

math

Joey Yu Hsu

2024-05-02

Contents

index	13
I by discipline	15
1 mathematics	17
1.1 tool	17
1.2 discipline	17
2 physics	19
2.1 discipline	19
3 plot	21
4 programming language	23
4.1 discipline	23
4.2 learning map	23
5 machine learning	25
5.1 Shai Ben-David	25
5.2 deep learning	25
II by date	33
6 A Minimal Book Example	35
6.1 About	35
6.1.1 Usage	35
6.1.2 Render book	35
6.1.3 Preview book	35
6.2 Hello bookdown	35
6.2.1 A section	35
An unnumbered section	36
6.3 Cross-references	36
6.3.1 Chapters and sub-chapters	36
6.3.2 Captioned figures and tables	36
6.4 Parts	37
6.5 Footnotes and citations	37
6.5.1 Footnotes	37
6.5.2 Citations	37
6.6 Blocks	37
6.6.1 Equations	37
6.6.2 Theorems and proofs	38
6.6.3 Callout blocks	38
6.7 Sharing your book	38
6.7.1 Publishing	38
6.7.2 404 pages	38
6.7.3 Metadata for sharing	38
7 test	39
7.1 RStudio	39
7.1.1 writer options	39

7.1.1.1	line wrapping	39
7.1.1.2	ensuring the same markdown between source / visual mode	39
7.1.2	Rtools	39
7.1.3	addins	39
7.1.4	Git	40
7.2	RMarkdown	40
7.2.1	Pandoc link	40
7.2.2	URL	40
7.2.3	arrow	40
7.2.4	superscript and subscript	40
7.2.4.1	LaTeX	41
7.2.5	equation	41
7.2.5.1	proof QED	41
7.2.6	image	41
7.2.6.1	DiagrammeR / mermaid flowchart	41
7.2.6.2	multiple images / figures in the same line	41
7.2.6.3	figure size	43
7.2.6.3.1	knitr	45
7.2.6.3.2	out.width vs. fig.width	45
7.2.6.4	dynamic knitr plot width and height	47
7.2.6.5	web image in PDF	47
7.2.6.6	SVG	47
7.2.7	horizontal rule	49
7.2.8	footnote	49
7.2.9	hyperlink	49
7.2.10	code chunk	50
7.2.10.1	code folding	50
7.2.11	xaringan	50
7.3	Bookdown	50
7.3.1	system locale	50
7.3.2	render_book()	50
7.3.3	serve_book()	50
7.3.4	LaTeX	50
7.3.4.1	hyperlink, URL, href	50
7.3.4.2	ugly mathptmx \sum	50
7.3.4.3	LaTeX package in HTML document	51
7.3.5	depth of table of contents toc_depth	51
7.3.6	multi-column layout / two columns	51
7.3.6.1	for both HTML and PDF	51
7.3.6.2	multi-column fig.cap must use fig.pos="H"	53
7.3.6.3	caption above figure	53
7.3.6.4	for only HTML	54
7.3.6.4.1	CSS flex	54
7.3.6.4.2	CSS grid	54
7.4	conditional block/chunk for either HTML or PDF, and Chinese issue	55
7.5	video embedding	55
7.5.1	timestamp	56
7.6	equation term coloring	56
7.6.1	font color	56
7.6.2	background color	56
7.7	link and reference	57
7.8	number and reference equations	58
7.9	footnote	58
7.10	citation	58
7.10.1	citation in fig.cap	59
7.10.2	backreference	59
7.11	environment for definition, theorem, proof	59
7.12	slide or presentation	60
7.12.1	Xaringan and Infinite Moon Reader	60
7.12.2	ioslides	60
7.12.3	PowerPoint	61

8.1 verbatim	63
9 partition	65
10 equivalence class	67
11 equivalence relation	69
12 Python	71
12.1 using Python in R / RMarkdown	71
12.2 SoloLearn	72
12.3 list comprehension	72
12.4 functional programming	73
12.5 object-oriented programming = OOP	73
13 TikZ	75
13.1 2D	76
13.2 3D	83
13.3 plots of functions	85
13.4 PGFplots	86
13.4.1 the axis environments	86
13.4.2 axis descriptions	87
13.4.2.1 placement of axis descriptions	87
13.4.2.1.1 coordinate system <code>axis description cs</code>	87
13.4.2.1.2 legend	89
13.4.2.1.3 tick option	89
13.4.3 declare function	89
13.4.3.1 <code>pgfmathparse</code>	89
13.4.3.2 <code>pgfmathdeclarefunction</code>	90
13.4.4 - and - in TikZ	91
13.4.5 <code>pgfplotsinvokeforeach</code>	91
13.4.6 interpolation dashed lines	92
13.4.7 Zewbie	92
13.4.7.1 coordinate axis/axes fine-tuing	93
13.4.7.1.1 range	93
13.4.7.1.2 scaling	93
13.4.7.1.3 direction vector	94
13.4.7.1.4 axis style	95
13.4.7.1.5 axis discontinuity	96
13.4.7.1.6 tick	96
13.4.7.2 <code>addplot+</code>	98
13.4.7.3 point	99
13.4.7.4 line	101
13.4.7.5 plane	103
13.4.7.6 polar coordinate	104
13.4.8 Arnav Bandekar: Using pgfplots to make economic graphs in LaTeX	104
13.4.9 PGFplots gallery	104
13.5 TikZplotLib / tikzplotlib	105
13.6 animation	107
14 xy-pic	109
15 statistics	111
15.1 Hung Hung	111
15.1.1 probability theory	113
15.1.1.1 probability function	113
15.1.1.2 conditional probability	115
15.1.2 univariable distribution	115
15.1.2.1 cumulative distribution function	116
15.1.2.2 probability density function	116
15.1.2.3 range vs. support	117
15.1.2.4 continuous monotone transformation	118
15.1.2.5 discrete monotone transformation	121
15.1.2.6 expected value	122

15.1.2.7 moment	129
15.1.2.7.1 moment generating function	132
15.1.2.7.2 characteristic function	133
15.1.2.8 common families of distributions	134
15.1.2.8.1 discrete distribution	134
15.1.2.8.2 continuous distribution	139
15.1.2.9 exponential family	140
15.1.2.9.1 binomial distribution with known n	141
15.1.2.9.2 continuous uniform distribution not in exponential family	141
15.1.2.9.3 normal distribution is in exponential family	142
15.1.2.9.4 normal distribution with unknown equal mean and standard deviation in curved exponential family	142
15.1.2.9.5 properties of exponential family	143
15.1.3 multivariable distribution	149
15.2 Chen, Lin-An	153
16 covariance matrix	155
16.1 vector direct product	155
16.1.1 vector direct product: increasing rank	155
16.1.2 vector inner product: decreasing rank	155
16.1.3 tensor direct product: increasing rank	156
16.2 covariance matrix and its properties	156
16.2.1 $V[\mathbf{X} + \mathbf{b}] = V[\mathbf{X}]$	156
16.2.2 $V[A\mathbf{X}] = A V[\mathbf{X}] A^\top$	156
16.2.3 $V[A\mathbf{X} + \mathbf{b}] = A V[\mathbf{X}] A^\top$	156
17 Gosper algorithm	157
18 Lorentz transformation	159
18.1 Einstein	159
18.2 Bondi k -calculus	159
18.3 wordline in Minkowski space	159
18.3.1 Wick rotation	159
18.3.1.1 Osterwalder-Schrader reconstruction theorem	159
19 R	161
19.1 TonyKuoYJ	161
19.1.1 quick intro	161
19.1.1.1 function	162
19.1.2 R style	162
19.1.3 data workflow or forward pipe	162
19.1.3.1 <code>%>%</code> operator	162
19.1.4 data processing with <code>dplyr</code>	163
19.1.5 visualization statically with <code>ggplot2</code>	165
19.1.6 loop	170
19.1.7 variable type	170
19.1.7.1 date	171
19.1.7.2 time	171
19.1.7.3 quotient <code>%%</code> operator	171
19.1.8 data type	172
19.1.8.1 vector	172
19.1.8.1.1 logic	173
19.1.8.1.2 <code>rep</code> repeat	173
19.1.8.1.3 <code>seq</code> sequence	173
19.1.8.2 factor	173
19.1.8.3 matrix	173
19.1.8.4 data frame	174
19.1.8.4.1 selecting variable or column	175
19.1.8.4.2 filtering observation or row	175
19.1.8.4.3 check mixed data type	175
19.2 W3School	176
19.2.1 same multiple variable	176
19.2.2 legal variable name	176
19.2.3 complex number	176

19.2.4 escape character	176
19.2.5 global assignment <<-	176
19.2.6 data type	177
19.2.6.1 array	177
19.2.6.2 list	177
19.3 Apan Liao	177
19.3.1 data input	178
19.3.2 descriptive statistics	178
20 Laplace transform	179
21 conic section	181
21.1 Cartesian coordinate: focus, directrix, eccentricity	181
21.2 two-definition equivalence for ellipse and hyperbola	183
21.2.1 first definition for conic sections including ellipses and hyperbolas	183
21.2.2 second definition for ellipses and hyperbolas	184
21.2.3 eccentricity and its equivalent representation	187
21.3 Cartesian coordinate: standard form / standard equation	187
21.4 parametric equation	189
21.5 polar coordinate	189
21.6 Cartesian coordinate: general form / quadratic equation	189
21.7 homogeneous coordinate	193
22 distance from a point to a line	195
22.1 by shortest $\overline{PP'}$	195
22.2 by perpendicular foot	196
22.3 by normal vector	197
22.4 by Cauchy inequality	197
23 real symmetric matrix diagonalizable	199
24 tangent half-angle formula	203
25 homogeneous coordinate	205
26 Archimedean property	207
26.1 integer Archimedean property	207
26.2 rational Archimedean property	207
26.3 real Archimedean property	207
27 MatPlotLib / matplotlib	209
27.1 Timothy H. Wu	209
27.1.1 funcitonal API	209
27.1.2 object-oriented API	213
27.1.2.1 Configure the figure size and DPI	214
27.1.2.2 subplot	215
27.1.2.3 color and linestyle	216
27.1.2.4 other inputs	216
27.1.2.5 legend	216
27.1.2.6 customize style	217
27.1.2.7 save to file	217
27.1.3 Seaborn	217
27.1.3.1 basic	219
27.1.3.2 data frame	219
27.1.3.3 axis-level plot and figure-level plot	221
27.1.3.4 accessing figure and axes objects	224
27.1.3.5 distribution plot	224
27.1.3.5.1 histogram	224
27.1.3.5.2 ECDF = empirical cumulative distrutive function	225
27.1.3.5.3 KDE = kernel density estimation	226
27.1.3.6 categorical plot	228
27.1.3.7 joint plot	228
27.1.3.8 pair plot	229
27.1.3.9 palette	229

27.1.3.10 volcano plot	230
27.1.3.11 heatmap	230
27.2 export	230
27.2.1 .svg	230
27.2.2 .eps	230
28 survival analysis	231
28.1 Python package <code>tableone</code>	231
28.2 Python package <code>lifelines</code>	231
29 Manim	233
29.1 VSCode extension: Manim Sideview	233
29.2 installation	233
29.2.1 Conda	233
29.3 quickstart	233
30 ggplot2	235
30.1 geom	235
30.1.1 basic plot system decomposition	239
30.2 statistic	241
30.3 scale	241
30.4 coordinate system	242
30.5 facet	243
30.6 jitter	244
30.7 font	244
31 notational system for design	245
31.1 graphic notation	245
31.2 pre-experimental design	245
31.2.1 one-shot case study	245
31.2.2 one-group pretest-posttest design	245
31.2.3 static-group comparison	246
31.3 true experimental design	246
31.3.1 posttest-only control group design	246
31.3.2 pretest-posttest control group design	249
31.3.3 Solomon four-group design	249
31.4 quasi-experimental design	250
31.5 correlational and ex post facto designs	250
31.6 graphic notation, advanced	250
32 design of experiment	251
32.1 notational system for design ^[31]	251
32.2 terminology	251
32.2.1 replication vs. subsample	252
32.2.2 replication vs. repeated measurements	252
32.2.3 replication, replicate	252
32.2.3.1 technical replicate, biological replicate	252
32.2.4 Latin square design	253
32.2.5 model assumption and experimental unit, measurement/observational unit	253
32.3 experiment structure	253
32.3.1 treatment structure	253
32.3.2 design structure	255
32.3.3 size of experimental unit	255
32.4 approach to experimentation	255
32.5 sample size estimation	255
32.6 statistical analysis plan	255
32.7 protocol	255
32.8 DoE course with six sigma and Minitab	256
32.8.1 evolution	256
33 quine	257
33.1 %r	257
34 quaternion	259

34.1 TaylorCatAlice	259
34.1.1 complex / dionion / bionion	259
34.1.2 trionion / triernion / triplex / ternion	259
34.1.3 quaternion	260
34.1.3.1 true origin of (dot product & cross product) / (inner product & outer product)	262
34.1.3.2 commutativity vs. anticommutativity	263
34.1.3.2.1 commutativity	263
34.1.3.2.2 anticommutativity	263
34.1.3.3 bracket	263
34.1.3.3.1 self-invented bracket	263
34.1.3.3.2 Poisson bracket	264
34.1.3.3.3 Lagrange bracket	264
34.1.3.3.4 Lie bracket	264
34.1.3.4 triple product	264
34.1.3.5 differential operator	266
34.1.3.5.1 4-differential operator	266
34.1.3.5.2 nabla	266
34.1.3.5.3 Laplace operator	267
34.1.3.5.4 d'Alembert operator	267
34.1.3.6 electromagnetism	267
34.1.3.6.1 4-potential	267
34.1.3.6.2 Maxwell compromise for both quaternion and 3-vector	268
34.1.3.7 Joule heat vs. Thomson heat (Kelvin heat?)	269
34.1.3.7.1 thermoelectric effect	269
34.1.3.8 source term	269
34.1.4 quaternion group	270
34.1.4.1 2D rotation	270
34.1.4.1.1 matrix	270
34.1.4.1.2 complex	271
34.1.4.1.3 $SO(2) \cong U(1)$	272
34.1.4.2 3D rotation	272
34.1.4.2.1 matrix	272
34.1.4.2.2 quaternion	275
34.1.5 octonion	277
34.2 Krasjet	277
34.2.1 Rodrigues rotation	277
34.2.2 quaternion operation	278
34.2.3 matrix form	285
34.2.4 exponential form	286
34.2.5 double cover	287
34.2.6 right-hand vs. left-hand coordinates	288
34.2.6.1 clock	288
34.2.6.2 Rodrigues rotation in right-hand vs. left-hand coordinates	288
34.2.7 gimbal lock	288
34.2.8 Lie group	289
34.2.8.1 special unitary group	289
34.2.9 geometric algebra and Clifford algebra	289
34.2.10 dual quaternion	289
34.2.11 interpolation	289
34.2.11.1 Lerp	289
34.2.11.2 Nlerp	289
34.2.11.3 Slerp	289
34.3 3Blue1Brown	289
34.3.1 Ben Eater	289
34.3.2 3B1B	289
34.3.3 Sutrable	289
34.4 CCJou: LA Revelation	289
35 DICOM	291
35.1 Innolitics: DICOM Standard Browser	291
35.2 David Clunie: Medical Image Format Site	291
35.3 Hsiao, Chia-Hung	291

36 tendon pathophysiology	293
36.1 histology	293
37 group theory	295
37.1 matrix group	295
37.2 group definition and basic theorem	295
37.3 Polya enumeration theorem	296
37.4 complex group representation	296
37.4.1 complex basis group	296
37.4.2 $\mathbb{C} \rightarrow \mathcal{M}_{2 \times 2}(\mathbb{R})$	296
37.4.3 (determinant of complex group representation) equivalent to (squared modulus of complex number)	297
37.4.3.1 Lagrange identity	298
37.4.4 Euler formula proved by complex group representation	298
37.5 quaternion group representation	299
37.5.1 $\mathbb{H} \rightarrow \mathcal{M}_{2 \times 2}(\mathbb{C})$	299
37.5.1.1 Euler identity	301
37.5.2 $\mathbb{H} \rightarrow \mathcal{M}_{4 \times 4}(\mathbb{R})$	303
38 tensor	307
38.1 Einstein summation convention	307
38.1.1 dummy index	307
38.1.2 free index	307
38.2 Elliot Schneider	307
38.2.1 Fundamentals of Cartesian Tensors	307
38.2.2 Fundamentals of Curvilinear Tensors	307
38.2.3 Fundamentals of Spacetime Tensors	307
39 dual space	309
39.1 linear equations	309
39.2 matrix multiplication	310
39.3 functional	311
39.4 definition of linear functional	313
39.5 set of all linear transformations is a vector space	314
39.6 definition of dual space	316
39.7 double dual	316
40 vector space	317
40.1 What is a vector?	317
40.2 vector space definition	317
40.2.1 commutative group structure of vector space	317
40.2.2 scalar distributivity	318
40.3 linearity	319
40.3.1 linear structure of vector space	319
40.3.2 linear transformation or linear map	319
40.4 vector space example	319
40.4.1 reciprocal space	320
40.4.1.1 Kronecker delta	320
40.4.1.2 double dual concept	322
40.5 field	323
40.6 module	323
40.7 subspace	323
41 df	325
41.1 df decomposed with partials as a set of basis in vector space	325
41.2 dual space of span of partials	326
41.3 directional derivative	326
41.4 coefficient of linear combination for vector space and dual space	327
41.5 change of basis / change of coordinate	327
42 determinant	329
42.1 induction of determinant axioms	329
42.2 determinant axioms	329
42.3 determinant theorems or properties	329
42.4 ellipse area	329

42.5 Cramer rule geometry perspective	329
43 hypergeometric function	331
43.1 linear space of function	331
43.2 beta function	332
43.2.1 Wallis product formula	334
43.3 gamma function	334
43.4 recursion	334
43.5 mean and variance of discrete probability distributions	334
44 Feynman method	335
44.1 Feynman method of differentiation / derivative	335
44.1.1 principle	335
44.1.2 examples	336
44.2 Feynman method of integration / integral	338
44.3 path integral	338
45 Hilbert space	339
45.1 Taylor expansion or Taylor series	339
45.1.1 Elliot Schneider	349
45.1.1.1 making complicated equation simple	350
45.1.1.2 non-relativistic limit	351
45.1.1.3 quantum momentum operator	351
45.2 Dirac delta function	351
45.2.1 complex delta function	355
45.2.2 3D delta function	358
45.2.3 quantum state	362
45.2.4 rigorous definition of delta function	362
45.3 Fourier expansion or Fourier series	363
45.3.1 Taylor vs. Fourier	363
45.4 Hilbert space	364
45.4.1 inner product space	365
45.5 Dirac bracket	366
45.5.1 Dirac bracket symmetry	368
45.6 quantum operator	370
45.7 Gram-Schmidt algorithm to find orthonormal basis	370
45.8 Hermite polynomials and Legendre polynomials	370
46 Lagrange inversion	371
47 quantum color dynamics	373
48 Lagrangian	375
49 Hamiltonian	377
50 Fourier analysis	379
50.1 basic calculation	379
50.2 Simon Xu	380
50.2.1 DFT = discrete Fourier transform	380
50.2.2 FFT = fast Fourier transform	380
50.2.3 wavelet	380
51 axiom of choice	381
51.1 Zorn lemma	381
52 linear algebra	383
52.1 CCJou	383
52.2 Chi, Chen-Yu	383
53 analysis	385
53.1 Chi, Chen-Yu	385
53.2 Chen, Jin-Tzu	385

54 matrix calculus	387
54.1 vector-by-scalar	387
54.2 scalar-by-vector	387
54.3 vector-by-vector	388
54.3.1 numerator-layout notation	388
54.3.2 denominator-layout notation	388
55 genetics	389
55.1 term	389
55.2 marker	389
55.3 genome project	389
55.4 linkage analysis	390
55.4.1 PLA = parametric linkage analysis	391
55.4.2 VCLA = variance component linkage analysis	391
55.5 association analysis	391
56 multivariate normal distribution	393
56.1 bivariate normal distribution	399
57 logic	401
57.1 Shai Ben-David	401
58 algorithm	403
58.1 Chen, Vivian	403
58.2 The Bubble Sort Curve	403
59 computer graphics	405
59.1 Cem Yuksel	405
60 autoregression in time series	407
60.1 AR(1) = 1 st -order autoregressive model = first-order autoregressive model	409
60.2 AR(2) = 2 nd -order autoregressive model = second-order autoregressive model	410
60.3 AR(p) = p^{th} -order autoregressive model	410
60.4 MA(q) = q^{th} -order moving-average model	410
60.5 ARMA(p,q) = $p^{\text{th}}, q^{\text{th}}$ -order autoregressive-moving-average model	411
61 conductivity	413
61.1 resistance	413
61.2 semiconductor	413
62 LaTeX annotation by TikZ	415
63 Shannon sampling	421
64 brachistochrone	423
65 magnetic resonance	425
65.1 NMR	425
65.2 MRI	425
65.2.1 Lin, Hsiu-Hau	425
66 OpenMMLab	427
66.1 Tongji TommyZihao	427
66.2 official	427
references	429

index

math on bookdown started on 2024/01/28

Part I

by discipline

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=jIX_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-usgymR7FTQ3ZRjfLs5jT4BO>

5.2 deep learning

我妻幸長

Esc = Einstein summation convention

$$\begin{aligned}
 W\mathbf{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} \stackrel{\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 \\
 \mathbf{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\mathbf{x} \\
 \mathbf{y}^\top &= \begin{pmatrix} y_0 \\ y_1 \\ \vdots \\ y_m \end{pmatrix}^\top = y_\mu^\top = (w_{\mu\nu} x_\nu)^\top = \begin{pmatrix} w_{0\nu} x_\nu \\ w_{1\nu} x_\nu \\ \vdots \\ w_{m\nu} x_\nu \end{pmatrix}^\top = \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]^\top = [W\mathbf{x}]^\top \\
 \mathbf{x}_\nu^\top w_{\mu\nu}^\top &= \begin{pmatrix} w_{0\nu} x_\nu \\ w_{1\nu} x_\nu \\ \vdots \\ w_{m\nu} x_\nu \end{pmatrix}^\top = \begin{pmatrix} x_0 \\ x_1 \\ \vdots \\ x_n \end{pmatrix}^\top \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}^\top = \mathbf{x}^\top W^\top \\
 \mathbf{x}^\top W^\top &= \begin{pmatrix} x_0 \\ x_1 \\ \vdots \\ x_n \end{pmatrix}^\top \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}^\top = \begin{pmatrix} w_{0\nu} x_\nu \\ w_{1\nu} x_\nu \\ \vdots \\ w_{m\nu} x_\nu \end{pmatrix}^\top = x_\nu^\top w_{\mu\nu}^\top = y_\mu^\top = \begin{pmatrix} y_0 \\ y_1 \\ \vdots \\ y_m \end{pmatrix}^\top = \mathbf{y}^\top
 \end{aligned}$$

$$\begin{aligned}
W\mathbf{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} \stackrel{\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\mathbf{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_{j=1}^n w_{0j} x_j \\ b_1 + \sum_{j=1}^n w_{1j} x_j \\ \vdots \\ b_m + \sum_{j=1}^n w_{mj} x_j \end{pmatrix} \stackrel{\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 \\
\mathbf{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\mathbf{x}, \begin{cases} x_0 = 1 \\ w_{\mu 0} = b_\mu \end{cases} \\
&= \mathbf{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\mathbf{x}, \begin{cases} 1 = x_0 \\ b_\mu = w_{\mu 0} \end{cases} \\
\mathbf{y}^\top &= \begin{pmatrix} y_0 \\ y_1 \\ \vdots \\ y_m \end{pmatrix}^\top = y_\mu^\top = (w_{\mu\nu} x_\nu)^\top = \begin{pmatrix} w_{0\nu} x_\nu \\ w_{1\nu} x_\nu \\ \vdots \\ w_{m\nu} x_\nu \end{pmatrix}^\top = \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]^\top = [W\mathbf{x}]^\top \\
&= x_\nu^\top w_{\mu\nu}^\top = \begin{pmatrix} w_{0\nu} x_\nu \\ w_{1\nu} x_\nu \\ \vdots \\ w_{m\nu} x_\nu \end{pmatrix}^\top = \begin{pmatrix} x_0 \\ x_1 \\ \vdots \\ x_n \end{pmatrix}^\top \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}^\top = \mathbf{x}^\top W^\top, \begin{cases} x_0 = 1 \\ w_{\mu 0} = b_\mu \end{cases} \\
&= b_\mu^\top + x_j^\top w_{\mu j}^\top = \begin{pmatrix} b_0 + w_{0j} x_j \\ b_1 + w_{1j} x_j \\ \vdots \\ b_m + w_{mj} x_j \end{pmatrix}^\top = \begin{pmatrix} 1 \\ x_1 \\ \vdots \\ x_n \end{pmatrix}^\top \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}^\top = \mathbf{x}^\top W^\top \\
\mathbf{x}^\top W^\top &= \begin{pmatrix} x_0 \\ x_1 \\ \vdots \\ x_n \end{pmatrix}^\top \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}^\top = \begin{pmatrix} w_{0\nu} x_\nu \\ w_{1\nu} x_\nu \\ \vdots \\ w_{m\nu} x_\nu \end{pmatrix}^\top = x_\nu^\top w_{\mu\nu}^\top = y_\mu^\top = \begin{pmatrix} y_0 \\ y_1 \\ \vdots \\ y_m \end{pmatrix}^\top = \mathbf{y}^\top \\
\mathbf{x}^\top W^\top &= \begin{pmatrix} 1 \\ x_1 \\ \vdots \\ x_n \end{pmatrix}^\top \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}^\top = \begin{pmatrix} b_0 + w_{0j} x_j \\ b_1 + w_{1j} x_j \\ \vdots \\ b_m + w_{mj} x_j \end{pmatrix}^\top = b_\mu^\top + x_j^\top w_{\mu j}^\top = y_\mu^\top = \begin{pmatrix} y_0 \\ y_1 \\ \vdots \\ y_m \end{pmatrix}^\top = \mathbf{y}^\top
\end{aligned}$$

$$\begin{aligned}
W\mathbf{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} \stackrel{\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\mathbf{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_{j=1}^n w_{0j} x_j \\ b_1 + \sum_{j=1}^n w_{1j} x_j \\ \vdots \\ b_m + \sum_{j=1}^n w_{mj} x_j \end{pmatrix} \stackrel{\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 \\
\mathbf{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\mathbf{x}, \begin{cases} x_0 = 1 \\ w_{\mu 0} = b_\mu \end{cases} \\
&= \mathbf{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\mathbf{x}, \begin{cases} 1 = x_0 \\ b_\mu = w_{\mu 0} \end{cases} \\
\mathbf{y}^\top &= \begin{pmatrix} y_0 \\ y_1 \\ \vdots \\ y_m \end{pmatrix}^\top = y_\mu^\top = (w_{\mu\nu} x_\nu)^\top = \begin{pmatrix} w_{0\nu} x_\nu \\ w_{1\nu} x_\nu \\ \vdots \\ w_{m\nu} x_\nu \end{pmatrix}^\top = \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]^\top = [W\mathbf{x}]^\top \\
&= x_\nu^\top w_{\mu\nu}^\top = x_\nu^\top w_{\nu\mu} = \begin{pmatrix} w_{0\nu} x_\nu \\ w_{1\nu} x_\nu \\ \vdots \\ w_{m\nu} x_\nu \end{pmatrix}^\top = \begin{pmatrix} x_0 \\ x_1 \\ \vdots \\ x_n \end{pmatrix}^\top \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}^\top = \mathbf{x}^\top W^\top \\
&= b_\mu^\top + x_j^\top w_{\mu j}^\top = b_\mu^\top + x_j^\top 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}^\top = \begin{pmatrix} 1 \\ x_1 \\ \vdots \\ x_n \end{pmatrix}^\top \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}^\top \\
\mathbf{x}^\top W^\top &= \begin{pmatrix} x_0 \\ x_1 \\ \vdots \\ x_n \end{pmatrix}^\top \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}^\top = \begin{pmatrix} w_{0\nu} x_\nu \\ w_{1\nu} x_\nu \\ \vdots \\ w_{m\nu} x_\nu \end{pmatrix}^\top = x_\nu^\top w_{\nu\mu} = y_\mu^\top = \begin{pmatrix} y_0 \\ y_1 \\ \vdots \\ y_m \end{pmatrix}^\top = \mathbf{y}^\top \\
\mathbf{x}^\top W^\top &= \begin{pmatrix} 1 \\ x_1 \\ \vdots \\ x_n \end{pmatrix}^\top \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}^\top = \begin{pmatrix} b_0 + w_{0j} x_j \\ b_1 + w_{1j} x_j \\ \vdots \\ b_m + w_{mj} x_j \end{pmatrix}^\top = b_\mu^\top + x_j^\top w_{j\mu} = y_\mu^\top = \begin{pmatrix} y_0 \\ y_1 \\ \vdots \\ y_m \end{pmatrix}^\top
\end{aligned}$$

$$\begin{aligned}
\mathbf{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\mathbf{x}, \begin{cases} x_0 = 1 \\ w_{\mu 0} = b_\mu \end{cases} \\
&= \mathbf{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\mathbf{x}, \begin{cases} 1 = x_0 \\ b_\mu = w_{\mu 0} \end{cases} \\
\mathbf{x}^\top W^\top &= \begin{pmatrix} x_0 \\ x_1 \\ \vdots \\ x_n \end{pmatrix}^\top \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}^\top = \begin{pmatrix} w_{0\nu} x_\nu \\ w_{1\nu} x_\nu \\ \vdots \\ w_{m\nu} x_\nu \end{pmatrix}^\top = x_\nu^\top w_{\mu\nu}^\top = x_\nu^\top w_{\nu\mu} = y_\mu^\top = \begin{pmatrix} y_0 \\ y_1 \\ \vdots \\ y_m \end{pmatrix}^\top = \mathbf{y}^\top \\
&= \mathbf{x}^\top W^\top = \begin{pmatrix} 1 \\ x_1 \\ \vdots \\ x_n \end{pmatrix}^\top \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}^\top = \begin{pmatrix} b_0 + w_{0j} x_j \\ b_1 + w_{1j} x_j \\ \vdots \\ b_m + w_{mj} x_j \end{pmatrix}^\top = b_\mu^\top + x_j^\top w_{j\mu} = y_\mu^\top = \begin{pmatrix} y_0 \\ y_1 \\ \vdots \\ y_m \end{pmatrix}^\top = \mathbf{y}^\top
\end{aligned}$$

$$\begin{aligned}
\sigma(\mathbf{y}) &= \sigma \begin{pmatrix} y_0 \\ y_1 \\ \vdots \\ y_\mu \end{pmatrix} = \sigma(y_\mu) = \sigma(w_{\mu\nu} x_\nu) = \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(W\mathbf{x}), \begin{cases} x_0 = 1 \\ w_{\mu 0} = b_\mu \end{cases} \\
&= \sigma(\mathbf{y}) = \sigma \begin{pmatrix} y_0 \\ y_1 \\ \vdots \\ y_\mu \end{pmatrix} = \sigma(y_\mu) = \sigma(w_{\mu j} x_j + b_\mu) = \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(W\mathbf{x}), \begin{cases} 1 = x_0 \\ b_\mu = w_{\mu 0} \end{cases} \\
(\mathbf{x}^\top W^\top) \varsigma &= \begin{pmatrix} x_0 \\ x_1 \\ \vdots \\ x_n \end{pmatrix}^\top \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}^\top \varsigma = \begin{pmatrix} w_{0\nu} x_\nu \\ w_{1\nu} x_\nu \\ \vdots \\ w_{m\nu} x_\nu \end{pmatrix}^\top \varsigma = (x_\nu^\top w_{\mu\nu}^\top) \varsigma = (x_\nu^\top w_{\nu\mu}) \varsigma = (y_\mu^\top) \varsigma = \begin{pmatrix} y_0 \\ y_1 \\ \vdots \\ y_m \end{pmatrix}^\top \varsigma = (\mathbf{y}^\top) \varsigma \\
&= (\mathbf{x}^\top W^\top) \varsigma = \begin{pmatrix} 1 \\ x_1 \\ \vdots \\ x_n \end{pmatrix}^\top \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}^\top \varsigma = \begin{pmatrix} b_0 + w_{0j} x_j \\ b_1 + w_{1j} x_j \\ \vdots \\ b_m + w_{mj} x_j \end{pmatrix}^\top \varsigma = (b_\mu^\top + x_j^\top w_{j\mu}) \varsigma = (y_\mu^\top) \varsigma = \begin{pmatrix} y_0 \\ y_1 \\ \vdots \\ y_m \end{pmatrix}^\top \varsigma = (\mathbf{y}^\top) \varsigma
\end{aligned}$$

$$\begin{aligned}
\mathbf{z} &= \sigma \mathbf{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 \mathbf{x}, \begin{cases} x_0 = 1 \\ w_{\mu 0} = b_\mu \end{cases} \\
&= \mathbf{z} = z_\mu = \sigma_\mu \begin{pmatrix} y_0 \\ y_1 \\ \vdots \\ y_\mu \end{pmatrix} = \sigma_\mu y_\mu = \sigma_\mu (w_{\mu j} x_j + b_\mu) = \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 \mathbf{x}, \begin{cases} 1 = x_0 \\ b_\mu = w_{\mu 0} \end{cases} \\
\mathbf{x}^\top W^\top \varsigma &= \begin{pmatrix} x_0 \\ x_1 \\ \vdots \\ x_n \end{pmatrix}^\top \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}^\top \varsigma = \begin{pmatrix} w_{0\nu} x_\nu \\ w_{1\nu} x_\nu \\ \vdots \\ w_{m\nu} x_\nu \end{pmatrix}^\top \varsigma = \mathbf{x}^\top w_{\mu\nu} \varsigma_\mu = x_\nu^\top w_{\nu\mu} \varsigma_\mu = y_\mu^\top \varsigma_\mu = \begin{pmatrix} y_0 \\ y_1 \\ \vdots \\ y_m \end{pmatrix}^\top \varsigma_\mu = \mathbf{y}^\top \varsigma = \mathbf{z}^\top \varsigma \\
&= \mathbf{x}^\top W^\top \varsigma = \begin{pmatrix} 1 \\ x_1 \\ \vdots \\ x_n \end{pmatrix}^\top \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}^\top \varsigma = \begin{pmatrix} b_0 + w_{0j} x_j \\ b_1 + w_{1j} x_j \\ \vdots \\ b_m + w_{mj} x_j \end{pmatrix}^\top \varsigma = (b_\mu^\top + x_j^\top w_{j\mu}) \varsigma_\mu = y_\mu^\top \varsigma_\mu = \begin{pmatrix} y_0 \\ y_1 \\ \vdots \\ y_m \end{pmatrix}^\top \varsigma_\mu = \mathbf{z}_\mu^\top = \mathbf{z}^\top \varsigma
\end{aligned}$$

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

matrix calculus^[54]

4-15

wrong or incompatible transpose

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

4-18

wrong or incompatible transpose

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

wrong or incompatible transpose

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

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)
```

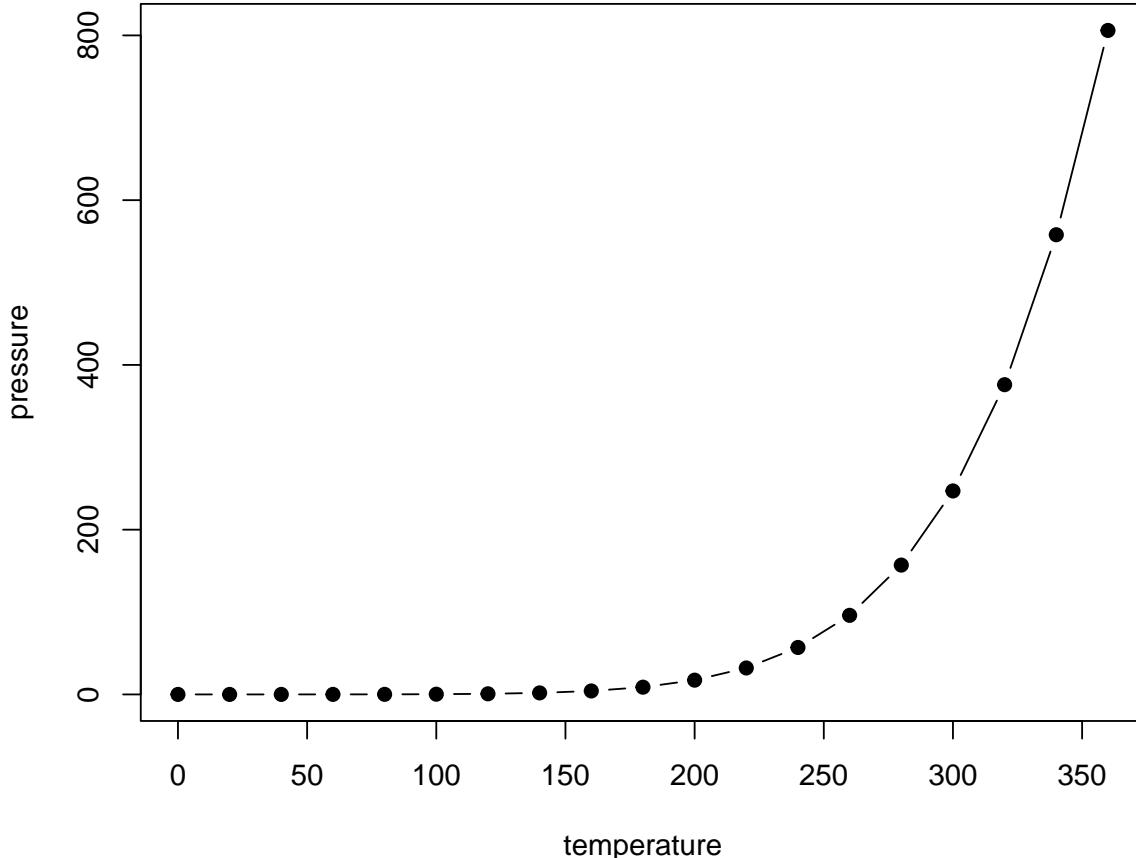


Figure 6.1: Here is a nice figure!

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** package¹ (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.

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

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

¹This 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

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 verbatim

<https://community.rstudio.com/t/continued-issues-with-new-verbatim-in-rstudio/139737>
<https://bookdown.org/yihui/rmarkdown-cookbook/verbatim-code-chunks.html>

7.2.2 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>

7.2.3 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)] (<https://www.google.com>)
[[I'm an inline-style link with title](https://www.google.com)] (<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>

7.2.4 arrow

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

Up arrow: ↑

Down arrow: ↓

Left arrow: ←

Right arrow: →

Double headed arrow: ↔

7.2.5 superscript and subscript

```
scriptsuperscriptsubscript
scriptsuperscriptsubscript
```

```
scriptsuperscript
~subscript~
```

```
scriptsubscript
```

7.2.5.1 LaTeX

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

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

7.2.6 equation

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

7.2.6.1 proof QED

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

