdot\right) & : & \downarrow & & \downarrow & \downarrow & \downarrow & \downarrow & \\ & \vdots & & F & \supseteq \{ & \left(f,\hat{f}_1\right) & \left(f,\hat{f}_2\right) & \left(f,\hat{f}_3\right) & \left(f,f\right) & \cdots \} \end{array}$$

$$\begin{array}{ccccccccccc} & & & V & = \{ & \hat{f}_1 & \hat{f}_2 & \hat{f}_3 & f & \cdots \} \\ & \hat{f}^1 = \left\langle \hat{f}_1 \middle| \cdot \right\rangle & : & \downarrow & & \downarrow & \downarrow & \downarrow & \downarrow & \\ V^* = \{ & \hat{f}^2 = \left\langle \hat{f}_2 \middle| \cdot \right\rangle & \} & F & \supseteq \{ & \left\langle \hat{f}_1 \middle| \hat{f}_1 \right\rangle = 1 & \left\langle \hat{f}_1 \middle| \hat{f}_2 \right\rangle = 0 & \left\langle \hat{f}_1 \middle| \hat{f}_3 \right\rangle = 0 & \left\langle \hat{f}_1 \middle| f \right\rangle & \cdots \} \\ & \hat{f}^3 = \left\langle \hat{f}_3 \middle| \cdot \right\rangle & & V & = \{ & \hat{f}_1 & \hat{f}_2 & \hat{f}_3 & f & \cdots \} \\ & f^* = \left\langle f|\cdot\right\rangle & : & \downarrow & & \downarrow & \downarrow & \downarrow & \downarrow & \\ & \vdots & & F & \supseteq \{ & \left\langle f \middle| \hat{f}_1 \right\rangle & \left\langle f \middle| \hat{f}_2 \right\rangle & \left\langle f \middle| \hat{f}_3 \right\rangle & \left\langle f|f\right\rangle & \cdots \} \end{array}$$

row vector is dual vector of column vector

$$f = \begin{pmatrix} a \\ b \end{pmatrix} \to v^*\left(f\right) = \eta_f = \begin{pmatrix} a^* & b^* \end{pmatrix} = \begin{pmatrix} \overline{a} & \overline{b} \end{pmatrix} = \overline{\begin{pmatrix} a \\ b \end{pmatrix}}^{\mathsf{T}} = \overline{\begin{pmatrix} a \\ b \end{pmatrix}^{\mathsf{T}}} = \begin{pmatrix} a \\ b \end{pmatrix}^{\dagger}$$

$$\eta_f \cdot f = \begin{pmatrix} a^* & b^* \end{pmatrix} \begin{pmatrix} a \\ b \end{pmatrix} = a^* a + b^* b$$

https://www.bilibili.com/video/BV1GN411m7A9/?t=2m40s

https://en.wikipedia.org/wiki/Riesz_representation_theorem

Theorem 45.5. *Riesz representation theorem*

to be proved

https://www.bilibili.com/video/BV1GN411m7A9/?t=6m20s

$$\text{Hilbert space } V \begin{cases} B = \left\{\hat{f}_i\right\}_{i\in I} = \left\{\cdots, \hat{f}_i, \cdots\right\} \\ \begin{cases} \sum_i c_i \hat{f}_i = 0 \Rightarrow c_i = 0 & \text{bases are linear independent} \\ f = \sum_i c_i \hat{f}_i, c_i \in \mathbb{C} & \text{expansion over bases} \end{cases} \\ \begin{cases} \left(\hat{f}_i, \hat{f}_j\right) = \left\langle \hat{f}_i \middle| \hat{f}_j \right\rangle = \delta_{ij} \\ \text{completeness} : \forall \hat{f}_i \in B, \nexists f \in V - B \left[\left(\hat{f}_i, f\right) = \left\langle \hat{f}_i \middle| f \right\rangle = 0\right] \end{cases} \end{cases}$$

$$f = \sum_i \left(\hat{f}_i \cdot f\right) \hat{f}_i = \sum_i \left(\hat{f}_i, \hat{f}_j\right) \hat{f}_i = \sum_i \left\langle \hat{f}_i \middle| \hat{f}_j \right\rangle \hat{f}_i$$

https://www.bilibili.com/video/BV1GN411m7A9/?t=7m55s

45.5.1 Dirac bracket symmetry

$$f = \sum_i \left(\hat{f}_i, f\right) \hat{f}_i = \sum_i \hat{f}_i \left(\hat{f}_i, f\right)$$

$$\begin{aligned} f^* = (f, \cdot) &= \left(\sum_i \left(\hat{f}_i, f\right) \hat{f}_i, \cdot\right) = \left(\sum_i \hat{f}_i \left(\hat{f}_i, f\right), \cdot\right) \\ &\overset{\text{antilinearity}}{=} \sum_i \overline{\left(\hat{f}_i, f\right)} \left(\hat{f}_i, \cdot\right) = \sum_i \overline{\left(\hat{f}_i, f\right)} \left(\hat{f}_i, \cdot\right) \\ &= \sum_i \left(f, \hat{f}_i\right) \left(\hat{f}_i, \cdot\right) \\ &\overset{\text{Dirac bracket}}{\rightarrow} \sum_i \left\langle f \middle| \hat{f}_i \right\rangle \left\langle \hat{f}_i \right| = \langle f| \\ &\Updownarrow \\ f &= \sum_i \hat{f}_i \left(\hat{f}_i, f\right) \\ &\overset{\text{Dirac bracket}}{\rightarrow} \sum_i \left|\hat{f}_i\right\rangle \left\langle \hat{f}_i \middle| f \right\rangle = |f\rangle \end{aligned}$$

$$\begin{cases} |f\rangle = \sum_i \left|\hat{f}_i\right\rangle \left\langle \hat{f}_i \middle| f \right\rangle = \sum_i \left\langle \hat{f}_i \middle| f \right\rangle \left|\hat{f}_i\right\rangle \\ \langle f| = \sum_i \left\langle f \middle| \hat{f}_i \right\rangle \left\langle \hat{f}_i \right| \\ \left\langle \hat{f}_i \middle| \hat{f}_j \right\rangle = \delta_{ij} \end{cases}$$

$$
\begin{array}{cccccccccc}
 & & V & = \{ & \left|\hat{f}_1\right\rangle & \left|\hat{f}_2\right\rangle & \left|\hat{f}_3\right\rangle & |f\rangle & \cdots\} \\
 & \left\langle\hat{f}_1\right| & : \downarrow & & \downarrow & \downarrow & \downarrow & \downarrow & \\
V^* = \{ & \left\langle\hat{f}_2\right| & \} \; F & \supseteq \{ & \left\langle\hat{f}_1\middle|\hat{f}_1\right\rangle = 1 & \left\langle\hat{f}_1\middle|\hat{f}_2\right\rangle = 0 & \left\langle\hat{f}_1\middle|\hat{f}_3\right\rangle = 0 & \left\langle\hat{f}_1\middle|f\right\rangle & \cdots\} \\
 & \left\langle\hat{f}_3\right| & V & = \{ & \left|\hat{f}_1\right\rangle & \left|\hat{f}_2\right\rangle & \left|\hat{f}_3\right\rangle & |f\rangle & \cdots\} \\
 & \langle f| & : \downarrow & & \downarrow & \downarrow & \downarrow & \downarrow & \\
 & \vdots & F & \supseteq \{ & \left\langle f\middle|\hat{f}_1\right\rangle & \left\langle f\middle|\hat{f}_2\right\rangle & \left\langle f\middle|\hat{f}_3\right\rangle & \langle f|f\rangle & \cdots\}
\end{array}
$$

$$
\begin{cases}
v^*(f+g) = v^*(f) + v^*(g) & \text{addivity = superposition} \\
v^*(\lambda f) = \lambda^* v^*(f) = \overline{\lambda} v^*(f) & \text{conjugate homogeneity = complex conjugate}
\end{cases}
$$

$$
\Leftrightarrow \begin{cases} v^*(f+g) = \lambda^* v^*(f) + v^*(g) = \overline{\lambda} v^*(f) + v^*(g) & \text{antilinearity} \end{cases}
$$

$$
\begin{cases}
(c\psi, \cdot) \to \langle c\psi| = c^* \langle\psi| & c\psi \to |c\psi\rangle = c|\psi\rangle \\
\langle\psi| \leftrightarrow |\psi\rangle & \\
\langle\psi|\phi\rangle = \overline{\langle\phi|\psi\rangle} & \langle\phi|\psi\rangle = \overline{\langle\psi|\phi\rangle} \\
\langle c\psi|\phi\rangle = c^* \langle\psi|\phi\rangle & \langle\phi|c\psi\rangle = c\langle\phi|\psi\rangle
\end{cases}
$$

$|f\rangle\langle g| =?$

$$
\begin{gathered}
|f\rangle\langle g|\,|h\rangle = |f\rangle\langle g|h\rangle = \langle g|h\rangle\,|f\rangle \\
\Downarrow \\
|h\rangle \overset{|f\rangle\langle g|}{\to} |f\rangle\langle g|\,|h\rangle = \langle g|h\rangle\,|f\rangle
\end{gathered}
$$

$$
\begin{gathered}
\langle h|\,|f\rangle\langle g| = \langle h|f\rangle\langle g| \\
\Downarrow \\
\langle h| \overset{|f\rangle\langle g|}{\to} \langle h|\,|f\rangle\langle g| = \langle h|f\rangle\langle g|
\end{gathered}
$$

$$
|f\rangle\langle g| \text{ is a linear transform or rank-2 tensor}
$$

$$
\begin{aligned}
|f\rangle &= \sum_i \left|\hat{f}_i\right\rangle\left\langle\hat{f}_i\middle|f\right\rangle \\
&= \sum_i \left|\hat{f}_i\right\rangle\left\langle\hat{f}_i\right|\,|f\rangle \\
&= \left(\sum_i \left|\hat{f}_i\right\rangle\left\langle\hat{f}_i\right|\right)|f\rangle = 1\,|f\rangle = |f\rangle \\
1 &= \sum_i \left|\hat{f}_i\right\rangle\left\langle\hat{f}_i\right| \\
\sum_i \left|\hat{f}_i\right\rangle\left\langle\hat{f}_i\right| &= 1
\end{aligned}
$$

$$
\sum_i \left|\hat{f}_i\right\rangle\left\langle\hat{f}_i\right| = 1
$$

the above is completeness relationship.

$$
\begin{cases}
|f\rangle = 1\,|f\rangle = \left(\sum_i \left|\hat{f}_i\right\rangle\left\langle\hat{f}_i\right|\right)|f\rangle = \sum_i \left|\hat{f}_i\right\rangle\left\langle\hat{f}_i\middle|f\right\rangle \\
\langle f| = \langle f|\,1 = \langle f|\left(\sum_i \left|\hat{f}_i\right\rangle\left\langle\hat{f}_i\right|\right) = \sum_i \left\langle f\middle|\hat{f}_i\right\rangle\left\langle\hat{f}_i\right|
\end{cases}
$$

> The sculpture is already complete within the marble block, before I start my work. It is already there, I just have to chisel away the superfluous material. Michelangelo

https://www.goodreads.com/quotes/1191114-the-sculpture-is-already-complete-within-the-marble-block-before

45.6 quantum operator

45.7 Gram-Schmidt algorithm to find orthonormal basis

45.8 Hermite polynomials and Legendre polynomials

Chapter 46

Lagrange inversion

Lagrange inversion theorem

拉格朗日反演

Chapter 47

quantum color dynamics

Chapter 48

Lagrangian

- Elliot Schneider: Physics with Elliot
  - Lagrangian fundamentals
  - Lagrangian mechanics

https://courses.physicswithelliot.com/products/fundamentals-of-lagrangian-mechanics/categories/2150492677/posts/2158458875

https://courses.physicswithelliot.com/products/fundamentals-of-lagrangian-mechanics/categories/2150492677/posts/2158458876

https://courses.physicswithelliot.com/products/fundamentals-of-lagrangian-mechanics/categories/2150492677/posts/2158458877

https://courses.physicswithelliot.com/products/fundamentals-of-lagrangian-mechanics/categories/2150492677/posts/2158458878

https://courses.physicswithelliot.com/products/fundamentals-of-lagrangian-mechanics/categories/2150492678/posts/2158458879

https://courses.physicswithelliot.com/products/fundamentals-of-lagrangian-mechanics/categories/2150492678/posts/2158458880

https://courses.physicswithelliot.com/products/fundamentals-of-lagrangian-mechanics/categories/2150492678/posts/2158458881

https://courses.physicswithelliot.com/products/fundamentals-of-lagrangian-mechanics/categories/2150492678/posts/2158458882

https://courses.physicswithelliot.com/products/fundamentals-of-lagrangian-mechanics/categories/2150492679/posts/2158458883

https://courses.physicswithelliot.com/products/fundamentals-of-lagrangian-mechanics/categories/2150492679/posts/2158458884

https://courses.physicswithelliot.com/products/fundamentals-of-lagrangian-mechanics/categories/2150492679/posts/2158458885

https://courses.physicswithelliot.com/products/fundamentals-of-lagrangian-mechanics/categories/2150492679/posts/2158458886

https://courses.physicswithelliot.com/products/fundamentals-of-lagrangian-mechanics/categories/2150492680/posts/2158458887

https://courses.physicswithelliot.com/products/fundamentals-of-lagrangian-mechanics/categories/2150492680/posts/2158458888

https://courses.physicswithelliot.com/products/fundamentals-of-lagrangian-mechanics/categories/2150492680/posts/2158458889

https://courses.physicswithelliot.com/products/fundamentals-of-lagrangian-mechanics/categories/2150492680/posts/2158458890

https://courses.physicswithelliot.com/products/fundamentals-of-lagrangian-mechanics/categories/2150492681/posts/2158458891 https://courses.physicswithelliot.com/products/fundamentals-of-lagrangian-mechanics/categories/2150492681/posts/2161912436

https://courses.physicswithelliot.com/products/fundamentals-of-lagrangian-mechanics/categories/2150492681/posts/2161912460

https://courses.physicswithelliot.com/products/fundamentals-of-lagrangian-mechanics/categories/2150492681/posts/2161912464