```
\tag*{$\Box$}
```

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

□

```
\tag*{$\blacksquare$}
```

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

■

7.2.7 image

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

7.2.7.1 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>

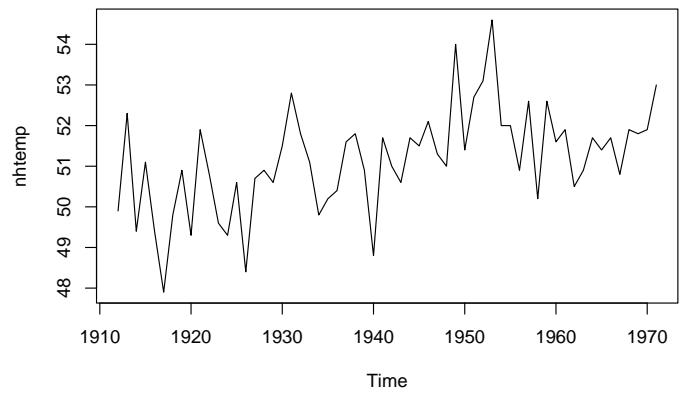
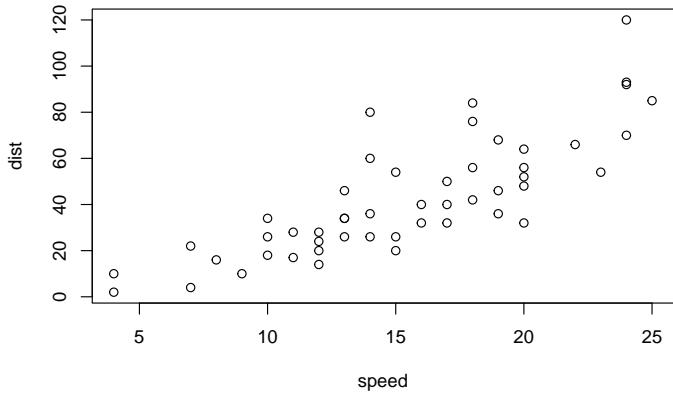
<https://stackoverflow.com/questions/55994210/how-to-solve-diagrammer-waste-of-space-issue-in-rmarkdown>

7.2.7.2 multiple images / figures in the same line

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

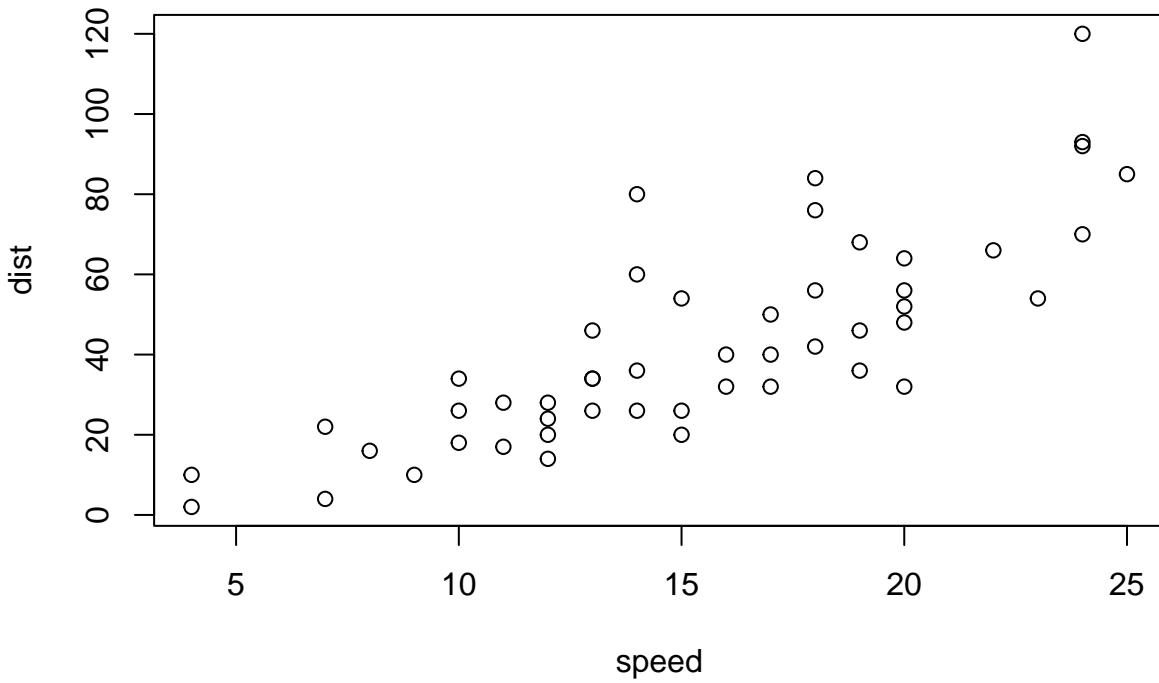
```
{r, fig.show = "hold", out.width = "50%"}
plot(cars)
plot(nhtemp)

plot(cars)
plot(nhtemp)
```

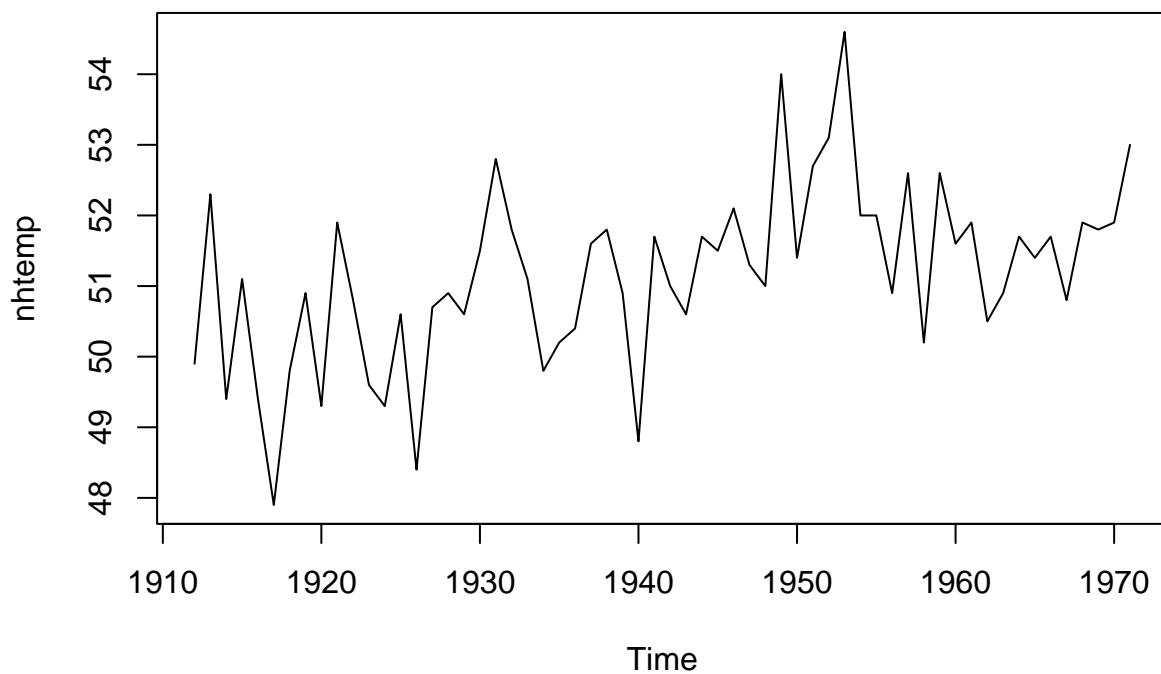


cf.

```
{r}
plot(cars)
plot(nhtemp)
plot(cars)
```



```
plot(nhtemp)
```



7.2.7.3 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

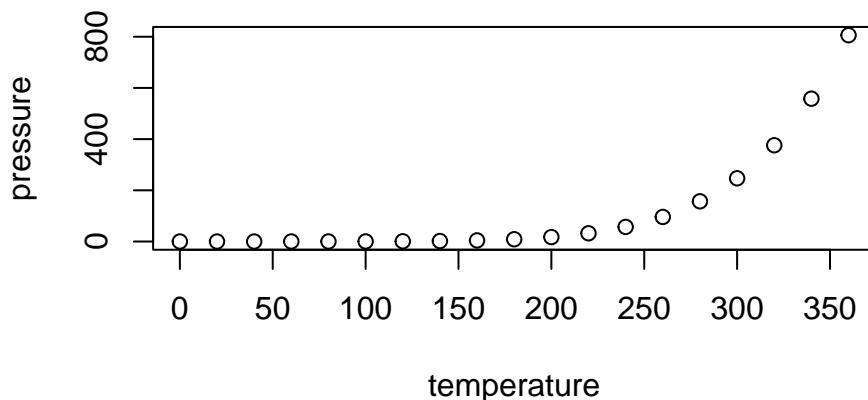
```
```{r fig.height = 3, fig.width = 5}
```

```
plot(pressure)
```

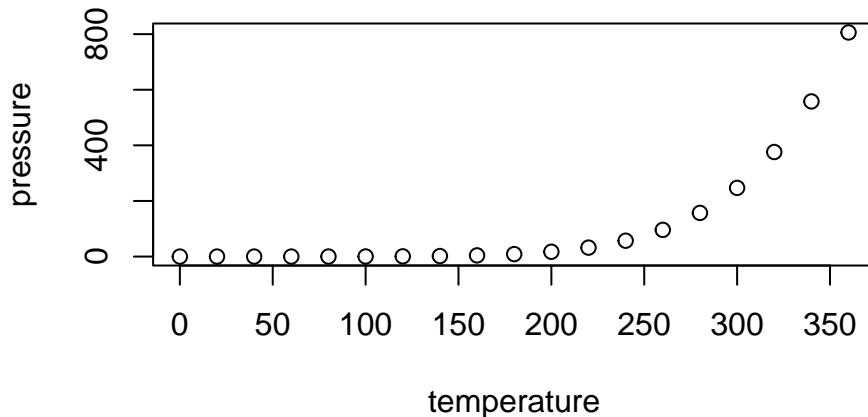
```
```
```

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

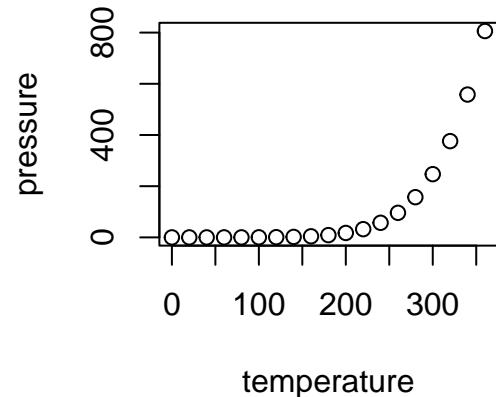
```
plot(pressure)
```



```
{r fig.height = 3, fig.width = 3, fig.align = "center"
plot(pressure)}
```



```
{r fig.width = 5, fig.asp = .62
plot(pressure)}
```

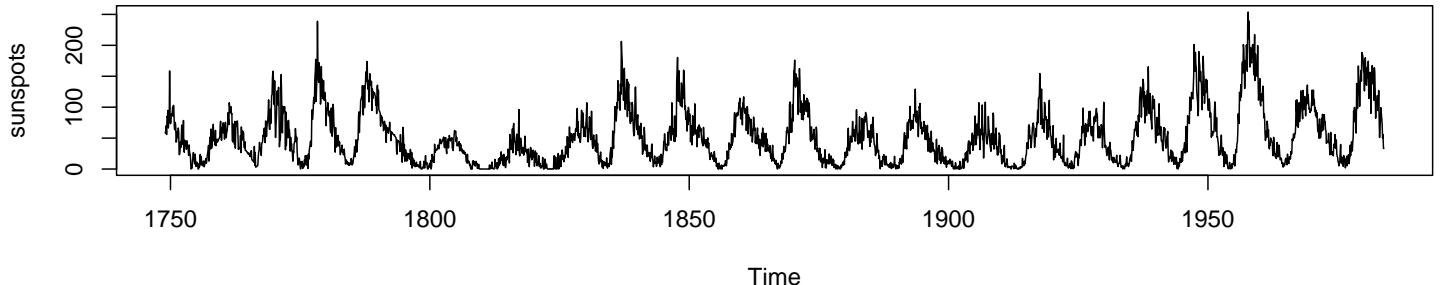


```
<center>
! [] (https://bookdown.org/yihui/rmarkdown-cookbook/images/cover.png){width=20%}
</center>
```

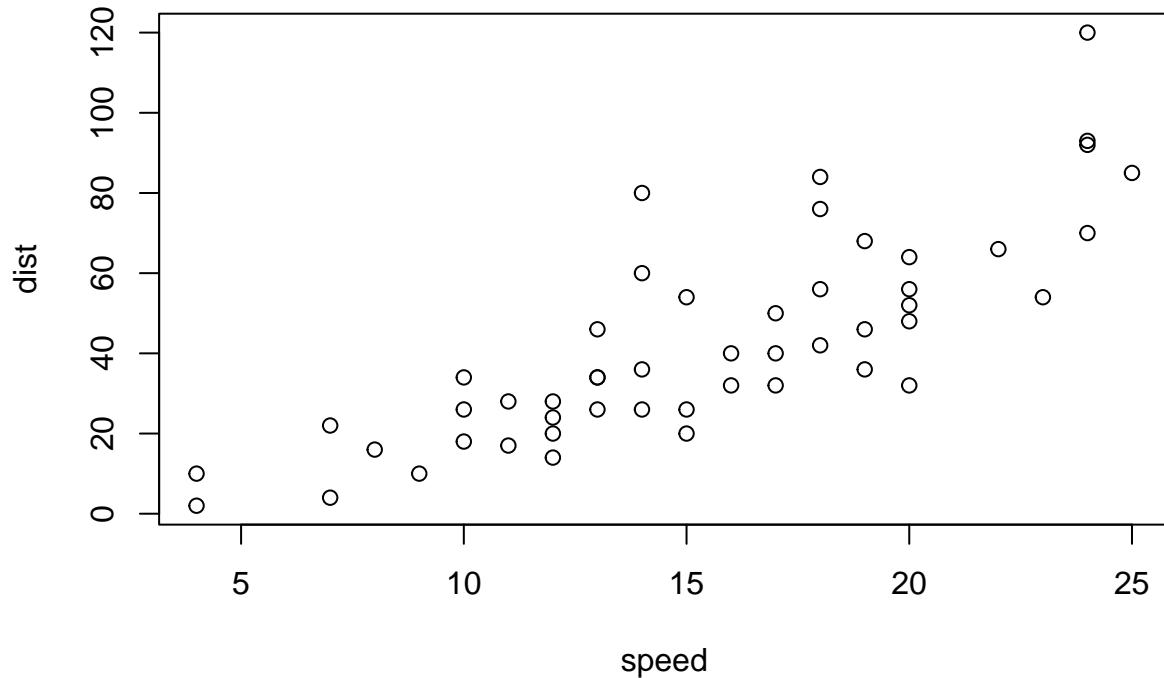
7.2.7.3.1 knitr <https://yihui.org/knitr/options/>

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

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



```
plot(cars)
```



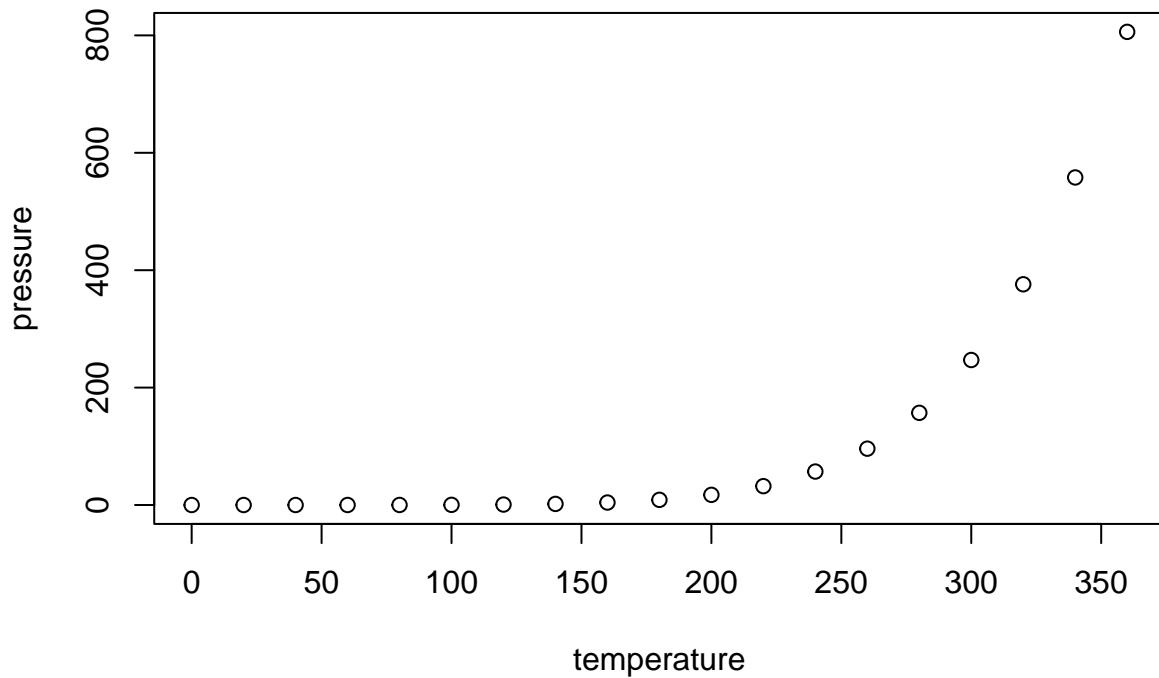
We know from the first fundamental theorem of calculus that
for x in $[a, b]$:

$$\frac{d}{dx} \left(\int_a^x f(u) du \right) = f(x).$$

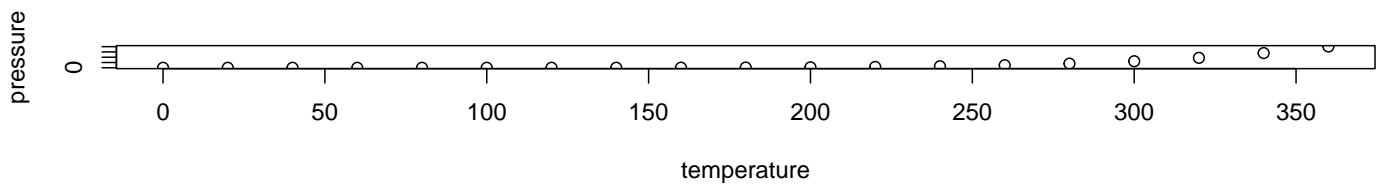
7.2.7.3.2 out.width vs. fig.width <https://stackoverflow.com/questions/29657777/how-to-make-fig-width-and-out-width-consistent-with-knitr>

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>



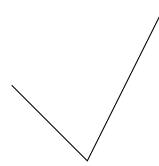
```
plot(pressure)
```



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

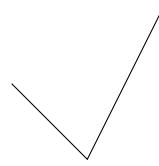
solution: `out.width=if (knitr:::is_html_output()) '100%'`

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



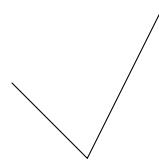
`fig.width=10, fig.height=2`

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



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

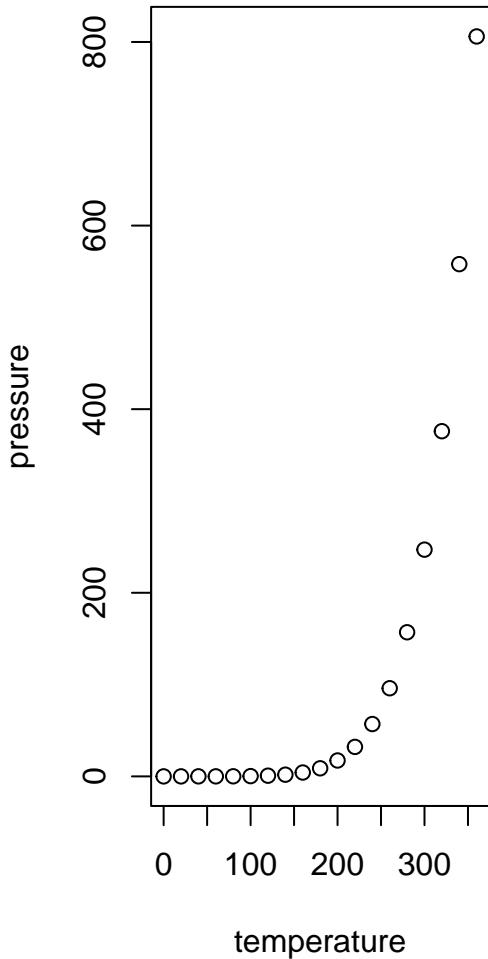
```
\begin{tikzpicture}
  \draw (-1,1)--(0,0)--(1,2);
\end{tikzpicture}
```



7.2.7.4 dynamic knitr plot width and height

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

```
plot(pressure)
```



7.2.7.5 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)
```

7.2.7.6 SVG

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

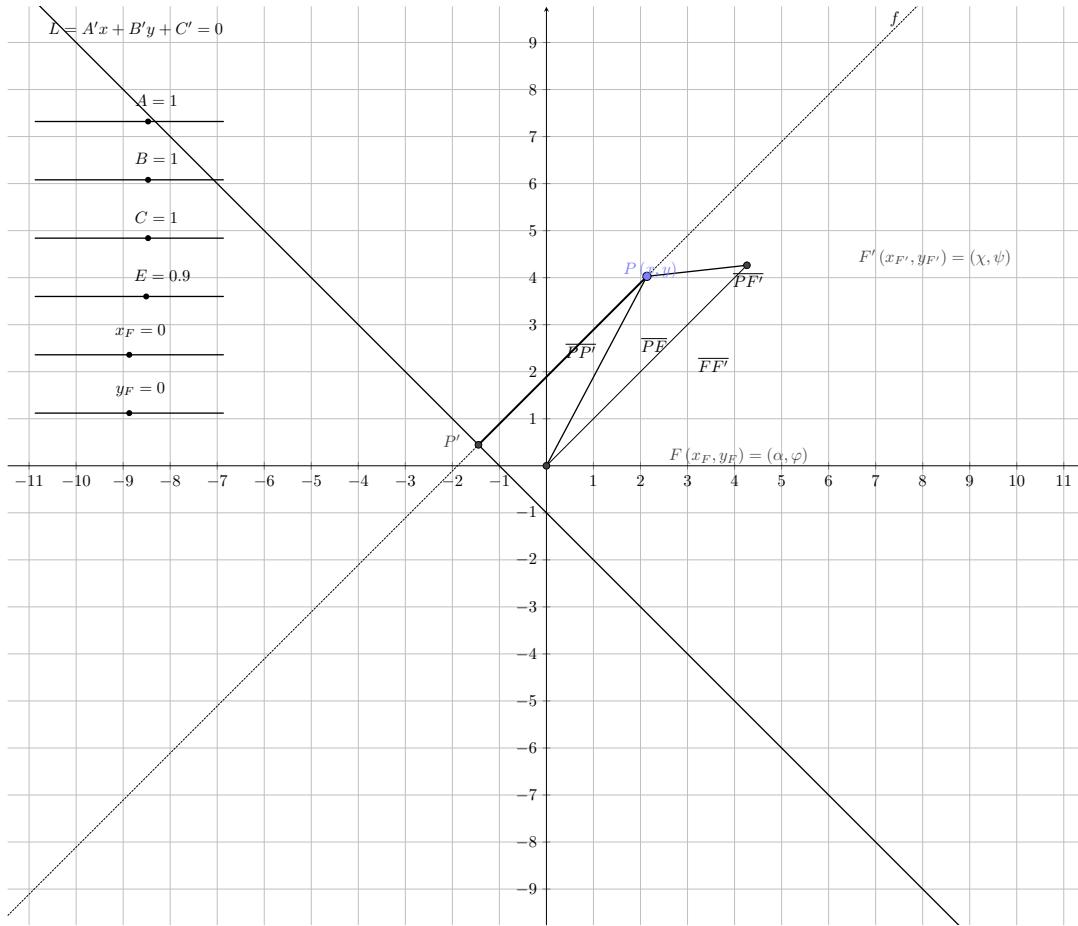
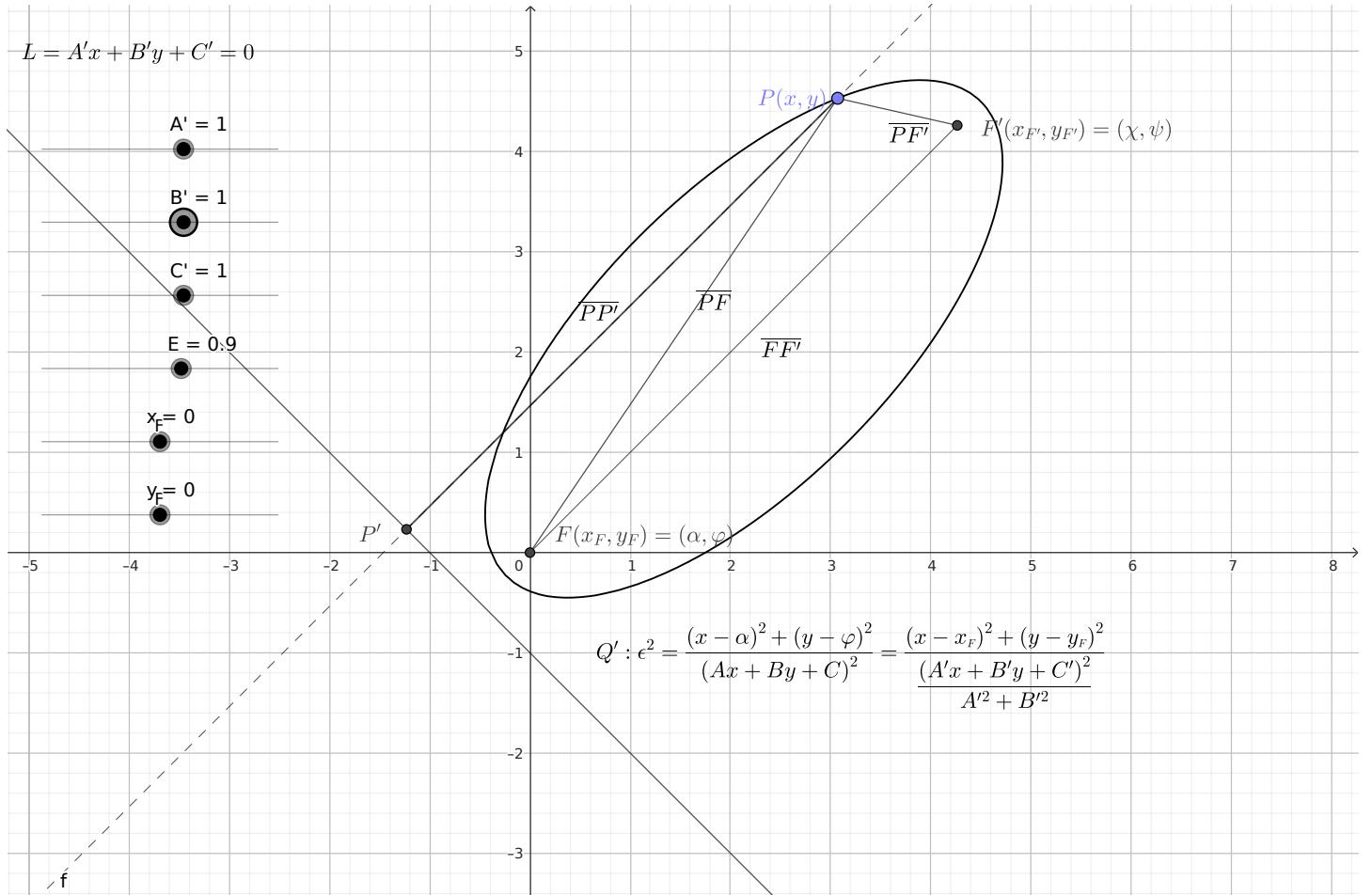


Figure 7.1: conic sections



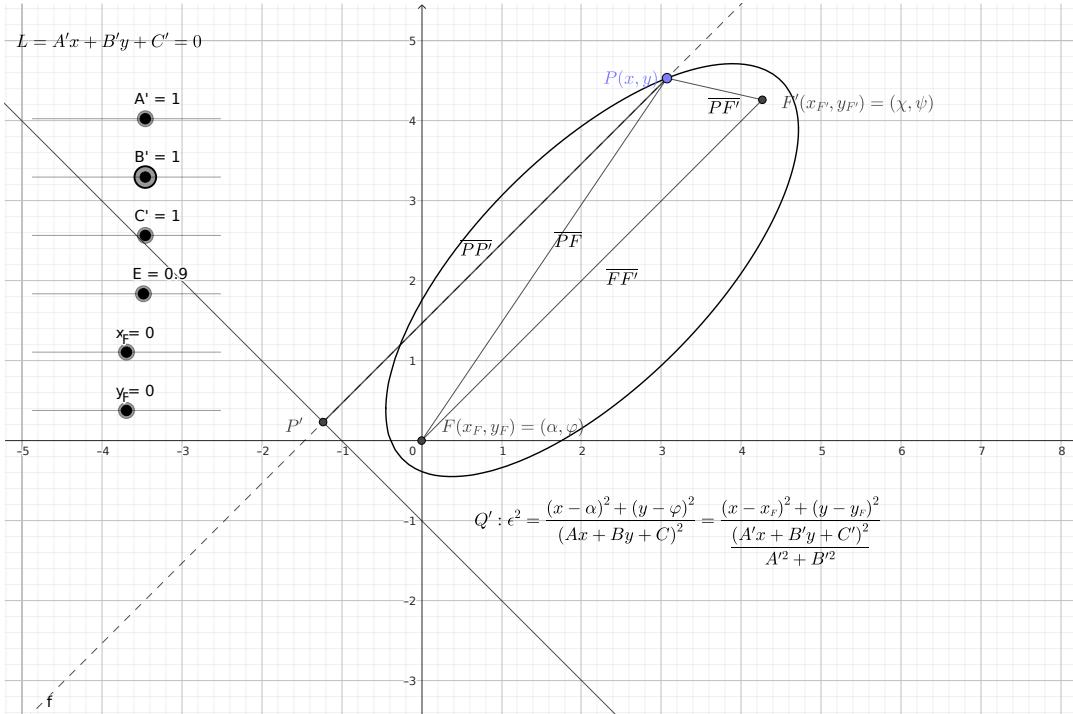


Figure 7.2: conic sections

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

7.2.8 horizontal rule

horizontal rule (or slide break)

dim(iris)

[1] 150 5

7.2.9 footnote

7.2.10 hyperlink

PDF pandoc internal link will lose focus

equivalence relation [11] equivalence relation¹ equivalence relation^[11]

equivalence class [10] equivalence class² equivalence class^[10]

partition [9] partition³ 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

7.2.11 code chunk

7.2.11.1 code folding

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

7.2.12 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>

7.2.13 embed a web page

7.2.13.1 iframe

7.2.13.2 include URL

<https://bookdown.org/yihui/rmarkdown-cookbook/include-url.html>

```
knitr::include_url("https://yihui.org")
```

I'm currently a freelancer, and was a software engineer at Posit, PBC (2013-2023). I earned my PhD from the Department of Statistics, Iowa State University. My thesis was *Dynamic Graphics and Reporting for Statistics*, advised by Di Cook and Heike Hofmann. I have developed a few R packages either seriously or for fun (or both), such as knitr, animation, bookdown, blogdown, pagedown, xaringan, and tinytex. I founded a Chinese website called "Capital of Statistics" in 2006, which has grown into a large online community on statistics. I initiated the China R conference in 2008. I'm a big fan of GitHub, LyX and Pandoc. I used to hate IE but no longer care since it has almost died. I fall asleep when I see beamer slides, and I yell at people who use \textbf{} to write \title{}. I know I cannot eat code, so I cook almost every day to stay away from my computer for two hours.

这是谢益辉的主页。2013年底我从 Ames 村办大学统计系毕业，终于解决了人生前三十年被问最多的问题：“你怎么还没毕业？”“目前为自由职业者。我支持开源，喜欢折腾网站和代码，是个高度自我驱动的人。打羽毛球爱勾对角，打乒乓球像太极，网球满场子捡球，篮球容易被撞飞，攀岩一次，腿软。宅，口重、嗜辣、屡教不改。智商中等偏下，对麻将和三国杀有不可逾越的认知障碍，实变函数课上曾被老师叫醒。略好读书与文，对诗词楹联感兴趣，目前较中意的一联是：千秋邈矣独留我；百战归来再读书。最喜欢的一首词是：

深情似海，问相逢初度，是何年纪？依约而今还记取，不是前生夙世。放学花前，题诗石上，春水园亭里。逢君一笑，人间无此欢喜。
无奈苍狗看云，红羊数劫，惆怅休提起。客气渐多真气少，汨没心灵何已。
千古声名，百年担负，事事违初意。心头阁住，儿时那种情味。

7.2.13.3 htmltools

<https://stackoverflow.com/questions/36524238/include-html-files-in-r-markdown-file>

however `out.height` etc. not works

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{largetsymbol}{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}"}
<!-- -->
:::

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

```
::: {.col data-latex="{0.4\textwidth}"}
```

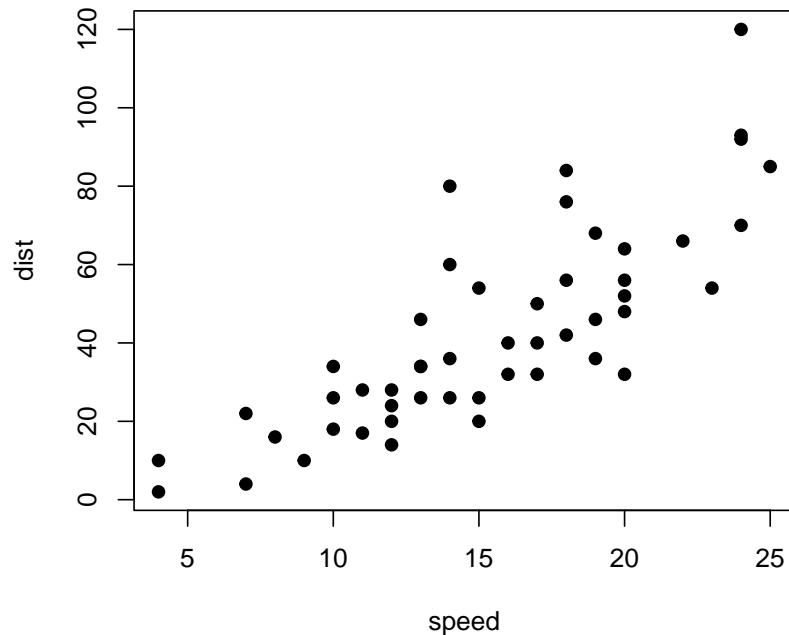
The figure on the left-hand side shows the `cars` data.

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.

```
:::
```

```
:::::::
```

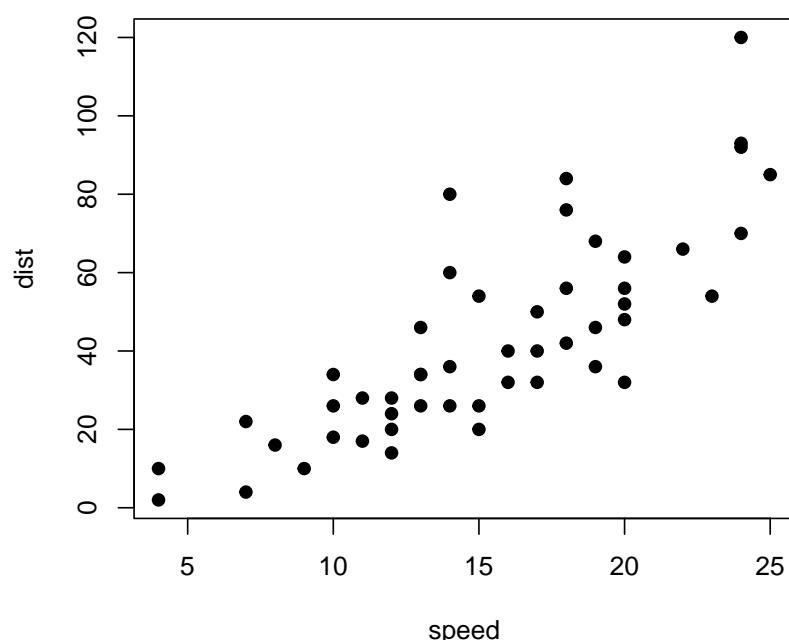
```
{r, echo=FALSE, fig.width=5, fig.height=4}
```



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.

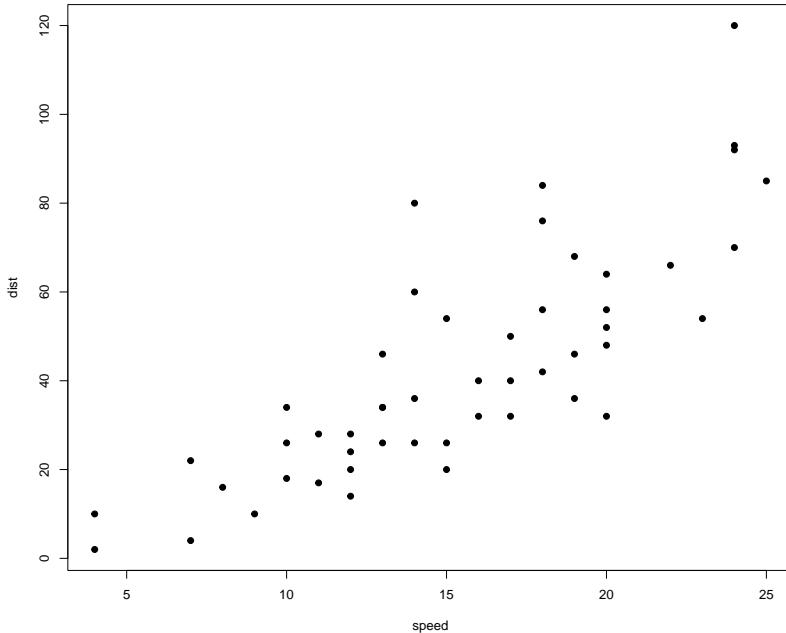
```
{r, echo=FALSE, fig.width=10, fig.height=8, out.width = "100%"}
```



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

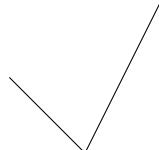
in output.yml add extra_dependencies: ["float"] under bookdown::pdf_book:

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

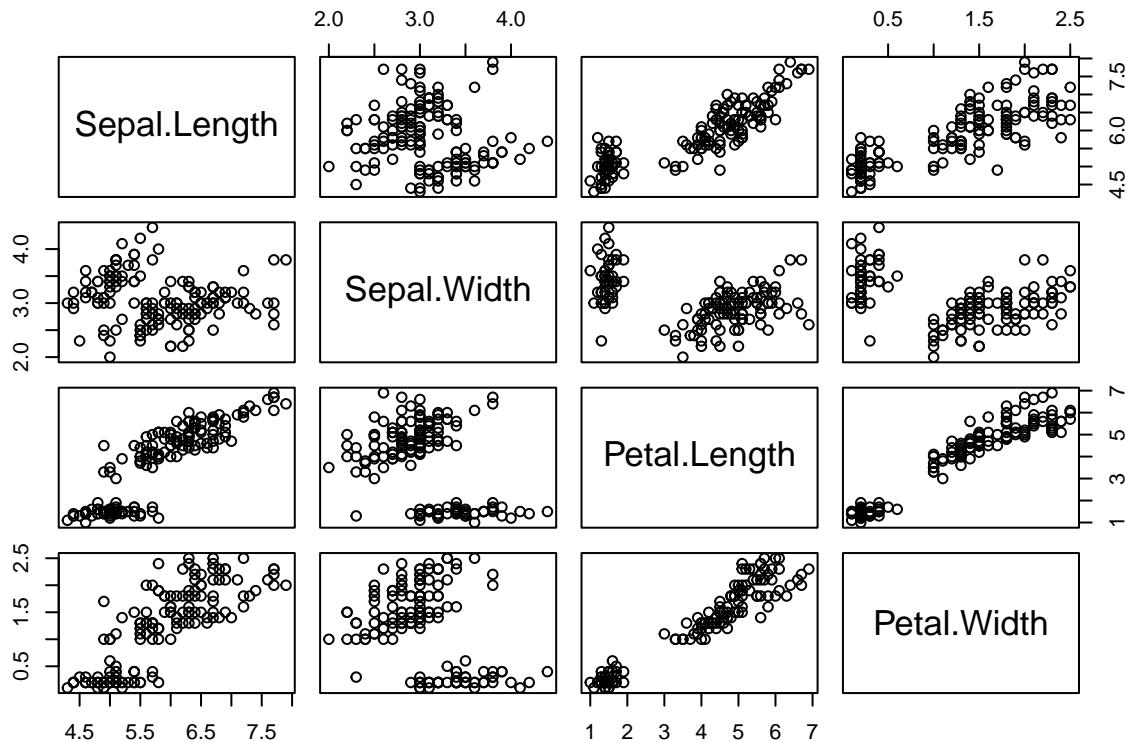
7.3.6.4.1 CSS flex Here is the first Div.

```
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])
```



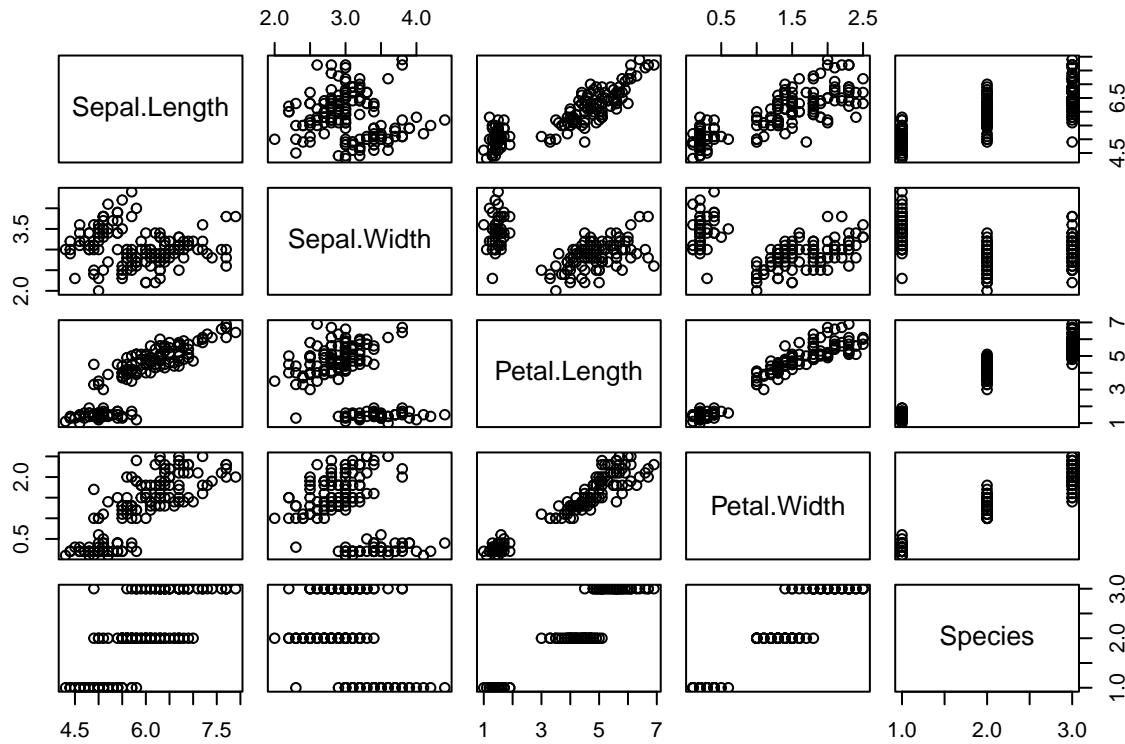
7.3.6.4.2 CSS grid <https://github.com/yihui/knitr/issues/1743>

<https://medium.com/enjoy-life-enjoy-coding/css-所以我说那个版能不能好切一点-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
```

```
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

R is an equivalence relation over $A \times B$

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

7.5 video embedding

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

always_allow_html: true

```
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`

$x = y$

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

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

$x = y$

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

$x = y$

```
:::: {show-in="html"}
$$
\frac{\colorbox{#FFD1DC}{$\epsilon^2}\left(y_{\{\scriptscriptstyle F\}}-y_{\{\scriptscriptstyle L\}}\right)^2}{1-\epsilon^2}
$$
::::
:::: {show-in="pdf"}
$$
\frac{\colorbox{red!50}{\text{\ensuremath{\epsilon^2}}}\left(y_{\{\scriptscriptstyle F\}}-y_{\{\scriptscriptstyle L\}}\right)^2}{1-\epsilon^2}
$$
::::
```

7.6.2 background color

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

LaTeX color

<https://latextcolor.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%20color%20package>

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>

$$F = ma$$

$$F = F$$

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

$$F = ma$$

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

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>

C is an equivalence class of a on A

$$\Leftrightarrow [a]_{\sim} = C = \left\{ x \left| \begin{array}{l} a \in A \\ x \in A \\ x \sim a \\ \sim \text{ is an equivalence relation over } A \times A = A^2 \end{array} \right. \right\} \subseteq A \neq \emptyset \quad (7.3)$$

$$\Leftrightarrow [a] = [a]_{\sim} = \left\{ x \left| \begin{array}{l} a \in A \\ x \in A \\ x \sim a \\ \sim \text{ is an equivalence relation on } A \end{array} \right. \right\} \subseteq A \neq \emptyset$$

$$\Rightarrow [a]_{\sim} = \{x | x \sim a\} \subseteq A \neq \emptyset$$

<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.
```

noun⁵

7.10 citation

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

citation 1³ citation 2³

citation 3⁴ citation 4⁴

⁵This 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	Maturity	Testing	Instrumentation	Regression	Selection	Mortality	Interaction of Selection and Maturity, etc.	Interaction of Testing and X	Interaction of Selection and X	Reactive Arrangements	Multiple-X Interference	
<i>Pre-Experimental Designs:</i>													
1. One-Shot Case Study <i>X O</i>	—	—			—	—			—				
2. One-Group Pretest-Posttest Design <i>O X O</i>	—	—	—	—	?	+	+	—	—	—	?		
3. Static-Group Comparison <i>X O</i> <i>O</i>	+	?	+	+	+	—	—	—	—				
<i>True Experimental Designs:</i>													
4. Pretest-Posttest Control Group Design <i>R O X O</i> <i>R O O</i>	+	+	+	+	+	+	+	+	—	?	?		
5. Solomon Four-Group Design <i>R O X O</i> <i>R O O</i> <i>R X O</i> <i>R O</i>	+	+	+	+	+	+	+	+	+	?	?		
6. Posttest-Only Control Group Design <i>R X O</i> <i>R O</i>	+	+	+	+	+	+	+	+	+	?	?		

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*

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

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

number and reference equations

(7.3)

(7.2)

7.2

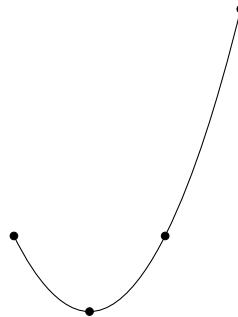


Figure 7.4: parabola arc with points

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=gkyjTepCITM>

<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

Chapter 9

partition

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

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

Chapter 10

equivalence class

C is an equivalence class of a on A

$$\Leftrightarrow [a]_{\sim} = C = \left\{ x \left| \begin{array}{l} a \in A \\ x \in A \\ x \sim a \\ \sim \text{ is an equivalence relation over } A \times A = A^2 \end{array} \right. \right\} \subseteq A \neq \emptyset$$
$$\Leftrightarrow [a] = [a]_{\sim} = \left\{ x \left| \begin{array}{l} a \in A \\ x \in A \\ x \sim a \\ \sim \text{ is an equivalence relation on } A \end{array} \right. \right\} \subseteq A \neq \emptyset$$
$$\Rightarrow [a]_{\sim} = \{x | x \sim a\} \subseteq A \neq \emptyset$$

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

Chapter 11

equivalence relation

等價關係 equivalence relation

R is an equivalence relation over $A \times B$

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

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"          "pgsql"        "Rscript"      "ruby"  
## [16] "sas"          "scala"        "sed"          "sh"          "stata"  
## [21] "zsh"          "asis"         "asy"          "block"        "block2"  
## [26] "bslib"        "c"            "cat"          "cc"          "comment"  
## [31] "css"          "dittaa"       "dot"          "embed"       "evviews"  
## [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

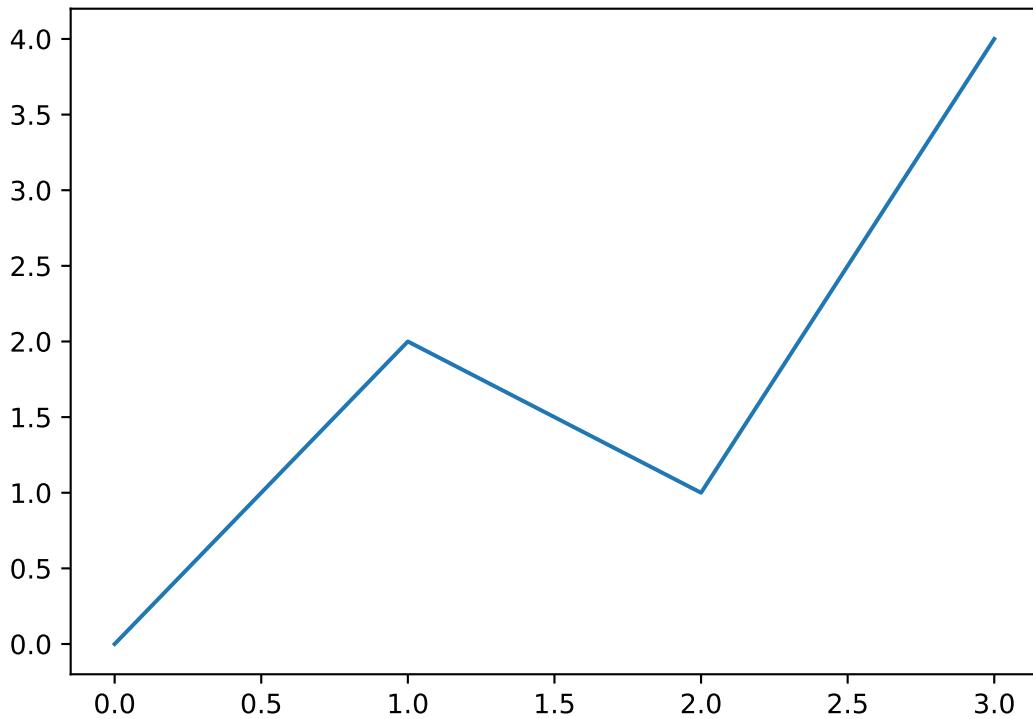
```
library(reticulate)  
version <- "3.9.12"  
# install_python(version)  
  
## create a new environment  
# virtualenv_create("r-reticulate", version = version)  
  
# use_virtualenv("r-reticulate")  
  
## install Matplotlib
```

```
# 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

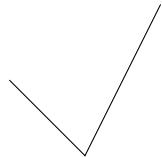
- TikZ
 - PGFplots^[13.4]
 - tikzplotlib^[13.5]: Python^[12] matplotlib^[27] export to TikZ^[13] .tex

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\aa{1.5} % amplitude
\def\bb{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 ({\aa * (\bb*\t -
sin(deg(\bb*\t)))},{-\aa * (1 -
cos(deg(\bb*\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}
```

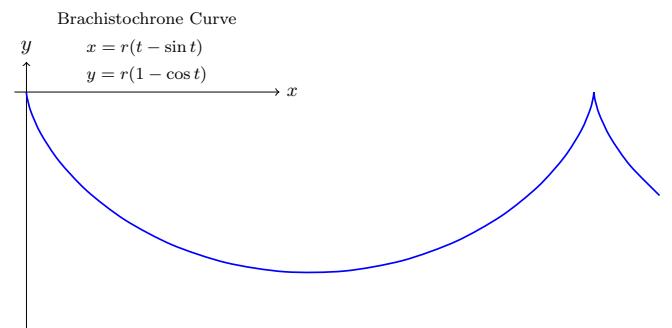


Figure 13.1: Brachistochrone Curve

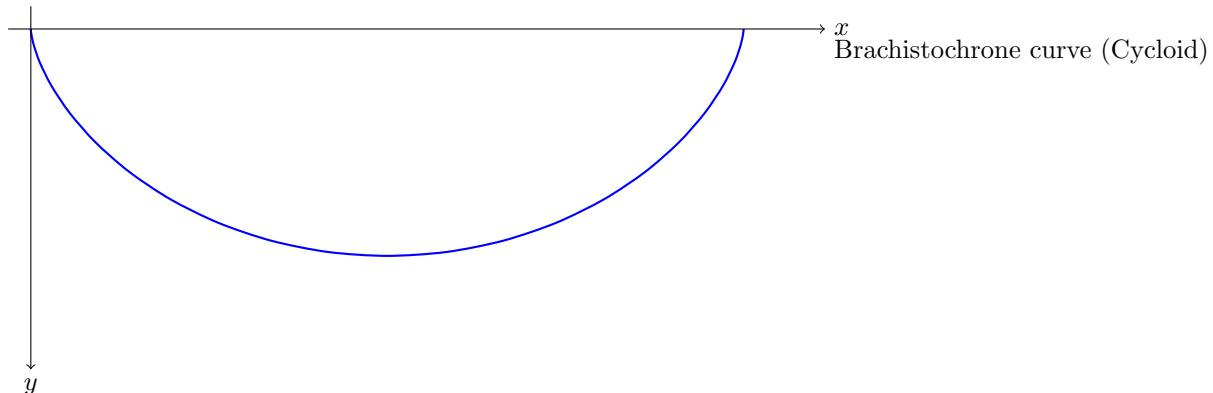


Figure 13.2: Brachistochrone Curve

13.1 2D

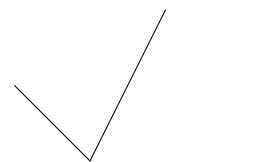
https://zhuanlan.zhihu.com/p/127155579?utm_psn=1741479950987960320

```
\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}
```

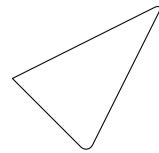


Figure 13.3: rounded corner pseudo-closed triangle

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

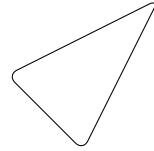


Figure 13.4: rounded corner triangle

```
\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}
```

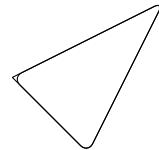


Figure 13.5: triangle vs. pseudo-closed triangle

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



Figure 13.6: rectangle

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

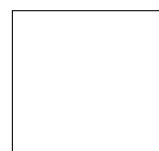


Figure 13.7: square

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

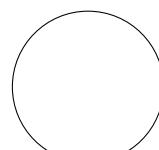


Figure 13.8: circle

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

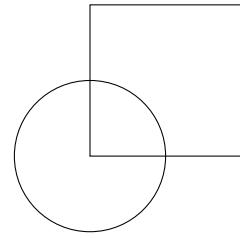


Figure 13.9: circle and square

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

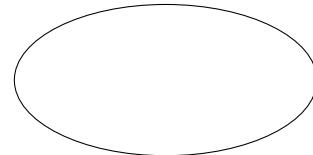


Figure 13.10: ellipse

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

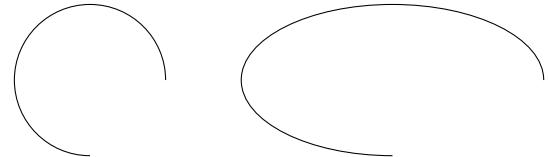


Figure 13.11: circle and ellipse arcs

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

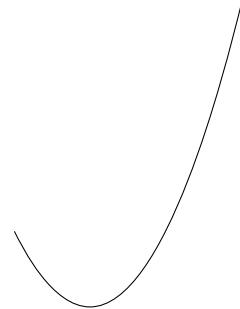


Figure 13.12: parabola arc

```
\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}
```

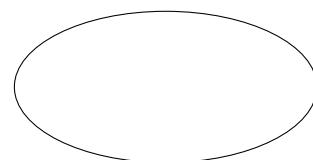


Figure 13.13: parabola arc with points

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

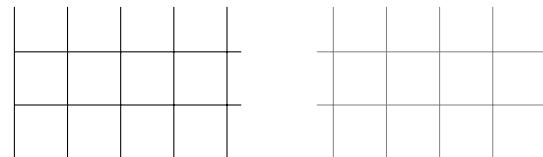


Figure 13.14: grid and help lines

```
\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}
```

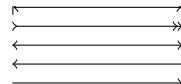


Figure 13.15: arrows

```
\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}
```

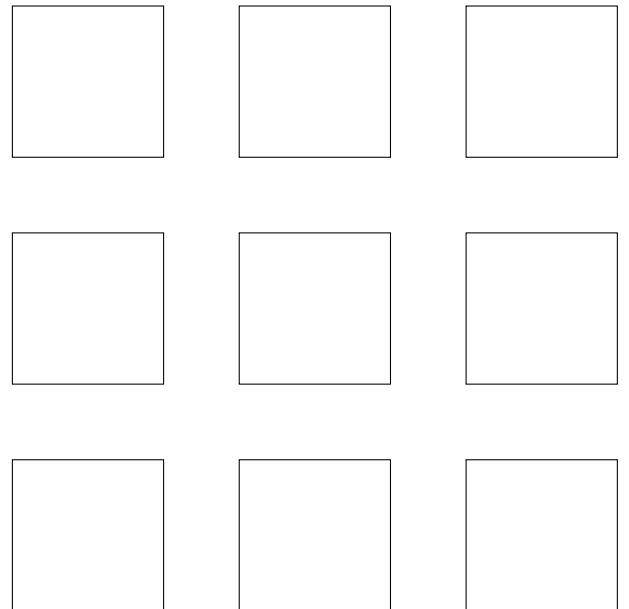


Figure 13.18: transform: shift

```
\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}
```

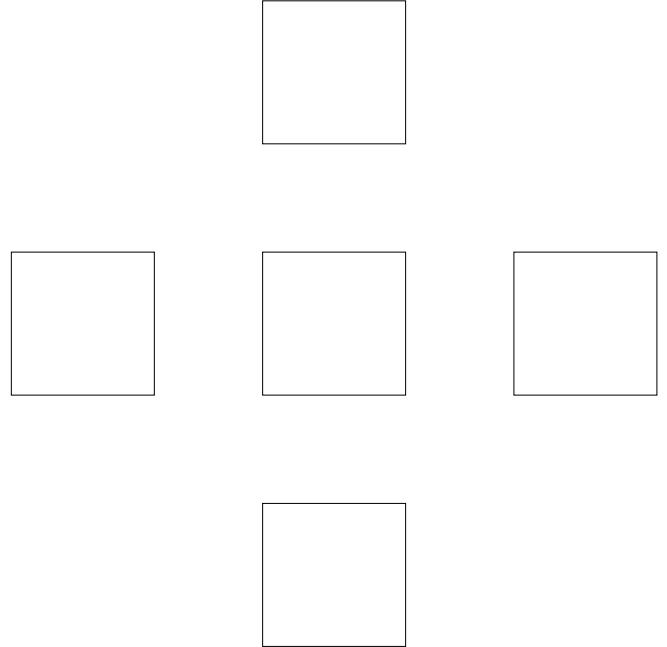


Figure 13.19: transform: shift x, y

```
\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}
```

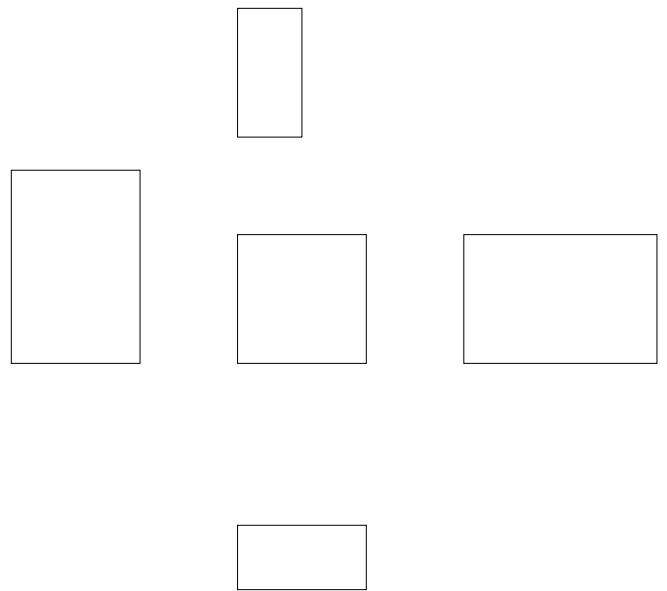


Figure 13.20: transform: scale x, y

```
\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}
```

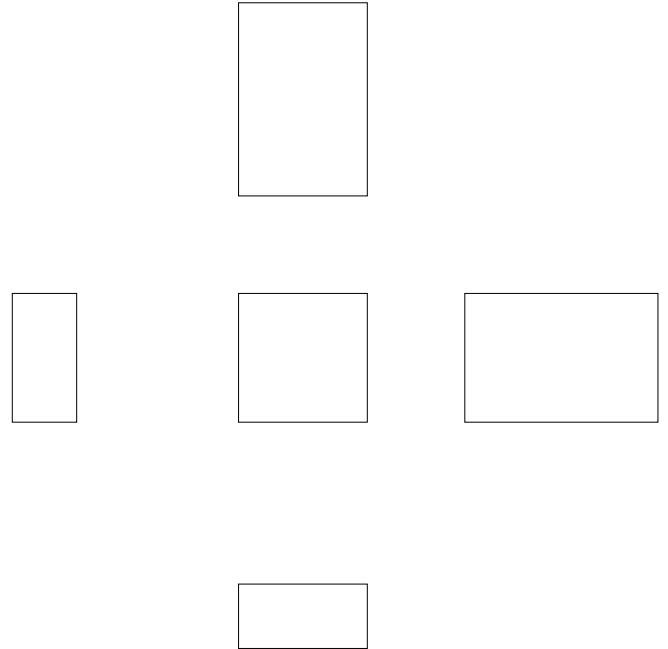


Figure 13.21: transform: scale

```
\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}
```

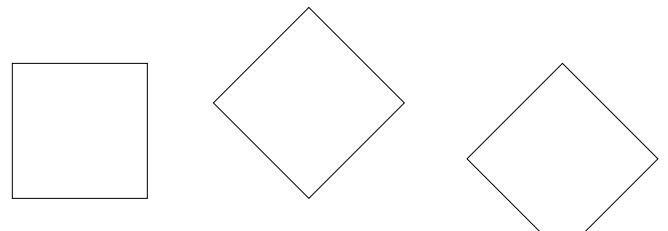


Figure 13.22: transform: rotate

```
\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}
```

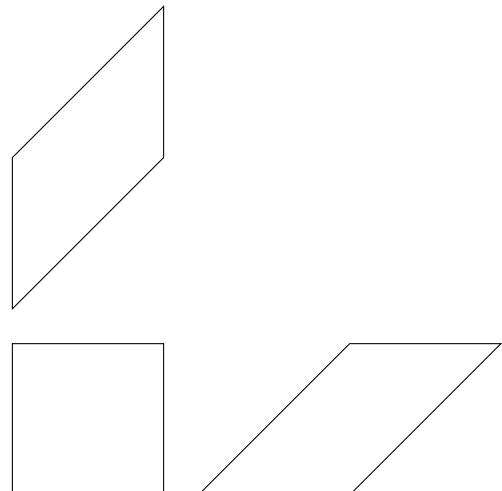


Figure 13.23: transform: slant

```
\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}
```

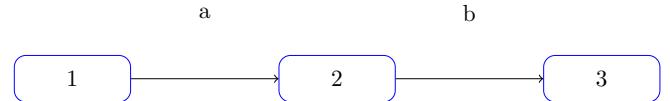


Figure 13.24: flowchart

```
\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}
```

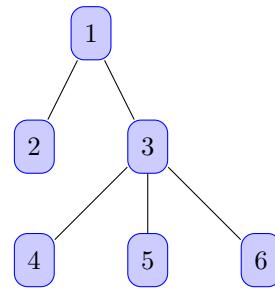


Figure 13.25: tree

```
\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}
```

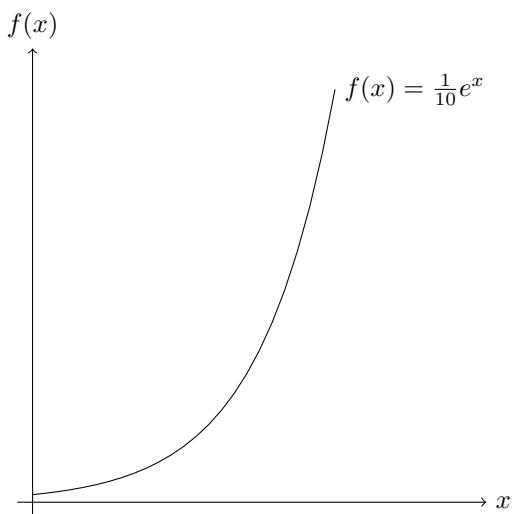


Figure 13.26: function plot

<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}
```

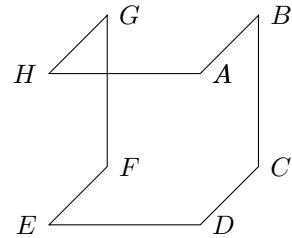


Figure 13.27: cube

```
\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}
```

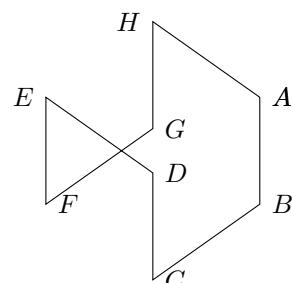


Figure 13.28: cube rotate

```
\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}
```

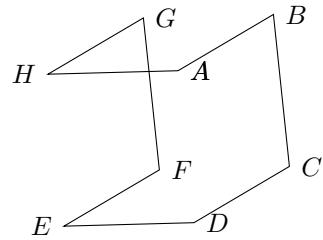


Figure 13.29: cube rotate around

<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}
```

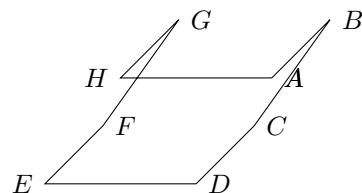


Figure 13.30: cube perspective slant

<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] {$f(x) = x$};
\draw[color=orange] plot (\x,{0.05*exp(\x)})
    node[right] {$f(x) = \frac{1}{20}e^x$};
\draw[color=blue] plot (\x,{sin(\x r)})
    node[right] {$f(x) = \sin x$};

\end{tikzpicture}
```

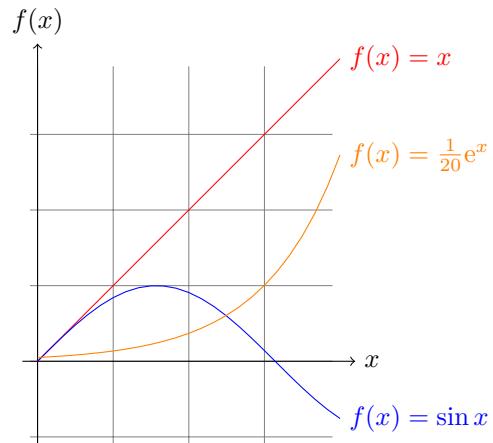


Figure 13.32: plots of functions

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

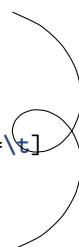


Figure 13.33: 2D parametric function

```
\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}
```

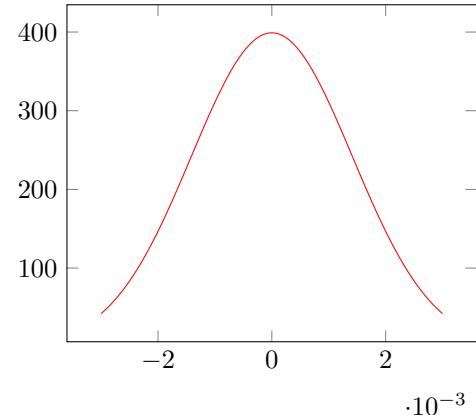


Figure 13.35: PGFplots: normal distribution

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}} $,
    $%
  ← \mu_{\scriptscriptstyle{0}} $,
    %extra x ticks={0},
    ytick = \emptyset,
    x = 1cm, y = 5cm, % x y scaling
    grid = major,
    domain = -10:10,
    samples = 1000
  ]
  \end{axis}
\end{tikzpicture}

```

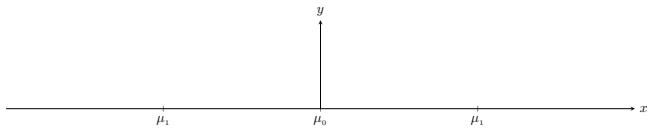


Figure 13.36: beginaxis

<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>

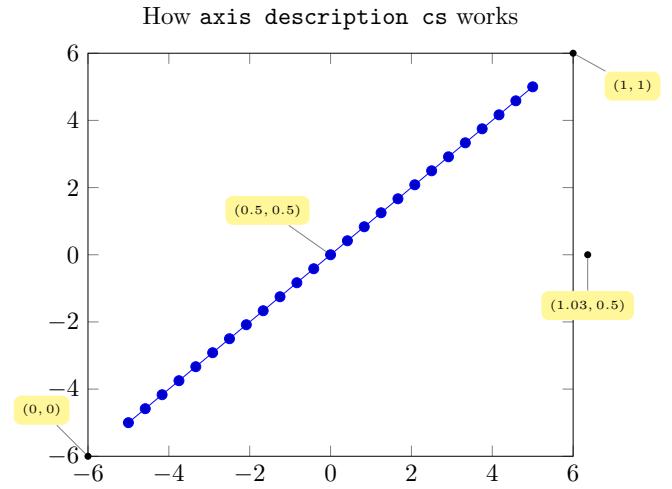
13.4.2.1 placement of axis descriptions

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

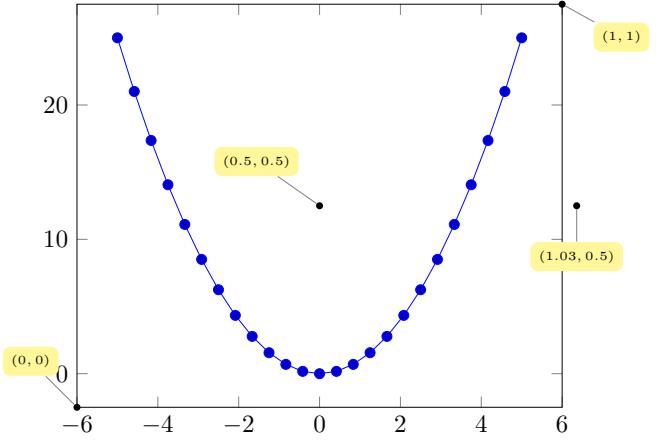
\addplot {x}; can change to \addplot {x^2}; still with auto blue dots

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}
```

How `axis description cs` worksFigure 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}
```

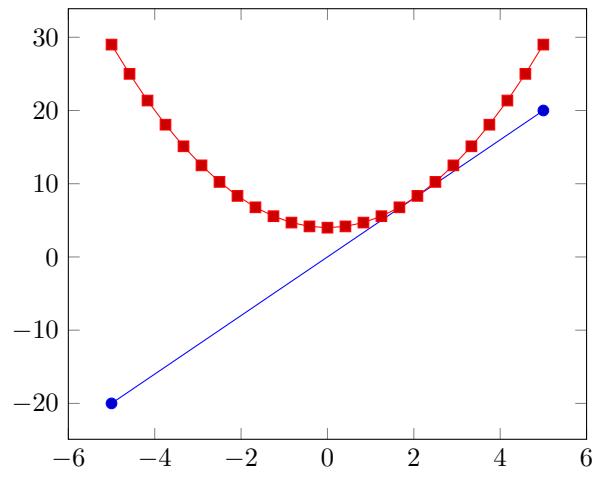


Figure 13.39: declare function

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 `\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 4} \$},
  %extra x ticks={0},
  ytick = \emptyset,
  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}
```

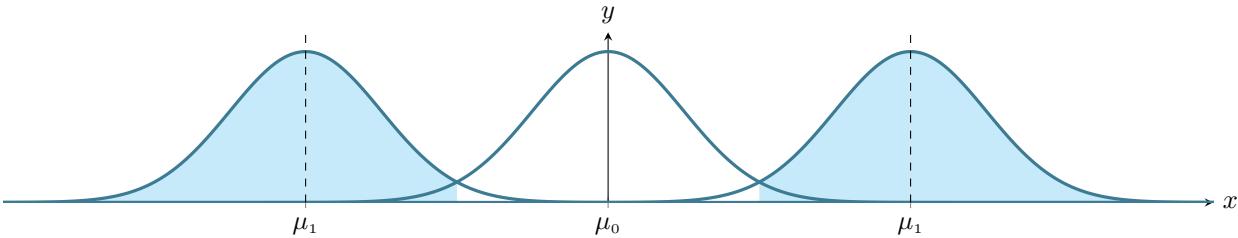


Figure 13.40: PGFmathDeclareFunction: normal distributions hypothesis testing

<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=\emptyset,
  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}
```

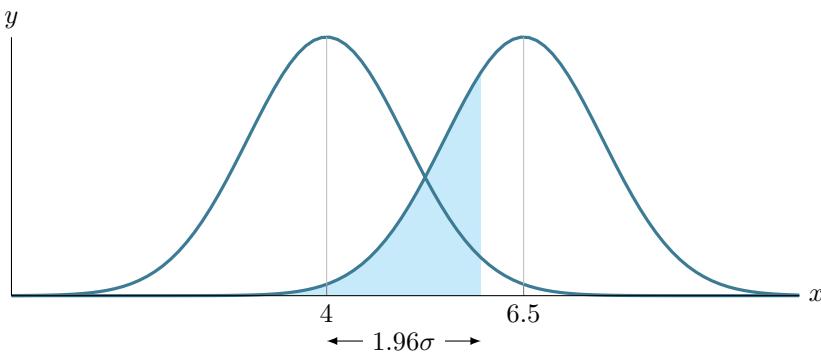


Figure 13.41: PGFmathDeclareFunction: normal distributions

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), 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), interp = interpolation
    coordinate[at start](interpa) coordinate[at end](interp);
\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}
```

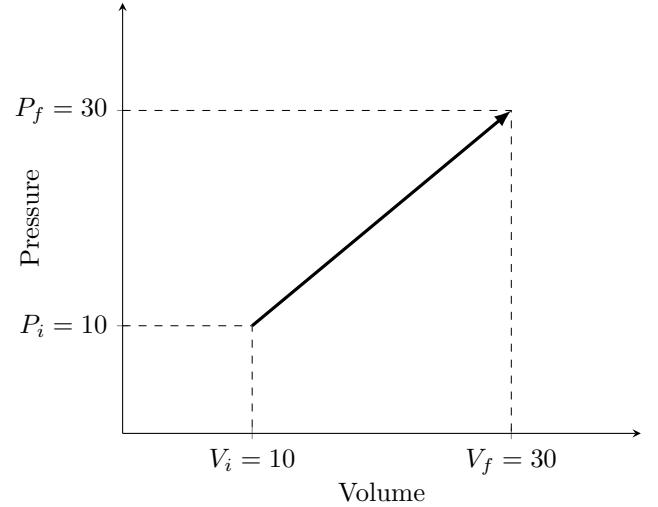


Figure 13.42: tick texts and interpolation dashed lines

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}
```

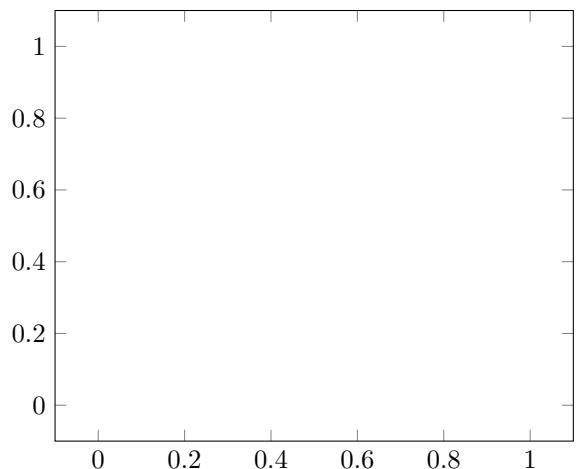


Figure 13.43: PGFplots: 2D axis/axes

13.4.7.1.1 range

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

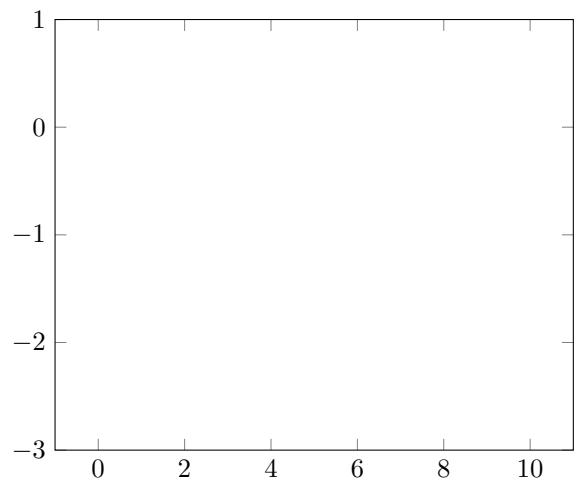


Figure 13.44: PGFplots: axis/axes range

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}
```

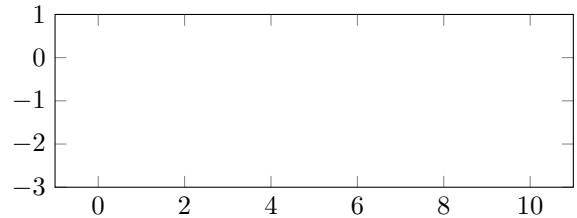


Figure 13.45: PGFplots: axis/axes range

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}
```

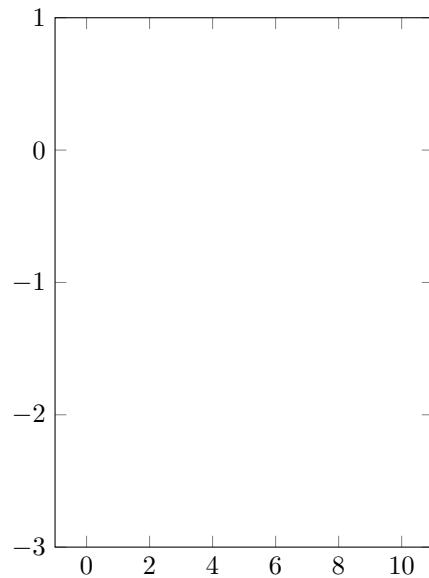


Figure 13.46: PGFplots: axis/axes range

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}
```

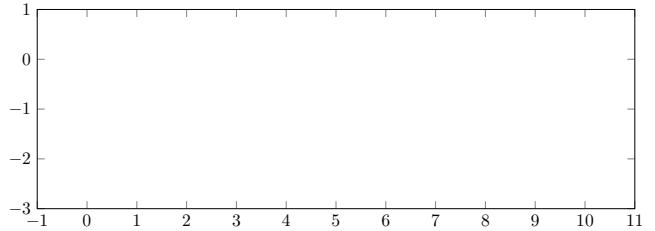


Figure 13.47: PGFplots: axis/axes range

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}
```

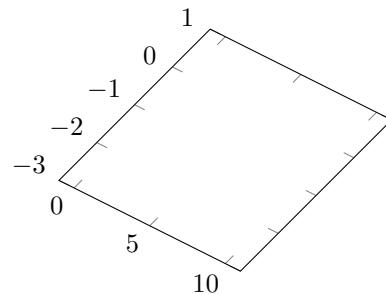


Figure 13.48: PGFplots: axis/axes range

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}
```

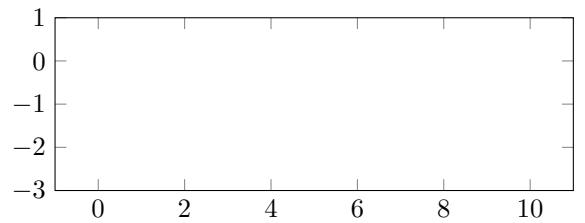


Figure 13.49: PGFplots: axis/axes range

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, `x:` bottom, top, `y:`), or `axis lines = none`(not shown), or even axis lines without 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}
```