Chapter 49

Hamiltonian

- Elliot Schneider: Physics with Elliot
  - Hamiltonian mechanics

Chapter 50

Fourier analysis

50.1 basic calculation

$$\int_{\frac{-\tau}{2}+\psi}^{\frac{\tau}{2}+\psi} \cos\left(n\frac{2\pi}{\tau}t+\phi\right)\mathrm{d}t = \left[\frac{\sin\left(n\frac{2\pi}{\tau}t+\phi\right)}{n\frac{2\pi}{\tau}}\right]_{\frac{-\tau}{2}+\psi}^{\frac{\tau}{2}+\psi} = 0$$

$$\begin{aligned}
\cos\alpha\cos\beta =& \frac{\cos(\alpha+\beta)+\cos(\alpha-\beta)}{2}\\
\sin\alpha\sin\beta =& \frac{-\cos(\alpha+\beta)+\cos(\alpha-\beta)}{2}\\
\sin\alpha\cos\beta =& \frac{\sin(\alpha+\beta)+\sin(\alpha-\beta)}{2}\\
\cos\alpha\sin\beta =& \frac{\sin(\alpha+\beta)-\sin(\alpha-\beta)}{2}
\end{aligned}$$

$$\begin{aligned}
&\int \cos\left(m\frac{2\pi}{\tau}t\right)\cos\left(n\frac{2\pi}{\tau}t+\phi_n\right)\mathrm{d}t\\
=&\begin{cases}\int \cos\phi_0 \mathrm{d}t & m=n=0\\ \int \frac{\cos\left((m+n)\frac{2\pi}{\tau}t+\phi_n\right)+\cos\left((m-n)\frac{2\pi}{\tau}t-\phi_n\right)\mathrm{d}t}{2} & mn\neq 0\end{cases}\\
=&\begin{cases}t\cos\phi_0 & m=n=0\\ \begin{cases}\int \frac{\cos\left(2n\frac{2\pi}{\tau}t+\phi_n\right)+\cos(-\phi_n)}{2}\mathrm{d}t & m=n\\ \int \frac{\cos\left((m+n)\frac{2\pi}{\tau}t+\phi_n\right)+\cos\left((m-n)\frac{2\pi}{\tau}t-\phi_n\right)}{2}\mathrm{d}t & m\neq n\end{cases} & mn\neq 0\end{cases}\\
=&\begin{cases}t\cos\phi_0 & m=n=0\\ \begin{cases}\int \frac{\cos\left(2n\frac{2\pi}{\tau}t+\phi_n\right)+\cos\phi_n}{2}\mathrm{d}t & m=n\\ \int \frac{\cos\left((m+n)\frac{2\pi}{\tau}t+\phi_n\right)+\cos\left((m-n)\frac{2\pi}{\tau}t-\phi_n\right)}{2}\mathrm{d}t & m\neq n\end{cases} & mn\neq 0\end{cases}\\
=&\begin{cases}t\cos\phi_0 & m=n=0\\ \begin{cases}\int \frac{\cos\left(2n\frac{2\pi}{\tau}t+\phi_n\right)}{2}\mathrm{d}t + \frac{t\cos\phi_n}{2} & m=n\\ \int \frac{\cos\left((m+n)\frac{2\pi}{\tau}t+\phi_n\right)+\cos\left((m-n)\frac{2\pi}{\tau}t-\phi_n\right)}{2}\mathrm{d}t & m\neq n\end{cases} & mn\neq 0\end{cases}
\end{aligned}$$

$$\begin{aligned}
\cos\alpha\cos\beta =& \frac{\cos(\alpha+\beta)+\cos(\alpha-\beta)}{2}\\
\sin\alpha\sin\beta =& \frac{-\cos(\alpha+\beta)+\cos(\alpha-\beta)}{2}\\
\sin\alpha\cos\beta =& \frac{\sin(\alpha+\beta)+\sin(\alpha-\beta)}{2}\\
\cos\alpha\sin\beta =& \frac{\sin(\alpha+\beta)-\sin(\alpha-\beta)}{2}
\end{aligned}$$

$$
\begin{aligned}
&\int_{\frac{-\tau}{2}+\psi}^{\frac{\tau}{2}+\psi} \cos\left(m\frac{2\pi}{\tau}t\right)\cos\left(n\frac{2\pi}{\tau}t+\phi_n\right)\mathrm{d}t \\
&=\begin{cases} \tau\cos\phi_0 & m=n=0 \\ \begin{cases} \frac{\tau\cos\phi_n}{2} & m=n \\ 0 & m\neq n \end{cases} & mn\neq 0 \end{cases} \\
&=\begin{cases} \begin{cases} \tau\cos\phi_0 & n=0 \\ \frac{\tau\cos\phi_n}{2} & n\neq 0 \end{cases} & m=n \\ 0 & m\neq n \end{cases}
\end{aligned}
$$

$$
\begin{aligned}
&\frac{2}{\tau\cos\phi_n}\int_{\frac{-\tau}{2}+\psi}^{\frac{\tau}{2}+\psi} \cos\left(m\frac{2\pi}{\tau}t\right)\cos\left(n\frac{2\pi}{\tau}t+\phi_n\right)\mathrm{d}t \\
&=\begin{cases} \begin{cases} 2 & n=0 \\ 1 & n\neq 0 \end{cases} & m=n \\ 0 & m\neq n \end{cases}
\end{aligned}
$$

$$
\begin{aligned}
&\int_{\frac{-\tau}{2}+\psi}^{\frac{\tau}{2}+\psi} \sin\left(m\frac{2\pi}{\tau}t\right)\cos\left(n\frac{2\pi}{\tau}t+\phi_n\right)\mathrm{d}t \\
&=\begin{cases} 0 & m=n \\ 0 & m\neq n \end{cases}
\end{aligned}
$$

$$
\begin{aligned}
&\int_{\frac{-\tau}{2}+\psi}^{\frac{\tau}{2}+\psi} \cos\left(m\frac{2\pi}{\tau}t\right)\sin\left(n\frac{2\pi}{\tau}t+\phi_n\right)\mathrm{d}t \\
&=\begin{cases} 0 & m=n \\ 0 & m\neq n \end{cases}
\end{aligned}
$$

$$
\begin{aligned}
\frac{2}{\tau\cos\phi_n}\int_{\frac{-\tau}{2}+\psi}^{\frac{\tau}{2}+\psi} \cos\left(m\frac{2\pi}{\tau}t\right)\cos\left(n\frac{2\pi}{\tau}t+\phi_n\right)\mathrm{d}t &= \begin{cases} \begin{cases} 2 & n=0 \\ 1 & n\neq 0 \end{cases} & m=n \\ 0 & m\neq n \end{cases} \\
\frac{2}{\tau\cos\phi_n}\int_{\frac{-\tau}{2}+\psi}^{\frac{\tau}{2}+\psi} \sin\left(m\frac{2\pi}{\tau}t\right)\sin\left(n\frac{2\pi}{\tau}t+\phi_n\right)\mathrm{d}t &= \begin{cases} 1 & m=n\neq 0 \\ 0 & \neg(m=n\neq 0) \end{cases} \\
\int_{\frac{-\tau}{2}+\psi}^{\frac{\tau}{2}+\psi} \sin\left(m\frac{2\pi}{\tau}t\right)\cos\left(n\frac{2\pi}{\tau}t+\phi_n\right)\mathrm{d}t &= \begin{cases} 0 & m=n \\ 0 & m\neq n \end{cases} \\
\int_{\frac{-\tau}{2}+\psi}^{\frac{\tau}{2}+\psi} \cos\left(m\frac{2\pi}{\tau}t\right)\sin\left(n\frac{2\pi}{\tau}t+\phi_n\right)\mathrm{d}t &= \begin{cases} 0 & m=n \\ 0 & m\neq n \end{cases}
\end{aligned}
$$

50.2 Simon Xu

50.2.1 DFT = discrete Fourier transform

50.2.2 FFT = fast Fourier transform

50.2.3 wavelet

Chapter 51

axiom of choice

51.1 Zorn lemma

Chapter 52

linear algebra

52.1 CCJou

https://www.youtube.com/playlist?list=PLP-JUp2VR1LsFtHT-i_vZ3oNFIAc3t_Ju

52.2 Chi, Chen-Yu

https://www.youtube.com/playlist?list=PLJWAeYEa8SXBej3kuQMz8vV41VabZUILb

Chapter 53

analysis

53.1 Chi, Chen-Yu

https://www.youtube.com/playlist?list=PLVJXJebpO4PhAc21JW-cYbzT3sq4s7Qg8

https://www.youtube.com/playlist?list=PLil-R4o6jmGihq7XzdNzb0d5hHqEJbr6L

https://www.youtube.com/playlist?list=PLil-R4o6jmGhUqtKbZf0LIFKd-xN___g_M

https://www.youtube.com/playlist?list=PLil-R4o6jmGhkuZPmKL_A5Y7N4HOsa1nX

53.2 Chen, Jin-Tzu

https://www.youtube.com/playlist?list=PLil-R4o6jmGjoxAWZurHXAY0q9yxwXv5F

Chapter 54

matrix calculus

[13]

$$\boldsymbol{x} = \langle x_1, x_2, \cdots, x_n \rangle = \begin{bmatrix} x_1 & x_2 & \cdots & x_n \end{bmatrix}^\mathsf{T} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}$$

$$f\left(x_1, x_2, \cdots, x_n\right) = f\left(\langle x_1, x_2, \cdots, x_n \rangle\right) = f\left(\boldsymbol{x}\right)$$

$$\boldsymbol{y} = \langle y_1, y_2, \cdots, y_m \rangle = \begin{bmatrix} y_1 & y_2 & \cdots & y_m \end{bmatrix}^\mathsf{T} = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_m \end{bmatrix}$$

54.1 vector-by-scalar

$$\frac{\partial \boldsymbol{y}}{\partial x} = \begin{bmatrix} \frac{\partial y_1}{\partial x} \\ \frac{\partial y_2}{\partial x} \\ \vdots \\ \frac{\partial y_m}{\partial x} \end{bmatrix}$$

54.2 scalar-by-vector

$$\boldsymbol{\nabla} f = \frac{\partial}{\partial \boldsymbol{x}} f = \frac{\partial f}{\partial \boldsymbol{x}} = \left\langle \frac{\partial f}{\partial x_1}, \frac{\partial f}{\partial x_2}, \cdots, \frac{\partial f}{\partial x_n} \right\rangle = \begin{bmatrix} \frac{\partial f}{\partial x_1} & \frac{\partial f}{\partial x_2} & \cdots & \frac{\partial f}{\partial x_n} \end{bmatrix}^\mathsf{T} = \begin{bmatrix} \frac{\partial f}{\partial x_1} \\ \frac{\partial f}{\partial x_2} \\ \vdots \\ \frac{\partial f}{\partial x_n} \end{bmatrix}$$

$$f_i = f_i\left(x_1, x_2, \cdots, x_n\right) = f_i\left(\boldsymbol{x}\right)$$

$$\boldsymbol{f} = \langle f_1, f_2, \cdots, f_m \rangle = \begin{bmatrix} f_1 & f_2 & \cdots & f_m \end{bmatrix}^\mathsf{T} = \begin{bmatrix} f_1 \\ f_2 \\ \vdots \\ f_m \end{bmatrix}$$

54.3 vector-by-vector

54.3.1 numerator-layout notation

分子布局

$$\frac{\partial \boldsymbol{y}}{\partial \boldsymbol{x}}=\begin{bmatrix}\frac{\partial y_1}{\partial \boldsymbol{x}}\\ \frac{\partial y_2}{\partial \boldsymbol{x}}\\ \vdots\\ \frac{\partial y_p}{\partial \boldsymbol{x}}\end{bmatrix}=\begin{bmatrix}\frac{\partial \boldsymbol{y}}{\partial x_1} & \frac{\partial \boldsymbol{y}}{\partial x_2} & \cdots & \frac{\partial \boldsymbol{y}}{\partial x_n}\end{bmatrix}=\begin{bmatrix}\frac{\partial y_1}{\partial x_1} & \frac{\partial y_1}{\partial x_2} & \cdots & \frac{\partial y_1}{\partial x_n}\\ \frac{\partial y_2}{\partial x_1} & \frac{\partial y_2}{\partial x_2} & \cdots & \frac{\partial y_2}{\partial x_n}\\ \vdots & \vdots & \ddots & \vdots\\ \frac{\partial y_m}{\partial x_1} & \frac{\partial y_m}{\partial x_2} & \cdots & \frac{\partial y_m}{\partial x_n}\end{bmatrix}=\begin{bmatrix}\frac{\partial y_i}{\partial x_j}\end{bmatrix}_{m\times n}$$

54.3.2 denominator-layout notation

分母布局

$$\frac{\partial \boldsymbol{y}}{\partial \boldsymbol{x}}=\begin{bmatrix}\frac{\partial \boldsymbol{y}}{\partial x_1}\\ \frac{\partial \boldsymbol{y}}{\partial x_2}\\ \vdots\\ \frac{\partial \boldsymbol{y}}{\partial x_n}\end{bmatrix}=\begin{bmatrix}\frac{\partial y_1}{\partial \boldsymbol{x}} & \frac{\partial y_2}{\partial \boldsymbol{x}} & \cdots & \frac{\partial y_m}{\partial \boldsymbol{x}}\end{bmatrix}=\begin{bmatrix}\frac{\partial y_1}{\partial x_1} & \frac{\partial y_2}{\partial x_1} & \cdots & \frac{\partial y_m}{\partial x_1}\\ \frac{\partial y_1}{\partial x_2} & \frac{\partial y_2}{\partial x_2} & \cdots & \frac{\partial y_m}{\partial x_2}\\ \vdots & \vdots & \ddots & \vdots\\ \frac{\partial y_1}{\partial x_n} & \frac{\partial y_2}{\partial x_n} & \cdots & \frac{\partial y_m}{\partial x_n}\end{bmatrix}=\begin{bmatrix}\frac{\partial y_j}{\partial x_i}\end{bmatrix}_{n\times m}$$

Chapter 55

genetics

genetics

55.1 term

- chromosome
- locus (sl.) loci (pl.)
- marker
- allele
- haplotype
- genotype
- phenotype / trait
  - endophenotype