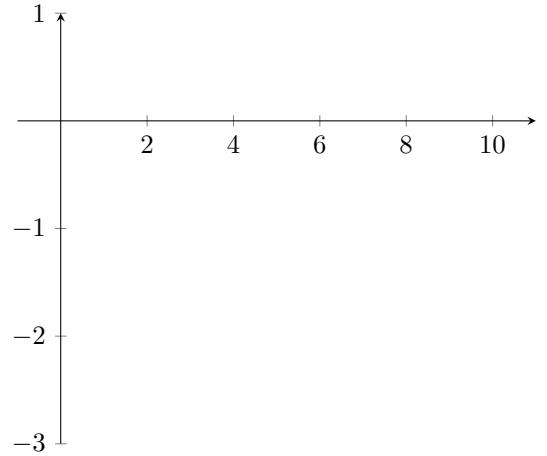


Figure 13.50: PGFplots: axis/axes range

`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}
```

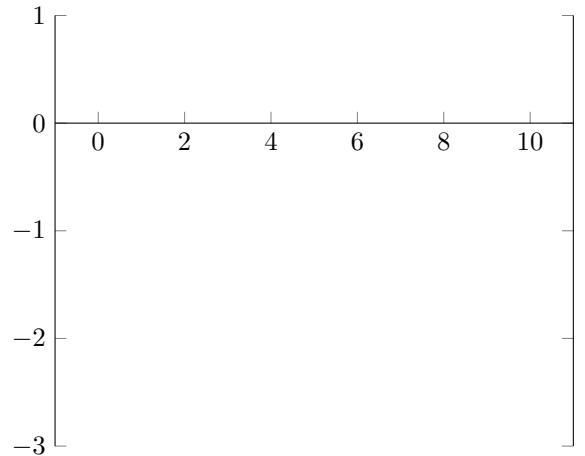


Figure 13.51: PGFplots: axis/axes range

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}
```

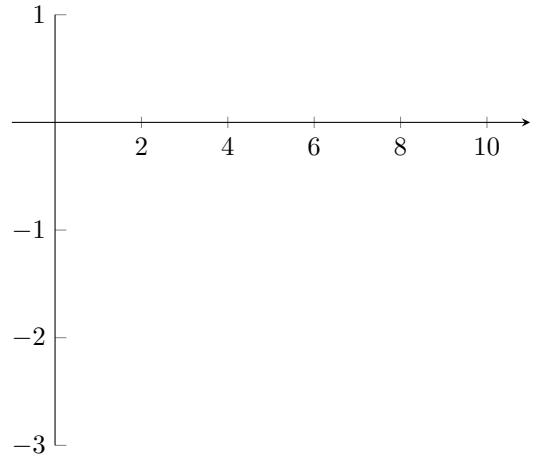


Figure 13.52: PGFplots: axis/axes range

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}
```

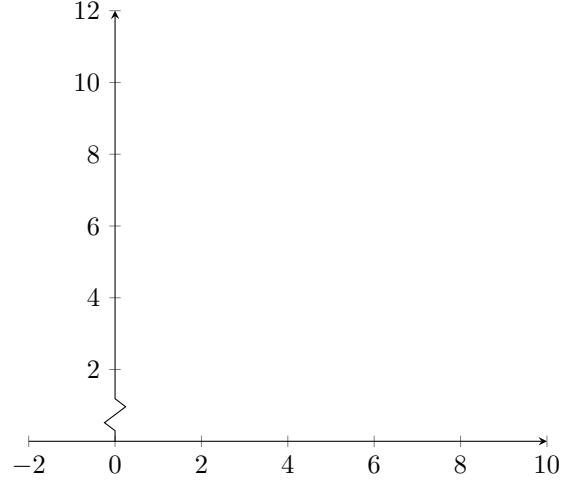


Figure 13.53: PGFplots: axis/axes range

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}
```

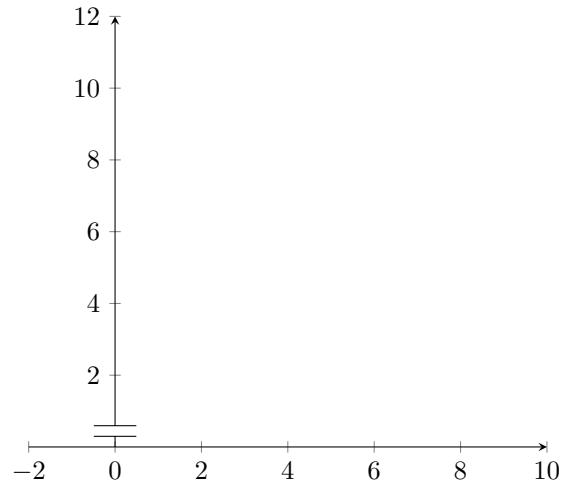


Figure 13.54: PGFplots: axis/axes range

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}
```

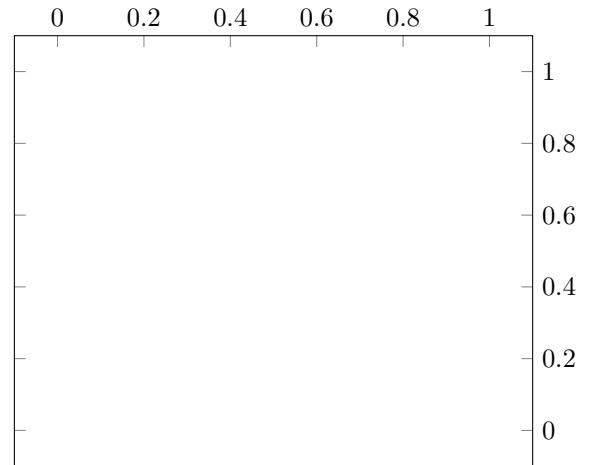


Figure 13.55: PGFplots: axis/axes range

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}
```

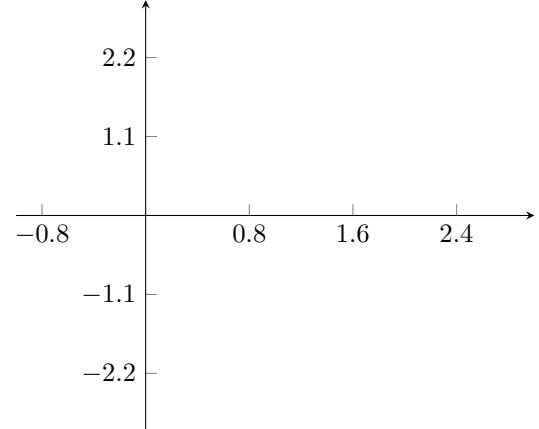


Figure 13.56: PGFplots: axis/axes range

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}
```

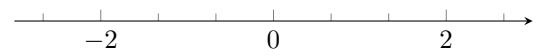


Figure 13.57: PGFplots: axis/axes range

xtick=\emptyset|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}
```

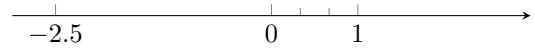


Figure 13.58: PGFplots: axis/axes range

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}
```

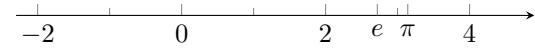


Figure 13.59: PGFplots: axis/axes range

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}
```

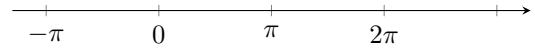


Figure 13.60: PGFplots: axis/axes range

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}
```

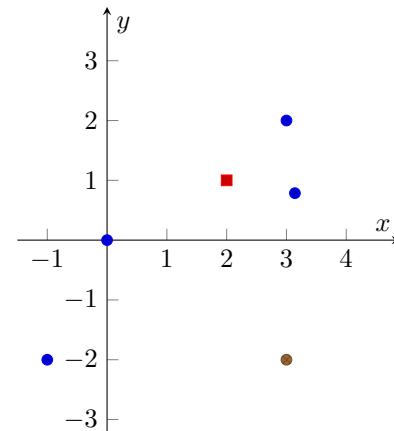


Figure 13.62: PGFplots: 2D points

```
\begin{tikzpicture}
\begin{axis}[axis equal image]
\addplot+ [only marks]
    table [x=xdata,y=ydata]
\end{axis}
\end{tikzpicture}
```

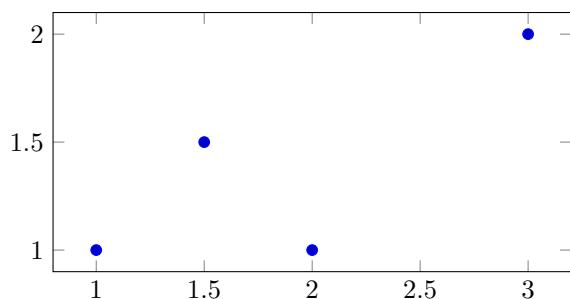


Figure 13.63: PGFplots: data points

```
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}
```

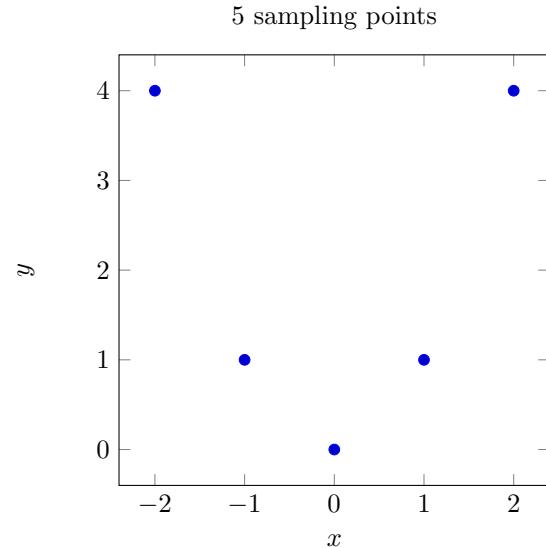


Figure 13.64: PGFplots: function sampling 5 points

```
\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}
```

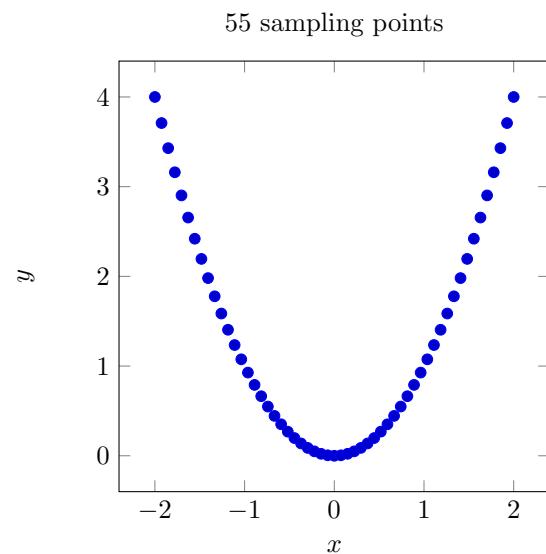


Figure 13.65: PGFplots: function sampling 55 points

```
\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}
```

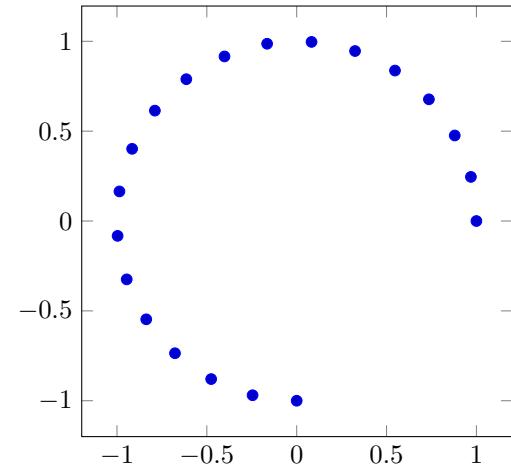


Figure 13.66: PGFplots: parametric function sampling

```
\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}
```

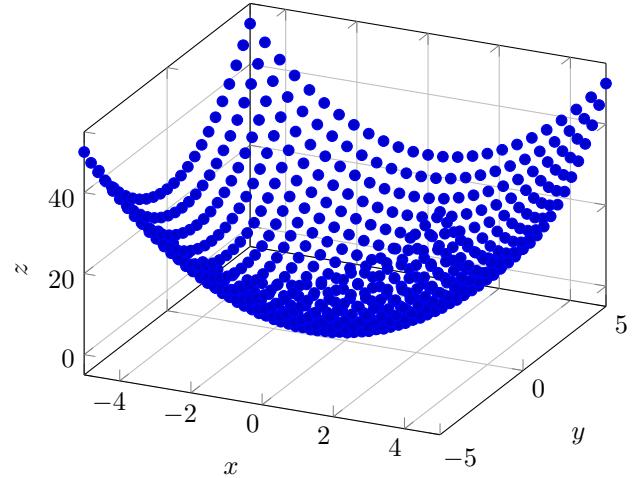


Figure 13.68: PGFplots: 3D function sampling points

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}
```

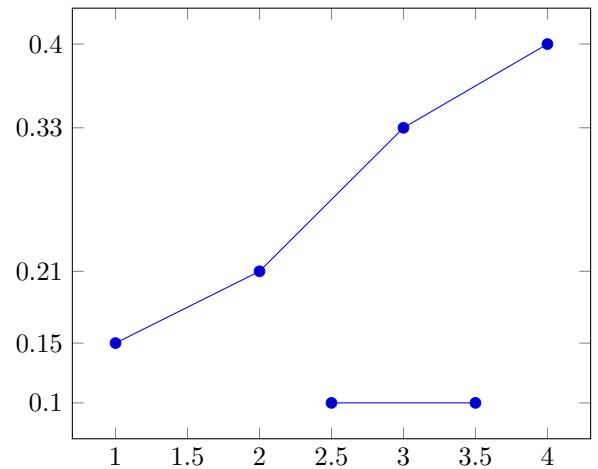


Figure 13.69: lines connecting adjacent points

`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}
```

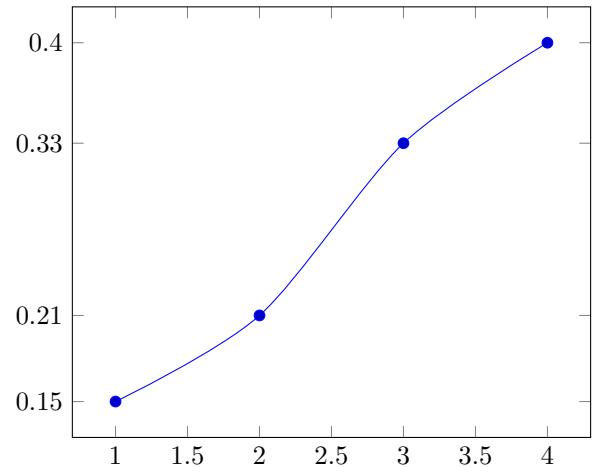
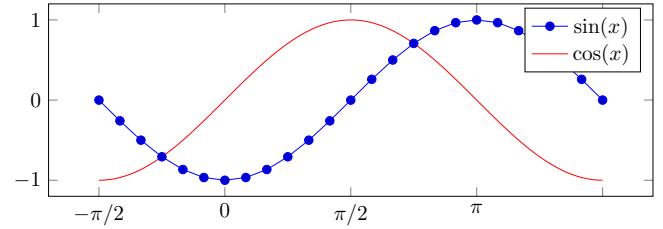


Figure 13.70: smooth lines connecting adjacent points

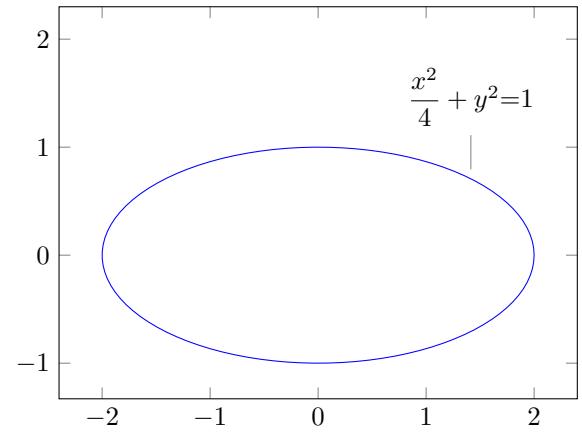
`no markers` to make no markers or points on the curves or lines

`\addlegendentry` to add legends

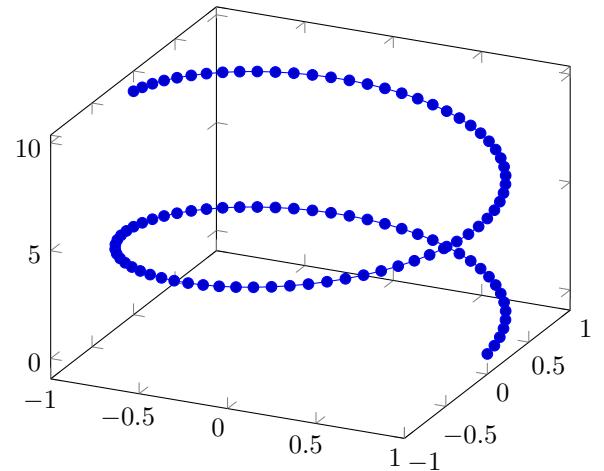
```
\begin{tikzpicture}
  \begin{axis}[
    trig format plots=rad,
    xtick distance=pi/2, ytick distance=1,
    xticklabels={-$-\pi$,$-\pi/2$,0,$\pi/2$,$\pi$},
    width=10cm, scale only axis,
    axis equal image]
    \addplot+ [domain=-pi:pi] {sin(x)};
    \addlegendentry{$\sin(x)$}
    \addplot+ [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:$\frac{x^2}{4}+y^2=1$]
      at ({2*cos(45)},{sin(45)}) {};
  \end{axis}
\end{tikzpicture}
```

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

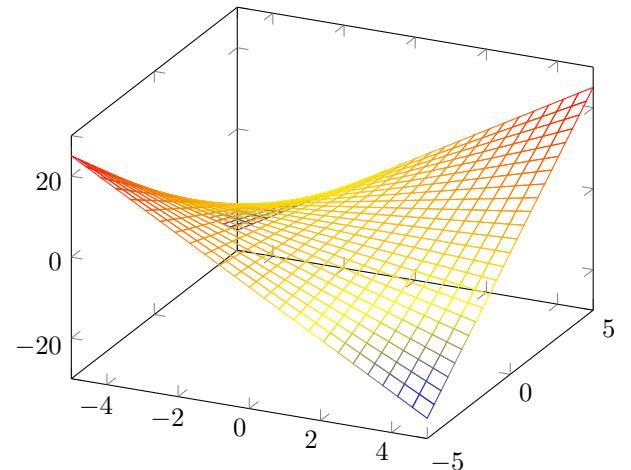
```
\begin{tikzpicture}
  \begin{axis}[trig format plots=rad]
    \addplot3+ [variable=t,
      domain=0:3*pi,
      samples=100,
      samples y=0]
      ({cos(t)},{sin(t)},{t});
  \end{axis}
\end{tikzpicture}
```

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

13.4.7.5 plane

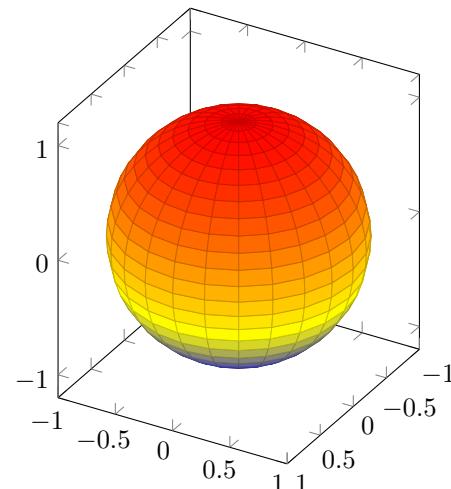
[mesh]

```
\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}
```

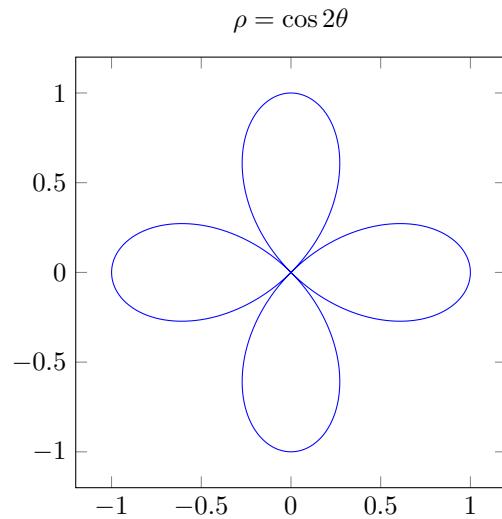


Figure 13.76: polar coordinate $\rho = \cos 2\theta$

\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}
```

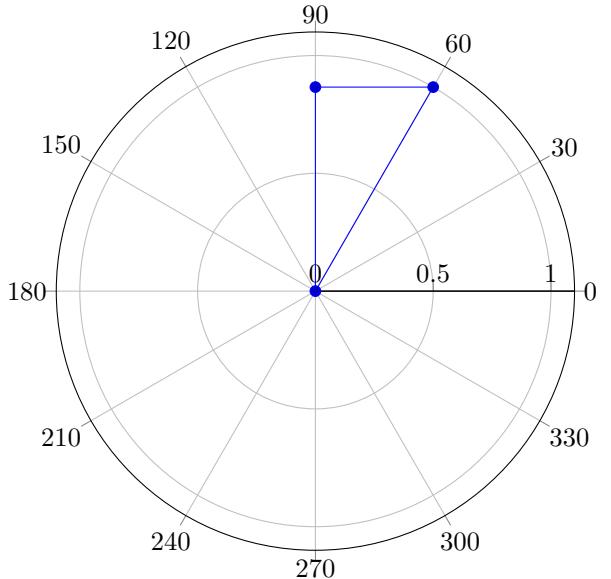


Figure 13.77: polar coordinate axes

<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-bcd8e27c0eb>

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

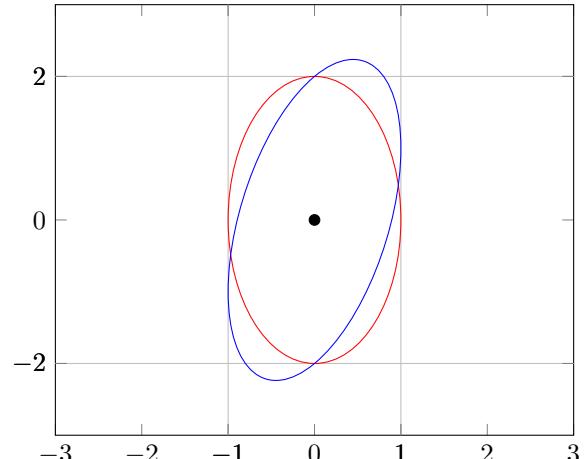


Figure 13.78: declare function

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()

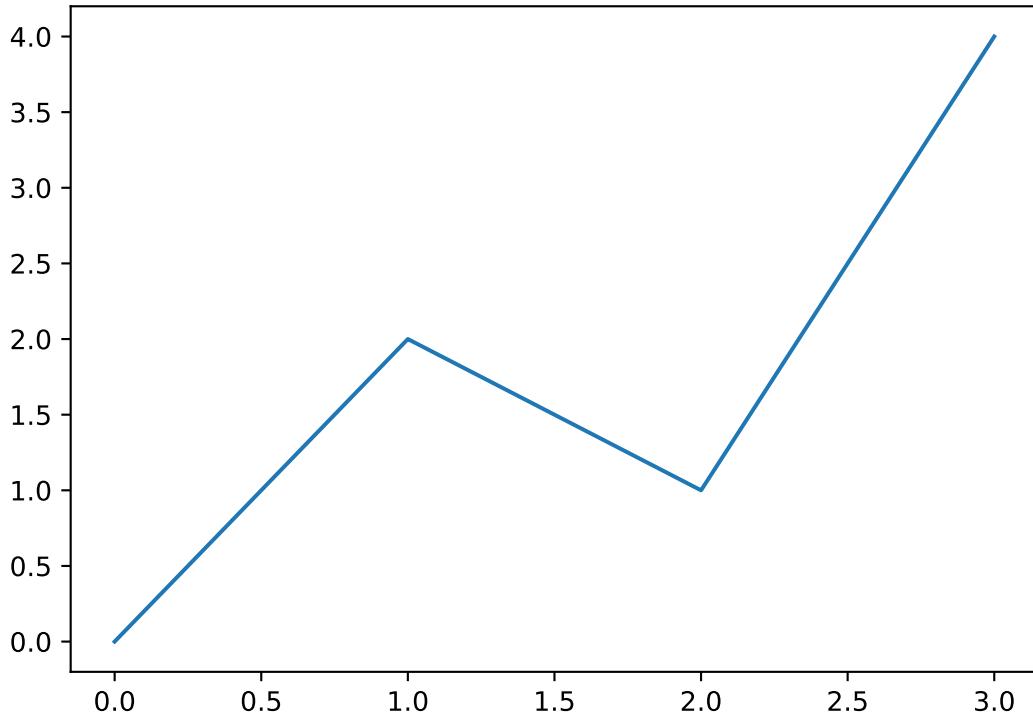
##          name           python
## 1      base      D:\\Anaconda3\\python.exe
## 2   fiftyone  D:\\Anaconda3\\envs\\fiftyone\\python.exe
## 3      keras      D:\\Anaconda3\\envs\\keras\\python.exe
## 4    labelme  D:\\Anaconda3\\envs\\labelme\\python.exe
## 5     manim      D:\\Anaconda3\\envs\\manim\\python.exe
## 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

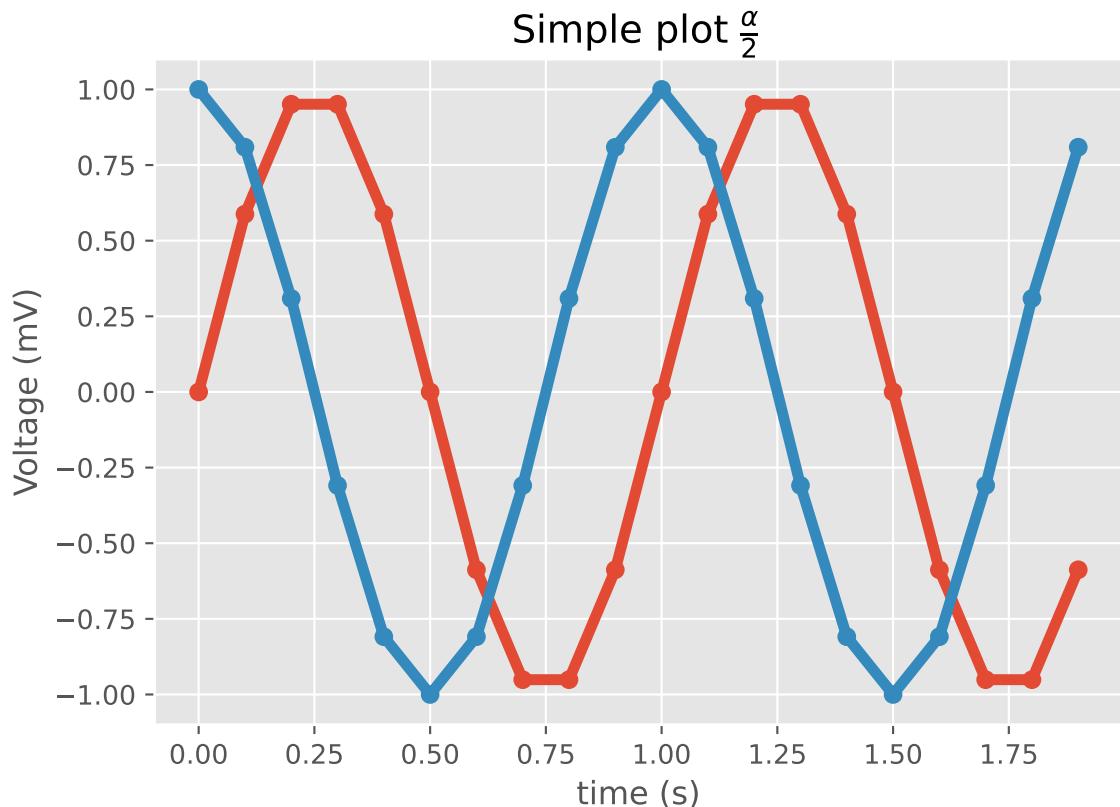
plt.close()

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

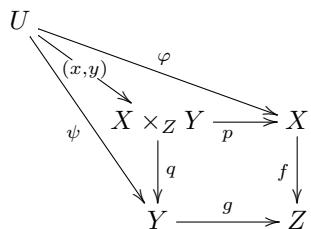
xy-pic

<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



Chapter 15

statistics

15.1 Hung Hung

population

普查 vs. 統計

random variable

X

sample has randomness

probability function

$$P_x(E) \in [0, 1]$$

event = subset of sample space

E

input event with output probability in 0 to 1

$$P_x : \{E_i\}_{i \in I} \rightarrow [0, 1]$$

target of interest

- probability function

but events are hard to be listed or enumerated

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

CDF = cumulative distribution function

$$F_x(x) = P_x((-\infty, x]) = P_x(X \leq x)$$

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

- $P_x(\cdot)$ PF = probability function
 - $f_x(x) = P_x(X = x)$ PMF = probability mass function
 - $f_x(x) = \frac{d}{dx} P_x(X \leq x)$ 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]

$$\begin{array}{ccccccc} X & \sim & F_x(x) & \xrightarrow{\text{FToC}} & f_x(x) & \leftrightarrow & P_x \\ \text{inversion formula :} & \downarrow & \forall \xi \approx 0 [M_x(\xi) \in \mathbb{R}] \wedge & \xleftarrow{\quad} & \uparrow \downarrow & \nwarrow \nearrow & \leftarrow \varphi \wedge \text{supp}(f_x) \text{ is bounded} \\ & \varphi_x(\xi) & & & M_x(\xi) & \rightarrow & \{\mu_n | n \in \mathbb{N}\} \end{array}$$

In population,

$$X \sim F_x(x)$$

by sampling,

$$X_1, \dots, X_i, \dots, X_n = X_i \sim F_x(x)$$

or

$$X_1, \dots, X_i, \dots, X_n = X_i \stackrel{\text{i.i.d.}}{\sim} F_x(x)$$

i.i.d. = independently identically distributed

and inference back

parametrically

$$\hat{X} \sim \hat{F}_x(x) = \hat{F}_x(x|\theta)$$

or nonparametrically

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

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

$$\hat{\Theta} \leftarrow (X_1, \dots, X_i, \dots, X_n) = T(X_1, \dots, X_n) = T(\dots, X_i, \dots) = T(X_i)$$

correspondng CDF for inference or estimation function of sampling random variables

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

wish to be unbiased and consistent

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

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

modeling or parameterizing with unknown parameter θ

$$F_x(x) \stackrel{M}{=} F_x(x|\theta) = F_x(x;\theta)$$

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

e.g. for normally distributed data

$$f_x(x) \stackrel{M}{=} f_x(x|\theta) = \frac{e^{-\frac{1}{2}(\frac{x-\mu}{\sigma})^2}}{\sqrt{2\pi\sigma^2}} = f_x(x|\mu, \sigma^2) = f_x(x|\mu, \sigma)$$

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. $f_x(x) \stackrel{M}{=} f_x(x|\theta) = f_x(x; \theta)$ various univariable distribution
3. $f_{\mathbf{x}}(\mathbf{x}) \stackrel{M}{=} f_{\mathbf{x}}(\mathbf{x}|\boldsymbol{\theta}) = f_{\mathbf{x}}(\mathbf{x}; \boldsymbol{\theta})$ multivariable distribution
4. $T(X_1, \dots, X_n)$ inference
 - point estimation

$$\hat{\mu} = \begin{cases} \bar{X} & \rightarrow \mu \\ \text{median}(X_1, \dots, X_n) & \rightarrow \mu \\ \vdots \end{cases}$$

- interval estimation = hypothesis testing

$$\begin{cases} H_0 : \theta = \theta_0 & \leftarrow T \in \{0, 1\} \\ H_1 : \theta \neq \theta_0 \end{cases}$$

5. how to find T
6. behavior of random function $T(X_1, \dots, X_n) = T \sim F_T(t)$
 - statistical properties of T
 - asymptotic properties

$$n \rightarrow \infty \begin{cases} \text{CLT} = \text{central limit theorem} \\ \text{LLN} = \text{law of large number} \end{cases}$$

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 probability⁶ p.72

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

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

tossing a dice

theorems

$$P(\emptyset) = 0$$

$$P(E) \leq 1$$

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

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

$$E_1 \subseteq E_2 \Rightarrow P(E_1) \leq P(E_2)$$

addition rule⁶ p.75 and extended addition rule⁶ p.76

inclusion-exclusion principle = sieve principle

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

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

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

symmetric difference⁶ p.75

union probability upper-bounded by sum of individual probability

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

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

Boole inequality

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

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

multiple hypothesis testing

How to control the family-wise error rate?

Ideally,

FWER = family-wise error rate

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

But condition H_0^j pairwise independent is too strong.

Practically,

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

Bonferroni correction

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

Bonferroni inequality⁶ p.77

Bonferroni inequality and Boole inequality are equivalent inequalities

birthday problem⁶ p.78

15.1.1.2 conditional probability

15.1.2 univariable distribution

$$\begin{cases} P_X(E) & \forall E \subseteq S \\ P_X(X \leq x) & \forall x \in \mathbb{R} \end{cases}$$

$$\begin{cases} P_x(X \in E) & \forall E \subseteq S \\ P_x(X \leq x) = P_x((-\infty, x]) = F_x(x) & \forall x \in \mathbb{R} \end{cases}$$

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

15.1.2.1 cumulative distribution function

CDF = cumulative distribution function

$$F_x(x) = P_x((-\infty, x]) = P_x(X \leq x)$$

$$X \sim P_x \leftrightarrow F_x(x)$$

$$X \sim F_x(x) \leftrightarrow P_x$$

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

$$F_x(x) = P_x((-\infty, x]) = P_x(X \leq x)$$

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

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

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

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

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

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

15.1.2.2 probability density function

$$\begin{cases} P_x(X \leq x) = P_x((-\infty, x]) = F_x(x) \\ P_x(X = x) = P_x(x) = ? \end{cases}$$

Definition 15.8. PDF = probability density function

PMF = probability mass function

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

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

$$\begin{cases} f_X(x) = \frac{d}{dx} F_X(x) & \Leftrightarrow F_X(x) = \int_{-\infty}^x f_X(t) dt \\ f_X(x) = F_X(x) - F_X(x^-) & \Leftrightarrow F_X(x) = \sum_{t \leq x} f_X(t) \end{cases}$$

$$\begin{cases} X \sim P_x \Leftrightarrow F_X(x) \leftrightarrow f_X(x) & \text{e.g. probability theory} \\ X \sim F_X(x) \leftrightarrow P_x & \Rightarrow F_X(x) \stackrel{M}{=} F_X(x|\theta) \text{ e.g. survival analysis} \\ X \sim f_X(x) \leftrightarrow F_X(x) \leftrightarrow P_x & \Rightarrow f_X(x) \stackrel{M}{=} f_X(x|\theta) \text{ e.g. general statistics} \end{cases}$$

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

$$\begin{cases} \forall x \in \mathbb{R} [f(x) \geq 0] \\ \begin{cases} \int_{-\infty}^{+\infty} f(x) dx = 1 & t \text{ continuous} \\ \sum_{x \in X(\Omega)} f(x) = 1 & t \text{ discrete} \end{cases} \end{cases}$$

$$\forall E \subseteq S \left[P_x(X \in E) = \begin{cases} \int_{x \in E} f_X(x) dx & X \text{ continuous} \\ \sum_{x \in E} f_X(x) & X \text{ discrete} \end{cases} \right]$$

$$\begin{aligned} P_x(X = x) &= \lim_{\epsilon \rightarrow 0} P_x([x - \epsilon, x + \epsilon]) \\ &= \lim_{\epsilon \rightarrow 0} P_x(x - \epsilon \leq X \leq x + \epsilon) \\ &= \lim_{\epsilon \rightarrow 0} [F_X(x + \epsilon) - F_X(x - \epsilon)] \\ &= \begin{cases} F_X(x) - F_X(x) = 0 & X \text{ continuous} \\ F_X(x) - F_X(x^-) = f_X(x) & X \text{ discrete} \end{cases} \end{aligned}$$

$$X \sim F_X(x) \leftrightarrow P_x$$

$$Y = g(X)$$

$$\begin{cases} Y \sim F_Y(y) \leftrightarrow f_Y(y) & \Rightarrow F_Y(y) \stackrel{M}{=} F_Y(y|\theta) \\ Y \sim f_Y(y) \leftrightarrow F_Y(y) \leftrightarrow P_Y & \Rightarrow f_Y(y) \stackrel{M}{=} f_Y(y|\theta) \end{cases}$$

15.1.2.3 range vs. support

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

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

Definition 15.10. support

$$\text{supp}(f) = \left\{ x \left| \begin{array}{l} f : D \rightarrow \mathcal{R} \\ x \in D \\ f_x(x) \neq 0 \end{array} \right. \right\}$$

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

$$\text{supp}(f_x) = \left\{ x \left| \begin{array}{l} x \in X(\Omega) \\ f_x(x) \neq 0 \end{array} \right. \right\} \stackrel{f_x(x) \geq 0}{=} \left\{ x \left| \begin{array}{l} x \in X(\Omega) \\ f_x(x) > 0 \end{array} \right. \right\}$$

15.1.2.4 continuous monotone transformation

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

$$f_Y(y) = f_x(g^{-1}(y)) \left| \frac{dg^{-1}(y)}{dy} \right|$$

Proof:

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

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

□

segment $g(X)$ into monotone functions

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

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

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

$$\begin{aligned}
F_Y(y) &= P_Y(Y \leq y) = P(X^2 \leq y) \\
&= P(\{X^2 \leq y\} \cap (\{X < 0\} \cup \{X \geq 0\})) \\
&= P((\{X^2 \leq y\} \cap \{X < 0\}) \cup (\{X^2 \leq y\} \cap \{X \geq 0\})) \\
&= P(\{X^2 \leq y\} \cap \{X < 0\}) + P(\{X^2 \leq y\} \cap \{X \geq 0\}) \\
&\quad - P((\{X^2 \leq y\} \cap \{X < 0\}) \cap (\{X^2 \leq y\} \cap \{X \geq 0\})) \\
&= P(\{X^2 \leq y\} \cap \{X < 0\}) + P(\{X^2 \leq y\} \cap \{X \geq 0\}) - P(\emptyset) \\
&= P(\{X^2 \leq y\} \cap \{X < 0\}) + P(\{X^2 \leq y\} \cap \{X \geq 0\}) - 0 \\
&= P(\{X^2 \leq y\} \cap \{X < 0\}) + P(\{X^2 \leq y\} \cap \{X \geq 0\}) \\
&= P(\{-X \leq \sqrt{y}\} \cap \{X < 0\}) + P(\{X \leq \sqrt{y}\} \cap \{X \geq 0\}) \\
&= P(\{X \geq -\sqrt{y}\} \cap \{X < 0\}) + P(\{X \leq \sqrt{y}\} \cap \{X \geq 0\}) \\
&= P_X(-\sqrt{y} \leq X < 0) + P_X(0 \leq X \leq \sqrt{y}) \\
&= [F_X(0) - F_X(-\sqrt{y})] + [F_X(\sqrt{y}) - F_X(0)] \\
&= F_X(\sqrt{y}) - F_X(-\sqrt{y})
\end{aligned}$$

□

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

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

$$\begin{aligned}
F_Y(y) &= P_Y(Y \leq y) = P(X^2 \leq y) \begin{cases} Y = g(X) = X^2 \\ X \in [-1, \infty) \end{cases} \\
&= P(\{X^2 \leq y\} \cap (\{X < 0\} \cup \{X \geq 0\})) = \dots \text{ as } \begin{cases} Y = g(X) = X^2 \\ X \in (-\infty, +\infty) \end{cases} \\
&= P(\{X^2 \leq y\} \cap \{X < 0\}) + P(\{X^2 \leq y\} \cap \{X \geq 0\}) \\
&= P(\{-X \leq \sqrt{y}\} \cap \{X < 0\}) + P(\{X \leq \sqrt{y}\} \cap \{X \geq 0\}) \\
&= \begin{cases} P(\{-X \leq \sqrt{y}\} \cap \{X < 0\} \cap \{X > 1\}) + P(\{X \leq \sqrt{y}\} \cap \{X \geq 0\} \cap \{X > 1\}) & Y \in (1, \infty) \Rightarrow X \in (1, \infty) \\ P(\{-X \leq \sqrt{y}\} \cap \{X < 0\} \cap \{X \geq -1\}) + P(\{X \leq \sqrt{y}\} \cap \{X \geq 0\} \cap \{X \leq 1\}) & Y \in [0, 1] \Rightarrow \begin{cases} X \in [0, 1] \\ X \in [-1, 0] \end{cases} \end{cases} \\
&= \begin{cases} P(\emptyset) + P_x(1 < X \leq \sqrt{y}) & Y \in (1, \infty) \Rightarrow X \in (1, \infty) \\ P(\{X \geq -\sqrt{y}\} \cap \{X < 0\} \cap \{X \geq -1\}) + P_x(0 \leq X \leq \min\{\sqrt{y}, 1\}) & Y \in [0, 1] \Rightarrow \begin{cases} X \in [0, 1] \\ X \in [-1, 0] \end{cases} \end{cases} \\
&= \begin{cases} 0 + [F_x(\sqrt{y}) - F_x(1)] & Y \in (1, \infty) \Rightarrow X \in (1, \infty) \\ P_x(\max\{-1, -\sqrt{y}\} \leq X < 0) + P_x(0 \leq X \leq \sqrt{y}) & Y \in [0, 1] \Rightarrow \begin{cases} X \in [0, 1] \\ X \in [-1, 0] \end{cases} \end{cases} \\
&= \begin{cases} F_x(\sqrt{y}) - F_x(1) & Y \in (1, \infty) \Rightarrow X \in (1, \infty) \\ P_x(-\sqrt{y} \leq X < 0) + [F_x(\sqrt{y}) - F_x(0)] & Y \in [0, 1] \Rightarrow \begin{cases} X \in [0, 1] \\ X \in [-1, 0] \end{cases} \end{cases} \\
&= \begin{cases} F_x(\sqrt{y}) - F_x(1) & y > 1 \\ F_x(\sqrt{y}) - F_x(-\sqrt{y}) & -1 \leq y \leq 1 \end{cases}
\end{aligned}$$

□

15.1.2.5 discrete monotone transformation

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

$$\begin{aligned}
f_Y(y) &= P_Y(Y = y) \\
&= P(X^2 = y) \\
&= P(\{X = \sqrt{y}\} \cup \{X = -\sqrt{y}\}) \\
&= P(\{X = \sqrt{y}\}) + P(\{X = -\sqrt{y}\}) - P(\{X = \sqrt{y}\} \cap \{X = -\sqrt{y}\}) \\
&= P_x(X = \sqrt{y}) + P_x(X = -\sqrt{y}) - P(\emptyset) \\
&= P_x(X = \sqrt{y}) + P_x(X = -\sqrt{y}) - 0 \\
&= P_x(X = \sqrt{y}) + P_x(X = -\sqrt{y}) \\
&= f_x(\sqrt{y}) + f_x(-\sqrt{y})
\end{aligned}$$

□

Theorem 15.4. *discrete monotone transformation*

$$\begin{aligned}
&\begin{cases} Y = g(X) \\ X \text{ discrete} \Rightarrow Y \text{ discrete} \end{cases} \\
&\Downarrow \\
f_Y(y) &= \sum_{\{x|g(x)=y\}} f_x(x) = \sum_{\{x|x=g^{-1}(y)\}} f_x(x)
\end{aligned}$$

Proof:

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

□

Theorem 15.5. probability integral transformation

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

Proof:

$$\begin{aligned}
F_Y(y) &= \text{P}_Y(Y \leq y) \stackrel{(t)}{=} \text{P}(F_X(X) \leq y), \forall x_1 < x_2 [F_X(x_1) < F_X(x_2)] \Rightarrow \left\{ \begin{array}{l} \exists F_X^{-1} : Y \rightarrow X \\ \forall y_1 < y_2 [F_X^{-1}(y_1) < F_X^{-1}(y_2)] \end{array} \right. \\
&= \text{P}_X(X \leq F_X^{-1}(y) = x) = \text{P}_X(X \leq x), x = F_X^{-1}(y) \\
&= \text{P}_X(X \leq x) = F_X(x) \stackrel{x=F_X^{-1}(y)}{\equiv} F_X(F_X^{-1}(y)) = y \\
F_Y(y) &= y
\end{aligned}$$

□

Note:

According to Theorem 15.5,

$$\begin{aligned}
U &= F_X(X) \stackrel{15.5}{\sim} U(u) \text{ on } [0, 1] \\
\Rightarrow X &= F_X^{-1}(U) \wedge X \sim F_X(x) \Rightarrow F_X^{-1}(U) = X \sim F_X(x) \Rightarrow F_X^{-1}(U) \sim F_X(x) \\
\Rightarrow X &= F_X^{-1}(U) \sim F_X(x) \\
&\text{i.e. uniform random variables substituted into the inverse of } F_X, \\
&\text{we can get random variables follow } F_X(x)
\end{aligned}$$

15.1.2.6 expected value

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

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

$$\text{E}(g(X)) = \text{E}[g(X)] = \begin{cases} \int_{-\infty}^{+\infty} g(x) f_X(x) dx & X \text{ continuous} \\ \sum_{x \in X(\Omega)} g(x) f_X(x) & X \text{ discrete} \end{cases}$$

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

$$E(X) = E[X] = \begin{cases} \int_{-\infty}^{+\infty} xf_X(x) dx = 1 & X \text{ continuous} \\ \sum_{x \in X(\Omega)} xf_X(x) = 1 & X \text{ discrete} \end{cases}$$

Theorem 15.6. *the rule of the lazy statistician*

the law of the unconscious statistician = the LOTUS

$$E(g(X)) = E[g(X)] = \begin{cases} \int_{-\infty}^{+\infty} g(x) f_X(x) dx & X \text{ continuous} \\ \sum_{x \in X(\Omega)} g(x) f_X(x) & X \text{ discrete} \end{cases}$$

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,

- $E[a_1g_1(X_1) + a_2g_2(X_2) + c] = a_1E[g_1(X_1)] + a_2E[g_2(X_2)] + c$
- $\forall x \in \mathbb{R} [g(x) \geq 0] \Rightarrow E[g(X)] \geq 0$
- $\forall x \in \mathbb{R} [g_1(x) \geq g_2(x)] \Rightarrow E[g_1(X)] \geq E[g_2(X)]$
- $\forall x \in \mathbb{R} [a \leq g(x) \leq b] \Rightarrow a \leq E[g(X)] \leq b$

Theorem 15.7. $E[X]$ minimizes Euclidean distance $E[(X - b)^2]$ over b , i.e.

$$E[X] = \arg \min_b E[(X - b)^2]$$

Proof:

$$\begin{aligned}
E[(X - b)^2] &= E[(X - E[X] + E[X] - b)^2] \\
&= E[\{(X - E[X]) + (E[X] - b)\}^2] \\
&= E[(X - E[X])^2 + 2(X - E[X])(E[X] - b) + (E[X] - b)^2] \\
&= E[(X - E[X])^2] + 2(E[X] - b)E[(X - E[X])] + E[(E[X] - b)^2] \\
&= E[(X - E[X])^2] + 2(E[X] - b)E[X - E[X]] + (E[X] - b)^2 \\
&= E[(X - E[X])^2] + 2(E[X] - b)(E[X] - E[X]) + (E[X] - b)^2 \\
&= E[(X - E[X])^2] + 2(E[X] - b)0 + (E[X] - b)^2 \\
&= E[(X - E[X])^2] + 0 + (E[X] - b)^2 \\
&= E[(X - E[X])^2] + (E[X] - b)^2 \stackrel{(E[X]-b)^2 \geq 0}{\geq} E[(X - E[X])^2]
\end{aligned}$$

\Downarrow

$$E[(X - b)^2] \geq E[(X - E[X])^2]$$

\Downarrow

$$E[(X - b)^2] = E[(X - E[X])^2] \text{ holds if } (E[X] - b)^2 = 0 \Rightarrow b = E[X] \Rightarrow E[X] = \arg \min_b E[(X - b)^2]$$

□

Note:

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

Theorem 15.8. median $[X]$ minimizes $E[|X - b|]$ over b , i.e.

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

Proof:

to be proved

□

Note:

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

Definition 15.14. indicator function

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

Note:

Theorem 15.9. probability as expected value

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

Proof:

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

□

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

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

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

negation = NOT

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

in set theory or domain of events,

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

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,

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

$$\mathbf{1}(E_1 \cup E_2) = \mathbf{1}(E_1) + \mathbf{1}(E_2) - \mathbf{1}(E_1)\mathbf{1}(E_2) = \mathbf{1}(E_1) + \mathbf{1}(E_2) - \mathbf{1}(E_1 \cap E_2)$$

□

implication = conditional

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

exclusive disjunction = XOR

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

biconditional = XNOR

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

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

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

floor and ceiling functions

floor function

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

ceiling function

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

Heaviside step function

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

or conveniently define “unit step function”

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

ramp function = rectified linear unit activation function = ReLU

$$\text{ReLU}(x) = \begin{cases} x & x \geq 0 \\ 0 & x < 0 \end{cases} = x [x \geq 0]$$

indicator function

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

$A, B \subseteq \Omega$,

$$A = B \Leftrightarrow 1_A = 1_B$$

$$A = \Omega \Leftrightarrow 1_A(x) = 1$$

$$A = \emptyset \Leftrightarrow 1_A(x) = 0$$

Theorem 15.10. *subset indicator order*

$$A \subset B \Rightarrow 1_A(x) \leq 1_B(x)$$

Proof:

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

□

in set theory or domain of events,

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

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

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

expectation in many perspectives

$$Y = g(X)$$

$$\int_{-\infty}^{+\infty} y f_Y(y) dy = E[Y] = E[g(X)] = \int_{-\infty}^{+\infty} g(x) f_X(x) dx$$

$$E_Y[Y] = \int_{-\infty}^{+\infty} y f_Y(y) dy = E[Y] = E[g(X)] = \int_{-\infty}^{+\infty} g(x) f_X(x) dx = E_X[g(X)]$$

$$E_Y[Y] = E_X[g(X)]$$

$$\begin{aligned}
& E[a_1 g_1(X_1) + a_2 g_2(X_2) + c], \begin{cases} Y_1 = g_1(X_1) \\ Y_2 = g_2(X_2) \end{cases} \\
= & E[a_1 Y_1 + a_2 Y_2 + c]
\end{aligned}$$

$$\begin{aligned}
E[a_1 g_1(X_1) + a_2 g_2(X_2) + c] &= a_1 E[g_1(X_1)] + a_2 E[g_2(X_2)] + c \\
&= E[a_1 Y_1 + a_2 Y_2 + c] = a_1 E[Y_1] + a_2 E[Y_2] + c
\end{aligned}$$

$$a_1 E[g_1(X_1)] + a_2 E[g_2(X_2)] + c = a_1 E_{X_1}[g_1(X_1)] + a_2 E_{X_2}[g_2(X_2)] + c$$

$$a_1 E[Y_1] + a_2 E[Y_2] + c = a_1 E_{Y_1}[Y_1] + a_2 E_{Y_2}[Y_2] + c$$

$$a_1 E_{X_1}[g_1(X_1)] + a_2 E_{X_2}[g_2(X_2)] + c = a_1 E_{Y_1}[Y_1] + a_2 E_{Y_1}[Y_2] + c$$

15.1.2.7 moment

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

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

$$\begin{aligned} E[X^n] &= \begin{cases} \int_{-\infty}^{+\infty} x^n f_X(x) dx & X \text{ continuous} \\ \sum_{x \in X(\Omega)} x^n f_X(x) & X \text{ discrete} \end{cases} \\ \mu = E[X^1] = E[X] &= \begin{cases} \int_{-\infty}^{+\infty} x^1 f_X(x) dx & X \text{ continuous} \\ \sum_{x \in X(\Omega)} x^1 f_X(x) & X \text{ discrete} \end{cases} = \begin{cases} \int_{-\infty}^{+\infty} x f_X(x) dx & X \text{ continuous} \\ \sum_{x \in X(\Omega)} x f_X(x) & X \text{ discrete} \end{cases} \\ \mu_n = E[(X - E[X])^n] &= \begin{cases} \int_{-\infty}^{+\infty} (x - \mu)^n f_X(x) dx & X \text{ continuous} \\ \sum_{x \in X(\Omega)} (x - \mu)^n f_X(x) & X \text{ discrete} \end{cases} \end{aligned}$$

1st moment of X = mean

$$\mu = E[X^1] = E[X] = \begin{cases} \int_{-\infty}^{+\infty} x^1 f_X(x) dx & X \text{ continuous} \\ \sum_{x \in X(\Omega)} x^1 f_X(x) & X \text{ discrete} \end{cases} = \begin{cases} \int_{-\infty}^{+\infty} x f_X(x) dx & X \text{ continuous} \\ \sum_{x \in X(\Omega)} x f_X(x) & X \text{ discrete} \end{cases}$$

1st central moment of X = 0

$$\begin{aligned} \mu_1 = E[(X - E[X])^1] &= \begin{cases} \int_{-\infty}^{+\infty} (x - \mu)^1 f_X(x) dx & X \text{ continuous} \\ \sum_{x \in X(\Omega)} (x - \mu)^1 f_X(x) & X \text{ discrete} \end{cases} = \begin{cases} \int_{-\infty}^{+\infty} (x - \mu) f_X(x) dx & X \text{ continuous} \\ \sum_{x \in X(\Omega)} (x - \mu) f_X(x) & X \text{ discrete} \end{cases} \\ &= \begin{cases} \int_{-\infty}^{+\infty} x f_X(x) dx - \int_{-\infty}^{+\infty} \mu f_X(x) dx & X \text{ continuous} \\ \sum_{x \in X(\Omega)} x f_X(x) - \sum_{x \in X(\Omega)} \mu f_X(x) & X \text{ discrete} \end{cases} \\ &= \begin{cases} E[X] - \mu \int_{-\infty}^{+\infty} f_X(x) dx & X \text{ continuous} \\ E[X] - \mu \sum_{x \in X(\Omega)} f_X(x) & X \text{ discrete} \end{cases} = \begin{cases} E[X] - \mu \cdot 1 & X \text{ continuous} \\ E[X] - \mu \cdot 1 & X \text{ discrete} \end{cases} \\ &= \begin{cases} E[X] - \mu & X \text{ continuous} \\ E[X] - \mu & X \text{ discrete} \end{cases} = \begin{cases} E[X] - E[X] & X \text{ continuous} \\ E[X] - E[X] & X \text{ discrete} \end{cases} \\ &= \begin{cases} 0 & X \text{ continuous} \\ 0 & X \text{ discrete} \end{cases} = 0 \end{aligned}$$

$$E[(X - E[X])] = 0$$

$$E[X - E[X]] = 0$$

$$\forall X (E[X - E[X]] = 0)$$

For normal distribution, actually for any distribution,

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

2nd central moment of X = variance

$$\begin{aligned}\mu_2 &= E[(X - E[X])^2] = \begin{cases} \int_{-\infty}^{+\infty} (x - \mu)^2 f_x(x) dx & X \text{ continuous} \\ \sum_{x \in X(\Omega)} (x - \mu)^2 f_x(x) & X \text{ discrete} \end{cases} \\ &= E[(X - E[X])^2] = V[X] = V(X)\end{aligned}$$

For normal distribution,

$$\begin{aligned}X &\sim n(0, 1) = \mathcal{N}(0, 1^2) = \mathcal{N}(\mu = 0, V^2[X] = 1^2) \\ &\Downarrow \\ V[X] &= V(X) = 1\end{aligned}$$

variance properties

$$V[aX + b] = a^2 V[X]$$

Proof:

to be proved

□

3rd central moment of X

$$\mu_3 = E[(X - E[X])^3] = \begin{cases} \int_{-\infty}^{+\infty} (x - \mu)^3 f_x(x) dx & X \text{ continuous} \\ \sum_{x \in X(\Omega)} (x - \mu)^3 f_x(x) & X \text{ discrete} \end{cases}$$

skewness

偏度

$$\begin{aligned}\text{skewness}[X] &= \frac{\mu_3}{\mu_2^{\frac{3}{2}}} = \frac{E[(X - E[X])^3]}{(V[X])^{\frac{3}{2}}} = \frac{E[(X - E[X])^3]}{\left(E[(X - E[X])^2]\right)^{\frac{3}{2}}} \\ &= \begin{cases} \frac{\int_{-\infty}^{+\infty} (x - \mu)^3 f_x(x) dx}{\left(\int_{-\infty}^{+\infty} (x - \mu)^2 f_x(x) dx\right)^{\frac{3}{2}}} & X \text{ continuous} \\ \frac{\sum_{x \in X(\Omega)} (x - \mu)^3 f_x(x)}{\left(\sum_{x \in X(\Omega)} (x - \mu)^2 f_x(x)\right)^{\frac{3}{2}}} & X \text{ discrete} \end{cases}\end{aligned}$$

For normal distribution,

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

Proof:

to be proved

□

4^{th} central moment of X

$$\mu_4 = \mathbb{E}[(X - \mathbb{E}[X])^4] = \begin{cases} \int_{-\infty}^{+\infty} (x - \mu)^4 f_x(x) dx & X \text{ continuous} \\ \sum_{x \in X(\Omega)} (x - \mu)^4 f_x(x) & X \text{ discrete} \end{cases}$$

kurtosis

峰度

$$\begin{aligned} \text{kurtosis}[X] &= \frac{\mu_4}{\mu_2^2} = \frac{\mathbb{E}[(X - \mathbb{E}[X])^4]}{(\mathbb{V}[X])^2} = \frac{\mathbb{E}[(X - \mathbb{E}[X])^4]}{\left(\mathbb{E}[(X - \mathbb{E}[X])^2]\right)^2} \\ &= \begin{cases} \frac{\int_{-\infty}^{+\infty} (x - \mu)^4 f_x(x) dx}{\left(\int_{-\infty}^{+\infty} (x - \mu)^2 f_x(x) dx\right)^2} & X \text{ continuous} \\ \frac{\sum_{x \in X(\Omega)} (x - \mu)^4 f_x(x)}{\left(\sum_{x \in X(\Omega)} (x - \mu)^2 f_x(x)\right)^2} & X \text{ discrete} \end{cases} \end{aligned}$$

For normal distribution,

$$\begin{aligned} X \sim n(0, 1) &= \mathcal{N}(0, 1^2) = \mathcal{N}(\mu = 0, \mathbb{V}^2[X] = 1^2) \\ &\Downarrow \\ \text{kurtosis}[X] &= \frac{\mathbb{E}[(X - \mathbb{E}[X])^4]}{(\mathbb{V}[X])^2} = \frac{[(X - \mathbb{E}[X])^4]}{1^2} = 3 \end{aligned}$$

Proof:

to be proved

□

For normal distribution,

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

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

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

$$X \sim F_x(x) \leftrightarrow f_x(x) \rightarrow \{\mu_n | n \in \mathbb{N}\} = \left\{ \mu_n \left| \begin{array}{l} n \in \mathbb{N} \\ \mu_n = \mathbb{E}[(X - \mathbb{E}[X])^n] \end{array} \right. \right\}$$

15.1.2.7.1 moment generating function

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

$$M(t) = M_x(t) = \mathbb{E}[e^{tX}]$$

$$M(\xi) = M_x(\xi) = \mathbb{E}[e^{\xi X}]$$

Theorem 15.11. moment generating function(MGF) generating moment

$$M_x^{(n)}(\xi) = \mathbb{E}[X^n]$$

where

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

Proof:

to be proved

□

$$X \sim F_x(x) \leftrightarrow f_x(x) \rightarrow \{\mu_n | n \in \mathbb{N}\} = \left\{ \mu_n \left| \begin{array}{l} n \in \mathbb{N} \\ \mu_n = \mathbb{E}[(X - \mathbb{E}[X])^n] \end{array} \right. \right\}$$

$$\begin{array}{ccccccc}
X & \sim & F_x(x) & \leftrightarrow & f_x(x) & \rightarrow & \{\mu_n | n \in \mathbb{N}\} \\
& & & & \downarrow & \nearrow & \\
& & & & M_x(\xi) & &
\end{array}$$

$$\begin{array}{ccccccc}
X & \sim & F_x(x) & \leftrightarrow & f_x(x) & & \\
& & & & \downarrow & \searrow & \\
& & & & M_x(\xi) & \rightarrow & \{\mu_n | n \in \mathbb{N}\}
\end{array}$$

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

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

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

Proof:

to be proved

□

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

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

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

Proof:

to be proved

□

$$\begin{array}{ccc} X & \sim & F_X(x) \\ & & \leftrightarrow f_X(x) \\ & \uparrow & \downarrow \\ \forall \xi \approx 0 [M_X(\xi) \in \mathbb{R}] & \wedge & M_X(\xi) \end{array} \rightarrow \begin{array}{c} \nwarrow \nearrow \\ \leftarrow \wedge \text{ supp}(f_X) \text{ is bounded} \\ \{\mu_n | n \in \mathbb{N}\} \end{array}$$

15.1.2.7.2 characteristic function

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

$$\varphi(t) = \varphi_X(t) = \mathbb{E}[e^{itX}]$$

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

Note:

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

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

2. moment generating function to characteristic function

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

3. inversion theorem or inversion formula

For $a < b$,

$$\begin{aligned} & \lim_{T \rightarrow \infty} \frac{1}{2\pi} \int_{-T}^{+T} \frac{e^{-ita} - e^{-itb}}{it} \varphi_x(t) dt \\ &= P(a < X < b) + \frac{1}{2} [P(X = a) + P(X = b)] \end{aligned}$$

i.e. $\varphi_x(\xi)$ determines $F_x(x)$.

$$\begin{array}{ccccccc} X & \sim & F_x(x) & \leftrightarrow & f_x(x) & \leftrightarrow & P_x \\ \uparrow & & \forall \xi \approx 0 [M_x(\xi) \in \mathbb{R}] \wedge & \uparrow & \downarrow & \leftarrow \rho \wedge \text{supp}(f_x) \text{ is bounded} \\ \varphi_x(\xi) & & \Downarrow & M_x(\xi) & \rightarrow & \{\mu_n | n \in \mathbb{N}\} \\ \\ X & \sim & F_x(x) & \xrightarrow{\text{FToC}} & f_x(x) & \leftrightarrow & P_x \\ \text{inversion formula : } & \uparrow & \forall \xi \approx 0 [M_x(\xi) \in \mathbb{R}] \wedge & \uparrow & \downarrow & \leftarrow \rho \wedge \text{supp}(f_x) \text{ is bounded} \\ \varphi_x(\xi) & & \Downarrow & M_x(\xi) & \rightarrow & \{\mu_n | n \in \mathbb{N}\} \end{array}$$

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

$$\begin{aligned} X &\sim \mathcal{DU}(1, N) \\ &\Updownarrow \\ \begin{cases} f_x(x|N) = \frac{1}{N} \\ x \in X(\Omega) = \{1, 2, \dots, N\} \end{cases} \\ &\Downarrow \\ \begin{cases} E[X] = \frac{N+1}{2} \\ V[X] = \frac{(N+1)(N-1)}{12} \end{cases} \end{aligned}$$

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

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

$$X \sim \mathcal{DU}(a, b) \Leftrightarrow \begin{cases} f_x(x|N) = \frac{1}{N} = \frac{1}{b-a+1} & N = b-a+1 \\ x \in X(\Omega) = \{a, a+1, \dots, b\} & \end{cases} \Leftrightarrow F_x(x|a, b) = \frac{\lfloor x \rfloor - a + 1}{b - a + 1}$$

15.1.2.8.1.2 hypergeometric distribution

$$\begin{aligned}
X &\sim \mathcal{HG}(N, M, K) \\
&\Updownarrow \\
\begin{cases} f_x(x|N, M, K) = \frac{\binom{N-M}{K-x} \binom{M}{x}}{\binom{N}{K}} \\ x \in X(\Omega) = \{\max\{0, K-(N-M)\}, \dots, \min\{K, M\}\} \end{cases} \\
&\Downarrow K \ll N, M \\
&x \in X(\Omega) = \{0, 1, \dots, K\} \\
&\Downarrow \\
\begin{cases} E[X] = \frac{KM}{N} \\ V[X] = \frac{KM}{N} \frac{(N-M)(N-K)}{N(N-1)} \end{cases}
\end{aligned}$$

15.1.2.8.1.3 Bernoulli distribution

$$\begin{aligned}
X &\sim \mathcal{B}(p), p = P(X=1) \\
&\Updownarrow \\
\begin{cases} f_x(x|p) = (1-p)^{1-x} p^x \\ x \in X(\Omega) = \{0, 1\} \end{cases} \\
&\Downarrow \\
\begin{cases} E[X] = p \\ V[X] = p(1-p) \\ M_x(\xi) = (1-p) + pe^\xi \end{cases}
\end{aligned}$$

15.1.2.8.1.4 binomial distribution independent and identical Bernoulli trials

$$\begin{aligned}
 X &\sim \text{b}(n, p), p = \text{P}(X = 1) \\
 &\Updownarrow \\
 \begin{cases} f_x(x|n, p) = \binom{n}{x} (1-p)^{n-x} p^x \\ x \in X(\Omega) = \{0, 1, \dots, n\} \end{cases} \\
 &\Downarrow \\
 \begin{cases} \text{E}[X] = np \\ \text{V}[X] = np(1-p) \\ M_x(\xi) = [(1-p) + pe^\xi]^n \end{cases}
 \end{aligned}$$

reparameterization technique

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

15.1.2.8.1.5 Poisson distribution count = number of events

an unbounded discrete distribution we first see or met

$$\begin{aligned}
 X &\sim \mathcal{P}(\lambda) \\
 &\Updownarrow \\
 \begin{cases} f_x(x|\lambda) = \frac{e^{-\lambda} \lambda^x}{x!} \\ x \in X(\Omega) = \{0, 1, \dots\} \end{cases} \\
 &\Downarrow \\
 \begin{cases} \text{E}[X] = \lambda \\ \text{V}[X] = \lambda = \text{E}[X] \\ M_x(\xi) = \exp[\lambda(e^\xi - 1)] = e^{\lambda(e^\xi - 1)} \end{cases}
 \end{aligned}$$

the Poisson postulates

$$\begin{cases}
 N_t & \text{a r.v. denoting the of events in } [0, t] \\
 N_0 = N_{t=0} = 0 & \text{reset the count at the initial point} \\
 \forall s < t [N_s \perp N_t - N_s] & \text{disjoint intervals independent} \\
 N_s = N_{t+s} - N_t & \text{depends on length instead of initial point} \\
 \lim_{t \rightarrow 0} \frac{P(N_t = 1)}{t} = \lambda & \Rightarrow \forall t \approx 0 [P(N_t = 1) \approx \lambda t] \\
 \lim_{t \rightarrow 0} \frac{P(N_t > 1)}{t} = 0 & \text{no coincidence for small } t \\
 \text{solve the differential equations } \Downarrow \text{ with probability axioms} \\
 f_x(x|\lambda t) = P(N_t = x) = \frac{e^{-\lambda t} (\lambda t)^x}{x!}
 \end{cases}$$

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.

$$\begin{aligned}
 X &\sim \mathcal{NB}(r, p) \\
 &\Downarrow \\
 \begin{cases} f_x(x|r, p) = \binom{x-1}{r-1} (1-p)^{x-r} p^r \\ x \in X(\Omega) = \{r, r+1, \dots\} \end{cases} \\
 &\Downarrow \\
 \begin{cases} E[X] = \\ V[X] = \\ M_X(\xi) = \end{cases}
 \end{aligned}$$

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

Note:

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

reparameterization technique

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

Note:

For sufficiently small t such that

$$0 \leq p^* = (1-p)e^{\xi} \leq 1$$

or else $t \gg 1$

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

15.1.2.8.1.7 geometric distribution

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

$r = 1$,

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

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

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

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

15.1.2.8.2 continuous distribution

15.1.2.8.2.1 uniform distribution = continuous uniform distribution

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

$$a = 0, b = 1$$

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

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

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

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

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

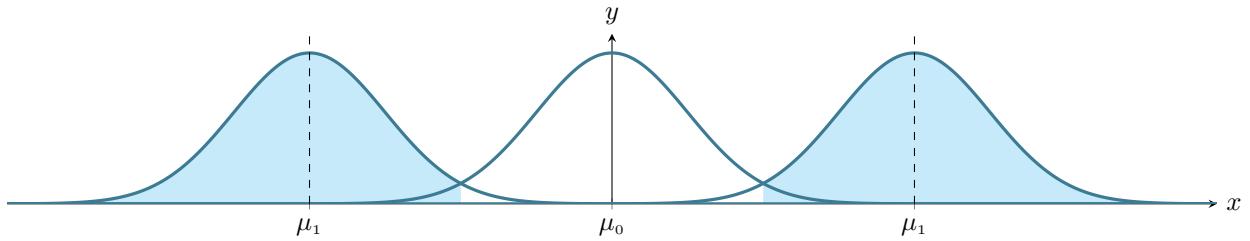


Figure 15.1: not completed: Cauchy distribution vs. normal distribution

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

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

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

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

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

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

- 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

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

not

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

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

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

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

$$\begin{aligned} f_x(x|a, b) &= \frac{1(x \in [a, b])}{b-a} \\ &= \frac{1}{b-a}, x \in [a, b] \end{aligned}$$

$$\frac{1}{b-a} = h(x) c(\boldsymbol{\theta}) e^{\sum_{j=1}^k w_j(\boldsymbol{\theta}) t_j(x)}, \begin{cases} h(x) = 1 \\ c(\boldsymbol{\theta}) = c(\theta_1, \theta_2) = c(a, b) = \frac{1}{b-a} \\ w_1(\theta_1) = w_2(\theta_2) = 0 \\ t_1(x) = t_2(x) = x \\ k = 2 \end{cases}$$

however,

$$1(x \in [a, b]) \neq h(x) c(\boldsymbol{\theta}) = h(x) c(a, b)$$

thus

$$f_x(x|a, b) = \frac{1(x \in [a, b])}{b - a} \notin \mathcal{E}^f = \left\{ f \mid f = f(x|\boldsymbol{\theta}) = h(x) c(\boldsymbol{\theta}) \exp \left(\sum_{j=1}^k w_j(\boldsymbol{\theta}) t_j(x) \right) \right\}$$

15.1.2.9.3 normal distribution is in exponential family

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

$$\begin{aligned} f_x(x|\mu, \sigma^2) &= f_x(x|\mu, \sigma) = \frac{1}{\sigma\sqrt{2\pi}} e^{\frac{-1}{2}(\frac{x-\mu}{\sigma})^2} \\ &= \frac{1}{\sigma\sqrt{2\pi}} e^{\frac{-1}{2\sigma^2}(x^2 - 2x\mu + \mu^2)} = \frac{e^{\frac{-1}{2}(\frac{\mu}{\sigma})^2}}{\sigma\sqrt{2\pi}} e^{\frac{-1}{2\sigma^2}x^2 + \frac{\mu}{\sigma^2}x} = \frac{e^{\frac{-1}{2}(\frac{\mu}{\sigma})^2}}{\sigma\sqrt{2\pi}} e^{\frac{\mu}{\sigma^2}x - \frac{1}{2\sigma^2}x^2} \\ &= h(x) c(\boldsymbol{\theta}) e^{\sum_{j=1}^k w_j(\boldsymbol{\theta}) t_j(x)}, \begin{cases} h(x) = 1 \\ c(\boldsymbol{\theta}) = c(\theta_1, \theta_2) = c(\mu, \sigma) = \frac{e^{\frac{-1}{2}(\frac{\mu}{\sigma})^2}}{\sigma\sqrt{2\pi}} \\ w_1(\theta_1, \theta_2) = w_1(\mu, \sigma) = \frac{\mu}{\sigma^2} \\ w_2(\theta_2) = w_2(\sigma) = \frac{-1}{2\sigma^2} \\ t_1(x) = x \\ t_2(x) = x^2 \\ k = 2 \end{cases} \\ &\Downarrow \end{aligned}$$

$$f_x(x|\mu, \sigma) = \frac{1}{\sigma\sqrt{2\pi}} e^{\frac{-1}{2}(\frac{x-\mu}{\sigma})^2} \in \mathcal{E}^f = \left\{ f \mid f = f(x|\boldsymbol{\theta}) = h(x) c(\boldsymbol{\theta}) \exp \left(\sum_{j=1}^k w_j(\boldsymbol{\theta}) t_j(x) \right) \right\}$$

15.1.2.9.4 normal distribution with unknown equal mean and standard deviation in curved exponential family

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

$$\begin{aligned}
f_x(x|\mu, \mu^2) &= f_x(x|\mu) = \frac{1}{\mu\sqrt{2\pi}} e^{\frac{-1}{2}(\frac{x-\mu}{\mu})^2} \\
&= \frac{1}{\mu\sqrt{2\pi}} e^{\frac{-1}{2\mu^2}(x^2 - 2x\mu + \mu^2)} = \frac{e^{\frac{-1}{2}}}{\mu\sqrt{2\pi}} e^{\frac{-1}{2\mu^2}x^2 + \frac{\mu}{\mu^2}x} = \frac{e^{\frac{-1}{2}}}{\mu\sqrt{2\pi}} e^{\frac{1}{\mu}x - \frac{1}{2\mu^2}x^2} \\
&= h(x) c(\boldsymbol{\theta}) \sum_{j=1}^k w_j(\boldsymbol{\theta}) t_j(x), \quad \begin{cases} h(x) = 1 \\ c(\boldsymbol{\theta}) = c(\theta_1) = c(\mu) = \frac{e^{\frac{-1}{2}}}{\mu\sqrt{2\pi}} \\ w_1(\theta_1) = w_1(\mu) = \frac{1}{\mu} \\ w_2(\theta_1) = w_2(\mu) = \frac{-1}{2\mu^2} \\ t_1(x) = x \\ t_2(x) = x^2 \\ k = 2 > 1 = p \end{cases} \\
&\Downarrow \begin{cases} p = \dim \boldsymbol{\theta} \\ k = \dim \mathbf{w} \end{cases} \\
f_x(x|\mu, \sigma) &= \frac{1}{\sigma\sqrt{2\pi}} e^{\frac{-1}{2}(\frac{x-\mu}{\sigma})^2} \in \mathcal{C}^f = \left\{ f \mid f = f(x|\boldsymbol{\theta}) = h(x) c(\boldsymbol{\theta}) \exp \left(\sum_{j=1}^{k>p} w_j(\boldsymbol{\theta}) t_j(x) \right) \right\} \\
&\subset \mathcal{E}^f = \left\{ f \mid f = f(x|\boldsymbol{\theta}) = h(x) c(\boldsymbol{\theta}) \exp \left(\sum_{j=1}^k w_j(\boldsymbol{\theta}) t_j(x) \right) \right\}
\end{aligned}$$

curved exponential family

$$\mathcal{C}^f = \left\{ f \mid f = f(x|\boldsymbol{\theta}) = h(x) c(\boldsymbol{\theta}) \exp \left(\sum_{j=1}^{k>p} w_j(\boldsymbol{\theta}) t_j(x) \right) \right\}$$

exponential family

$$\mathcal{E}^f = \left\{ f \mid f = f(x|\boldsymbol{\theta}) = h(x) c(\boldsymbol{\theta}) \exp \left(\sum_{j=1}^k w_j(\boldsymbol{\theta}) t_j(x) \right) \right\}$$

<https://tex.stackexchange.com/questions/145969/filling-specified-area-by-random-dots-in-tikz>

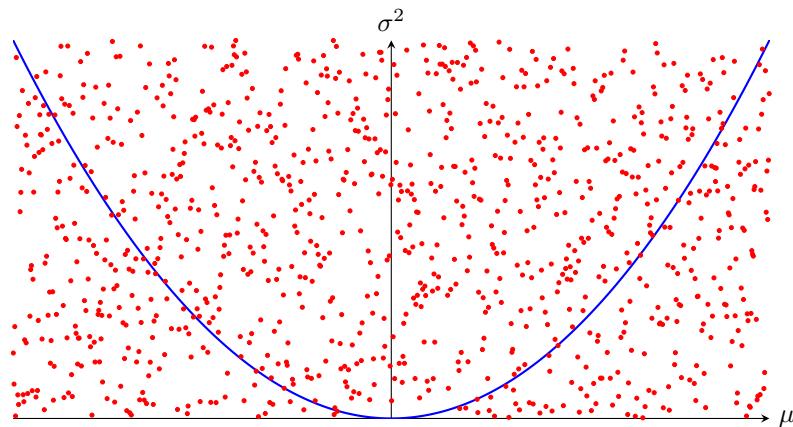


Figure 15.2: curved exponential family vs. exponential family

15.1.2.9.5 properties of exponential family

15.1.2.9.5.1 fundamentals of statistical inference

Lemma 15.1. *Leibnitz rule*

$$\begin{aligned}
 \frac{d}{d\theta} \int_{a(\theta)}^{b(\theta)} g(x, \theta) dx &= \frac{d \int_{a(\theta)}^{b(\theta)} g(x, \theta) dx}{d\theta} \\
 &= g(x, b(\theta)) \frac{db(\theta)}{d\theta} - g(x, a(\theta)) \frac{da(\theta)}{d\theta} + \int_{a(\theta)}^{b(\theta)} \frac{d}{d\theta} g(x, \theta) dx \\
 &= g(x, b(\theta)) \frac{db(\theta)}{d\theta} - g(x, a(\theta)) \frac{da(\theta)}{d\theta} + \int_{a(\theta)}^{b(\theta)} \frac{dg(x, \theta)}{d\theta} dx \\
 \frac{d}{d\theta} \int_{a(\theta)}^{b(\theta)} g(x, \theta) dx &= g(x, b(\theta)) \frac{db(\theta)}{d\theta} - g(x, a(\theta)) \frac{da(\theta)}{d\theta} + \int_{a(\theta)}^{b(\theta)} \frac{dg(x, \theta)}{d\theta} dx \\
 \hline
 \frac{d}{d\theta} \int_{a(\theta)=a}^{b(\theta)=b} g(x, \theta) dx &= g(x, b(\theta)) \frac{db(\theta)}{d\theta} - g(x, a(\theta)) \frac{da(\theta)}{d\theta} + \int_{a(\theta)}^{b(\theta)} \frac{dg(x, \theta)}{d\theta} dx \\
 &= g(x, b) \frac{db}{d\theta} - g(x, a) \frac{da}{d\theta} + \int_a^b \frac{dg(x, \theta)}{d\theta} dx \\
 &= g(x, b) 0 - g(x, a) 0 + \int_a^b \frac{dg(x, \theta)}{d\theta} dx = 0 + 0 + \int_a^b \frac{dg(x, \theta)}{d\theta} dx \\
 &= \int_a^b \frac{dg(x, \theta)}{d\theta} dx \\
 \frac{d}{d\theta} \int_{X(\Omega) \perp \theta} g(x, \theta) dx &= \frac{d}{d\theta} \int_{a(\theta)=a}^{b(\theta)=b} g(x, \theta) dx = \int_a^b \frac{dg(x, \theta)}{d\theta} dx = \int_{X(\Omega) \perp \theta} \frac{dg(x, \theta)}{d\theta} dx \\
 \frac{d}{d\theta} \int_{X(\Omega) \perp \theta} g(x, \theta) dx &= \int_{X(\Omega) \perp \theta} \frac{dg(x, \theta)}{d\theta} dx \\
 \frac{d}{d\theta} \int_{X(\Omega) \perp \theta} g(x, \theta) dx &= \int_{X(\Omega) \perp \theta} \frac{dg(x, \theta)}{d\theta} dx
 \end{aligned}$$

point estimation

Lemma 15.2. *parameter-independent expectation:*

Assume the domain of X independent of θ , then

$$E \left[\frac{\partial \ln f(X|\boldsymbol{\theta})}{\partial \theta_i} \right] = 0, \forall i = 1, \dots, p$$

Proof:

$$\begin{aligned}
 E \left[\frac{\partial \ln f(X|\boldsymbol{\theta})}{\partial \theta_i} \right] &= \int \frac{\partial \ln f(x|\boldsymbol{\theta})}{\partial \theta_i} f(x|\boldsymbol{\theta}) dx \\
 &= \int \left[\frac{\partial \ln f(x|\boldsymbol{\theta})}{\partial \theta_i} \right] f(x|\boldsymbol{\theta}) dx \\
 &= \int \left[\frac{1}{f(x|\boldsymbol{\theta})} \frac{\partial f(x|\boldsymbol{\theta})}{\partial \theta_i} \right] f(x|\boldsymbol{\theta}) dx = \int \frac{1}{f(x|\boldsymbol{\theta})} \frac{\partial f(x|\boldsymbol{\theta})}{\partial \theta_i} f(x|\boldsymbol{\theta}) dx \\
 &= \int \frac{f(x|\boldsymbol{\theta})}{f(x|\boldsymbol{\theta})} \frac{\partial f(x|\boldsymbol{\theta})}{\partial \theta_i} dx = \int 1 \frac{\partial f(x|\boldsymbol{\theta})}{\partial \theta_i} dx = \int \frac{\partial f(x|\boldsymbol{\theta})}{\partial \theta_i} dx \\
 &\stackrel{\text{Leibnitz rule}}{=} \frac{\partial}{\partial \theta_i} \int f(x|\boldsymbol{\theta}) dx
 \end{aligned}$$