55.2 marker

- large marker: STRP = short tandem repeat polymorphism, STRs = short tandem repeats
  - linkage analysis
  - paternity testing
  - taxonomy
- small marker: SNP = single-nucleotide polymorphism
  - association analysis
  - disease diagnosis
  - pharmacodynamics? drug design / pharmacogenomics = custom drug
- RFLP = restriction fragment length polymorphism
- platform
  - customized: MALDI-TOF MS(= mass spectrometry)
  - whole-genome gene chip
    * Affymetrix
      · Taiwan BioBank: TWB2.0
    * Illumina
  - NGS = next-generation sequencing

55.3 genome project

- HGP = Human Genome Project
  - 20~25K genes
  - 3 billion bps(base pairs)
  - ELSI = ethical, legal, and social issues
  - 99.9% bps are exactly the same in all people
  - germline mutation
    * male : female = 2 : 1
- HapMap = International HapMap Project
  - SNP, haplotype, tag SNP
    * haplotypes are combination of SNPs
    * tag SNPs can identify unique haplotypes

  - HapMap 1 & 2
    - between-ancestry SNP: 1 common SNP per 5 Kb to 1 common SNP per 1 Kb
      - African
      - European
      - East Asian
        - Han Chinese
        - Japanese
  - HapMap 3: more ancestries
- 1000 Genomes Project
  - NGS
  - identify 95% genetic variants with frequencies at least 1%
  - final phase 77M SNPs
  - browser
    - Ensembl GRCh37
    - Ensembl GRCh38 http://asia.ensembl.org/Homo_sapiens/Info/Index
- TWB = Taiwan BioBank
  - browser: TaiwanView https://taiwanview.twbiobank.org.tw/index
  - pricing: https://www.biobank.org.tw/about_price.php
- TPMI = Taiwan Precision Medicine Initiative https://tpmi.ibms.sinica.edu.tw/www/precision-medicine/

55.4 linkage analysis

- Mendel 1^{st} & 2^{nd} laws
  - law of segregation ~ 3 : 1
  - law of assortment ~ 9 : 3 : 3 : 1
- phenotypic model by R.A. Fisher
  - $$P = G + E$$
    - G is the genotypic component
    - E is the environmental component
  - $$P = G + E + G \cdot E$$
    - $G \cdot E$ is the interaction between the genotypic component and environmental component
- linkage = co-segregation = cosegregation
  - θ = recombination fraction: 1% recombination = 1 crossover per 100 meioses = 1 cM(centiMorgan) on genetic map
- statistical hypothesis testing for in linkage mapping
  - statistical hypothesis testing for categorical trait in linkage mapping: PLA = parametric linkage analysis
    - H_0: no linkage $\theta = 0.5$
    - H_1: linkage $\theta < 0.5$
  - statistical hypothesis testing for quantitative trait in linkage mapping: VCLA = variance component linkage analysis
    - H_0: no linkage $\sigma_q^2 = 0$
    - H_1: linkage $\sigma_q^2 > 0$
- study design
  - case control
  - trio
  - affected / discordant sib-pair
  - extended pedigree
- data format: linkage format
  - family-based
    - PID = pedigree Id
    - IID = individual Id
    - FID = father Id
    - MID = mother Id
    - gender
    - affection
    - marker
      - M1 = marker 1
      - M2 = marker 2
      - ...
- single major locus model
  - a two-allele A and a locus influences a dichotomous trait

- allele frequency
  - $p = \mathrm{P}(A)$
  - $q = \mathrm{P}(a) = 1 - p$
- penetrance
  - $f_{AA} = \mathrm{P}(\text{affected} \mid AA)$
  - $f_{Aa} = \mathrm{P}(\text{affected} \mid Aa) = f_{aA}$
  - $f_{aa} = \mathrm{P}(\text{affected} \mid aa)$
- disease mode of inheritance
  - dominant model
    - $\begin{cases} f_{AA} = 1 \\ f_{Aa} = 1 \\ f_{aa} = 0 \end{cases}$
  - recessive model
    - $\begin{cases} f_{AA} = 0 \\ f_{Aa} = 0 \\ f_{aa} = 1 \end{cases}$
  - additive model
    - $\begin{cases} f_{AA} = 1 \\ f_{Aa} = \frac{1}{2} \\ f_{aa} = 0 \end{cases}$
  - phenocopy model
    - $f_{aa} > 0$ perhaps due to environmental cause
  - liability model
    - e.g. f_{AA}, f_{Aa}, f_{aa} are age-dependent

55.4.1 PLA = parametric linkage analysis

- LOD score

55.4.2 VCLA = variance component linkage analysis

- allele-sharing: IBS and IBD
  - IBS = identity-by-state
  - IBD = identity-by-descent

55.5 association analysis

- LD = linkage disequilibrium
- genotype & allele frequency
  - diallelic marker
    - p_{AA}
    - p_{Aa}
    - p_{aA}
    - p_{aa}
    - p_A
    - p_a
- HWC = Hardy-Weinberg condition
  - HWE = Hardy-Weinberg equilibrium
  - HWD = Hardy-Weinberg disequilibrium
- gametic or haplotype frequency
  - LE = linkage equilibrium
  - LD = linkage disequilibrium

Chapter 56

multivariate normal distribution

Definition 56.1. probability density function (PDF) of normal distribution (= Gaussian distribution)

$$\mathcal{N}\left(x \mid \mu, \sigma^2\right) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left\{-\frac{(x-\mu)^2}{2\sigma^2}\right\}$$

If a continuous random variable X follows a normal distribution with mean 0 and variance σ^2

$$\begin{aligned} X &\sim \mathrm{n}\left(\mu, \sigma^2\right) = \mathcal{N}\left(\mu, \sigma^2\right) \\ \Leftrightarrow f_X(x) &= \frac{1}{\sigma\sqrt{2\pi}} \mathrm{e}^{\frac{-1}{2}\left(\frac{x-\mu}{\sigma}\right)^2} = \frac{\mathrm{e}^{\frac{-1}{2}\left(\frac{x-\mu}{\sigma}\right)^2}}{\sigma\sqrt{2\pi}} = \mathrm{n}\left(x \mid \mu, \sigma^2\right) \\ &= \frac{1}{\sqrt{2\pi}\sigma} \exp\left\{-\frac{(x-\mu)^2}{2\sigma^2}\right\} = \mathcal{N}\left(x \mid \mu, \sigma^2\right) \end{aligned}$$

A random variable X can be standardized by subtracting the mean μ and dividing by the standard deviation σ, resulting in the standardized random variable Z

$$Z = \frac{X-\mu}{\sigma} \text{ or } z = \frac{x-\mu}{\sigma}$$

The standardized random variable Z follows the standard normal distribution

$$\begin{aligned} Z &\sim \mathrm{n}\left(0, 1^2\right) = \mathcal{N}\left(0, 1^2\right) \\ \Leftrightarrow f_Z(z) &= \frac{\mathrm{e}^{\frac{-1}{2}\left(\frac{x-0}{1}\right)^2}}{1 \cdot \sqrt{2\pi}} = \frac{1}{\sqrt{2\pi}} \exp\left\{-\frac{z^2}{2}\right\} \\ &= \mathrm{n}\left(z \mid 0, 1^2\right) = \mathcal{N}\left(z \mid 0, 1^2\right) \end{aligned}$$

To generalize from univariate random variables to multivariate random vectors, a random vector[14]

$$\boldsymbol{Z} = \langle Z_1, Z_2, \cdots, Z_p \rangle = \begin{bmatrix} Z_1 & Z_2 & \cdots & Z_p \end{bmatrix}^\intercal = \begin{bmatrix} Z_1 \\ Z_2 \\ \vdots \\ Z_p \end{bmatrix}$$

with p random variable components is said to follow the standard multivariate normal distribution if and only if its joint PDF is given by

$$f_{\boldsymbol{Z}}(\boldsymbol{z}) = \frac{1}{(2\pi)^{p/2}} \exp\left\{-\frac{\boldsymbol{z}^\intercal \boldsymbol{z}}{2}\right\} = \frac{1}{(2\pi)^{p/2}} \exp\left\{-\frac{\boldsymbol{z} \cdot \boldsymbol{z}}{2}\right\} \tag{56.1}$$

(56.1) can be rewritten as the following

$$
\begin{aligned}
f_{\boldsymbol{Z}}(\boldsymbol{z}) &= \underbrace{\frac{1}{\sqrt{2\pi}}\frac{1}{\sqrt{2\pi}}\cdots\frac{1}{\sqrt{2\pi}}}_{p\text{ times}}\exp\left\{-\frac{z_1^2}{2}-\frac{z_2^2}{2}-\cdots-\frac{z_p^2}{2}\right\} \\
&= \frac{1}{\sqrt{2\pi}}\exp\left\{-\frac{z_1^2}{2}\right\}\frac{1}{\sqrt{2\pi}}\exp\left\{-\frac{z_2^2}{2}\right\}\cdots\frac{1}{\sqrt{2\pi}}\exp\left\{-\frac{z_p^2}{2}\right\} \\
&= f(z_1)\,f(z_2)\cdots f(z_p)
\end{aligned}
$$

where

$$
f_{Z_i}(z_i) = \frac{1}{\sqrt{2\pi}}\exp\left\{-\frac{z_i^2}{2}\right\} = f(z_i) \Rightarrow Z_i \sim \mathcal{N}\left(0,1^2\right) = \mathrm{n}\left(0,1^2\right) \tag{56.2}
$$

$$
\begin{aligned}
f_{Z_i}(z_i) &= \int_{-\infty}^{\infty}\cdots\int_{-\infty}^{\infty} f_{\boldsymbol{Z}}(z_1,\ldots,z_{i-1},z_i,z_{i+1},\ldots,z_p)\,\mathrm{d}z_1\cdots\mathrm{d}z_{i-1}\mathrm{d}z_{i+1}\cdots\mathrm{d}z_p \\
&= \int_{-\infty}^{\infty}\cdots\int_{-\infty}^{\infty} f(z_1)\cdots f(z_{i-1})\,f(z_i)\,f(z_{i+1})\cdots f(z_p)\,\mathrm{d}z_1\cdots\mathrm{d}z_{i-1}\mathrm{d}z_{i+1}\cdots\mathrm{d}z_p \\
&= f(z_i)\int_{-\infty}^{\infty} f(z_1)\,\mathrm{d}z_1\cdots\int_{-\infty}^{\infty} f(z_{i-1})\,\mathrm{d}z_{i-1}\int_{-\infty}^{\infty} f(z_{i+1})\,\mathrm{d}z_{i+1}\cdots\int_{-\infty}^{\infty} f(z_p)\,\mathrm{d}z_p \\
&= f(z_i) \overset{(56.2)}{=} \frac{1}{\sqrt{2\pi}}\exp\left\{-\frac{z_i^2}{2}\right\}
\end{aligned}
$$

Definition 56.2. covariance matrix of a random vector[8]

$$
\mathrm{C}[\boldsymbol{X}] = \mathrm{Cov}[\boldsymbol{X}] = \mathrm{V}[\boldsymbol{X}] = \mathrm{E}\left[[\boldsymbol{X}-\mathrm{E}(\boldsymbol{X})]\,[\boldsymbol{X}-\mathrm{E}(\boldsymbol{X})]^{\mathsf{T}}\right]
$$

$$
\boldsymbol{X} = \langle X_1, X_2, \cdots, X_p\rangle = \begin{bmatrix} X_1 & X_2 & \cdots & X_p\end{bmatrix}^{\mathsf{T}} = \begin{bmatrix} X_1 \\ X_2 \\ \vdots \\ X_p\end{bmatrix}
$$

$$
\mathrm{E}[\boldsymbol{X}] = \langle \mathrm{E}[X_1], \mathrm{E}[X_2], \cdots, \mathrm{E}[X_p]\rangle = \begin{bmatrix} \mathrm{E}[X_1] & \mathrm{E}[X_2] & \cdots & \mathrm{E}[X_p]\end{bmatrix}^{\mathsf{T}} = \begin{bmatrix} \mathrm{E}[X_1] \\ \mathrm{E}[X_2] \\ \vdots \\ \mathrm{E}[X_p]\end{bmatrix}
$$

$$
\boldsymbol{X} - \mathrm{E}[\boldsymbol{X}] = \begin{bmatrix} X_1-\mathrm{E}[X_1] \\ X_2-\mathrm{E}[X_2] \\ \vdots \\ X_p-\mathrm{E}[X_p]\end{bmatrix}
$$

$$
\begin{aligned}
&[\boldsymbol{X}-\mathrm{E}(\boldsymbol{X})]\,[\boldsymbol{X}-\mathrm{E}(\boldsymbol{X})]^{\mathsf{T}} \\
&= \begin{bmatrix} X_1-\mathrm{E}[X_1] \\ X_2-\mathrm{E}[X_2] \\ \vdots \\ X_p-\mathrm{E}[X_p]\end{bmatrix}\begin{bmatrix} X_1-\mathrm{E}[X_1] & X_2-\mathrm{E}[X_2] & \cdots & X_p-\mathrm{E}[X_p]\end{bmatrix} \\
&= \begin{bmatrix}
(X_1-\mathrm{E}[X_1])(X_1-\mathrm{E}[X_1]) & (X_1-\mathrm{E}[X_1])(X_2-\mathrm{E}[X_2]) & \cdots & (X_1-\mathrm{E}[X_1])(X_p-\mathrm{E}[X_p]) \\
(X_2-\mathrm{E}[X_2])(X_1-\mathrm{E}[X_1]) & (X_2-\mathrm{E}[X_2])(X_2-\mathrm{E}[X_2]) & \cdots & (X_2-\mathrm{E}[X_2])(X_p-\mathrm{E}[X_p]) \\
\vdots & \vdots & \ddots & \vdots \\
(X_p-\mathrm{E}[X_p])(X_1-\mathrm{E}[X_1]) & (X_p-\mathrm{E}[X_p])(X_2-\mathrm{E}[X_2]) & \cdots & (X_p-\mathrm{E}[X_p])(X_p-\mathrm{E}[X_p])
\end{bmatrix} \\
&= \begin{bmatrix}
(X_1-\mathrm{E}[X_1])^2 & \cdots & (X_1-\mathrm{E}[X_1])(X_p-\mathrm{E}[X_p]) \\
\vdots & \ddots & \vdots \\
(X_p-\mathrm{E}[X_p])(X_1-\mathrm{E}[X_1]) & \cdots & (X_p-\mathrm{E}[X_p])^2
\end{bmatrix}
\end{aligned}
$$

$$\begin{aligned}
&\mathrm{E}\left[\left[\boldsymbol{X}-\mathrm{E}\left(\boldsymbol{X}\right)\right]\left[\boldsymbol{X}-\mathrm{E}\left(\boldsymbol{X}\right)\right]^{\intercal}\right] && (56.3)\\
=&\mathrm{E}\begin{bmatrix}\left(X_1-\mathrm{E}\left[X_1\right]\right)^2 & \cdots & \left(X_1-\mathrm{E}\left[X_1\right]\right)\left(X_p-\mathrm{E}\left[X_p\right]\right)\\ \vdots & \ddots & \vdots\\ \left(X_p-\mathrm{E}\left[X_p\right]\right)\left(X_1-\mathrm{E}\left[X_1\right]\right) & \cdots & \left(X_p-\mathrm{E}\left[X_p\right]\right)^2\end{bmatrix} && (56.4)\\
=&\begin{bmatrix}\mathrm{E}\left[\left(X_1-\mathrm{E}\left[X_1\right]\right)^2\right] & \cdots & \mathrm{E}\left[\left(X_1-\mathrm{E}\left[X_1\right]\right)\left(X_p-\mathrm{E}\left[X_p\right]\right)\right]\\ \vdots & \ddots & \vdots\\ \mathrm{E}\left[\left(X_p-\mathrm{E}\left[X_p\right]\right)\left(X_1-\mathrm{E}\left[X_1\right]\right)\right] & \cdots & \mathrm{E}\left[\left(X_p-\mathrm{E}\left[X_p\right]\right)^2\right]\end{bmatrix} && (56.5)\\
=&\begin{bmatrix}\mathrm{V}\left(X_1,X_1\right) & \cdots & \mathrm{V}\left(X_1,X_p\right)\\ \vdots & \ddots & \vdots\\ \mathrm{V}\left(X_p,X_1\right) & \cdots & \mathrm{V}\left(X_p,X_p\right)\end{bmatrix}=\begin{bmatrix}\mathrm{V}\left(X_1,X_1\right) & \mathrm{V}\left(X_1,X_2\right) & \cdots & \mathrm{V}\left(X_1,X_p\right)\\ \mathrm{V}\left(X_2,X_1\right) & \mathrm{V}\left(X_2,X_2\right) & \cdots & \mathrm{V}\left(X_2,X_p\right)\\ \vdots & \vdots & \ddots & \vdots\\ \mathrm{V}\left(X_p,X_1\right) & \mathrm{V}\left(X_p,X_2\right) & \cdots & \mathrm{V}\left(X_p,X_p\right)\end{bmatrix} && (56.6)\\
=&\begin{bmatrix}\mathrm{V}\left(X_1\right) & \cdots & \mathrm{V}\left(X_1,X_p\right)\\ \vdots & \ddots & \vdots\\ \mathrm{V}\left(X_p,X_1\right) & \cdots & \mathrm{V}\left(X_p\right)\end{bmatrix}=\begin{bmatrix}\mathrm{V}\left(X_1\right) & \mathrm{V}\left(X_1,X_2\right) & \cdots & \mathrm{V}\left(X_1,X_p\right)\\ \mathrm{V}\left(X_2,X_1\right) & \mathrm{V}\left(X_2\right) & \cdots & \mathrm{V}\left(X_2,X_p\right)\\ \vdots & \vdots & \ddots & \vdots\\ \mathrm{V}\left(X_p,X_1\right) & \mathrm{V}\left(X_p,X_2\right) & \cdots & \mathrm{V}\left(X_p\right)\end{bmatrix} && (56.7)\\
=&\begin{bmatrix}\mathrm{V}\left(X_1\right) & \cdots & \mathrm{C}\left(X_1,X_p\right)\\ \vdots & \ddots & \vdots\\ \mathrm{C}\left(X_p,X_1\right) & \cdots & \mathrm{V}\left(X_p\right)\end{bmatrix}=\begin{bmatrix}\mathrm{V}\left(X_1\right) & \mathrm{C}\left(X_1,X_2\right) & \cdots & \mathrm{C}\left(X_1,X_p\right)\\ \mathrm{C}\left(X_2,X_1\right) & \mathrm{V}\left(X_2\right) & \cdots & \mathrm{C}\left(X_2,X_p\right)\\ \vdots & \vdots & \ddots & \vdots\\ \mathrm{C}\left(X_p,X_1\right) & \mathrm{C}\left(X_p,X_2\right) & \cdots & \mathrm{V}\left(X_p\right)\end{bmatrix} && (56.8)\\
=&\begin{bmatrix}\sigma_1^2 & \cdots & \sigma_{1p}\\ \vdots & \ddots & \vdots\\ \sigma_{p1} & \cdots & \sigma_p^2\end{bmatrix}=\begin{bmatrix}\sigma_1^2 & \sigma_{12} & \cdots & \sigma_{1p}\\ \sigma_{21} & \sigma_2^2 & \cdots & \sigma_{2p}\\ \vdots & \vdots & \ddots & \vdots\\ \sigma_{p1} & \sigma_{p2} & \cdots & \sigma_p^2\end{bmatrix}=\begin{bmatrix}\sigma_{11} & \sigma_{12} & \cdots & \sigma_{1p}\\ \sigma_{21} & \sigma_{22} & \cdots & \sigma_{2p}\\ \vdots & \vdots & \ddots & \vdots\\ \sigma_{p1} & \sigma_{p2} & \cdots & \sigma_{pp}\end{bmatrix}=\left[\sigma_{ij}\right]_{p\times p}=\mathit{\Sigma} && (56.9)
\end{aligned}$$

$$\boldsymbol{X}\sim\mathcal{D}\left(\boldsymbol{\mu},\mathit{\Sigma}\right)=\mathrm{d}\left(\boldsymbol{\mu}_{\boldsymbol{X}},\mathit{\Sigma}_{\boldsymbol{X}}\right)=\mathrm{d}\left(\mathrm{E}\left[\boldsymbol{X}\right],\mathrm{C}\left[\boldsymbol{X}\right]\right)=\mathrm{d}\left(\mathrm{E}\left[\boldsymbol{X}\right],\mathrm{V}\left[\boldsymbol{X}\right]\right)$$

$$\boldsymbol{Z}\sim\mathcal{N}\left(\boldsymbol{\mu}_{\boldsymbol{Z}},\mathit{\Sigma}_{\boldsymbol{Z}}\right)=\mathrm{n}\left(\mathrm{E}\left[\boldsymbol{Z}\right],\mathrm{V}\left[\boldsymbol{Z}\right]\right)$$

$$\begin{aligned}
&\mathrm{E}\left[\boldsymbol{Z}\right]=\begin{bmatrix}\mathrm{E}\left[Z_1\right]\\ \mathrm{E}\left[Z_2\right]\\ \vdots\\ \mathrm{E}\left[Z_p\right]\end{bmatrix}=\left[\mathrm{E}\left[Z_i\right]\right]_{p\times 1}\\
\Rightarrow&\mathrm{E}\left[Z_i\right]=\int_{-\infty}^{\infty}z_i f_{Z_i}\left(z_i\right)\mathrm{d}z_i\overset{(56.2)}{=}\int_{-\infty}^{\infty}z_i\frac{\mathrm{e}^{\frac{-1}{2}z_i^2}}{\sqrt{2\pi}}\mathrm{d}z_i=0\\
\Rightarrow&\mathrm{E}\left[\boldsymbol{Z}\right]=\boldsymbol{0}\\
\Rightarrow&\boldsymbol{Z}\sim\mathcal{N}\left(\boldsymbol{\mu}_{\boldsymbol{Z}}=\boldsymbol{0},\mathit{\Sigma}_{\boldsymbol{Z}}\right)=\mathrm{n}\left(\boldsymbol{0},\mathrm{V}\left[\boldsymbol{Z}\right]\right)
\end{aligned}$$

$$\mathrm{V}\left(Z_i\right)=\int_{-\infty}^{\infty}\left(z_i-\mu_{Z_i}\right)^2 f_{Z_i}\left(z_i\right)\mathrm{d}z_i\overset{(56.2)}{=}\int_{-\infty}^{\infty}\left(z_i-0\right)^2\frac{\mathrm{e}^{\frac{-1}{2}z_i^2}}{\sqrt{2\pi}}\mathrm{d}z_i=1$$

$$\mathrm{V}\left(Z_i,Z_j\right)\overset{i\neq j\Rightarrow Z_i,Z_j\text{are independent}}{=}0 \tag{56.10}$$

$$
\begin{aligned}
\mathrm{V}[\boldsymbol{Z}] &= \begin{bmatrix} \mathrm{V}(Z_1) & \mathrm{V}(Z_1,Z_2) & \cdots & \mathrm{V}(Z_1,Z_p) \\ \mathrm{V}(Z_2,Z_1) & \mathrm{V}(Z_2) & \cdots & \mathrm{V}(Z_2,Z_p) \\ \vdots & \vdots & \ddots & \vdots \\ \mathrm{V}(Z_p,Z_1) & \mathrm{V}(Z_p,Z_2) & \cdots & \mathrm{V}(Z_p) \end{bmatrix} = \begin{bmatrix} \sigma_{11} & \sigma_{12} & \cdots & \sigma_{1p} \\ \sigma_{21} & \sigma_{22} & \cdots & \sigma_{2p} \\ \vdots & \vdots & \ddots & \vdots \\ \sigma_{p1} & \sigma_{p2} & \cdots & \sigma_{pp} \end{bmatrix} \\
&= \begin{bmatrix} \sigma_1^2 & \sigma_{12} & \cdots & \sigma_{1p} \\ \sigma_{21} & \sigma_2^2 & \cdots & \sigma_{2p} \\ \vdots & \vdots & \ddots & \vdots \\ \sigma_{p1} & \sigma_{p2} & \cdots & \sigma_p^2 \end{bmatrix} \overset{56.10}{=} \begin{bmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 \end{bmatrix} = I_{p\times p} = I_p = I
\end{aligned}
$$

$$
\boldsymbol{Z} \sim \mathcal{N}(\boldsymbol{\mu}_{\boldsymbol{Z}}, \Sigma_{\boldsymbol{Z}}) = \mathrm{n}(\mathrm{E}[\boldsymbol{Z}], \mathrm{V}[\boldsymbol{Z}]) = \mathcal{N}(\boldsymbol{0}, I) \Leftrightarrow \begin{cases} \boldsymbol{\mu}_{\boldsymbol{Z}} = \mathrm{E}[\boldsymbol{Z}] = \boldsymbol{0} = [0]_p = [0]_{p\times 1} \\ \Sigma_{\boldsymbol{Z}} = \mathrm{V}[\boldsymbol{Z}] = I = I_p = I_{p\times p} \end{cases}
$$

$$
\begin{aligned}
\boldsymbol{Z} = \begin{bmatrix} Z_1 \\ Z_2 \\ \vdots \\ Z_p \end{bmatrix} &= \begin{bmatrix} \dfrac{X_1-\mu_1}{\sigma_1} \\ \dfrac{X_2-\mu_2}{\sigma_2} \\ \vdots \\ \dfrac{X_p-\mu_p}{\sigma_p} \end{bmatrix} = \begin{bmatrix} \dfrac{1}{\sigma_1} & 0 & \cdots & 0 \\ 0 & \dfrac{1}{\sigma_2} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \dfrac{1}{\sigma_p} \end{bmatrix} \begin{bmatrix} X_1-\mu_1 \\ X_2-\mu_2 \\ \vdots \\ X_p-\mu_p \end{bmatrix} \\
&= \begin{bmatrix} \dfrac{1}{\sigma_1} & 0 & \cdots & 0 \\ 0 & \dfrac{1}{\sigma_2} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \dfrac{1}{\sigma_p} \end{bmatrix} \left(\begin{bmatrix} X_1 \\ X_2 \\ \vdots \\ X_p \end{bmatrix} - \begin{bmatrix} \mu_1 \\ \mu_2 \\ \vdots \\ \mu_p \end{bmatrix} \right) = B^{-1}(\boldsymbol{X}-\boldsymbol{\mu}) \\
\Rightarrow \boldsymbol{X} &= B\boldsymbol{Z} + \boldsymbol{\mu}
\end{aligned}
$$

$$
\boldsymbol{X} = B\boldsymbol{Z} + \boldsymbol{\mu} = T(\boldsymbol{Z})
$$

by $\mathrm{V}[A\boldsymbol{X}+\boldsymbol{b}] = A\mathrm{V}[\boldsymbol{X}]A^\intercal$

$$
\Sigma = \Sigma_{\boldsymbol{X}} = \mathrm{V}[\boldsymbol{X}] = \mathrm{V}[B\boldsymbol{Z}+\boldsymbol{\mu}] = B\mathrm{V}[\boldsymbol{Z}]B^\intercal = BIB^\intercal = BB^\intercal \tag{56.11}
$$

Consider two infinitesimal volumes of p-dimensional parallelepipeds in the different $\mathbb{R}^p$ spaces[15]

$$
V_{\boldsymbol{x}} = [x_1, x_1+\mathrm{d}x_1] \times [x_2, x_2+\mathrm{d}x_2] \times \cdots \times [x_p, x_p+\mathrm{d}x_p]
$$

and

$$
V_{\boldsymbol{z}} = [z_1, z_1+\mathrm{d}z_1] \times [z_2, z_2+\mathrm{d}z_2] \times \cdots \times [z_p, z_p+\mathrm{d}z_p]
$$