□

Note:

$$\begin{aligned}
 \mathbb{E} \left[\frac{\partial \ln f(X|\boldsymbol{\theta})}{\partial \theta_i} \right] &= \int \frac{\partial \ln f(x|\boldsymbol{\theta})}{\partial \theta_i} f(x|\boldsymbol{\theta}) dx = \int \left(\frac{\frac{\partial f(x|\boldsymbol{\theta})}{\partial \theta_i}}{f(x|\boldsymbol{\theta})} \right) f(x|\boldsymbol{\theta}) dx \\
 X \sim f_x(x|\boldsymbol{\theta}) \Rightarrow \mathbb{E} \left[\frac{\partial \ln f_x(X|\boldsymbol{\theta})}{\partial \theta_i} \right] &= \int \frac{\partial \ln f_x(x|\boldsymbol{\theta})}{\partial \theta_i} f_x(x|\boldsymbol{\theta}) dx = \int \left(\frac{\frac{\partial f_x(x|\boldsymbol{\theta})}{\partial \theta_i}}{f_x(x|\boldsymbol{\theta})} \right) f_x(x|\boldsymbol{\theta}) dx \\
 &= \int \frac{\partial f(x|\boldsymbol{\theta})}{\partial \theta_i} dx \\
 X \sim f_x(x|\boldsymbol{\theta}) \Rightarrow \mathbb{E} \left[\frac{\partial \ln f_x(X|\boldsymbol{\theta}^*)}{\partial \theta_i} \right] &= \int \frac{\partial \ln f_x(x|\boldsymbol{\theta}^*)}{\partial \theta_i} f_x(x|\boldsymbol{\theta}) dx = \int \left(\frac{\frac{\partial f_x(x|\boldsymbol{\theta}^*)}{\partial \theta_i}}{f_x(x|\boldsymbol{\theta}^*)} \right) f(x|\boldsymbol{\theta}) dx \\
 &= \int \frac{\partial f_x(x|\boldsymbol{\theta}^*)}{\partial \theta_i} \left(\frac{f_x(x|\boldsymbol{\theta})}{f_x(x|\boldsymbol{\theta}^*)} \right) dx
 \end{aligned}$$

$$\mathbb{E} \left[\frac{\partial \ln f_x(X|\boldsymbol{\theta}^*)}{\partial \theta_i} \right] = \mathbb{E}_{\boldsymbol{\theta}^*} \left[\frac{\partial \ln f_x(X|\boldsymbol{\theta}^*)}{\partial \theta_i} \right] = \int \frac{\partial \ln f_x(x|\boldsymbol{\theta}^*)}{\partial \theta_i} f_x(x|\boldsymbol{\theta}^*) dx$$

$$\mathbb{E} \left[\frac{\partial \ln f_x(X|\boldsymbol{\theta}^*)}{\partial \theta_i} \right] = \mathbb{E}_{\boldsymbol{\theta}} \left[\frac{\partial \ln f_x(X|\boldsymbol{\theta}^*)}{\partial \theta_i} \right] = \int \frac{\partial \ln f_x(x|\boldsymbol{\theta}^*)}{\partial \theta_i} f_x(x|\boldsymbol{\theta}) dx$$

$$\mathbb{E}_{\boldsymbol{\theta}} \left[\frac{\partial \ln f_x(X|\boldsymbol{\theta}^*)}{\partial \theta_i} \right] = \int \frac{\partial \ln f_x(x|\boldsymbol{\theta}^*)}{\partial \theta_i} f_x(x|\boldsymbol{\theta}) dx$$

$$\mathbb{E}_{\boldsymbol{\theta}} \left[\frac{\partial \ln f_x(X|\boldsymbol{\theta}^*)}{\partial \theta_i} \right] \not\equiv 0$$

$$\mathbb{E}_{\boldsymbol{\theta}} \left[\frac{\partial \ln f_x(X|\boldsymbol{\theta}^*)}{\partial \theta_i} \right] \underset{\boldsymbol{\theta}^*=\boldsymbol{\theta}}{\equiv} 0$$

as the fundamental to estimate parameters.

interval estimation

Lemma 15.3. *parameter-independent variance:*

Assume the domain of X independent of θ , then

$$\mathbb{E} \left[\frac{\partial^2 \ln f(X|\boldsymbol{\theta})}{\partial \theta_i^2} \right] = -\mathbb{E} \left[\left(\frac{\partial \ln f(X|\boldsymbol{\theta})}{\partial \theta_i} \right)^2 \right], \forall i = 1, \dots, p$$

Proof:

$$\begin{aligned}
 \mathbb{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}) dx = \int \frac{\partial}{\partial \theta_i} \frac{\partial \ln f(x|\boldsymbol{\theta})}{\partial \theta_i} f(x|\boldsymbol{\theta}) dx \\
 &= \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}) dx = \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}) dx \\
 &= \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}) dx = \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} dx \\
 &= \int \frac{\partial^2 f(x|\boldsymbol{\theta})}{\partial \theta_i^2} dx - \int \left(\frac{\frac{\partial f(x|\boldsymbol{\theta})}{\partial \theta_i}}{f(x|\boldsymbol{\theta})} \right)^2 f(x|\boldsymbol{\theta}) dx \stackrel{\text{Leibnitz rule}}{=} \frac{\partial^2}{\partial \theta_i^2} \int f(x|\boldsymbol{\theta}) dx - \int \left(\frac{\frac{\partial f(x|\boldsymbol{\theta})}{\partial \theta_i}}{f(x|\boldsymbol{\theta})} \right)^2 f(x|\boldsymbol{\theta}) dx \\
 &= \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}) dx = 0 - \int \left(\frac{\frac{\partial f(x|\boldsymbol{\theta})}{\partial \theta_i}}{f(x|\boldsymbol{\theta})} \right)^2 f(x|\boldsymbol{\theta}) dx = - \int \left(\frac{\frac{\partial f(x|\boldsymbol{\theta})}{\partial \theta_i}}{f(x|\boldsymbol{\theta})} \right)^2 f(x|\boldsymbol{\theta}) dx \\
 &= - \int \left(\frac{\partial \ln f(x|\boldsymbol{\theta})}{\partial \theta_i} \right)^2 f(x|\boldsymbol{\theta}) dx = -\mathbb{E} \left[\left(\frac{\partial \ln f(X|\boldsymbol{\theta})}{\partial \theta_i} \right)^2 \right]
 \end{aligned}$$

□

$$\begin{cases} \mathbb{E} \left[\frac{\partial \ln f(X|\boldsymbol{\theta})}{\partial \theta_i} \right] = 0 \\ \mathbb{E} \left[\frac{\partial^2 \ln f(X|\boldsymbol{\theta})}{\partial \theta_i^2} \right] = -\mathbb{E} \left[\left(\frac{\partial \ln f(X|\boldsymbol{\theta})}{\partial \theta_i} \right)^2 \right], \forall i = 1, \dots, p \end{cases}$$

exponential family expectation and variance

Theorem 15.15. *exponential family expectation*

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

$\Downarrow \forall i = 1, \dots, p$

$$\mathbb{E} \left[\sum_{j=1}^k \frac{\partial w_j(\boldsymbol{\theta})}{\partial \theta_i} t_j(X) \right] = \frac{-\partial \ln c(\boldsymbol{\theta})}{\partial \theta_i}$$

Proof:

$$\begin{aligned} f(x|\boldsymbol{\theta}) &= h(x) c(\boldsymbol{\theta}) 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}) e^{\sum_{j=1}^k w_j(\boldsymbol{\theta}) t_j(x)} \\ &= \ln h(x) + \ln c(\boldsymbol{\theta}) + \ln 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}$$

$$\begin{aligned}
\frac{\partial \ln f(X|\boldsymbol{\theta})}{\partial \theta_i} &= \frac{\partial \ln c(\boldsymbol{\theta})}{\partial \theta_i} + \sum_{j=1}^k \left\{ \frac{\partial w_j(\boldsymbol{\theta})}{\partial \theta_i} t_j(X) \right\} \\
0 &\stackrel{\text{lemma: } E\left[\frac{\partial \ln f(X|\boldsymbol{\theta})}{\partial \theta_i}\right] = 0}{=} E\left[\frac{\partial \ln f(X|\boldsymbol{\theta})}{\partial \theta_i}\right] = E\left[\frac{\partial \ln c(\boldsymbol{\theta})}{\partial \theta_i} + \sum_{j=1}^k \left\{ \frac{\partial w_j(\boldsymbol{\theta})}{\partial \theta_i} t_j(X) \right\}\right] \\
&= E\left[\frac{\partial \ln c(\boldsymbol{\theta})}{\partial \theta_i}\right] + E\left[\sum_{j=1}^k \frac{\partial w_j(\boldsymbol{\theta})}{\partial \theta_i} t_j(X)\right] \\
&\stackrel{E[c]=c}{=} \frac{\partial \ln c(\boldsymbol{\theta})}{\partial \theta_i} + E\left[\sum_{j=1}^k \frac{\partial w_j(\boldsymbol{\theta})}{\partial \theta_i} t_j(X)\right] \\
&= \frac{\partial \ln c(\boldsymbol{\theta})}{\partial \theta_i} + \sum_{j=1}^k E\left[\frac{\partial w_j(\boldsymbol{\theta})}{\partial \theta_i} t_j(X)\right] \\
&= \frac{\partial \ln c(\boldsymbol{\theta})}{\partial \theta_i} + \sum_{j=1}^k \frac{\partial w_j(\boldsymbol{\theta})}{\partial \theta_i} E[t_j(X)] \\
0 &= \frac{\partial \ln c(\boldsymbol{\theta})}{\partial \theta_i} + E\left[\sum_{j=1}^k \frac{\partial w_j(\boldsymbol{\theta})}{\partial \theta_i} t_j(X)\right] \\
E\left[\sum_{j=1}^k \frac{\partial w_j(\boldsymbol{\theta})}{\partial \theta_i} t_j(X)\right] &= -\frac{\partial \ln c(\boldsymbol{\theta})}{\partial \theta_i}
\end{aligned}$$

□

Theorem 15.16. exponential family variance

$$\begin{aligned}
X \in \mathcal{E}^f &= \left\{ f \middle| f = h(x) c(\boldsymbol{\theta}) \exp \left(\sum_{j=1}^k w_j(\boldsymbol{\theta}) t_j(x) \right) \right\} \\
&\Downarrow \forall i = 1, \dots, p \\
V\left[\sum_{j=1}^k \frac{\partial w_j(\boldsymbol{\theta})}{\partial \theta_i} t_j(X)\right] &= \frac{-\partial^2 \ln c(\boldsymbol{\theta})}{\partial \theta_i^2} - E\left[\sum_{j=1}^k \frac{\partial^2 w_j(\boldsymbol{\theta})}{\partial \theta_i^2} t_j(X)\right]
\end{aligned}$$

Proof:

same as the above

$$\begin{aligned}
f(x|\boldsymbol{\theta}) &= h(x) c(\boldsymbol{\theta}) 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}) e^{\sum_{j=1}^k w_j(\boldsymbol{\theta}) t_j(x)} \\
&= \ln h(x) + \ln c(\boldsymbol{\theta}) + \ln 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}$$

$$\begin{aligned}
\frac{\partial \ln f(X|\boldsymbol{\theta})}{\partial \theta_i} &= \frac{\partial \ln c(\boldsymbol{\theta})}{\partial \theta_i} + \sum_{j=1}^k \frac{\partial w_j(\boldsymbol{\theta})}{\partial \theta_i} t_j(X) \\
V \left[\frac{\partial \ln f(X|\boldsymbol{\theta})}{\partial \theta_i} \right] &= V \left[\frac{\partial \ln c(\boldsymbol{\theta})}{\partial \theta_i} + \sum_{j=1}^k \frac{\partial w_j(\boldsymbol{\theta})}{\partial \theta_i} t_j(X) \right] \\
V[aX+b] &\stackrel{=} {=} a^2 V[X] V \left[\sum_{j=1}^k \frac{\partial w_j(\boldsymbol{\theta})}{\partial \theta_i} t_j(X) \right] \\
V \left[\sum_{j=1}^k \frac{\partial w_j(\boldsymbol{\theta})}{\partial \theta_i} t_j(X) \right] &= V \left[\frac{\partial \ln f(X|\boldsymbol{\theta})}{\partial \theta_i} \right]
\end{aligned}$$

$$\begin{aligned}
& \text{V} \left[\sum_{j=1}^k \frac{\partial w_j(\boldsymbol{\theta})}{\partial \theta_i} t_j(X) \right] = \text{V} \left[\frac{\partial \ln f(X|\boldsymbol{\theta})}{\partial \theta_i} \right] = \text{E} \left[\left(\frac{\partial \ln f(X|\boldsymbol{\theta})}{\partial \theta_i} - \text{E} \left[\frac{\partial \ln f(X|\boldsymbol{\theta})}{\partial \theta_i} \right] \right)^2 \right] \\
& \stackrel{\text{lemma: } \text{E} \left[\frac{\partial \ln f(X|\boldsymbol{\theta})}{\partial \theta_i} \right] = 0}{=} \text{E} \left[\left(\frac{\partial \ln f(X|\boldsymbol{\theta})}{\partial \theta_i} - 0 \right)^2 \right] = \text{E} \left[\left(\frac{\partial \ln f(X|\boldsymbol{\theta})}{\partial \theta_i} \right)^2 \right] \\
& \stackrel{\text{lemma: } \text{E} \left[\frac{\partial^2 \ln f(X|\boldsymbol{\theta})}{\partial \theta_i^2} \right] = -\text{E} \left[\left(\frac{\partial \ln f(X|\boldsymbol{\theta})}{\partial \theta_i} \right)^2 \right]}{=} -\text{E} \left[\frac{\partial^2 \ln f(X|\boldsymbol{\theta})}{\partial \theta_i^2} \right] = -\text{E} \left[\frac{\partial}{\partial \theta_i} \frac{\partial \ln f(X|\boldsymbol{\theta})}{\partial \theta_i} \right] \\
& \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) \\
& = -\text{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] \\
& = -\text{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} - \text{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} - \text{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} - \text{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} - \text{E} \left[\sum_{j=1}^k \frac{\partial^2 w_j(\boldsymbol{\theta})}{\partial \theta_i^2} t_j(X) \right] \\
& \text{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} - \text{E} \left[\sum_{j=1}^k \frac{\partial^2 w_j(\boldsymbol{\theta})}{\partial \theta_i^2} t_j(X) \right]
\end{aligned}$$

□

exponential family expectation and variance

sense of downgrading by differentiation instead of integration

$$\begin{aligned}
X \in \mathcal{E}^f &= \left\{ f \mid f = h(x) c(\boldsymbol{\theta}) \exp \left(\sum_{j=1}^k w_j(\boldsymbol{\theta}) t_j(x) \right) \right\} \\
&\Rightarrow \begin{cases} \text{E} \left[\sum_{j=1}^k \frac{\partial w_j(\boldsymbol{\theta})}{\partial \theta_i} t_j(X) \right] = \frac{-\partial \ln c(\boldsymbol{\theta})}{\partial \theta_i} \\ \text{V} \left[\sum_{j=1}^k \frac{\partial w_j(\boldsymbol{\theta})}{\partial \theta_i} t_j(X) \right] = \frac{-\partial^2 \ln c(\boldsymbol{\theta})}{\partial \theta_i^2} - \text{E} \left[\sum_{j=1}^k \frac{\partial^2 w_j(\boldsymbol{\theta})}{\partial \theta_i^2} t_j(X) \right], \forall i = 1, \dots, p \end{cases}
\end{aligned}$$

15.1.3 multivariable distribution

univariable random vector

discrete: PMF equals probability function

$$f_x(x) = \text{P}(X = x)$$

continuous: PDF equals probability intensity

$$f_X(x) dx = dP(X \leq x)$$

Given $(X, Y) = (X_1, X_2) = \langle X, Y \rangle = \langle X_1, X_2 \rangle \sim f_{XY} = f_{XY}(x, y) = f_{X_1 X_2} = f_{X_1 X_2}(x_1, x_2)$

discrete:

Definition 15.19. JPMF = joint probability mass function: the JPMF of $(X_1, X_2) = \langle X_1, X_2 \rangle$ is

$$f_{X_1 X_2} = f_{X_1 X_2}(x_1, x_2) = P(X_1 = x_1, X_2 = x_2)$$

joint vs. marginal

Theorem 15.17. joint probability mass function can inference marginal probability mass function:

The marginal PMF of X_1 , $f_{X_1}(x_1) = P(X_1 = x_1)$ is given by

$$f_{X_1}(x_1) = \sum_{x_2 \in X_2(\Omega)} f_{X_1 X_2}(x_1, x_2) = \sum_{x_2 \in X_2(\Omega)} P(X_1 = x_1, X_2 = x_2)$$

Proof:

$$\begin{aligned} f_{X_1}(x_1) &= P(X_1 = x_1) \\ &= P(\{X_1 = x_1\} \cap \{X_2 \in (-\infty, \infty)\}) \\ &= P\left(\bigcup_{x_2 \in (-\infty, \infty)} \{X_1 = x_1 \wedge X_2 = x_2\}\right) \\ &= \bigcup_{x_2 \in (-\infty, \infty)} P(X_1 = x_1 \wedge X_2 = x_2) = \bigcup_{x_2 \in X_2(\Omega)} P(X_1 = x_1, X_2 = x_2) \\ &= \bigcup_{x_2 \in X_2(\Omega)} P(X_1 = x_1, X_2 = x_2) = \sum_{x_2 \in X_2(\Omega)} f_{X_1 X_2}(x_1, x_2) \end{aligned}$$

□

continuous:

Definition 15.20. JCDF = joint cumulative distribution function: the JCDF of $(X_1, X_2) = \langle X_1, X_2 \rangle$ is

$$F_{X_1 X_2} = F_{X_1 X_2}(x_1, x_2) = P(X_1 \leq x_1, X_2 \leq x_2)$$

Definition 15.21. JPDF = joint probability density function: the JPDF of $(X_1, X_2) = \langle X_1, X_2 \rangle$ is

$$f_{X_1 X_2} = f_{X_1 X_2}(x_1, x_2) = \frac{\partial^2}{\partial x_1 \partial x_2} F_{X_1 X_2}(x_1, x_2)$$

joint vs. marginal

Theorem 15.18. joint cumulative distribution function can inference marginal cumulative distribution function:

The marginal CDF of X_1 is

$$F_{X_1}(x_1) = F_{X_1 X_2}(x_1, \infty)$$

Proof:

$$\begin{aligned} F_{X_1}(x_1) &= P(X_1 \leq x_1) \\ &= P(X_1 \leq x_1 \wedge X_2 \leq \infty) \\ &= P(X_1 \leq x_1, X_2 \leq \infty) \\ &= F_{X_1 X_2}(x_1, \infty) \end{aligned}$$

□

Note:

$$F_{X_1}(x_1)$$

$$F_{X_1}(x_1) = F_{X_1 X_2}(x_1, \infty) = \int_{-\infty}^{\infty} \int_{-\infty}^{x_1} f_{X_1 X_2}(u_1, u_2) du_1 du_2$$

$$F_{X_1}(x_1) = \int_{-\infty}^{\infty} \int_{-\infty}^{x_1} f_{X_1 X_2}(u_1, u_2) du_1 du_2$$

$$f_{X_1}(x_1)$$

According to the fundamental theorem of calculus,

$$f_{X_1}(x_1) = \frac{d}{dx_1} F_{X_1}(x_1) = \frac{d}{dx_1} F_{X_1 X_2}(x_1, \infty) = \frac{d}{dx_1} \int_{-\infty}^{\infty} \int_{-\infty}^{x_1} f_{X_1 X_2}(u_1, u_2) du_1 du_2 = \int_{-\infty}^{\infty} f_{X_1 X_2}(x_1, u_2) du_2$$

$$f_{X_1}(x_1) = \int_{-\infty}^{\infty} f_{X_1 X_2}(x_1, u_2) du_2$$

$$f_{X_1}(x_1) = \int_{-\infty}^{\infty} f_{X_1 X_2}(x_1, x_2) dx_2$$

$$f_{X_1}(x_1) = \begin{cases} \int_{-\infty}^{\infty} f_{X_1 X_2}(x_1, x_2) dx_2 & \text{continuous} \\ \sum_{x_2 \in X_2(\Omega)} f_{X_1 X_2}(x_1, x_2) & \text{discrete} \end{cases}$$

Theorem 15.19. A function $f(x, y)$ is a joint PDF/PMF or JPDF/JPMF iff

$$\begin{cases} f(x, y) \geq 0 & \text{(ne) non-negative} \\ \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(x, y) dx dy = 1 & \text{continuous} \\ \sum_{x \in X(\Omega_X)} \sum_{y \in Y(\Omega_Y)} f(x, y) = 1 & \text{discrete} \end{cases} \quad (1) \text{ total event}$$

Definition 15.22. expected value: The expected value of a random vector $g(X, Y)$ is

$$E_{X,Y}[g(X, Y)] = E[g(X, Y)] = \begin{cases} \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} g(x, y) f_{XY}(x, y) dx dy & \text{continuous} \\ \sum_{x \in X(\Omega_X)} \sum_{y \in Y(\Omega_Y)} g(x, y) f_{XY}(x, y) & \text{discrete} \end{cases}$$

Def: 15.14

$$\begin{aligned} P((X, Y) \in E) = E[1((X, Y) \in E)] &= \begin{cases} \int_{y \in Y(\Omega_Y)} \int_{x \in X(\Omega_X)} 1((X, Y) \in E) f_{XY}(x, y) dx dy & \text{continuous} \\ \sum_{x \in X(\Omega_X)} \sum_{y \in Y(\Omega_Y)} 1((X, Y) \in E) g(x, y) f_{XY}(x, y) & \text{discrete} \end{cases} \\ &= \begin{cases} \iint_{(X, Y) \in E} f_{XY}(x, y) dx dy & \text{continuous} \\ \sum_{(X, Y) \in E} f_{XY}(x, y) & \text{discrete} \end{cases} \end{aligned}$$

example

$$f_{XY}(x, y) = e^{-y} 1(0 < x < y < \infty) \Rightarrow P(X + Y \leq 1)$$

Check if

$$f_{XY}(x, y) = e^{-y} 1(0 < x < y < \infty)$$

is a JPDF:

$$\left\{ \begin{array}{l} f_{XY}(x, y) = e^{-y} 1(0 < x < y < \infty) \\ \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f_{XY}(x, y) dx dy = ? \end{array} \right. \quad \left\{ \begin{array}{l} e^{-y} > 0 \\ 1(0 < x < y < \infty) \geq 0 \end{array} \right. = \begin{cases} 0 & (ne) \text{ non-negative} \\ 1 & (1) \text{ total event} \end{cases}$$

$$\begin{aligned} & \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f_{XY}(x, y) dx dy = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} e^{-y} 1(0 < x < y < \infty) dx dy, x \in (0, \infty) \\ &= \int_{-\infty}^{\infty} \int_0^{\infty} e^{-y} 1(0 < x < y < \infty) dx dy, y > x \in (0, \infty) \\ &= \int_0^{\infty} \int_0^{\infty} e^{-y} 1(0 < x < y < \infty) dx dy \stackrel{\text{Fubini}}{=} \int_0^{\infty} \int_0^{\infty} e^{-y} 1(0 < x < y < \infty) dy dx, y > x \\ &= \int_0^{\infty} \int_x^{\infty} e^{-y} 1 dy dx = \int_0^{\infty} \int_x^{\infty} e^{-y} dy dx = \int_0^{\infty} [-e^{-y}]_{y=x}^{\infty} dx = \int_0^{\infty} [-e^{-\infty} - (-e^{-x})] dx \\ &= \int_0^{\infty} [-0 - (-e^{-x})] dx = \int_0^{\infty} e^{-x} dx = [-e^{-x}]_{x=0}^{\infty} = [-e^{-\infty} - (-e^{-0})] = [-0 - (-1)] = 1 \end{aligned}$$

$$\left\{ \begin{array}{l} f_{XY}(x, y) = e^{-y} 1(0 < x < y < \infty) \geq 0 \\ \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f_{XY}(x, y) dx dy = 1 \end{array} \right. \quad \left\{ \begin{array}{l} (ne) \text{ non-negative} \\ (1) \text{ total event} \end{array} \right.$$

$$P(X + Y \leq 1)$$

<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]

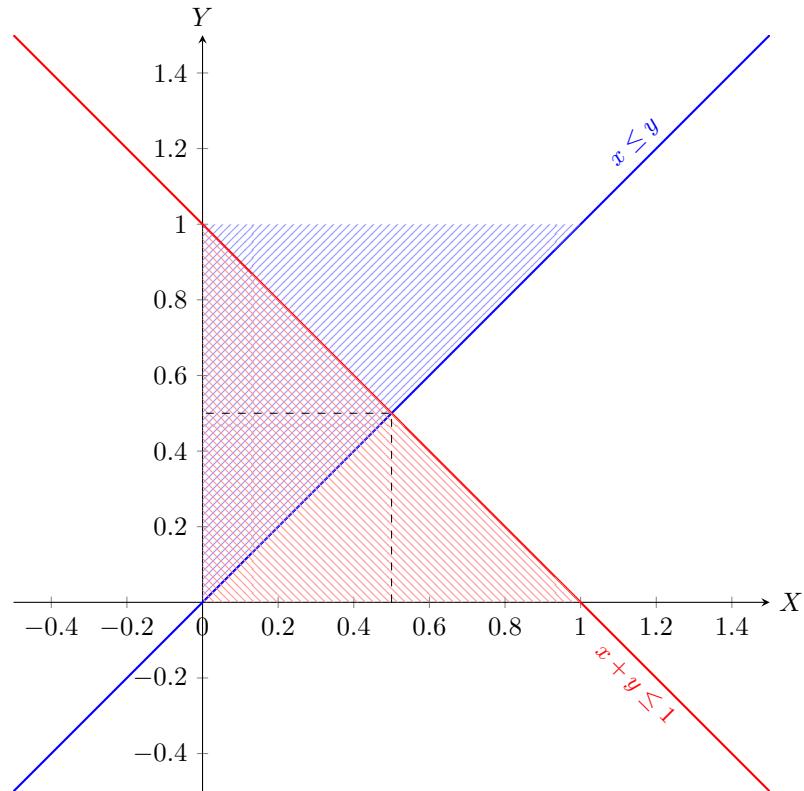


Figure 15.3: $X + Y \leq 1$

$$\begin{aligned}
P(X + Y \leq 1) &= \iint_{(X,Y) \in E} f_{XY}(x,y) dx dy, E = \{X + Y \leq 1\} \\
&= \iint_E e^{-y} \mathbf{1}(0 < x < y < \infty) dx dy, E = \{X + Y \leq 1\} \\
&= \int_0^{0.5} \int_x^{1-x} e^{-y} dy dx = \int_0^{0.5} [-e^{-y}]_{y=x}^{1-x} dx = \int_0^{0.5} [-e^{-(1-x)} - (-e^{-x})] dx \\
&= \int_0^{0.5} [e^{-x} - e^{x-1}] dx = [-e^{-x} - e^{x-1}]_{x=0}^{0.5} = [-e^{-0.5} - e^{0.5-1}] - [-e^0 - e^{0-1}] \\
&= 1 + e^{-1} - 2e^{-0.5} \approx 0.154818\ldots
\end{aligned}$$

□

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

$$\begin{aligned} U \otimes V &= UV^\top = U_i V_j \\ &= \begin{pmatrix} U_1 \\ U_2 \\ U_3 \end{pmatrix} \otimes \begin{pmatrix} V_1 \\ V_2 \\ V_3 \end{pmatrix} = \begin{pmatrix} U_1 \\ U_2 \\ U_3 \end{pmatrix} (V_1 \quad V_2 \quad V_3) = \begin{pmatrix} U_1 V_1 & U_1 V_2 & U_1 V_3 \\ U_2 V_1 & U_2 V_2 & U_2 V_3 \\ U_3 V_1 & U_3 V_2 & U_3 V_3 \end{pmatrix} \\ &= \begin{pmatrix} U_1 V^\top \\ U_2 V^\top \\ U_3 V^\top \end{pmatrix} = (UV_1 \quad UV_2 \quad UV_3) \end{aligned}$$

16.1.2 vector inner product: decreasing rank

vector inner product = rank-1 tensor times rank-1 tensor equals rank-0 tensor: decreasing rank

$$\begin{aligned} U \cdot V &= V^\top U = V_i U_i \\ &= \begin{pmatrix} U_1 \\ U_2 \\ U_3 \end{pmatrix} \cdot \begin{pmatrix} V_1 \\ V_2 \\ V_3 \end{pmatrix} = (V_1 \quad V_2 \quad V_3) \begin{pmatrix} U_1 \\ U_2 \\ U_3 \end{pmatrix} = V_1 U_1 + V_2 U_2 + V_3 U_3 \end{aligned}$$

16.1.3 tensor direct product: increasing rank

$$S \otimes T = S_{ik}T_{jl} \quad (16.1)$$

$$(i, j), (k, l) \in \{(1, 1), (1, 2), (2, 1), (2, 2)\} \quad (16.2)$$

$$\begin{pmatrix} S_{11} & S_{12} \\ S_{21} & S_{22} \end{pmatrix} \otimes \begin{pmatrix} T_{11} & T_{12} \\ T_{21} & T_{22} \end{pmatrix} = \begin{pmatrix} S_{11}T_{11} & S_{11}T_{12} & S_{12}T_{11} & S_{12}T_{12} \\ S_{11}T_{21} & S_{11}T_{22} & S_{12}T_{21} & S_{12}T_{22} \\ S_{21}T_{11} & S_{21}T_{12} & S_{22}T_{11} & S_{22}T_{12} \\ S_{21}T_{21} & S_{21}T_{22} & S_{22}T_{21} & S_{22}T_{22} \end{pmatrix} \quad (16.3)$$

$$\begin{pmatrix} S_{11} & S_{12} \\ S_{21} & S_{22} \end{pmatrix} \begin{pmatrix} T_{11} & T_{12} \\ T_{21} & T_{22} \end{pmatrix} = \begin{pmatrix} S_{11}T & S_{12}T \\ S_{21}T & S_{22}T \end{pmatrix} \quad (16.4)$$

$$= \begin{pmatrix} S_{11} \begin{pmatrix} T_{11} & T_{12} \\ T_{21} & T_{22} \end{pmatrix} & S_{12} \begin{pmatrix} T_{11} & T_{12} \\ T_{21} & T_{22} \end{pmatrix} \\ S_{21} \begin{pmatrix} T_{11} & T_{12} \\ T_{21} & T_{22} \end{pmatrix} & S_{22} \begin{pmatrix} T_{11} & T_{12} \\ T_{21} & T_{22} \end{pmatrix} \end{pmatrix} \quad (16.5)$$

16.2 covariance matrix and its properties

8

$$\begin{aligned} C[\mathbf{X}] = \text{Cov}[\mathbf{X}] = V[\mathbf{X}] &= E[(\mathbf{X} - E(\mathbf{X}))(\mathbf{X} - E(\mathbf{X}))^\top] \\ &= E[(\mathbf{X} - E(\mathbf{X}))(\mathbf{X}^\top - E(\mathbf{X})^\top)] \\ &= E[\mathbf{X}\mathbf{X}^\top - E(\mathbf{X})\mathbf{X}^\top - E\mathbf{X}(E(\mathbf{X})^\top) + E(\mathbf{X})E(\mathbf{X})^\top] \\ &= E[\mathbf{X}\mathbf{X}^\top] - E[E(\mathbf{X})\mathbf{X}^\top] - E[\mathbf{X}E(\mathbf{X})^\top] + E[E(\mathbf{X})E(\mathbf{X})^\top] \\ &= E[\mathbf{X}\mathbf{X}^\top] - E(\mathbf{X})E[\mathbf{X}^\top] - E[\mathbf{X}]E(\mathbf{X})^\top + E(\mathbf{X})E(\mathbf{X})^\top \\ &= E[\mathbf{X}\mathbf{X}^\top] - E(\mathbf{X})E(\mathbf{X})^\top - E(\mathbf{X})E(\mathbf{X})^\top + E(\mathbf{X})E(\mathbf{X})^\top \\ &= E[\mathbf{X}\mathbf{X}^\top] - E(\mathbf{X})E(\mathbf{X})^\top \end{aligned}$$

$$\begin{aligned} \mathbf{X} = [X]_{1 \times 1} = X \Rightarrow C(X) = C[\mathbf{X}] &= E[\mathbf{X}\mathbf{X}^\top] - E(\mathbf{X})E(\mathbf{X})^\top \\ &= E[XX] - E(X)E(X) \\ &= E(X^2) - [E(X)]^2 = V(X) \end{aligned}$$

16.2.1 $V[\mathbf{X} + \mathbf{b}] = V[\mathbf{X}]$

$$\begin{aligned} V[\mathbf{X} + \mathbf{b}] &= E[((\mathbf{X} + \mathbf{b}) - E(\mathbf{X} + \mathbf{b}))((\mathbf{X} + \mathbf{b}) - E(\mathbf{X} + \mathbf{b}))^\top] \\ &\stackrel{E(\mathbf{X}+\mathbf{b})=E(\mathbf{X})+\mathbf{b}}{=} E[(\mathbf{X} + \mathbf{b} - E(\mathbf{X}) - \mathbf{b})(\mathbf{X} + \mathbf{b} - E(\mathbf{X}) - \mathbf{b})^\top] \\ &= E[(\mathbf{X} - E(\mathbf{X}))(\mathbf{X} - E(\mathbf{X}))^\top] = V[\mathbf{X}] \end{aligned}$$

16.2.2 $V[A\mathbf{X}] = AV[\mathbf{X}]A^\top$

$$\begin{aligned} V[A\mathbf{X}] &= E[((A\mathbf{X}) - E(A\mathbf{X}))((A\mathbf{X}) - E(A\mathbf{X}))^\top] \\ &\stackrel{E(A\mathbf{X})=AE(\mathbf{X})}{=} E[(A\mathbf{X} - AE(\mathbf{X}))(A\mathbf{X} - AE(\mathbf{X}))^\top] \\ &= E[A(\mathbf{X} - E(\mathbf{X}))(A(\mathbf{X} - E(\mathbf{X})))^\top] \\ &= E[A(\mathbf{X} - E(\mathbf{X}))(\mathbf{X} - E(\mathbf{X}))^\top A^\top] \\ &= AE[(\mathbf{X} - E(\mathbf{X}))(\mathbf{X} - E(\mathbf{X}))^\top] A^\top = AV[\mathbf{X}]A^\top \end{aligned}$$

16.2.3 $V[A\mathbf{X} + \mathbf{b}] = AV[\mathbf{X}]A^\top$

$$V[A\mathbf{X} + \mathbf{b}] = V[A\mathbf{X}] = AV[\mathbf{X}]A^\top$$

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.packages()
```

```
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
```

```
# ?c
# help(c)
```

Ctrl + L to clean R console
path with slash / in R, differing backslash \ in M\$ Windows

19.1.1.1 function

```
add <- function(x, y) {
  return(x + y)
}

add(11, 13)
```

```
## [1] 24
```

$$BMI = \frac{BW \text{ [Kg]}}{BH \text{ [m]}^2}$$

```
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

```
abs(-5:5)
```

```
## [1] 5 4 3 2 1 0 1 2 3 4 5
```

```
# install.packages("magrittr")
```

```
library(magrittr)
```

```
##
## Attaching package: 'magrittr'
```

```
## The following object is masked _by_ '.GlobalEnv':
```

```
##
```

```
##      add
```

```
-5:5 %>% abs()
```

```
## [1] 5 4 3 2 1 0 1 2 3 4 5
```

```
# 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
# less line but also less readability
sys_date_num <- as.numeric(format(Sys.Date(), format = "%Y"))
sys_date_num

## [1] 2024
# 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**

```
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
# install.packages("gapminder")

library(gapminder)

## Warning: package 'gapminder' was built under R version 4.2.3
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.

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
```

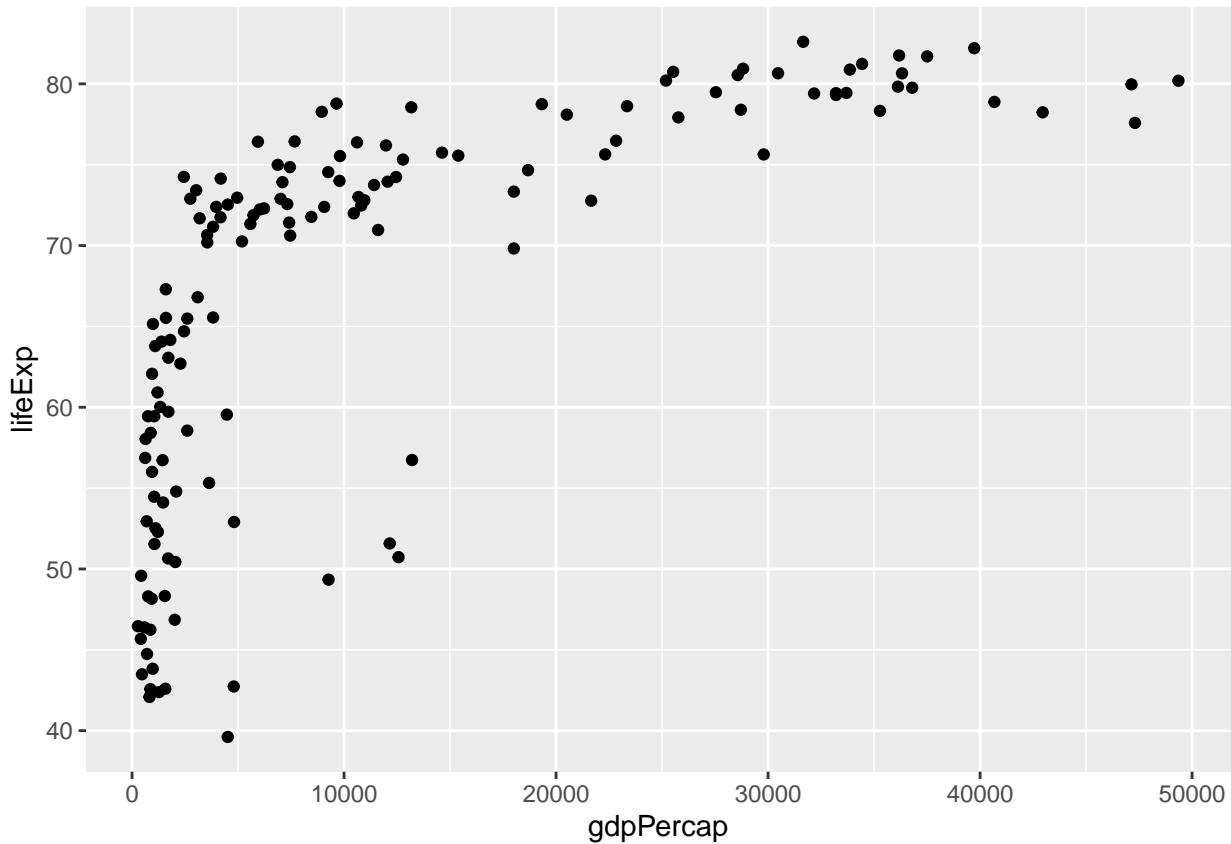
19.1.5 visualization statically with ggplot2

```
library(ggplot2)