Their relationship under linear transformation is

$$
\begin{aligned}
V_{\boldsymbol{x}} = T(V_{\boldsymbol{z}}) = &[T(z_1), T(z_1)+T(\mathrm{d}z_1)] \\
&\times [T(z_2), T(z_2)+T(\mathrm{d}z_2)] \\
&\times \cdots \\
&\times [T(z_p), T(z_p)+T(\mathrm{d}z_p)]
\end{aligned}
$$

and

$$
\mathrm{d}x_i = \sum_j \frac{\partial x_i}{\partial z_j}\mathrm{d}z_j
$$

For examples in 2 dimension,

$$\begin{bmatrix}\mathrm{d}x_1\\\mathrm{d}x_2\end{bmatrix}=\begin{bmatrix}\dfrac{\partial x_1}{\partial z_1}&\dfrac{\partial x_1}{\partial z_2}\\\dfrac{\partial x_2}{\partial z_1}&\dfrac{\partial x_2}{\partial z_2}\end{bmatrix}\begin{bmatrix}\mathrm{d}z_1\\\mathrm{d}z_2\end{bmatrix}$$

Two element infinitesimal one-directional vectors of $\boldsymbol{Z}$ transformed into another space of $\boldsymbol{X}$ are

$$T\left(\mathrm{d}\boldsymbol{z}_1\right)=\begin{bmatrix}\dfrac{\partial x_1}{\partial z_1}&\dfrac{\partial x_1}{\partial z_2}\\\dfrac{\partial x_2}{\partial z_1}&\dfrac{\partial x_2}{\partial z_2}\end{bmatrix}\begin{bmatrix}\mathrm{d}z_1\\0\end{bmatrix}=\begin{bmatrix}\dfrac{\partial x_1}{\partial z_1}\mathrm{d}z_1\\\dfrac{\partial x_2}{\partial z_1}\mathrm{d}z_1\end{bmatrix}$$

and

$$T\left(\mathrm{d}\boldsymbol{z}_2\right)=\begin{bmatrix}\dfrac{\partial x_1}{\partial z_1}&\dfrac{\partial x_1}{\partial z_2}\\\dfrac{\partial x_2}{\partial z_1}&\dfrac{\partial x_2}{\partial z_2}\end{bmatrix}\begin{bmatrix}0\\\mathrm{d}z_2\end{bmatrix}=\begin{bmatrix}\dfrac{\partial x_1}{\partial z_2}\mathrm{d}z_2\\\dfrac{\partial x_2}{\partial z_2}\mathrm{d}z_2\end{bmatrix}$$

Their corresponding area(volume) in the space of $\boldsymbol{X}$ is

$$\begin{aligned}&\int_{A_{\boldsymbol{x}}}\mathrm{d}A_{\boldsymbol{x}}=\int_{A_{\boldsymbol{x}}}\mathrm{d}x_1\mathrm{d}x_2=\int_{T(A_{\boldsymbol{z}})}\mathrm{d}x_1\mathrm{d}x_2\\&=\int_{A_{\boldsymbol{z}}}\left|\begin{bmatrix}T\left(\mathrm{d}\boldsymbol{z}_1\right)&T\left(\mathrm{d}\boldsymbol{z}_2\right)\end{bmatrix}\right|=\int_{A_{\boldsymbol{z}}}\begin{vmatrix}\dfrac{\partial x_1}{\partial z_1}\mathrm{d}z_1&\dfrac{\partial x_1}{\partial z_2}\mathrm{d}z_2\\\dfrac{\partial x_2}{\partial z_1}\mathrm{d}z_1&\dfrac{\partial x_2}{\partial z_2}\mathrm{d}z_2\end{vmatrix}\\&=\int_{A_{\boldsymbol{z}}}\begin{vmatrix}\dfrac{\partial x_1}{\partial z_1}&\dfrac{\partial x_1}{\partial z_2}\\\dfrac{\partial x_2}{\partial z_1}&\dfrac{\partial x_2}{\partial z_2}\end{vmatrix}\mathrm{d}z_1\mathrm{d}z_2=\int_{A_{\boldsymbol{z}}}|J|\,\mathrm{d}A_{\boldsymbol{z}}\end{aligned}$$

To generalize for volumes in p dimension,

$$\begin{aligned}&\int_{V_{\boldsymbol{x}}}\mathrm{d}V_{\boldsymbol{x}}=\int_{V_{\boldsymbol{x}}}\mathrm{d}x_1\mathrm{d}x_2\cdots\mathrm{d}x_p=\int_{T(V_{\boldsymbol{z}})}\mathrm{d}x_1\mathrm{d}x_2\cdots\mathrm{d}x_p\\&=\int_{A_{\boldsymbol{z}}}\left|\begin{bmatrix}T\left(\mathrm{d}\boldsymbol{z}_1\right)&T\left(\mathrm{d}\boldsymbol{z}_2\right)&\cdots&T\left(\mathrm{d}\boldsymbol{z}_p\right)\end{bmatrix}\right|=\int_{V_{\boldsymbol{z}}}\left|\begin{bmatrix}\dfrac{\partial x_i}{\partial z_j}\mathrm{d}z_j\end{bmatrix}_{p\times p}\right|\\&=\int_{V_{\boldsymbol{z}}}\left|\begin{bmatrix}\dfrac{\partial x_i}{\partial z_j}\end{bmatrix}_{p\times p}\right|\mathrm{d}z_1\mathrm{d}z_2\cdots\mathrm{d}z_p=\int_{V_{\boldsymbol{z}}}|J|\,\mathrm{d}V_{\boldsymbol{z}}\end{aligned}$$

i.e.

$$\int_{V_{\boldsymbol{x}}}\mathrm{d}V_{\boldsymbol{x}}=\int_{V_{\boldsymbol{z}}}|J|\,\mathrm{d}V_{\boldsymbol{z}}\tag{56.12}$$

where J is a Jacobian matrix

$$J=\begin{bmatrix}\dfrac{\partial x_i}{\partial z_j}\end{bmatrix}_{p\times p}=\frac{\partial\boldsymbol{x}}{\partial\boldsymbol{z}}$$

or $|J|$ is a Jacobian determinant or simply Jacobian

$$|J|=\left|\frac{\partial x_i}{\partial z_j}\right|_{p\times p}=\left|\frac{\partial\boldsymbol{x}}{\partial\boldsymbol{z}}\right|$$

The probability of the same event should be invariant under transformation.

$$\begin{aligned}
&\int_{V_{\boldsymbol{x}}} f_{\boldsymbol{X}}(\boldsymbol{x})\,\mathrm{d}V_{\boldsymbol{x}} = \int_{V_{\boldsymbol{x}}} f_{\boldsymbol{X}}(\boldsymbol{x})\,\mathrm{d}x_1\mathrm{d}x_2\cdots\mathrm{d}x_p \\
=&\int_{T(V_{\boldsymbol{z}})} f_{\boldsymbol{X}}(\boldsymbol{x})\,\mathrm{d}x_1\mathrm{d}x_2\cdots\mathrm{d}x_p \\
=&\int_{V_{\boldsymbol{z}}} f_{\boldsymbol{Z}}(\boldsymbol{z})\,\mathrm{d}V_{\boldsymbol{z}} = \int_{V_{\boldsymbol{z}}} f_{\boldsymbol{Z}}(\boldsymbol{z})\,\mathrm{d}z_1\mathrm{d}z_2\cdots\mathrm{d}z_p
\end{aligned}$$

i.e.

$$\int_{V_{\boldsymbol{x}}} f_{\boldsymbol{X}}(\boldsymbol{x})\,\mathrm{d}V_{\boldsymbol{x}} = \int_{V_{\boldsymbol{z}}} f_{\boldsymbol{Z}}(\boldsymbol{z})\,\mathrm{d}V_{\boldsymbol{z}} \tag{56.13}$$

$$\begin{cases}\int_{V_{\boldsymbol{x}}} f_{\boldsymbol{X}}(\boldsymbol{x})\,\mathrm{d}V_{\boldsymbol{x}} = \int_{V_{\boldsymbol{z}}} f_{\boldsymbol{Z}}(\boldsymbol{z})\,\mathrm{d}V_{\boldsymbol{z}} & 56.13\\ \int_{V_{\boldsymbol{x}}} \mathrm{d}V_{\boldsymbol{x}} = \int_{V_{\boldsymbol{z}}} |J|\,\mathrm{d}V_{\boldsymbol{z}} & 56.12\end{cases}$$

$$\begin{aligned}
\boldsymbol{Z} =& B^{-1}(\boldsymbol{X}-\boldsymbol{\mu}) \\
\boldsymbol{z} =& B^{-1}(\boldsymbol{x}-\boldsymbol{\mu}) \\
\boldsymbol{X} =& B\boldsymbol{Z}+\boldsymbol{\mu} \\
\boldsymbol{x} =& B\boldsymbol{z}+\boldsymbol{\mu}
\end{aligned} \tag{56.14}$$

$$\begin{aligned}
J = \left[\frac{\partial x_i}{\partial z_j}\right]_{p\times p} = \frac{\partial \boldsymbol{x}}{\partial \boldsymbol{z}} =& B \\
|J| = \left|\frac{\partial x_i}{\partial z_j}\right|_{p\times p} = \left|\frac{\partial \boldsymbol{x}}{\partial \boldsymbol{z}}\right| =& |B|
\end{aligned} \tag{56.15}$$

$$\begin{aligned}
\int_{V_{\boldsymbol{z}}} f_{\boldsymbol{Z}}(\boldsymbol{z})\,\mathrm{d}V_{\boldsymbol{z}} \overset{56.13}{=}& \int_{V_{\boldsymbol{x}}} f_{\boldsymbol{X}}(\boldsymbol{x})\,\mathrm{d}V_{\boldsymbol{x}} = \int_{V_{\boldsymbol{x}}} \mathrm{d}V_{\boldsymbol{x}} f_{\boldsymbol{X}}(\boldsymbol{x}) \\
\overset{56.12}{=} \int_{V_{\boldsymbol{z}}} |J|\,\mathrm{d}V_{\boldsymbol{z}} f_{\boldsymbol{X}}(\boldsymbol{x}(\boldsymbol{z})) =& \int_{V_{\boldsymbol{z}}} f_{\boldsymbol{X}}(\boldsymbol{x}(\boldsymbol{z}))\,|J|\,\mathrm{d}V_{\boldsymbol{z}} \\
f_{\boldsymbol{Z}}(\boldsymbol{z}) \overset{\Downarrow}{=}& f_{\boldsymbol{X}}(\boldsymbol{x}(\boldsymbol{z}))\,|J| \\
f_{\boldsymbol{X}}(\boldsymbol{x}(\boldsymbol{z})) \overset{\Downarrow}{=}& |J|^{-1} f_{\boldsymbol{Z}}(\boldsymbol{z}) \overset{56.1}{=} |J|^{-1}\frac{1}{(2\pi)^{p/2}}\exp\left\{\frac{-\boldsymbol{z}^\intercal\boldsymbol{z}}{2}\right\} \\
f_{\boldsymbol{X}}(\boldsymbol{x}) \overset{\Downarrow}{=}& |J|^{-1} f_{\boldsymbol{Z}}(\boldsymbol{z}(\boldsymbol{x})) \overset{56.15,56.14}{=} |B|^{-1} f_{\boldsymbol{Z}}\left(B^{-1}(\boldsymbol{x}-\boldsymbol{\mu})\right) \\
=& |B|^{-1}(2\pi)^{-p/2}\exp\left\{\frac{-1}{2}\left[B^{-1}(\boldsymbol{x}-\boldsymbol{\mu})\right]^\intercal\left[B^{-1}(\boldsymbol{x}-\boldsymbol{\mu})\right]\right\} \\
=& |B|^{-1/2}|B|^{-1/2}(2\pi)^{-p/2}\exp\left\{\frac{-1}{2}(\boldsymbol{x}-\boldsymbol{\mu})^\intercal\left(B^{-1}\right)^\intercal B^{-1}(\boldsymbol{x}-\boldsymbol{\mu})\right\} \\
=& |B|^{-1/2}|B^\intercal|^{-1/2}(2\pi)^{-p/2}\exp\left\{\frac{-1}{2}(\boldsymbol{x}-\boldsymbol{\mu})^\intercal(B^\intercal)^{-1}B^{-1}(\boldsymbol{x}-\boldsymbol{\mu})\right\} \\
=& |BB^\intercal|^{-1/2}(2\pi)^{-p/2}\exp\left\{\frac{-1}{2}(\boldsymbol{x}-\boldsymbol{\mu})^\intercal(BB^\intercal)^{-1}(\boldsymbol{x}-\boldsymbol{\mu})\right\} \\
\overset{56.11}{=}& |\Sigma|^{-1/2}(2\pi)^{-p/2}\exp\left\{\frac{-1}{2}(\boldsymbol{x}-\boldsymbol{\mu})^\intercal\Sigma^{-1}(\boldsymbol{x}-\boldsymbol{\mu})\right\} \\
=& \left(|\Sigma|(2\pi)^p\right)^{-1/2}\exp\left\{\frac{-1}{2}(\boldsymbol{x}-\boldsymbol{\mu})^\intercal\Sigma^{-1}(\boldsymbol{x}-\boldsymbol{\mu})\right\}
\end{aligned}$$

Definition 56.3. probability density function (PDF) of multivariate normal distribution (= multivariate Gaussian distribution)

$$\mathcal{N}(\boldsymbol{x}\mid\boldsymbol{\mu},\Sigma) = f_{\boldsymbol{X}}(\boldsymbol{x}) = \left(|\Sigma|(2\pi)^p\right)^{-1/2}\exp\left\{\frac{-1}{2}(\boldsymbol{x}-\boldsymbol{\mu})^\intercal\Sigma^{-1}(\boldsymbol{x}-\boldsymbol{\mu})\right\}$$

Definition 56.4. correlation coefficient

$$\rho_{ij} = \frac{\sigma_{ij}}{\sqrt{\sigma_{ii}}\sqrt{\sigma_{jj}}} = \frac{\sigma_{ij}}{\sqrt{\sigma_i^2}\sqrt{\sigma_j^2}} = \frac{\sigma_{ij}}{\sigma_i \sigma_j} = \frac{\mathrm{V}(X_i, X_j)}{\sqrt{\mathrm{V}(X_i)}\sqrt{\mathrm{V}(X_j)}} = \mathrm{R}(X_i, X_j)$$