## Warning: package 'ggplot2' was built under R version 4.2.3

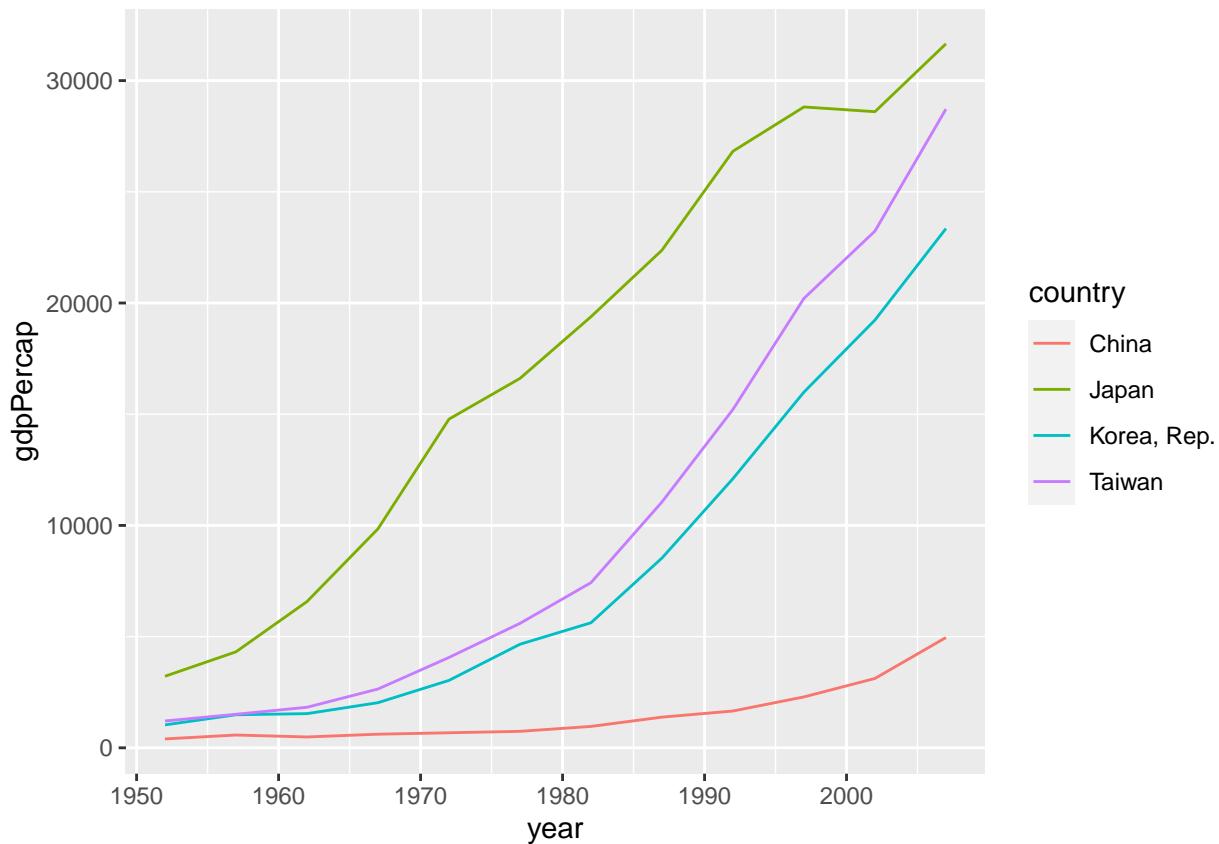
library(gapminder)

gapminder_2007 <- gapminder %>%
  filter(year == 2007)
scatter_plot <- ggplot(gapminder_2007, aes(x = gdpPercap, y = lifeExp)) +
  geom_point()
scatter_plot
```



```
library(ggplot2)
library(gapminder)

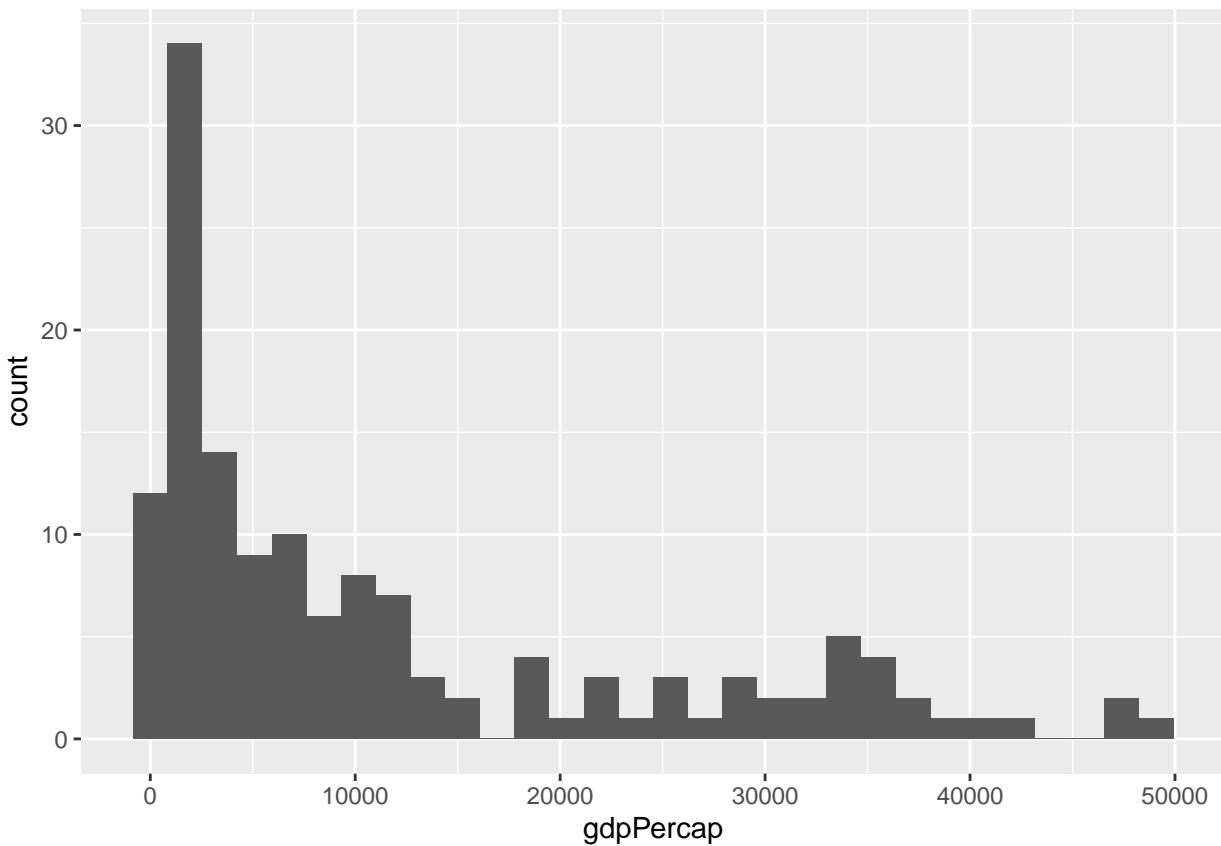
north_asia <- gapminder %>%
  filter(country %in% c("China", "Japan", "Taiwan", "Korea, Rep."))
line_plot <- ggplot(north_asia, aes(x = year, y = gdpPercap, colour = country)) +
  geom_line()
line_plot
```



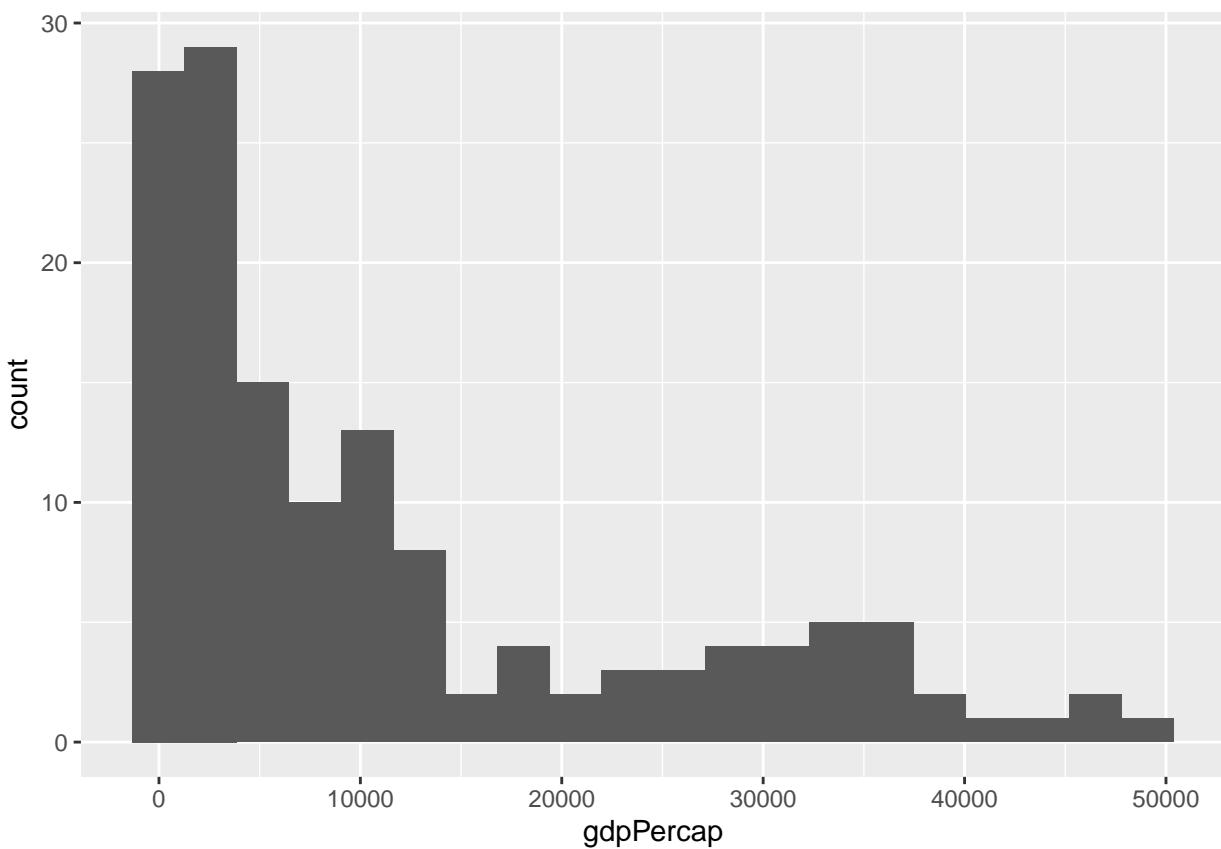
```
library(ggplot2)
library(gapminder)

hist_plot <- ggplot(gapminder_2007, aes(x = gdpPercap)) +
  geom_histogram()
hist_plot
```

```
## `stat_bin()` using `bins = 30`. Pick better value with `binwidth`.
```

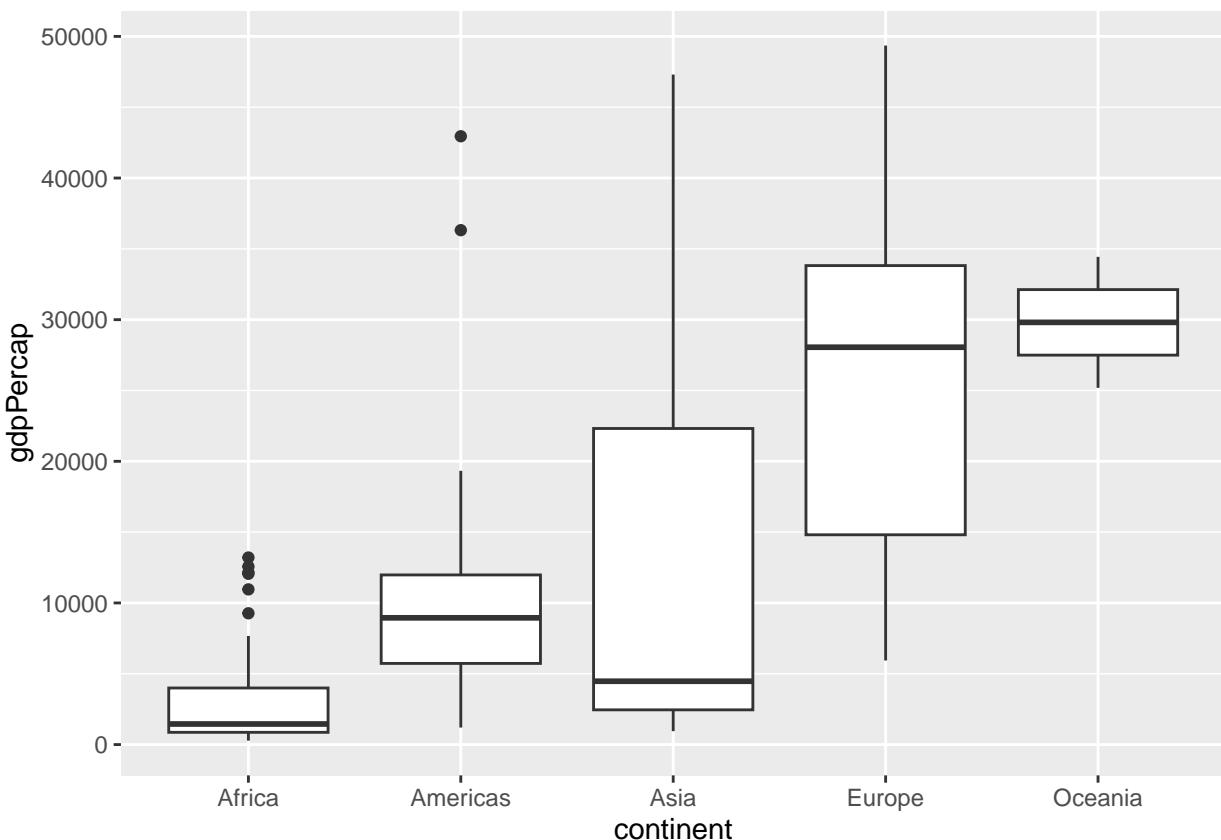


```
hist_plot <- ggplot(gapminder_2007, aes(x = gdpPercap)) +  
  geom_histogram(bins = 20)  
hist_plot
```



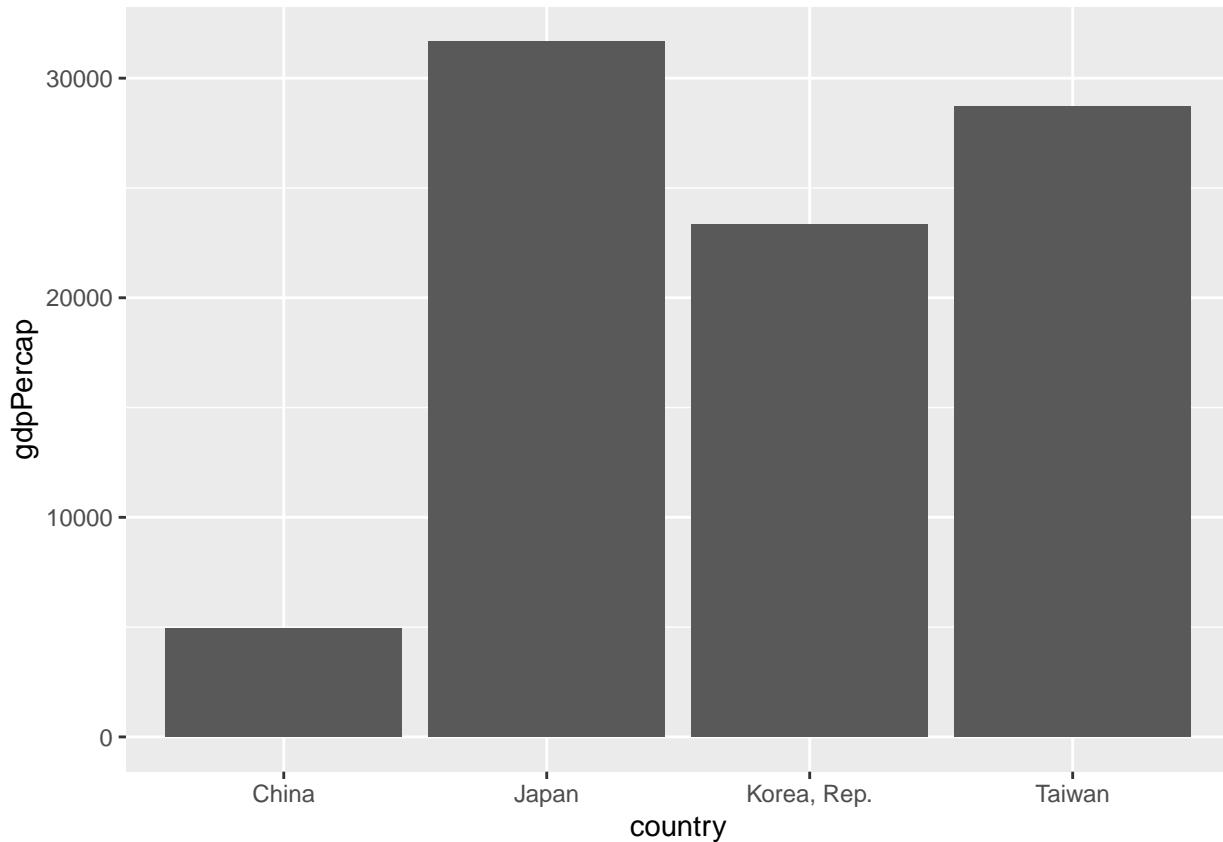
```
library(ggplot2)  
library(gapminder)
```

```
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

```
class(2L)
```

```
## [1] "integer"
```

```
class(2.0L)
```

```
## [1] "integer"
```

```
class(2.3L)
```

```
## [1] "numeric"
```

time: POSIXct POSIXt

```
class(Sys.time())
```

```
## [1] "POSIXct" "POSIXt"
```

```
0 %in% -5:5
```

```
## [1] TRUE
```

19.1.7.1 date

1970-01-01 = 0L

```
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

```
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

```
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]
- n D
 - `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"

rep(7L, times = 8)
```

19.1.8.1.2 rep repeat

```
## [1] 7 7 7 7 7 7 7 7

rep("R", times = 10)

## [1] "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

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>

```
four_seasons <- c("spring", "summer", "autumn", "winter")
four_seasons
```

```
## [1] "spring" "summer" "autumn" "winter"
four_seasons_factor <- factor(four_seasons)
four_seasons_factor
```

```
## [1] spring summer autumn winter
## Levels: autumn spring summer winter
```

```
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
```

```
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

```
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>

```

my_mat <- matrix(1:6, nrow = 2)
my_mat

##      [,1] [,2] [,3]
## [1,]    1    3    5
## [2,]    2    4    6
class(my_mat)

## [1] "matrix" "array"
my_mat2 <- matrix(1:6, nrow = 2, byrow = TRUE)
my_mat2

##      [,1] [,2] [,3]
## [1,]    1    2    3
## [2,]    4    5    6
my_mat2[2, 3]

## [1] 6
my_mat2[2, ]

## [1] 4 5 6
my_mat2[, 3]

## [1] 3 6
filter <- my_mat2 < 6 & my_mat2 > 1
my_mat2[filter]

## [1] 4 2 5 3

```

boolean will become value in a matrix, like vector

```

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
class(my_mat3[, 2])

## [1] "numeric"

```

19.1.8.4 data frame

- variable: column
- observation: row
- value: cell

```

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

```

```

great_nba_teams[1, 1]

## [1] "Chicago Bulls"
great_nba_teams[1, ]

##      team_name wins losses is_champion season
## 1 Chicago Bulls    72     10        TRUE 1995-96
great_nba_teams[, 1]

## [1] "Chicago Bulls"           "Golden State Warriors"
stringsAsFactors = TRUE

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

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"

great_nba_teams$team_name

```

19.1.8.4.1 selecting variable or column

```

## [1] "Chicago Bulls"           "Golden State Warriors"
great_nba_teams[, "team_name"]

## [1] "Chicago Bulls"           "Golden State Warriors"

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

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)
}
```

```
my_function()

## [1] "R is fantastic"
print(txt)

## [1] "fantastic"

txt <- "awesome"
my_function <- function() {
  txt <<- "fantastic"
  paste("R is", txt)
}

my_function()

## [1] "R is fantastic"
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

```
# 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

# 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

multiarray[2, 3, 2]

## [1] 22
```

19.2.6.2 list

https://www.w3schools.com/r/r_lists.asp

19.3 Apan Liao

<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>

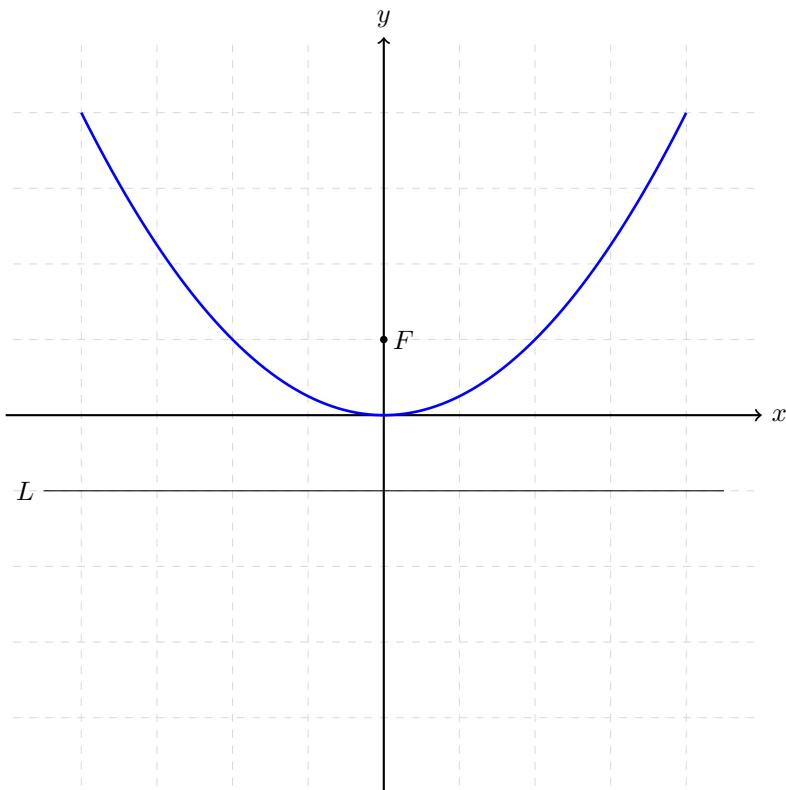


Figure 21.1: parabola defined by focus, directrix, eccentricity

21.1 Cartesian coordinate: focus, directrix, eccentricity

focus, directrix, eccentricity 焦點, 準線, 離心率

$$\begin{cases} F = (0, y_F) \\ L = y - y_L = 0 \\ \epsilon = \frac{\overline{PF}}{d(P, L)} = \frac{\|(x, y) - (0, y_F)\|}{\|y - y_L\|} \end{cases}$$

F : focus L : directrix
 ϵ : eccentricity

$$0 \leq \epsilon = \frac{\overline{PF}}{d(P, L)} = \frac{\overline{PF}}{\overline{PP'}} = \frac{\|(x, y) - (0, y_F)\|}{\|(x, y) - (x, y_L)\|} = \frac{\|(x, y - y_F)\|}{\|(0, y - y_L)\|} = \frac{\sqrt{x^2 + (y - y_F)^2}}{\sqrt{(y - y_L)^2}} \quad (21.1)$$

$$\epsilon^2 = \frac{x^2 + (y - y_F)^2}{(y - y_L)^2} = \frac{x^2 + y^2 - 2y_F y + y_F^2}{y^2 - 2y_L y + y_L^2} \quad (21.2)$$

$$0 = x^2 + (1 - \epsilon^2) y^2 - 2(y_F - \epsilon^2 y_L) y + (y_F^2 - \epsilon^2 y_L^2) \quad (21.3)$$

$$\stackrel{\epsilon \neq 1}{=} x^2 + (1 - \epsilon^2) \left[y^2 - \frac{2(y_F - \epsilon^2 y_L)}{1 - \epsilon^2} y + \frac{y_F^2 - \epsilon^2 y_L^2}{1 - \epsilon^2} \right] \quad (21.4)$$

$$= x^2 + (1 - \epsilon^2) \quad (21.5)$$

$$\left[y^2 - \frac{2(y_F - \epsilon^2 y_L)}{1 - \epsilon^2} y + \left(\frac{y_F - \epsilon^2 y_L}{1 - \epsilon^2} \right)^2 - \left(\frac{y_F - \epsilon^2 y_L}{1 - \epsilon^2} \right)^2 + \frac{y_F^2 - \epsilon^2 y_L^2}{1 - \epsilon^2} \right] \quad (21.6)$$

$$= x^2 + (1 - \epsilon^2) \left[\left(y - \frac{y_F - \epsilon^2 y_L}{1 - \epsilon^2} \right)^2 + \frac{(y_F^2 - \epsilon^2 y_L^2)(1 - \epsilon^2) - (y_F - \epsilon^2 y_L)^2}{(1 - \epsilon^2)^2} \right] \quad (21.7)$$

$$= x^2 + (1 - \epsilon^2) \left(y - \frac{y_F - \epsilon^2 y_L}{1 - \epsilon^2} \right)^2 + \frac{(y_F^2 - \epsilon^2 y_L^2)(1 - \epsilon^2) - (y_F - \epsilon^2 y_L)^2}{1 - \epsilon^2} \quad (21.8)$$

$$\begin{aligned} & (y_F^2 - \epsilon^2 y_L^2)(1 - \epsilon^2) - (y_F - \epsilon^2 y_L)^2 \\ &= (1 - \epsilon^2) y_F^2 - (\epsilon^2 - \epsilon^4) y_L^2 - y_F^2 + 2\epsilon^2 y_F y_L - \epsilon^4 y_L^2 \\ &= -\epsilon^2 y_F^2 - \epsilon^2 y_L^2 + 2\epsilon^2 y_F y_L = -\epsilon^2 (y_F - y_L)^2 \end{aligned}$$

$$\begin{aligned} \frac{\epsilon^2 (y_F - y_L)^2}{1 - \epsilon^2} &\stackrel{\epsilon \neq 1}{=} x^2 + (1 - \epsilon^2) \left(y - \frac{y_F - \epsilon^2 y_L}{1 - \epsilon^2} \right)^2 \\ &\stackrel{\epsilon \neq 0, 1}{=} 1 \begin{cases} \left(\frac{x - 0}{\epsilon (y_F - y_L)} \right)^2 + \left(\frac{y - \frac{y_F - \epsilon^2 y_L}{1 - \epsilon^2}}{\epsilon (y_F - y_L)} \right)^2 & 1 - \epsilon^2 > 0 \stackrel{\epsilon > 0}{\Rightarrow} 0 < \epsilon < 1 \\ -\left(\frac{x - 0}{\epsilon (y_F - y_L)} \right)^2 + \left(\frac{y - \frac{y_F - \epsilon^2 y_L}{1 - \epsilon^2}}{\epsilon (y_F - y_L)} \right)^2 & 1 - \epsilon^2 < 0 \stackrel{\epsilon > 0}{\Rightarrow} \epsilon > 1 \end{cases} \end{aligned}$$

$$\epsilon = 0 \text{ or } \lim_{|y_L| \rightarrow \infty} \epsilon = 0$$

$$r = \overline{PF} = \|(x, y) - (0, y_F)\| = \|(x, y - y_F)\| = \sqrt{x^2 + (y - y_F)^2}$$

$$\epsilon = \frac{r}{d(P, L)} = \frac{\overline{PF}}{\overline{PP'}} = \frac{\|(x, y) - (0, y_F)\|}{\|(x, y) - (x, y_L)\|} = \frac{\|(x, y - y_F)\|}{\|(0, y - y_L)\|} = \frac{\sqrt{x^2 + (y - y_F)^2}}{|y - y_L|}$$

$$\lim_{|y_L| \rightarrow \infty} \epsilon = \lim_{|y_L| \rightarrow \infty} \frac{r}{d(P, L)} = \lim_{|y_L| \rightarrow \infty} \frac{\sqrt{x^2 + (y - y_F)^2}}{|y - y_L|} = 0$$

$$\epsilon = 1$$

$$\begin{aligned} 0 &= x^2 + (1 - \epsilon^2) y^2 - 2(y_F - \epsilon^2 y_L) y + (y_F^2 - \epsilon^2 y_L^2) \\ &\stackrel{\epsilon = 1}{=} x^2 + (1 - 1^2) y^2 - 2(y_F - 1^2 y_L) y + (y_F^2 - 1^2 y_L^2) \\ &= x^2 - 2(y_F - y_L) y + (y_F^2 - y_L^2) \\ &= x^2 - 2(y_F - y_L) y + (y_F + y_L)(y_F - y_L) \\ x^2 &= 2(y_F - y_L) \left(y - \frac{y_F + y_L}{2} \right) \end{aligned}$$

Let one curve vertex $P = V = (0, 0)$ on the curve, and fix the directrix L or y_L ,

$$\epsilon \neq 1$$

$$\begin{aligned} 1 &\stackrel{P(x,y)=V(0,0)}{=} 0 + \left(\frac{0 - \frac{y_F - \epsilon^2 y_L}{1 - \epsilon^2}}{\frac{\epsilon(y_F - y_L)}{1 - \epsilon^2}} \right)^2 \\ &\Rightarrow y_F - \epsilon^2 y_L = \pm \epsilon (y_F - y_L) \\ &\Rightarrow \begin{cases} (1 - \epsilon) y_F = \epsilon (\epsilon - 1) y_L & + \\ (1 + \epsilon) y_F = \epsilon (\epsilon + 1) y_L & - \end{cases} \\ &\Rightarrow y_F = \begin{cases} -\epsilon y_L & + \\ \epsilon y_L & - \end{cases} \end{aligned}$$

$$\epsilon = 1$$

$$\begin{aligned} x^2 &= 2(y_F - y_L) \left(y - \frac{y_F + y_L}{2} \right) \\ &\stackrel{P(x,y)=V(0,0)}{=} 0^2 = 2(y_F - y_L) \left(0 - \frac{y_F + y_L}{2} \right) \\ &\Rightarrow 0 = (y_F - y_L)(y_F + y_L) \\ &\Rightarrow y_F = \mp y_L \end{aligned}$$

or by definition of eccentricity (21.1)

$$\begin{aligned} 0 \leq \epsilon &= \frac{\overline{PF}}{d(P, L)} = \frac{\overline{PF}}{\overline{PP'}} = \frac{\|(x, y) - (0, y_F)\|}{\|(x, y) - (x, y_L)\|} = \frac{\|(x, y - y_F)\|}{\|(0, y - y_L)\|} = \frac{\sqrt{x^2 + (y - y_F)^2}}{\sqrt{(y - y_L)^2}} \\ &\stackrel{P(x,y)=V(0,0)}{=} \frac{\sqrt{0^2 + (0 - y_F)^2}}{\sqrt{(0 - y_L)^2}} = \sqrt{\left(\frac{y_F}{y_L}\right)^2} \\ \epsilon^2 &= \left(\frac{y_F}{y_L}\right)^2 \Rightarrow y_F = \mp \epsilon y_L \end{aligned}$$

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^]

$$0 \leq \epsilon = \frac{\overline{PF}}{d(P, L)} = \frac{\sqrt{(x - x_F)^2 + (y - y_F)^2}}{\frac{|A'x + B'y + C'|}{\sqrt{A'^2 + B'^2}}} = \frac{\sqrt{(x - \alpha)^2 + (y - \varphi)^2}}{|Ax + By + C|}, \begin{cases} A = \frac{A'}{\sqrt{A'^2 + B'^2}} \\ B = \frac{B'}{\sqrt{A'^2 + B'^2}} \\ C = \frac{C'}{\sqrt{A'^2 + B'^2}} \end{cases}$$

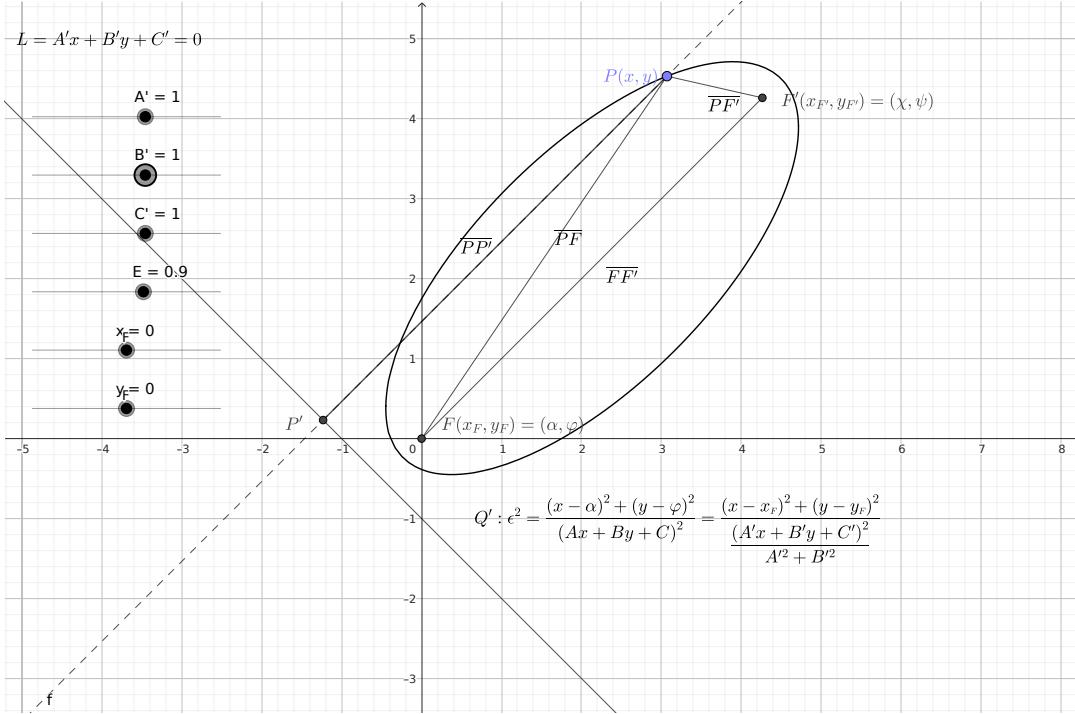


Figure 21.2: conic sections

$$A^2 + B^2 = \left(\frac{A'}{\sqrt{A'^2 + B'^2}} \right)^2 + \left(\frac{B'}{\sqrt{A'^2 + B'^2}} \right)^2 = 1$$

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

$$2c = \overline{FF'} = \|(x_F, y_F) - (x_{F'}, y_{F'})\| = \|(\alpha, \varphi) - (\chi, \psi)\| \\ = \sqrt{(\alpha - \chi)^2 + (\chi - \psi)^2}$$

$$D = \begin{cases} \sqrt{(x - x_F)^2 + (y - y_F)^2} + \sqrt{(x - x_{F'})^2 + (y - y_{F'})^2} & \text{ellipse} \\ \sqrt{(x - x_F)^2 + (y - y_F)^2} - \sqrt{(x - x_{F'})^2 + (y - y_{F'})^2} & \text{hyperbola} \end{cases} \\ = \sqrt{(x - x_F)^2 + (y - y_F)^2} \pm \sqrt{(x - x_{F'})^2 + (y - y_{F'})^2} \\ = \sqrt{(x - \alpha)^2 + (y - \varphi)^2} \pm \sqrt{(x - \chi)^2 + (y - \psi)^2}$$

$$(x - \alpha)^2 + (y - \varphi)^2 = \left(D \mp \sqrt{(x - \chi)^2 + (y - \psi)^2} \right)^2 \\ = D^2 \mp 2D\sqrt{(x - \chi)^2 + (y - \psi)^2} \\ + (x - \chi)^2 + (y - \psi)^2$$

$$\begin{aligned}
D^2 &= (x - \alpha)^2 + (y - \varphi)^2 + (x - \chi)^2 + (y - \psi)^2 \\
&\quad \pm 2\sqrt{\left[(x - \alpha)^2 + (y - \varphi)^2\right]\left[(x - \chi)^2 + (y - \psi)^2\right]} \\
&\quad (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]} \\
&\quad \left[(x - \alpha)^2 + (y - \varphi)^2 + (x - \chi)^2 + (y - \psi)^2\right]^2 + D^4 \\
&\quad - 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] \\
&\quad \left[(x - \alpha)^2 + (y - \varphi)^2\right]^2 + \left[(x - \chi)^2 + (y - \psi)^2\right]^2 \\
&\quad + 2 \left[(x - \alpha)^2 + (y - \varphi)^2\right] \left[(x - \chi)^2 + (y - \psi)^2\right] + D^4 \\
&\quad - 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 \\
&\quad - 2 \left[(x - \alpha)^2 + (y - \varphi)^2\right] \left[(x - \chi)^2 + (y - \psi)^2\right] + D^4 \\
&\quad - 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 \\
&\quad - 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 \\
&\quad - 2D^2 \left\{ \left[(x - \chi)^2 + (y - \psi)^2\right] - \left[(x - \alpha)^2 + (y - \varphi)^2\right] \right\} \\
&\quad - 4D^2 \left[(x - \alpha)^2 + (y - \varphi)^2\right] \\
&\quad (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 - (\alpha^2 - \chi^2) + 2(\varphi - \psi)y - (\varphi^2 - \psi^2) - D^2 \right\}^2 \\
&= \left\{ 2(\alpha - \chi)x + 2(\varphi - \psi)y - [(\alpha^2 - \chi^2) + (\varphi^2 - \psi^2) + D^2] \right\}^2
\end{aligned}$$