56.1 bivariate normal distribution

$p = 2$ is the case of bivariate normal distribution

$$\mathit{\Sigma} = [\sigma_{ij}]_{2\times 2} = \begin{bmatrix} \sigma_{11} & \sigma_{12} \\ \sigma_{21} & \sigma_{22} \end{bmatrix} = \begin{bmatrix} \sigma_1^2 & \sigma_{12} \\ \sigma_{21} & \sigma_2^2 \end{bmatrix} = \begin{bmatrix} \sigma_1^2 & \sigma_1\sigma_2\rho_{12} \\ \sigma_2\sigma_1\rho_{21} & \sigma_2^2 \end{bmatrix} = \begin{bmatrix} \sigma_1^2 & \sigma_1\sigma_2\rho \\ \sigma_2\sigma_1\rho & \sigma_2^2 \end{bmatrix}$$

$$\rho_{12} = \rho = \rho_{21}$$

$$|\mathit{\Sigma}| = \begin{vmatrix} \sigma_1^2 & \sigma_1\sigma_2\rho_{12} \\ \sigma_2\sigma_1\rho_{21} & \sigma_2^2 \end{vmatrix} = \sigma_1^2\sigma_2^2(1-\rho_{12}\rho_{21}) = \sigma_1^2\sigma_2^2\left(1-\rho^2\right)$$

$$\begin{bmatrix} a & b \\ c & d \end{bmatrix}^{-1} = \frac{1}{\begin{vmatrix} a & b \\ c & d \end{vmatrix}} \begin{bmatrix} d & -b \\ -c & a \end{bmatrix}$$

$$\begin{aligned} \mathit{\Sigma}^{-1} =& \frac{1}{|\mathit{\Sigma}|} \begin{bmatrix} \sigma_2^2 & -\sigma_1\sigma_2\rho \\ -\sigma_2\sigma_1\rho & \sigma_1^2 \end{bmatrix} \\ =& \frac{1}{\sigma_1^2\sigma_2^2(1-\rho^2)} \begin{bmatrix} \sigma_2^2 & -\sigma_1\sigma_2\rho \\ -\sigma_2\sigma_1\rho & \sigma_1^2 \end{bmatrix} \\ =& \frac{1}{(1-\rho^2)} \begin{bmatrix} \frac{1}{\sigma_1^2} & \frac{-\rho}{\sigma_1\sigma_2} \\ \frac{-\rho}{\sigma_2\sigma_1} & \frac{1}{\sigma_2^2} \end{bmatrix} \end{aligned}$$

$$\begin{aligned} & \mathcal{N}\left(\boldsymbol{x} = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \middle| \boldsymbol{\mu} = \begin{bmatrix} \mu_1 \\ \mu_2 \end{bmatrix}, \mathit{\Sigma} = \begin{bmatrix} \sigma_1^2 & \sigma_1\sigma_2\rho \\ \sigma_2\sigma_1\rho & \sigma_2^2 \end{bmatrix}\right) \\ =& \left(|\mathit{\Sigma}|\,(2\pi)^{p=2}\right)^{-1/2} \exp\left\{\frac{-1}{2}(\boldsymbol{x}-\boldsymbol{\mu})^\mathsf{T}\mathit{\Sigma}^{-1}(\boldsymbol{x}-\boldsymbol{\mu})\right\} \\ =& \left(\sigma_1^2\sigma_2^2\left(1-\rho^2\right)(2\pi)^2\right)^{-1/2} \exp\left\{\frac{-1}{2}\begin{bmatrix} x_1-\mu_1 \\ x_2-\mu_2 \end{bmatrix}^\mathsf{T} \mathit{\Sigma}^{-1} \begin{bmatrix} x_1-\mu_1 \\ x_2-\mu_2 \end{bmatrix}\right\} \\ =& \frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-\rho^2}} \exp\left\{\frac{-1}{2}\begin{bmatrix} x_1-\mu_1 \\ x_2-\mu_2 \end{bmatrix}^\mathsf{T} \frac{1}{(1-\rho^2)} \begin{bmatrix} \frac{1}{\sigma_1^2} & \frac{-\rho}{\sigma_1\sigma_2} \\ \frac{-\rho}{\sigma_2\sigma_1} & \frac{1}{\sigma_2^2} \end{bmatrix} \begin{bmatrix} x_1-\mu_1 \\ x_2-\mu_2 \end{bmatrix}\right\} \\ =& \frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-\rho^2}} \exp\left\{\frac{-1}{2(1-\rho^2)}\begin{bmatrix} x_1-\mu_1 \\ x_2-\mu_2 \end{bmatrix}^\mathsf{T} \begin{bmatrix} \frac{1}{\sigma_1^2} & \frac{-\rho}{\sigma_1\sigma_2} \\ \frac{-\rho}{\sigma_2\sigma_1} & \frac{1}{\sigma_2^2} \end{bmatrix} \begin{bmatrix} x_1-\mu_1 \\ x_2-\mu_2 \end{bmatrix}\right\} \\ =& \frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-\rho^2}} \exp\left\{\frac{-1}{2(1-\rho^2)}\left[\left(\frac{x_1-\mu_1}{\sigma_1}\right)^2 - 2\rho\left(\frac{x_1-\mu_1}{\sigma_1}\right)\left(\frac{x_2-\mu_2}{\sigma_2}\right) + \left(\frac{x_2-\mu_2}{\sigma_2}\right)^2\right]\right\} \end{aligned}$$

Definition 56.5. probability density function (PDF) of bivariate normal distribution (= bivariate Gaussian distribution)

$$\begin{aligned} & \mathcal{N}\left(\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \middle| \begin{bmatrix} \mu_1 \\ \mu_2 \end{bmatrix}, \begin{bmatrix} \sigma_1^2 & \sigma_1\sigma_2\rho \\ \sigma_2\sigma_1\rho & \sigma_2^2 \end{bmatrix}\right) \\ =& \frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-\rho^2}} \exp\left\{\frac{-1}{2(1-\rho^2)}\left[\left(\frac{x_1-\mu_1}{\sigma_1}\right)^2 - 2\rho\left(\frac{x_1-\mu_1}{\sigma_1}\right)\left(\frac{x_2-\mu_2}{\sigma_2}\right) + \left(\frac{x_2-\mu_2}{\sigma_2}\right)^2\right]\right\} \end{aligned}$$

Chapter 57

logic

57.1 Shai Ben-David

https://www.youtube.com/playlist?list=PLPW2keNyw-utXOOzLR-Wp1p0eE5LEtv3N

Chapter 58

algorithm

58.1 Chen, Vivian

https://www.youtube.com/playlist?list=PLTpF-A8hKVUMcBaEuGyWbBXBn21seO3vA

58.2 The Bubble Sort Curve

https://www.youtube.com/watch?v=Gm8v_MR7TGk

https://linesthatconnect.github.io/blog/a-rigorous-derivation-of-the-bubble-sort-curve/

Chapter 59

computer graphics

59.1 Cem Yuksel

https://www.youtube.com/playlist?list=PLplnkTzzqsZTfYh4UbhLGpI5kGd5oW_Hh

https://www.youtube.com/playlist?list=PLplnkTzzqsZS3R5DjmCQsqupu43oS9CFN

Chapter 60

autoregression in time series

張翔老師. 2015. "ARMA Part1." https://www.youtube.com/watch?v=G-0dR57W-fo.

張翔老師. 2015. "ARMA Part2." https://www.youtube.com/watch?v=fQaZzO7E6FE.

張翔老師. 2015. "ARMA Part3." https://www.youtube.com/watch?v=Ocw4NXoO8Xo.

time series [data]

$$\cdots, Y_{t-2}, Y_{t-1}, Y_t, Y_{t+1}, Y_{t+2}, \cdots$$

- lag
  - 1^{st} lag = lag 1

    $$Y_{t-1}$$

  - k^{th} lag = lag k

    $$Y_{t-k}$$

1^{st} difference

$$\Delta Y_t = Y_t - Y_{t-1}$$

approximation for RoR = rate of return

$$\begin{aligned}\Delta \ln Y_t = \ln Y_t - \ln Y_{t-1} =& \ln \frac{Y_t}{Y_{t-1}} = \ln\left(1 + \frac{Y_t - Y_{t-1}}{Y_{t-1}}\right) \\ =& \frac{Y_t - Y_{t-1}}{Y_{t-1}} + \mathrm{O}\left(\left(\frac{Y_t - Y_{t-1}}{Y_{t-1}}\right)^2\right) \\ \approx& \frac{Y_t - Y_{t-1}}{Y_{t-1}} = \mathrm{RoR}\end{aligned}$$

$$\begin{cases}\ln(1+x) = \sum\limits_{n=1}^{\infty} (-1)^{n+1} \dfrac{x^n}{n} = x - \frac{x^2}{2} + \frac{x^3}{3} - \cdots \\ \ln(1-x) = -\sum\limits_{n=1}^{\infty} \dfrac{x^n}{n} = -x - \frac{x^2}{2} - \frac{x^3}{3} - \cdots\end{cases}$$

$$\begin{aligned}\frac{1}{1+t} =& 1 - t + t^2 - t^3 + \cdots \\ \ln(1+x) =& \int_0^x \frac{1}{1+t} dt \\ =& x - \frac{x^2}{2} + \frac{x^3}{3} - \frac{x^4}{4} + \cdots\end{aligned}$$

Definition 60.1. autocorrelation = serial correlation

$$\exists t_1 \neq t_2 \left[\mathrm{V}\left(Y_{t_1}, Y_{t_2}\right) \neq 0\right]$$

$$\begin{cases} \mathrm{E}\left(Y_t\right) \triangleq \mu_t \\ \mathrm{V}\left(Y_t\right) \triangleq \sigma_t^2 \end{cases}$$

Definition 60.2. k^{th} order autocovariance

$$\mathrm{V}\left(Y_t, Y_{t-k}\right) \triangleq \gamma\left(t, k\right)$$

$$\sigma_t^2 \triangleq \mathrm{V}\left(Y_t\right) = \mathrm{V}\left(Y_t, Y_t\right) = \mathrm{V}\left(Y_t, Y_{t-0}\right) = \gamma\left(t, k=0\right) = \gamma\left(t, 0\right)$$

Definition 60.3. k^{th} order autocorrelation

$$\rho_{t,t-k} = \frac{\sigma_{t,t-k}}{\sqrt{\sigma_{tt}}\sqrt{\sigma_{t-k,t-k}}} = \frac{\sigma_{t,t-k}}{\sqrt{\sigma_t^2}\sqrt{\sigma_{t-k}^2}} = \frac{\sigma_{t,t-k}}{\sigma_t \sigma_{t-k}} = \frac{\mathrm{V}\left(Y_t, Y_{t-k}\right)}{\sqrt{\mathrm{V}\left(Y_t\right)}\sqrt{\mathrm{V}\left(Y_{t-k}\right)}} \triangleq \mathrm{R}\left(Y_t, Y_{t-k}\right) \triangleq \rho\left(t, k\right)$$

Definition 60.4. stationary time series

$$\begin{cases} \mathrm{E}\left(Y_t\right) \triangleq \mu_t = \mu & (1)\,\text{independent of } t \\ \mathrm{V}\left(Y_t\right) \triangleq \sigma_t^2 = \sigma^2 < \infty & (2)\,\text{independent of } t \\ \mathrm{V}\left(Y_t, Y_{t-k}\right) \triangleq \gamma\left(t, k\right) = \gamma\left(k\right) & (3)\,\text{independent of } t \end{cases}$$

properties

$$\begin{cases} \mathrm{V}\left(Y_t\right) \triangleq \sigma_t^2 = \gamma\left(t, k=0\right) = \gamma\left(k=0\right) \overset{(3)}{=} \gamma\left(0\right) \triangleq \gamma_0 \overset{(2)}{=} \sigma^2 & (4) \Rightarrow \mathrm{V}\left(Y_{t-k}\right) = \gamma\left(0\right) \\ \mathrm{R}\left(Y_t, Y_{t-k}\right) \triangleq \rho\left(t, k\right) \triangleq \dfrac{\mathrm{V}\left(Y_t, Y_{t-k}\right)}{\sqrt{\mathrm{V}\left(Y_t\right)}\sqrt{\mathrm{V}\left(Y_{t-k}\right)}} & \\ = \dfrac{\gamma\left(t, k\right)}{\sqrt{\gamma\left(0\right)}\sqrt{\gamma\left(0\right)}} \underset{(4)}{\overset{(3)}{=}} \dfrac{\gamma\left(k\right)}{\gamma\left(0\right)} \triangleq \dfrac{\gamma_k}{\gamma_0} \triangleq \rho\left(k\right) \triangleq \rho_k & (5) \\ \gamma\left(k\right) = \gamma\left(-k\right) & (6) \Rightarrow \rho\left(k\right) = \rho\left(-k\right) \end{cases}$$

$$\gamma\left(k\right) \overset{(3)}{=} \mathrm{V}\left(Y_t, Y_{t-k}\right) = \mathrm{V}\left(Y_{t-k}, Y_t\right) = \mathrm{V}\left(Y_{t'}, Y_{t'+k}\right) = \mathrm{V}\left(Y_{t'}, Y_{t'-(-k)}\right) \overset{(3)}{=} \gamma\left(-k\right) \Rightarrow (6)$$

point estimation

$$\begin{cases} \widehat{\mathrm{E}}\left(Y_t\right) \triangleq \overline{Y} \triangleq \widehat{\mu} = \dfrac{1}{T}\sum_{t=1}^{T} Y_t & \to \mathrm{E}\left(Y_t\right) = \mu \\ \widehat{\mathrm{V}}\left(Y_t\right) \triangleq \widehat{\gamma}_0 = \dfrac{1}{T}\sum_{t=1}^{T}\left(Y_t - \overline{Y}\right)^2 & \to \mathrm{V}\left(Y_t\right) = \sigma^2 = \gamma_0 \\ \widehat{\mathrm{V}}\left(Y_t, Y_{t-k}\right) \triangleq \widehat{\gamma}_k & \\ = \dfrac{1}{T}\sum_{t=k+1}^{T}\left(Y_t - \overline{Y}\right)\left(Y_{t-k} - \overline{Y}\right) = \dfrac{1}{T}\sum_{t=1}^{T-k}\left(Y_t - \overline{Y}\right)\left(Y_{t+k} - \overline{Y}\right) & \to \mathrm{V}\left(Y_t, Y_{t-k}\right) = \gamma_k \\ \dfrac{\widehat{\mathrm{V}}\left(Y_t, Y_{t-k}\right)}{\sqrt{\widehat{\mathrm{V}}\left(Y_t\right)}\sqrt{\widehat{\mathrm{V}}\left(Y_{t-k}\right)}} \triangleq \widehat{\rho}_k = \dfrac{\widehat{\gamma}_k}{\widehat{\gamma}_0} = \dfrac{\dfrac{1}{T}\sum_{t=k+1}^{T}\left(Y_t - \overline{Y}\right)\left(Y_{t-k} - \overline{Y}\right)}{\dfrac{1}{T}\sum_{t=1}^{T}\left(Y_t - \overline{Y}\right)^2} & \to \mathrm{R}\left(Y_t, Y_{t-k}\right) = \rho_k \end{cases}$$

$$Y_1 Y_{1+k} + Y_2 Y_{2+k} + \cdots + Y_t Y_{t+k} + \cdots + Y_{T-k} Y_T$$

1^{st}-order autocorrelation estimation

$$\widehat{\rho}_1 = \frac{\widehat{\gamma}_1}{\widehat{\gamma}_0} = \frac{\dfrac{1}{T}\sum_{t=1+1=2}^{T}\left(Y_t - \overline{Y}\right)\left(Y_{t-1} - \overline{Y}\right)}{\dfrac{1}{T}\sum_{t=1}^{T}\left(Y_t - \overline{Y}\right)^2}$$

60.1 AR(1) = 1^{st}-order autoregressive model = first-order autoregressive model

張翔老師. 2015. "ARMA Part3." https://www.youtube.com/watch?v=Ocw4NXoO8Xo.

Definition 60.5. AR(1) = 1^{st}-order autoregressive model$

$$Y_t = \beta_0 + \beta_1 Y_{t-1} + \mathcal{E}_t$$

$$\mathcal{E}_t \sim \mathcal{N}\left(0, \sigma_{\mathcal{E}}^2\right)$$

$$\begin{aligned}
Y_t =& \beta_0 + \beta_1 Y_{t-1} + \mathcal{E}_t \\
=& \beta_0 + \beta_1 \left(\beta_0 + \beta_1 Y_{t-2} + \mathcal{E}_{t-1}\right) + \mathcal{E}_t \\
=& \beta_0 \left(1 + \beta_1\right) + \beta_1^2 Y_{t-2} + \mathcal{E}_t + \beta_1 \mathcal{E}_{t-1} \\
=& \beta_0 \left(1 + \beta_1\right) + \beta_1^2 \left(\beta_0 + \beta_1 Y_{t-3} + \mathcal{E}_{t-2}\right) + \mathcal{E}_t + \beta_1 \mathcal{E}_{t-1} \\
=& \beta_0 \left(1 + \beta_1 + \beta_1^2\right) + \beta_1^3 Y_{t-3} + \mathcal{E}_t + \beta_1 \mathcal{E}_{t-1} + \beta_1^2 \mathcal{E}_{t-2} \\
&\vdots \\
=& \beta_0 \left(1 + \beta_1 + \beta_1^2 + \cdots\right) + \mathcal{E}_t + \beta_1 \mathcal{E}_{t-1} + \beta_1^2 \mathcal{E}_{t-2} + \cdots \\
\overset{|\beta_1|<1}{=}& \frac{\beta_0}{1-\beta_1} + \sum_{k=0}^{\infty} \beta_1^k \mathcal{E}_{t-k}
\end{aligned}$$

$$\begin{cases}
\mu = \mathrm{E}\left(Y_t\right) = \mathrm{E}\left(\frac{\beta_0}{1-\beta_1} + \sum_{i=0}^{\infty} \beta_1^i \mathcal{E}_{t-i}\right) & \\
= \frac{\beta_0}{1-\beta_1} + \sum_{i=0}^{\infty} \beta_1^i \mathrm{E}\left(\mathcal{E}_{t-i}\right) \overset{\mathcal{E}_t \sim \mathcal{N}\left(0,\sigma_{\mathcal{E}}^2\right)}{=} \frac{\beta_0}{1-\beta_1} & \mu = \frac{\beta_0}{1-\beta_1} \\
\gamma_0 = \sigma^2 = \mathrm{V}\left(Y_t\right) = \mathrm{V}\left(\mu + \sum_{i=0}^{\infty} \beta_1^i \mathcal{E}_{t-i}\right) = \mathrm{V}\left(\sum_{i=0}^{\infty} \beta_1^i \mathcal{E}_{t-i}\right) & \\
= \sum_{i=0}^{\infty} \left(\beta_1^i\right)^2 \mathrm{V}\left(\mathcal{E}_{t-i}\right) \overset{\mathcal{E}_t \sim \mathcal{N}\left(0,\sigma_{\mathcal{E}}^2\right)}{=} \sum_{i=0}^{\infty} \left(\beta_1^2\right)^i \sigma_{\mathcal{E}}^2 = \frac{\sigma_{\mathcal{E}}^2}{1-\beta_1^2} & \gamma_0 = \sigma^2 = \frac{\sigma_{\mathcal{E}}^2}{1-\beta_1^2} \\
\gamma_k = \mathrm{V}\left(Y_t, Y_{t-k}\right) = \mathrm{V}\left(\mu + \sum_{i=0}^{\infty} \beta_1^i \mathcal{E}_{t-i}, \mu + \sum_{i=0}^{\infty} \beta_1^i \mathcal{E}_{t-k-i}\right) & \\
= \mathrm{V}\left(\sum_{i=0}^{\infty} \beta_1^i \mathcal{E}_{t-i}, \sum_{i=0}^{\infty} \beta_1^i \mathcal{E}_{t-k-i}\right) = \cdots & \gamma_k = \frac{\beta_1^k \sigma_{\mathcal{E}}^2}{1-\beta_1^2} \\
\rho_k = \mathrm{R}\left(Y_t, Y_{t-k}\right) = \frac{\gamma_k}{\gamma_0} = \frac{\frac{\beta_1^k \sigma_{\mathcal{E}}^2}{1-\beta_1^2}}{\frac{\sigma_{\mathcal{E}}^2}{1-\beta_1^2}} = \beta_1^k & \rho_k = \frac{\gamma_k}{\gamma_0} = \beta_1^k
\end{cases}$$

$$\rho_1 = \frac{\gamma_1}{\gamma_0} = \beta_1^1 = \beta_1$$

$$\begin{aligned}
\gamma_k = \mathrm{V}\left(Y_t, Y_{t-k}\right) =& \mathrm{V}\left(\mu + \sum_{i=0}^{\infty} \beta_1^i \mathcal{E}_{t-i}, \mu + \sum_{i=0}^{\infty} \beta_1^i \mathcal{E}_{t-k-i}\right) \\
=& \mathrm{V}\left(\sum_{i=0}^{\infty} \beta_1^i \mathcal{E}_{t-i}, \sum_{i=0}^{\infty} \beta_1^i \mathcal{E}_{t-k-i}\right) = \cdots
\end{aligned}$$

$$\boldsymbol{Y} = \langle Y_1, Y_2, \cdots, Y_T \rangle = \begin{bmatrix} Y_1 & Y_2 & \cdots & Y_T \end{bmatrix}^{\mathrm{T}} = \begin{bmatrix} Y_1 \\ Y_2 \\ \vdots \\ Y_T \end{bmatrix}$$

AR(1) covariance matrix

by (56.3)

$$\begin{aligned}
\mathrm{V}(\boldsymbol{Y}) =& \mathrm{E}\left[[\boldsymbol{Y} - \mathrm{E}(\boldsymbol{Y})][\boldsymbol{Y} - \mathrm{E}(\boldsymbol{Y})]^{\mathrm{T}}\right] \\
=& \begin{bmatrix} \sigma_1^2 & \sigma_{12} & \cdots & \sigma_{1T} \\ \sigma_{21} & \sigma_2^2 & \cdots & \sigma_{2T} \\ \vdots & \vdots & \ddots & \vdots \\ \sigma_{T1} & \sigma_{T2} & \cdots & \sigma_T^2 \end{bmatrix} = \begin{bmatrix} \sigma_{11} & \sigma_{12} & \cdots & \sigma_{1T} \\ \sigma_{21} & \sigma_{22} & \cdots & \sigma_{2T} \\ \vdots & \vdots & \ddots & \vdots \\ \sigma_{T1} & \sigma_{T2} & \cdots & \sigma_{TT} \end{bmatrix} = [\sigma_{ij}]_{T\times T} = \varSigma \\
=& \begin{bmatrix} \gamma_0 & \gamma_1 & \cdots & \gamma_{T-1} \\ \gamma_1 & \gamma_0 & \cdots & \gamma_{T-2} \\ \vdots & \vdots & \ddots & \vdots \\ \gamma_{T-1} & \gamma_{T-2} & \cdots & \gamma_0 \end{bmatrix} = \begin{bmatrix} \frac{\sigma_{\mathcal{E}}^2}{1-\beta_1^2} & \frac{\beta_1^1\sigma_{\mathcal{E}}^2}{1-\beta_1^2} & \cdots & \frac{\beta_1^{T-1}\sigma_{\mathcal{E}}^2}{1-\beta_1^2} \\ \frac{\beta_1^1\sigma_{\mathcal{E}}^2}{1-\beta_1^2} & \frac{\sigma_{\mathcal{E}}^2}{1-\beta_1^2} & \cdots & \frac{\beta_1^{T-2}\sigma_{\mathcal{E}}^2}{1-\beta_1^2} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\beta_1^{T-1}\sigma_{\mathcal{E}}^2}{1-\beta_1^2} & \frac{\beta_1^{T-2}\sigma_{\mathcal{E}}^2}{1-\beta_1^2} & \cdots & \frac{\sigma_{\mathcal{E}}^2}{1-\beta_1^2} \end{bmatrix} \\
=& \frac{\sigma_{\mathcal{E}}^2}{1-\beta_1^2} \begin{bmatrix} 1 & \beta_1 & \cdots & \beta_1^{T-1} \\ \beta_1 & 1 & \cdots & \beta_1^{T-2} \\ \vdots & \vdots & \ddots & \vdots \\ \beta_1^{T-1} & \beta_1^{T-2} & \cdots & 1 \end{bmatrix} = \sigma^2 \begin{bmatrix} 1 & \beta_1 & \beta_1^2 & \cdots & \beta_1^{T-1} \\ \beta_1 & 1 & \beta_1 & \cdots & \beta_1^{T-2} \\ \beta_1^2 & \beta_1 & 1 & \cdots & \beta_1^{T-3} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \beta_1^{T-1} & \beta_1^{T-2} & \beta_1^{T-3} & \cdots & 1 \end{bmatrix} \\
=& \sigma^2 \begin{bmatrix} 1 & \rho_1 & \rho_2 & \cdots & \rho_{T-1} \\ \rho_1 & 1 & \rho_1 & \cdots & \rho_{T-2} \\ \rho_2 & \rho_1 & 1 & \cdots & \rho_{T-3} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \rho_{T-1} & \rho_{T-2} & \rho_{T-3} & \cdots & 1 \end{bmatrix} = \sigma^2 \begin{bmatrix} 1 & \rho & \rho^2 & \cdots & \rho^{T-1} \\ \rho & 1 & \rho & \cdots & \rho^{T-2} \\ \rho^2 & \rho & 1 & \cdots & \rho^{T-3} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \rho^{T-1} & \rho^{T-2} & \rho^{T-3} & \cdots & 1 \end{bmatrix}
\end{aligned}$$

$$\boldsymbol{Y} \sim \mathcal{D}(\boldsymbol{\mu}, \varSigma) = \mathrm{d}(\boldsymbol{\mu}_{\boldsymbol{Y}}, \varSigma_{\boldsymbol{Y}}) = \mathrm{d}(\mathrm{E}[\boldsymbol{Y}], \mathrm{V}[\boldsymbol{Y}])$$

where

$$\varSigma = \begin{bmatrix} \gamma_0 & \gamma_1 & \cdots & \gamma_{T-1} \\ \gamma_1 & \gamma_0 & \cdots & \gamma_{T-2} \\ \vdots & \vdots & \ddots & \vdots \\ \gamma_{T-1} & \gamma_{T-2} & \cdots & \gamma_0 \end{bmatrix} = \sigma^2 \begin{bmatrix} 1 & \rho & \rho^2 & \cdots & \rho^{T-1} \\ \rho & 1 & \rho & \cdots & \rho^{T-2} \\ \rho^2 & \rho & 1 & \cdots & \rho^{T-3} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \rho^{T-1} & \rho^{T-2} & \rho^{T-3} & \cdots & 1 \end{bmatrix}$$

60.2 AR(2) = 2^{nd}-order autoregressive model = second-order autoregressive model

Definition 60.6. AR(2) = 2^{nd}-order autoregressive model

$$Y_t = \beta_0 + \beta_1 Y_{t-1} + \beta_2 Y_{t-2} + \mathcal{E}_t$$

$$\mathcal{E}_t \sim \mathcal{N}\left(0, \sigma_{\mathcal{E}}^2\right)$$

60.3 AR(p) = p^{th}-order autoregressive model

$$Y_t = \beta_0 + \sum_{i=1}^{p} \beta_i Y_{t-i} + \mathcal{E}_t$$

60.4 MA(q) = q^{th}-order moving-average model

$$Y_t = \mu + \mathcal{E}_t + \sum_{i=1}^{q} \alpha_i \mathcal{E}_{t-i}$$

60.5 ARMA(p,q) = $p^{\mathbf{th}}, q^{\mathbf{th}}$-order autoregressive-moving-average model

$$Y_t = \mathcal{E}_t + \sum_{i=1}^{q} \alpha_i \mathcal{E}_{t-i} + \sum_{i=1}^{p} \beta_i Y_{t-i}$$

Chapter 61

conductivity

61.1 resistance

https://www.bilibili.com/video/BV1tN41147SK/?spm_id_from=333.999.section.playall&vd_source=2b462dd516a0bbedcb7d97d0cfc60b38

61.2 semiconductor

https://www.youtube.com/watch?v=-lHXZk5M6cI

Chapter 62

LaTeX annotation by TikZ

```r
knitr::opts_chunk$set(fig.pos = "H", out.extra = "")
```

$$
\boxed{
\begin{aligned}
f\left(x\right) =& x^{x} \\
\ln f\left(x\right) =& x \ln x \\
\frac{\mathrm{d}}{\mathrm{d}x} \ln f\left(x\right) =& \frac{\mathrm{d}}{\mathrm{d}x}\left[x \ln x\right] \\
\frac{f'\left(x\right)}{f\left(x\right)} =& \left[x \ln x\right]' \\
f'\left(x\right) =& f\left(x\right)\left[x \ln x\right]' = x^{x}\left[\left(x\right)' \ln x + x\left(\ln x\right)'\right] \\
=& x^{x}\left[1 \ln x + x\frac{1}{x}\right] = x^{x}\left[\ln x + 1\right] = x^{x}\left[1 + \ln x\right] \\
=& x^{x} + x^{x} \ln x
\end{aligned}
}
$$

Figure 62.1: test

```
\begin{tikzpicture}
  \begin{axis}[
        %axis x line = center,
        %axis y line = center,
        %xlabel = {$ x $}, xlabel style =
        ↪  {right},
        %ylabel = {$ y=f\left(x\right) $},
        ↪  ylabel style = {above},
        xmin = -1, xmax = 1,
        ymin = -0.5, ymax = 0.5,
        %hide obscured x ticks=false, % for
        ↪  origin x tick label i.e. xtick =
        ↪  {0}
        xtick= \empty,
        xticklabels= \empty,
        %extra x ticks={0},
        ytick = \empty,
        yticklabels= \empty,
        x = 5cm, y = 5cm, % x y scaling
        %grid = major,
        domain = -1:1,
        %samples = 1000
    ]
    \node at (axis cs: 0,0) {
      $\begin{aligned}
      f\left(x\right)= & x^{x}\\
      \ln f\left(x\right)= & x\ln x\\
      \dfrac{\mathrm{d}}{\mathrm{d}x}\ln
↪   f\left(x\right)= &
↪   \dfrac{\mathrm{d}}{\mathrm{d}x}\left[x\ln
↪   x\right]\\

↪   \dfrac{f^{\prime}\left(x\right)}{f\left(x\right)}=
↪   & \left[x\ln x\right]^{\prime}\\
      f^{\prime}\left(x\right)= &
↪   f\left(x\right)\left[x\ln
↪   x\right]^{\prime}=x^{x}\left[\left(x\right)^{\prime}\ln
↪   x+x\left(\ln x\right)^{\prime}\right]\\
      = & x^{x}\left[1\ln
↪   x+x\dfrac{1}{x}\right]=x^{x}\left[\ln
↪   x+1\right]=x^{x}\left[1+\ln x\right]\\
      = & x^{x}+x^{x}\ln x
      \end{aligned}$
      };
  \end{axis}
\end{tikzpicture}
```

$$\begin{aligned}
f\left(x\right)= & x^{x}\\
\ln f\left(x\right)= & x\ln x\\
\dfrac{\mathrm{d}}{\mathrm{d}x}\ln f\left(x\right)= & \dfrac{\mathrm{d}}{\mathrm{d}x}\left[x\ln x\right]\\
\dfrac{f^{\prime}\left(x\right)}{f\left(x\right)}= & \left[x\ln x\right]^{\prime}\\
f^{\prime}\left(x\right)= & f\left(x\right)\left[x\ln x\right]^{\prime}=x^{x}\left[\left(x\right)^{\prime}\ln x+x\left(\ln x\right)^{\prime}\right]\\
= & x^{x}\left[1\ln x+x\dfrac{1}{x}\right]=x^{x}\left[\ln x+1\right]=x^{x}\left[1+\ln x\right]\\
= & x^{x}+x^{x}\ln x
\end{aligned}$$

Figure 62.2: test

https://tex.stackexchange.com/questions/670153/how-to-annotate-calculations

WRONG!

This is important!

$$\frac{4}{2} + \frac{6}{3} = \frac{(3\times4)+\cancel{(2\times6)}}{\cancel{2\times3}} = \boxed{\frac{12+12}{6}} = \frac{24}{6} = 4 \quad \text{RESULT}$$

This multiplication is not necessary Try to simplify.

Figure 62.3: test

```
\begin{tikzpicture}
  \begin{axis}[
        %axis x line = center,
        %axis y line = center,
        %xlabel = {$ x $}, xlabel style = {right},
        %ylabel = {$ y=f\left(x\right) $}, ylabel style = {above},
        xmin = -0.7, xmax = 0.9,
        ymin = -0.3, ymax = 0.3,
        %hide obscured x ticks=false, % for origin x tick label i.e. xtick = {0}
        xtick= \empty,
        xticklabels= \empty,
        %extra x ticks={0},
        ytick = \empty,
        yticklabels= \empty,
        x = 5cm, y = 5cm, % x y scaling
        %grid = major,
        domain = -1:1,
        %samples = 1000
    ]
    \node at (axis cs:0,0) {
      $\tikzmarknode{a1}{\frac{4}{2}}+\tikzmarknode{a2}{\frac{6}{3}}=\frac{\tikzmarknode{b1}{(3\times 4)}+\tikzmarknode{b2}{(2\times 6)}}{\tikzmarknode{b3}{2\times 3}}=\tikzmarknode{c1}{\frac{12+12}{6}}=\tikzmarknode{c2}{\frac{24}{6}}=\tikzmarknode{d1}{4}$};

    \draw[red, thick, ->, shorten <=1em, , shorten >=1em](a2.south east)--node[below, pos=.45, scale=.4]{1}(a1.north west);
    \node[draw=red, thick, ellipse, fit=(b1), inner ysep=0, inner xsep=-1mm](ell){};
    \draw[thick, red] (ell)--++(0,.5)node[above, red, scale=.4]{This is important!};
    \draw[red, thick] (b2.south west)--node[above=2mm, red, scale=.5]{\bfseries WRONG!}(b2.north east);
    \draw[green!70!black, thick] (b3.east)--node[below]{$\uparrow$}node[below=5mm, align=left, green!70!black, scale=.5]{This multiplication is not\\necessary Try to simplify.}(b3.west) ([yshift=2pt]b3.east)--([yshift=2pt]b3.west);
    \node[draw=blue, thick, fit=(c1), inner xsep=1pt](box){};
    \draw[blue, thick, ->, shorten >=3pt] (box) to[out=90, in=90] (c2.north);
    \node[draw=yellow!70!orange, thick, circle, fit=(d1), inner sep=1pt](cir){};
    \node[draw=yellow!70!orange, thick, text=yellow!70!orange, single arrow, shape border rotate=180, anchor=west, scale=.4, single arrow tip angle=40, minimum height=30mm] at (cir.east){\quad RESULT};
  \end{axis}
\end{tikzpicture}
```

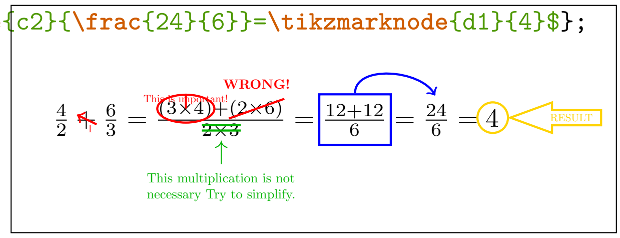


Figure 62.4: test

https://tex.stackexchange.com/questions/494884/anchor-alignment-in-tikzmarknode

$$\sum_{i=1}^{3}\sum_{j=1}^{4} a_{ij} = \boxed{\sum_{j=1}^{4} a_{1j}} + \boxed{\sum_{j=1}^{4} a_{2j}} + \boxed{\sum_{j=1}^{4} a_{3j}}$$

$$= \boxed{a_{11} + a_{12} + a_{13} + a_{14}} + \boxed{a_{21} + a_{22} + a_{23} + a_{24}} + \boxed{a_{31} + a_{32} + a_{33} + a_{34}}$$

Figure 62.5: test

```
\tikzset{every tikzmarknode/.append
↪  style={inner sep=3pt,rounded
corners}}

\begin{tikzpicture}
  \begin{axis}[
        %axis x line = center,
        %axis y line = center,
        %xlabel = {$ x $}, xlabel style =
        ↪  {right},
        %ylabel = {$ y=f\left(x\right) $},
        ↪  ylabel style = {above},
        xmin = -1.5, xmax = 1.5,
        ymin = -0.4, ymax = 0.4,
        %hide obscured x ticks=false, % for
        ↪  origin x tick label i.e. xtick =
        ↪  {0}
        xtick= \empty,
        xticklabels= \empty,
        %extra x ticks={0},
        ytick = \empty,
        yticklabels= \empty,
        x = 5cm, y = 5cm, % x y scaling
        %grid = major,
        domain = -1:1,
        %samples = 1000
    ]
    \node at (axis cs: 0,0) {
      $\begin{aligned}
  \sum\limits_{i=1}^{3}\sum\limits_{j=1}^{4}
↪  a_{ij}&=

↪  \tikzmarknode[fill=red!20]{red1}{\sum\limits_{j=1}^{4}a_{1j}}
↪  +

↪  \tikzmarknode[fill=green!20]{green1}{\sum\limits_{j=1}^{4}a_{2j}}
↪  +

↪  \tikzmarknode[fill=blue!20]{blue1}{\sum\limits_{j=1}^{4}a_{3j}}\\
  \\
  &=
  \tikzmarknode[fill=red!20]{red2}{a_{11} +
↪  a_{12} + a_{13} + a_{14}} +
  \tikzmarknode[fill=green!20]{green2}{a_{21}
↪  + a_{22} + a_{23} + a_{24}} +
  \tikzmarknode[fill=blue!20]{blue2}{a_{31} +
↪  a_{32} + a_{33} + a_{34}}
\end{aligned}$
      };
    \draw[->,red!20] (red1.south)
    ↪  to[out=-90,in=120,looseness=0.3]
    ↪  (red2.north);
    \draw[->,green!20] (green1)
    ↪  to[out=-90,in=120,looseness=0.3]
    ↪  (green2);
    \draw[->,blue!20] (blue1)
    ↪  to[out=-90,in=135,looseness=0.3]
    ↪  (blue2);
  \end{axis}
\end{tikzpicture}
```

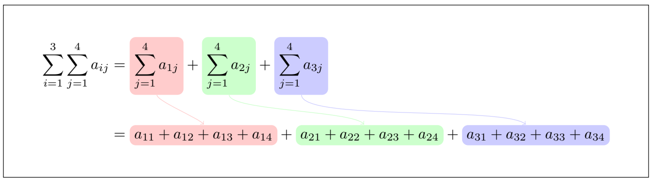


Figure 62.6: test

https://github.com/synercys/annotated_latex_equations

https://github.com/synercys/annotated_latex_equations/blob/main/example_prob.tex

$$
\boxed{
\begin{aligned}
f\left(x\right) =& x^{x} \\
\ln f\left(x\right) =& x\ln x \\
\frac{\mathrm{d}}{\mathrm{d}x}\ln f\left(x\right) =& \frac{\mathrm{d}}{\mathrm{d}x}\left[x\ln x\right] \\
\frac{f'\left(x\right)}{f\left(x\right)} =& \left[x\ln x\right]' \\
f'\left(x\right) =& f\left(x\right)\left[x\ln x\right]' = x^{x}\left[\left(x\right)'\ln x + x\left(\ln x\right)'\right] \\
=& x^{x}\left[1\ln x + x\frac{1}{x}\right] = x^{x}\left[\ln x + 1\right] = x^{x}\left[1+\ln x\right] \\
=& x^{x} + x^{x}\ln x
\end{aligned}
}
$$

Figure 62.7: test

Chapter 63

Shannon sampling

https://www.youtube.com/watch?v=ePGDQpJAvjE

Chapter 64

brachistochrone

https://www.youtube.com/watch?v=2c_bdVC9KS8

Chapter 65

magnetic resonance

65.1 NMR

65.2 MRI

65.2.1 Lin, Hsiu-Hau

Lin, Hsiu-Hau / Porcupine Lin / Hedgehog Note

https://www.youtube.com/watch?v=Y-z-1XCu7fE&list=PLS0SUwlYe8czNqxfQq2XWeAHDqT8vYjmC&index=60

Chapter 66

OpenMMLab

MMLab

66.1 Tongji TommyZihao

https://github.com/TommyZihao

https://space.bilibili.com/1900783/channel/series

66.2 official

https://github.com/open-mmlab

https://space.bilibili.com/1293512903/channel/series

references

1. Xie, Y. *Bookdown: Authoring Books and Technical Documents with r Markdown*. (2023).
2. Xie, Y. *Dynamic Documents with R and Knitr*. (Chapman; Hall/CRC, Boca Raton, Florida, 2015).
3. Bookdown books on the web: Downloading and converting to pdf - R Markdown. *Posit Community* (2019).
4. ccjou. 二次型與正定矩陣. (2009).
5. Campbell, D. T. & Stanley, J. *Experimental and Quasi-Experimental Designs for Research*. (Cengage Learning, Belomt, CA, 1963).
6. 張翔. & 廖崇智. 提綱挈領學統計. (大碩教育, 2021).
7. Mittelhammer, R. C. *Mathematical Statistics for Economics and Business*. (Springer, New York Heidelberg Dordrecht London, 2013).
8. ccjou. 共變異數矩陣的性質. (2014).
9. Hu, F.-C. *Statistical Design and Analysis of Biomedical Research*. (International-Harvard (I-H) Statistical Consulting Company, Taipei, Taiwan, 2022).
10. Milliken, G. A. & Johnson, D. E. *Analysis of Messy Data Volume 1: Designed Experiments, Second Edition*. (Chapman; Hall/CRC, Boca Raton, Fla., 2004).
11. Buchanan, C. D. *et al.* Sorghum bicolor's Transcriptome Response to Dehydration, High Salinity and ABA. *Plant Molecular Biology* **58**, 699–720 (2005).
12. Casler, M. D., Vermerris, W. & Dixon, R. A. Replication Concepts for Bioenergy Research Experiments. *BioEnergy Research* **8**, 1–16 (2015).
13. ccjou. 矩陣導數. (2013).
14. ccjou. 多變量常態分布. (2014).
15. ccjou. Jacobian 矩陣與行列式. (2012).