$D \neq 0$

$$\begin{aligned}
&(x - \alpha)^2 + (y - \varphi)^2 \\
&= \left[\frac{\alpha - \chi}{D}x + \frac{\varphi - \psi}{D}y - \left(\frac{\alpha^2 - \chi^2}{2D} + \frac{\varphi^2 - \psi^2}{2D} + \frac{D}{2} \right) \right]^2
\end{aligned}$$

$$\begin{cases} (x - \alpha)^2 + (y - \varphi)^2 = [\epsilon(Ax + By + C)]^2 \\ (x - \alpha)^2 + (y - \varphi)^2 = \left[\frac{\alpha - \chi}{D}x + \frac{\varphi - \psi}{D}y - \left(\frac{\alpha^2 - \chi^2}{2D} + \frac{\varphi^2 - \psi^2}{2D} + \frac{D}{2} \right) \right]^2 \end{cases}$$

$$(A, B, C) \rightleftharpoons (\chi, \psi, D)$$

$$\begin{cases}
\epsilon A = \pm \frac{\alpha - \chi}{D} & \chi \pm \epsilon AD = \alpha \\
\epsilon B = \pm \frac{\varphi - \psi}{D} & \psi \pm \epsilon BD = \varphi \\
\epsilon C = \mp \left(\frac{\alpha^2 - \chi^2}{2D} + \frac{\varphi^2 - \psi^2}{2D} + \frac{D}{2} \right) & \\
\end{cases}$$

$$\begin{aligned}
2\epsilon C &= \mp \left(\frac{\alpha - \chi}{D} (\alpha + \chi) + \frac{\varphi - \psi}{D} (\varphi + \psi) + D \right) \\
&= \mp (\pm \epsilon A(\alpha + \chi) \pm \epsilon B(\varphi + \psi) + D) \\
\mp \epsilon (A\alpha + B\varphi + 2C) &= \pm \epsilon A\chi \pm \epsilon B\psi + D
\end{aligned}$$

$$\begin{pmatrix} 1 & 0 & \pm \epsilon A \\ 0 & 1 & \pm \epsilon B \\ \pm \epsilon A & \pm \epsilon B & 1 \end{pmatrix} \begin{pmatrix} \chi \\ \psi \\ D \end{pmatrix} = \begin{pmatrix} \alpha \\ \varphi \\ \mp \epsilon (A\alpha + B\varphi + 2C) \end{pmatrix}$$

$$\begin{pmatrix} 1 & 0 & \pm \epsilon A & \alpha \\ 0 & 1 & \pm \epsilon B & \varphi \\ 0 & \pm \epsilon B & 1 \mp \epsilon^2 A^2 & \mp \epsilon (2A\alpha + B\varphi + 2C) \end{pmatrix}$$

$$\begin{pmatrix} 1 & 0 & \pm \epsilon A & \alpha \\ 0 & 1 & \pm \epsilon B & \varphi \\ 0 & 0 & 1 \mp \epsilon^2 A^2 \mp \epsilon^2 B^2 & \mp \epsilon (2A\alpha + 2B\varphi + 2C) \end{pmatrix}$$

$$\begin{pmatrix} 1 & 0 & \pm \epsilon A & \alpha \\ 0 & 1 & \pm \epsilon B & \varphi \\ 0 & 0 & 1 & \frac{\mp 2\epsilon (A\alpha + B\varphi + C)}{1 \mp \epsilon^2 (A^2 + B^2)} \end{pmatrix}$$

$$A^2 + B^2 = \left(\frac{A'}{\sqrt{A'^2 + B'^2}} \right)^2 + \left(\frac{B'}{\sqrt{A'^2 + B'^2}} \right)^2 = 1$$

$$\begin{cases}
\chi = \alpha \mp \epsilon AD = \alpha \mp \epsilon \frac{A'}{\sqrt{A'^2 + B'^2}} D \\
\psi = \varphi \mp \epsilon BD = \varphi \mp \epsilon \frac{B'}{\sqrt{A'^2 + B'^2}} D \\
D = \frac{\mp 2\epsilon (A\alpha + B\varphi + C)}{1 \mp \epsilon^2 (A^2 + B^2)} = \frac{\mp 2\epsilon}{1 \mp \epsilon^2} \frac{A'\alpha + B'\varphi + C'}{\sqrt{A'^2 + B'^2}} \quad A^2 + B^2 = 1
\end{cases}$$

actually, only one of two solutions is true

$$\begin{cases}
\chi = \alpha - \epsilon AD = \alpha - \epsilon \frac{A'}{\sqrt{A'^2 + B'^2}} D = \alpha - \frac{2\epsilon^2}{\epsilon^2 - 1} \frac{A'^2 \alpha + A'B'\varphi + A'C'}{A'^2 + B'^2} \\
\psi = \varphi - \epsilon BD = \varphi - \epsilon \frac{B'}{\sqrt{A'^2 + B'^2}} D = \varphi - \frac{2\epsilon^2}{\epsilon^2 - 1} \frac{A'B'\alpha + B'^2 \varphi + B'C'}{A'^2 + B'^2} \\
D = \frac{-2\epsilon (A\alpha + B\varphi + C)}{1 - \epsilon^2 (A^2 + B^2)} = \frac{-2\epsilon}{1 - \epsilon^2} \frac{A'\alpha + B'\varphi + C'}{\sqrt{A'^2 + B'^2}} = \frac{2\epsilon}{\epsilon^2 - 1} \frac{A'\alpha + B'\varphi + C'}{\sqrt{A'^2 + B'^2}}
\end{cases}$$

$$\begin{cases}
\chi = \frac{(\epsilon^2 - 1)(A'^2 + B'^2)\alpha - 2\epsilon^2(A'^2\alpha + A'B'\varphi + A'C')}{(\epsilon^2 - 1)(A'^2 + B'^2)} \\
\psi = \frac{(\epsilon^2 - 1)(A'^2 + B'^2)\varphi - 2\epsilon^2(A'B'\alpha + B'^2\varphi + B'C')}{(\epsilon^2 - 1)(A'^2 + B'^2)} \\
\left| \frac{D}{d(F, L)} \right| = \left| \frac{2\epsilon}{1 - \epsilon^2} \right| \Rightarrow \left(\frac{D}{d(F, L)} \right)^2 = \left(\frac{2\epsilon}{1 - \epsilon^2} \right)^2
\end{cases}$$

$$\begin{aligned}
&(\epsilon^2 - 1)(A'^2 + B'^2)\alpha - 2\epsilon^2(A'^2\alpha + A'B'\varphi + A'C') \\
&= (-(\epsilon^2 + 1)A'^2 + (\epsilon^2 - 1)B'^2)\alpha - 2\epsilon^2(A'B'\varphi + A'C') \\
&= (-(\epsilon^2 + 1)A'^2 + (\epsilon^2 - 1)B'^2)\alpha - 2\epsilon^2(A'B'\varphi + A'C')
\end{aligned}$$

Can the above be more simplified?

$$\begin{aligned}
 \overline{FF'}^2 &= (\alpha - \chi)^2 + (\varphi - \psi)^2 \\
 &= (\alpha - (\alpha - \epsilon AD))^2 + (\varphi - (\varphi - \epsilon BD))^2 \\
 &= (\epsilon D)^2 (A^2 + B^2) \\
 &= (\epsilon D)^2
 \end{aligned}$$

21.2.3 eccentricity and its equivalent representation

$$\left(\frac{c}{a}\right)^2 = \left(\frac{\overline{PF}}{d(P, L)}\right)^2 = \epsilon^2 = \left(\frac{\overline{FF'}}{D}\right)^2 = \left(\frac{2c}{D}\right)^2 \Rightarrow D = 2a$$

$$\left(\frac{D}{d(F, L)}\right)^2 = \left(\frac{2\epsilon}{1 - \epsilon^2}\right)^2$$

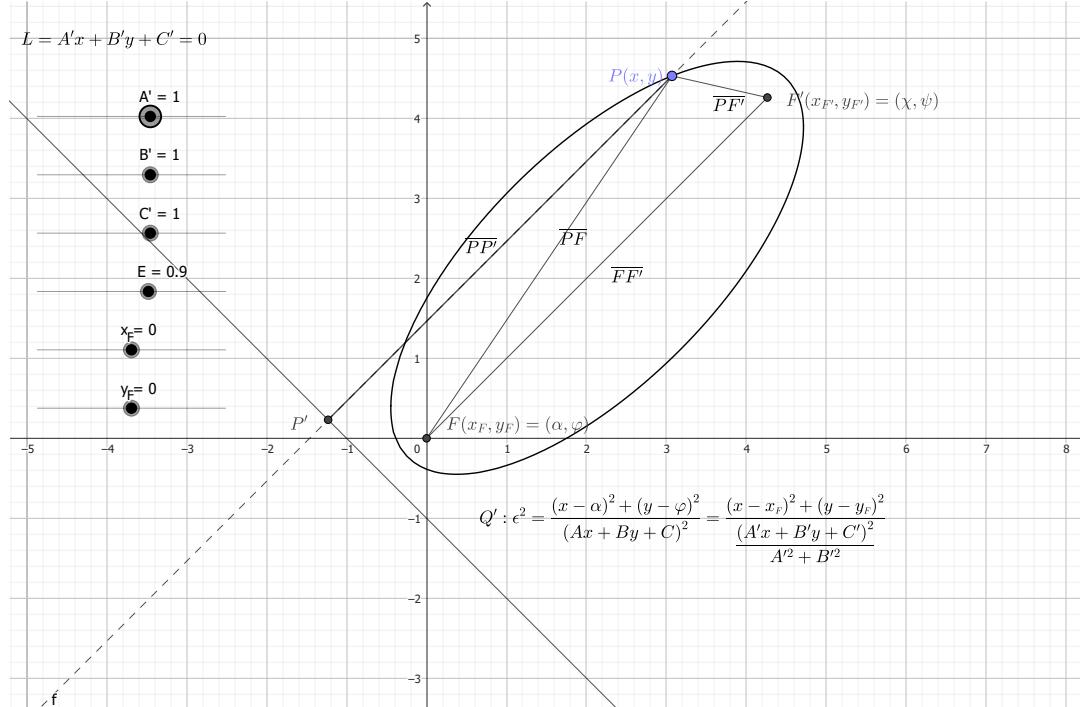


Figure 21.3: conic sections: ellipse

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 $\frac{y - k}{b} = 0 \Rightarrow \frac{x - h}{a} = \pm 1$

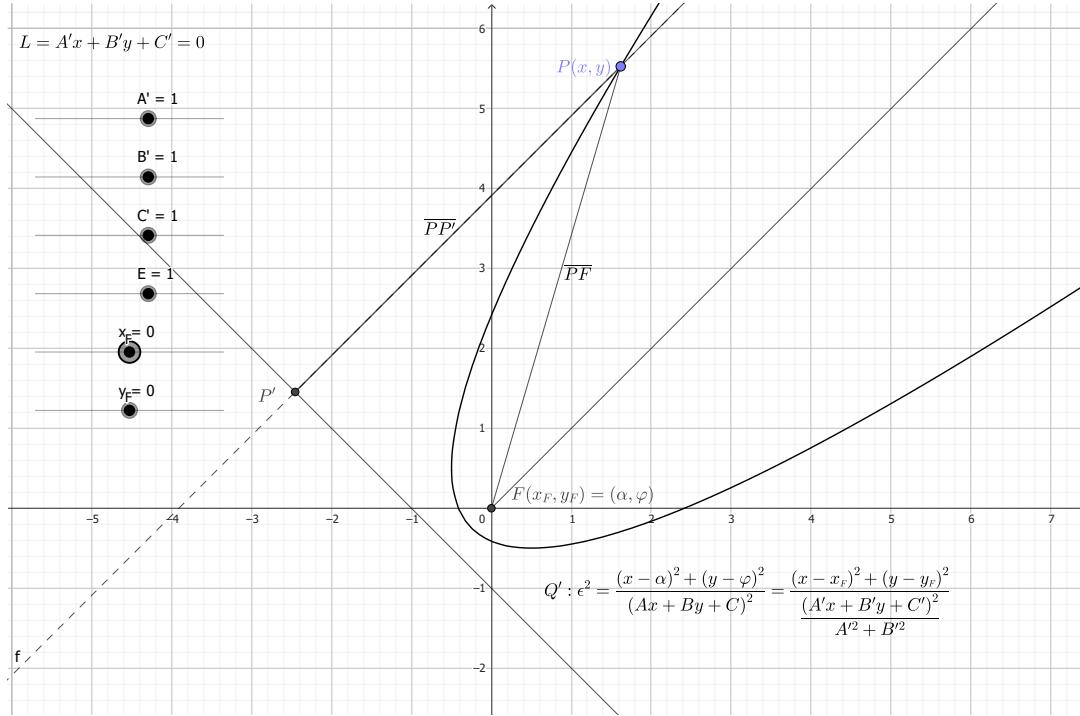


Figure 21.4: conic sections: parabola

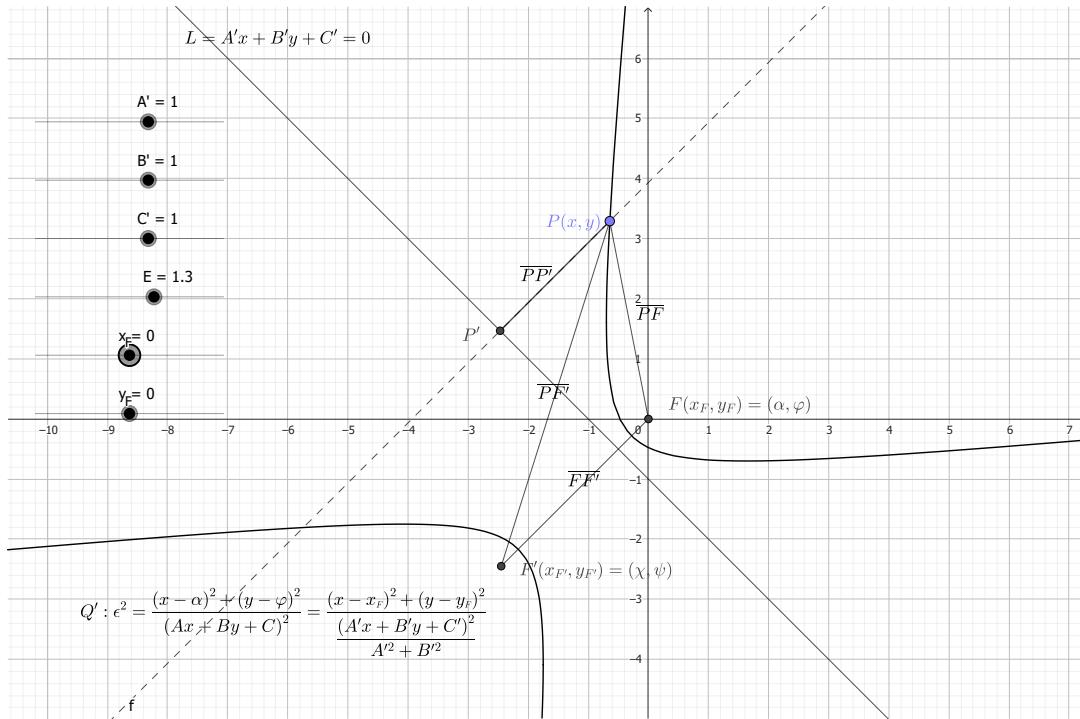


Figure 21.5: conic sections: hyperbola

21.4 parametric equation

circle	$\left(\frac{y-k}{a}\right)^2 + \left(\frac{x-h}{a}\right)^2 = 1$	$\begin{pmatrix} x \\ y \\ 1 \end{pmatrix} = \begin{pmatrix} a & 0 & h \\ 0 & a & k \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} \cos t \\ \sin t \\ 1 \end{pmatrix} = \begin{pmatrix} \cos t & 0 & h \\ 0 & \sin t & k \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} a \\ a \\ 1 \end{pmatrix}$
ellipse	$\left(\frac{y-k}{b}\right)^2 + \left(\frac{x-h}{a}\right)^2 = 1$	$\begin{pmatrix} x \\ y \\ 1 \end{pmatrix} = \begin{pmatrix} a & 0 & h \\ 0 & b & k \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} \cos t \\ \sin t \\ 1 \end{pmatrix} = \begin{pmatrix} \cos t & 0 & h \\ 0 & \sin t & k \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} a \\ b \\ 1 \end{pmatrix}$
parabola	$(y-k) - 4c(x-h)^2 = 0$	$\begin{pmatrix} x \\ y \\ 1 \end{pmatrix} = \begin{pmatrix} 1 & 0 & h \\ 0 & 4c & k \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} t \\ t^2 \\ 1 \end{pmatrix} = \begin{pmatrix} t & 0 & h \\ 0 & t^2 & k \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 \\ 4c \\ 1 \end{pmatrix}$
	$-4c(y-k)^2 + (x-h) = 0$	$\begin{pmatrix} x \\ y \\ 1 \end{pmatrix} = \begin{pmatrix} 4c & 0 & h \\ 0 & 1 & k \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} t \\ t^2 \\ 1 \end{pmatrix} = \begin{pmatrix} t^2 & 0 & h \\ 0 & t & k \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 \\ 4c \\ 1 \end{pmatrix}$
hyperbola	$\left(\frac{y-k}{b}\right)^2 - \left(\frac{x-h}{a}\right)^2 = 1$	$\begin{pmatrix} x \\ y \\ 1 \end{pmatrix} = \begin{pmatrix} a & 0 & h \\ 0 & b & k \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} \pm \cosh t \\ \sinh t \\ 1 \end{pmatrix} = \begin{pmatrix} \tan t & 0 & h \\ 0 & \sec t & k \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} a \\ b \\ 1 \end{pmatrix}$
	$-\left(\frac{y-k}{b}\right)^2 + \left(\frac{x-h}{a}\right)^2 = 1$	$\begin{pmatrix} x \\ y \\ 1 \end{pmatrix} = \begin{pmatrix} a & 0 & h \\ 0 & b & k \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} \pm \cosh t \\ \sinh t \\ 1 \end{pmatrix} = \begin{pmatrix} \sec t & 0 & h \\ 0 & \tan t & k \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} a \\ b \\ 1 \end{pmatrix}$

tangent half-angle formula^[24]

21.5 polar coordinate

$$(x - \alpha)^2 + (y - \varphi)^2 = [\epsilon(Ax + By + C)]^2$$

$$\begin{cases} x = r \cos \theta \\ y = r \sin \theta \end{cases}$$

$$(r \cos \theta - \alpha)^2 + (r \sin \theta - \varphi)^2 = [\epsilon(Ar \cos \theta + Br \sin \theta + C)]^2$$

$$\text{If } \begin{cases} F = (x_F, y_F) = (\alpha, \varphi) = (0, 0) \\ L = Ax + By + C = x + p = 0 \end{cases}$$

$$\begin{aligned} (r \cos \theta)^2 + (r \sin \theta)^2 &= [\epsilon(r \cos \theta + p)]^2 \\ r^2 &= \\ r &= \pm \epsilon(r \cos \theta + p) \\ &= \pm(r\epsilon \cos \theta + \epsilon p) \\ r(1 \mp \epsilon \cos \theta) &= \epsilon p \\ r &= \frac{\epsilon p}{1 \mp \epsilon \cos \theta} \end{aligned}$$

<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/%E9%99%90%E9%9D%A2%E9%8C%88/>

https://en.wikipedia.org/wiki/Matrix_representation_of_conic_sections

$$ax^2 + bxy + cy^2 + dx + ey + f = 0$$

$$(x \ y) \begin{pmatrix} a & b/2 \\ b/2 & c \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = (x \ y) \begin{pmatrix} ax + (b/2)y \\ (b/2)x + cy \end{pmatrix} = ax^2 + bxy + cy^2$$

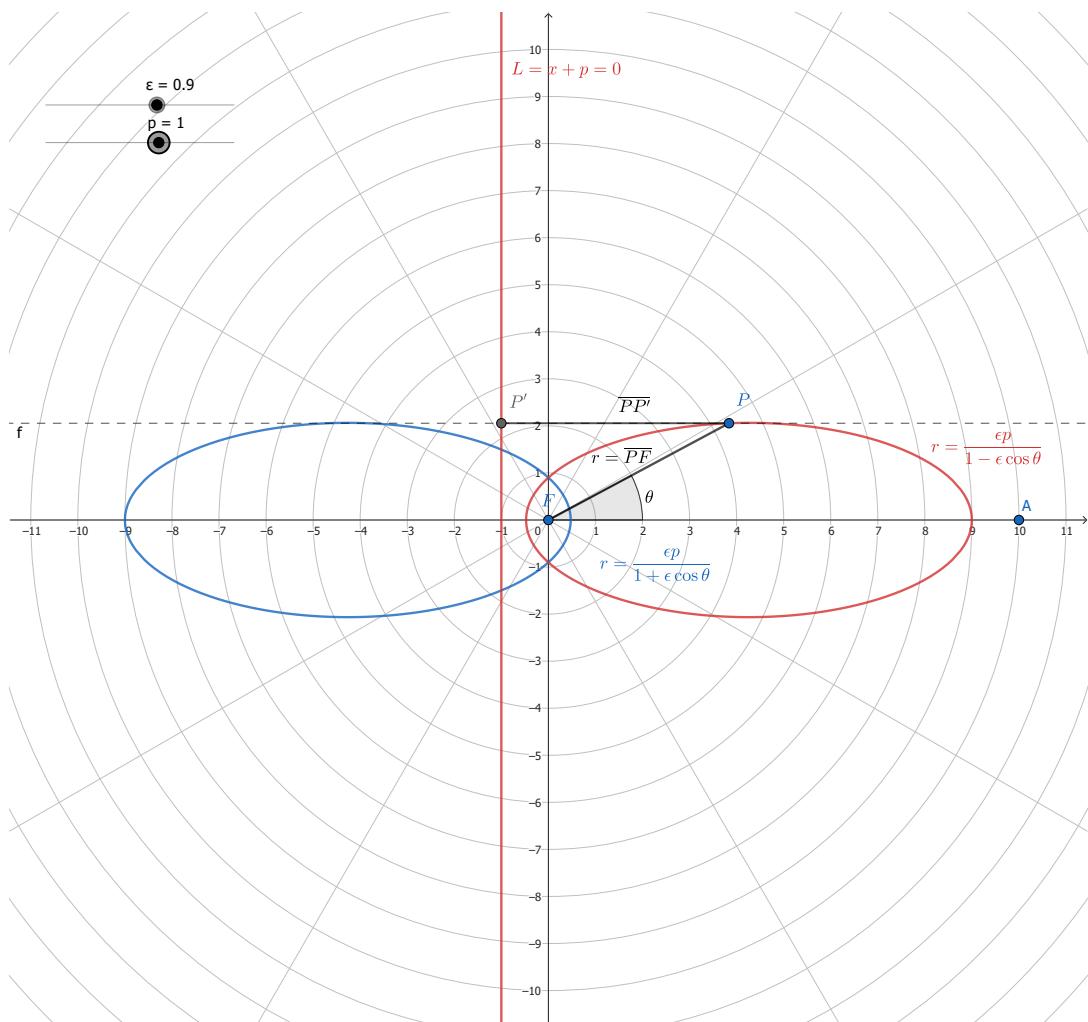


Figure 21.6: polar conic sections: ellipse

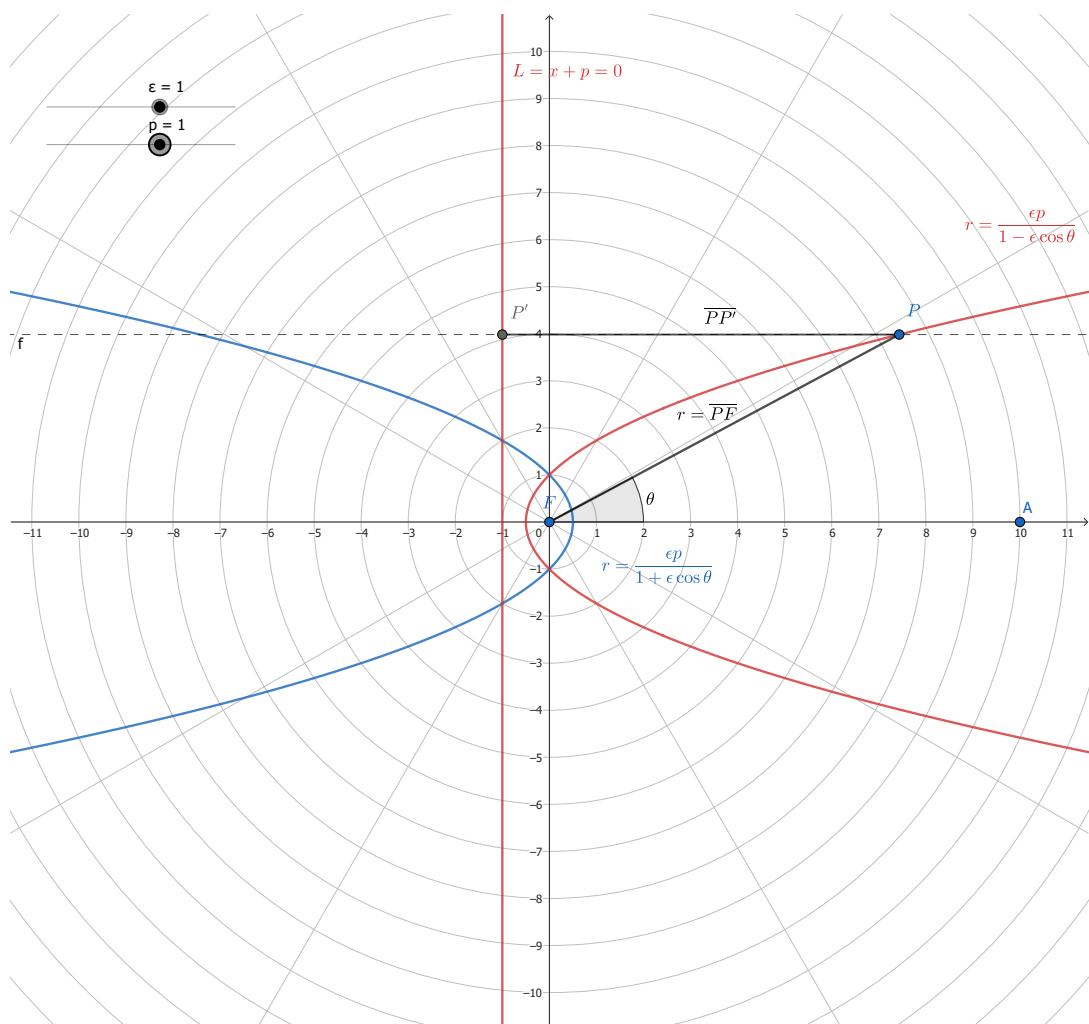


Figure 21.7: polar conic sections: parabola

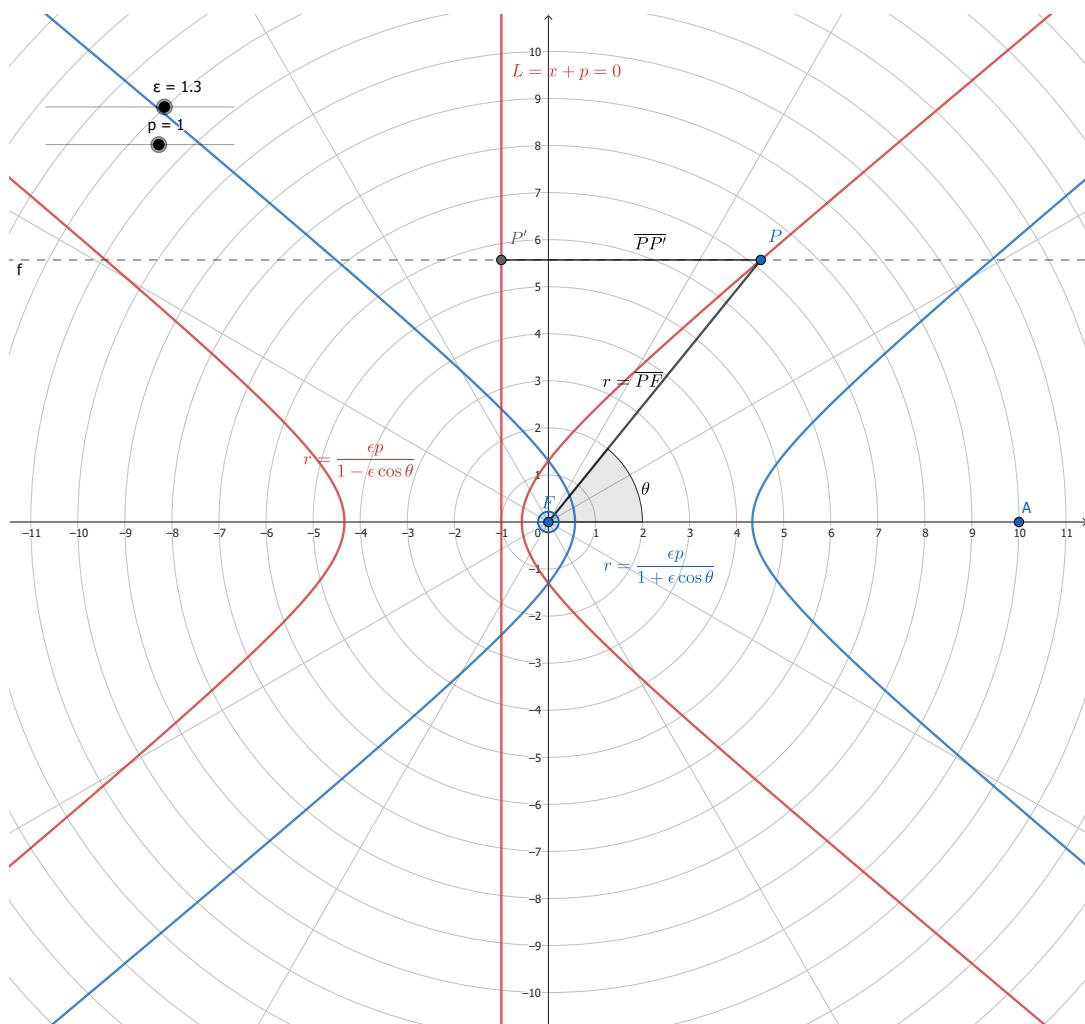


Figure 21.8: polar conic sections: hyperbola

$$0 = (x \ y) \begin{pmatrix} a & b/2 \\ b/2 & c \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} + (d \ e) \begin{pmatrix} x \\ y \end{pmatrix} + f$$

$$= \mathbf{x}^\top A \mathbf{x} + \mathbf{b}^\top \mathbf{x} + f, \quad \begin{cases} A = \begin{pmatrix} a & b/2 \\ b/2 & c \end{pmatrix} & A \text{ real symmetric} \\ \mathbf{b} = \begin{pmatrix} d \\ e \end{pmatrix} \\ \mathbf{x} = \begin{pmatrix} x \\ y \end{pmatrix} \end{cases}$$

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/%E9%9E%A4%E9%9D%A2%E9%9D%A2/>

$$(x \ y) \begin{pmatrix} a & b/2 \\ b/2 & c \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = (x \ y \ 1) \begin{pmatrix} a & b/2 & 0 \\ b/2 & c & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} x \\ y \\ 1 \end{pmatrix} = (x \ y \ 1) \begin{pmatrix} a & b/2 & 0 \\ b/2 & c & 0 \\ 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} x \\ y \\ 1 \end{pmatrix}$$

$$(d \ e) \begin{pmatrix} x \\ y \end{pmatrix} = (x \ y \ 1) \begin{pmatrix} \alpha & \beta & \gamma \\ \delta & \epsilon & \zeta \\ \eta & \theta & \kappa \end{pmatrix} \begin{pmatrix} x \\ y \\ 1 \end{pmatrix} = (x \ y \ 1) \begin{pmatrix} \alpha x + \beta y + \gamma \\ \delta x + \epsilon y + \zeta \\ \eta x + \theta y + \kappa \end{pmatrix}, \quad \begin{cases} \gamma + \eta = d \\ \zeta + \theta = e \end{cases}$$

$$= (x \ y \ 1) \begin{pmatrix} 0 & 0 & \gamma \\ 0 & 0 & \zeta \\ \eta & \theta & 0 \end{pmatrix} \begin{pmatrix} x \\ y \\ 1 \end{pmatrix} = (x \ y \ 1) \begin{pmatrix} 0 & 0 & d/2 \\ 0 & 0 & e/2 \\ d/2 & e/2 & 0 \end{pmatrix} \begin{pmatrix} x \\ y \\ 1 \end{pmatrix}$$

$$0 = ax^2 + bxy + cy^2 + dx + ey + f$$

$$= (x \ y) \begin{pmatrix} a & b/2 \\ b/2 & c \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} + (d \ e) \begin{pmatrix} x \\ y \end{pmatrix} + f = \mathbf{x}^\top A \mathbf{x} + \mathbf{b}^\top \mathbf{x} + f$$

$$= (x \ y \ 1) \begin{pmatrix} a & b/2 & d/2 \\ b/2 & c & e/2 \\ d/2 & e/2 & f \end{pmatrix} \begin{pmatrix} x \\ y \\ 1 \end{pmatrix} = (\mathbf{x}^\top \ 1) M \begin{pmatrix} x \\ y \\ 1 \end{pmatrix}, M = \begin{pmatrix} a & b/2 & d/2 \\ b/2 & c & e/2 \\ d/2 & e/2 & f \end{pmatrix}$$

$$0 = ax^2 + bxy + cy^2 + dx + ey + f$$

$$= (x \ y) \begin{pmatrix} a & b/2 \\ b/2 & c \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} + (d \ e) \begin{pmatrix} x \\ y \end{pmatrix} + f = \mathbf{x}^\top A \mathbf{x} + \mathbf{b}^\top \mathbf{x} + f$$

$$= (x \ y \ 1) \begin{pmatrix} a & b/2 & d/2 \\ b/2 & c & e/2 \\ d/2 & e/2 & f \end{pmatrix} \begin{pmatrix} x \\ y \\ 1 \end{pmatrix} = (\mathbf{x}^\top \ 1) M \begin{pmatrix} x \\ y \\ 1 \end{pmatrix}, M = \begin{pmatrix} a & b/2 & d/2 \\ b/2 & c & e/2 \\ d/2 & e/2 & f \end{pmatrix}$$

https://en.wikipedia.org/wiki/Matrix_representation_of_conic_sections

$$0 = Q = Ax^2 + Bxy + Cy^2 + Dx + Ey + F$$

$$= [x \ y \ 1] \begin{bmatrix} A & B/2 & D/2 \\ B/2 & C & E/2 \\ D/2 & E/2 & F \end{bmatrix} \begin{bmatrix} x \\ y \\ 1 \end{bmatrix} = \mathbf{x}_h^\top A_Q \mathbf{x}_h$$

$$= [x \ y] \begin{bmatrix} A & B/2 \\ B/2 & C \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} + [D \ E] \begin{bmatrix} x \\ y \end{bmatrix} + F = \mathbf{x}^\top A_{Q,33} \mathbf{x} + \mathbf{b}^\top \mathbf{x} + F$$

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 $\overline{PP'}$

$$\begin{aligned} P' &= P'(x, y) \in L = Ax + By + C = 0 \\ \Rightarrow y &= \frac{-1}{B}(Ax + C) \end{aligned}$$

$$\begin{aligned} \overline{PP'}^2(x, y) &= (x_0 - x)^2 + (y_0 - y)^2 \\ &= (x_0 - x)^2 + \left(y_0 - \frac{-1}{B}(Ax + C) \right)^2 \\ &= (x - x_0)^2 + \left(\frac{A}{B}x + \frac{C}{B} + y_0 \right)^2 = \overline{PP'}^2(x) \end{aligned}$$

$$\begin{aligned} 0 &= \frac{\partial}{\partial x} \overline{PP'}^2(x) = 2(x - x_0) + 2 \left(\frac{A}{B}x + \frac{C}{B} + y_0 \right) \frac{A}{B} \\ &= \frac{2}{B^2} (B^2(x - x_0) + A^2x + AC + ABy_0) \\ &= \frac{2}{B^2} [(A^2 + B^2)x - (B^2x_0 - ABy_0 - AC)] \\ x &= \frac{B^2x_0 - ABy_0 - AC}{A^2 + B^2} \end{aligned}$$

or by completing the square to find x .

$$\begin{aligned}
& \overline{PP'}^2 \left(x = \frac{B^2 x_0 - AB y_0 - AC}{A^2 + B^2} \right) \\
&= \left(\frac{B^2 x_0 - AB y_0 - AC}{A^2 + B^2} - x_0 \right)^2 + \left(\frac{A B^2 x_0 - AB y_0 - AC}{A^2 + B^2} + \frac{C}{B} + y_0 \right)^2 \\
&= \left(\frac{-A^2 x_0 - AB y_0 - AC}{A^2 + B^2} \right)^2 + \left(\frac{A (B^2 x_0 - AB y_0 - AC) + C (A^2 + B^2) + B (A^2 + B^2) y_0}{B (A^2 + B^2)} \right)^2 \\
&= \left(\frac{-A (Ax_0 + By_0 + C)}{A^2 + B^2} \right)^2 + \left(\frac{AB^2 x_0 + B^3 y_0 + B^2 C}{B (A^2 + B^2)} \right)^2 \\
&= \frac{A^2 (Ax_0 + By_0 + C)^2}{(A^2 + B^2)^2} + \frac{B^2 (Ax_0 + By_0 + C)^2}{(A^2 + B^2)^2} \\
&= \frac{(Ax_0 + By_0 + C)^2}{A^2 + B^2} \\
\overline{PP'} &= \overline{PP'} \left(x = \frac{B^2 x_0 - AB y_0 - AC}{A^2 + B^2} \right) = \frac{|Ax_0 + By_0 + C|}{\sqrt{A^2 + B^2}}
\end{aligned}$$

22.2 by perpendicular foot

$$y = \frac{-A}{B}x - \frac{C}{B} = \frac{-1}{B}(Ax + C), \text{ if } B \neq 0$$

$$L_{\perp} : \left(y = \frac{B}{A}x + K \right) \perp \left(y = \frac{-A}{B}x - \frac{C}{B} \right) : L$$

$$L_{\perp} = L_{\perp}(x, y) = Bx - Ay + K = 0$$

$$P = P(x_0, y_0) \in L_{\perp} = B(x - x_0) - A(y - y_0) = 0$$

$$L_{\perp} = Bx - Ay - (Bx_0 - Ay_0) = 0$$

perpendicular foot = foot of the perpendicular P'

$$\begin{aligned}
P' \in (L_{\perp} \cap L) &= \begin{cases} L = Ax + By + C = 0 \\ L_{\perp} = Bx - Ay - (Bx_0 - Ay_0) = 0 \end{cases} \\
&= \begin{cases} Ax + By = -C \\ Bx - Ay = Bx_0 - Ay_0 \end{cases} \\
P' = P'(x, y) &= \left(\frac{\begin{vmatrix} -C & B \\ Bx_0 - Ay_0 & -A \end{vmatrix}}{\begin{vmatrix} A & B \\ B & -A \end{vmatrix}}, \frac{\begin{vmatrix} A & -C \\ B & Bx_0 - Ay_0 \end{vmatrix}}{\begin{vmatrix} A & B \\ B & -A \end{vmatrix}} \right) \\
&= \left(\frac{\begin{vmatrix} C & B \\ -Bx_0 + Ay_0 & -A \end{vmatrix}}{\begin{vmatrix} A & -B \\ B & A \end{vmatrix}}, \frac{\begin{vmatrix} A & C \\ B & -Bx_0 + Ay_0 \end{vmatrix}}{\begin{vmatrix} A & -B \\ B & A \end{vmatrix}} \right) \\
&= \left(\frac{B^2 x_0 - AB y_0 - AC}{A^2 + B^2}, \frac{-AB x_0 + A^2 y_0 - BC}{A^2 + B^2} \right)
\end{aligned}$$

$$\begin{aligned}
d(P, L) &= \overline{PP'} \\
&= \left\| (x_0, y_0) - \left(\frac{B^2 x_0 - AB y_0 - AC}{A^2 + B^2}, \frac{-AB x_0 + A^2 y_0 - BC}{A^2 + B^2} \right) \right\| \\
&= \sqrt{\left(x_0 - \frac{B^2 x_0 - AB y_0 - AC}{A^2 + B^2} \right)^2 + \left(y_0 - \frac{-AB x_0 + A^2 y_0 - BC}{A^2 + B^2} \right)^2} \\
&= \sqrt{\left(\frac{A^2 x_0 + AB y_0 + AC}{A^2 + B^2} \right)^2 + \left(\frac{AB x_0 + B^2 y_0 + BC}{A^2 + B^2} \right)^2} \\
&= \sqrt{\frac{A^2 (Ax_0 + By_0 + C)^2 + B^2 (Ax_0 + By_0 + C)^2}{(A^2 + B^2)^2}} = \sqrt{\frac{(Ax_0 + By_0 + C)^2}{A^2 + B^2}} \\
&= \frac{|Ax_0 + By_0 + C|}{\sqrt{A^2 + B^2}}
\end{aligned}$$

22.3 by normal vector

$$\begin{cases} \vec{n} = (A, B) \perp L = Ax + By + C = 0 \\ \vec{PP'} = P' - P = (x - x_0, y - y_0) \end{cases}$$

P 到 L 的距離 $d(P, L)$ 即為 L 線上一點 P' 對應之 $\vec{PP'}$ 在 L 法向量 \vec{n} 方向上的投影長

$$\begin{aligned}
\vec{PP'} \cdot \vec{n} &= \left\| \vec{PP'} \right\| \left\| \vec{n} \right\| \cos \theta \\
\left| \vec{PP'} \cdot \vec{n} \right| &= \left\| \vec{PP'} \right\| \left\| \vec{n} \right\| |\cos \theta| \\
\left\| \vec{PP'} \right\| |\cos \theta| &= \left| \vec{PP'} \cdot \hat{n} \right| = \frac{\left| \vec{PP'} \cdot \vec{n} \right|}{\left\| \vec{n} \right\|} = \frac{|(x - x_0, y - y_0) \cdot (A, B)|}{\|(A, B)\|} \\
&= \frac{|A(x - x_0) + B(y - y_0)|}{\sqrt{A^2 + B^2}} = \frac{|-Ax_0 - By_0 + Ax + By|}{\sqrt{A^2 + B^2}} \\
&\stackrel{Ax+By+C=0}{=} \frac{|-Ax_0 - By_0 - C|}{\sqrt{A^2 + B^2}} = \frac{|Ax_0 + By_0 + C|}{\sqrt{A^2 + B^2}}
\end{aligned}$$

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

```

\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

$$\begin{aligned}
Ax + By + C &= 0 \\
Ax + By &= -C \\
(Ax + By) - (Ax_0 + By_0) &= -C - (Ax_0 + By_0) \\
A(x - x_0) + B(y - y_0) &= -(Ax_0 + By_0 + C)
\end{aligned}$$

$$\begin{aligned}\overline{PP'}^2 &= (x_0 - x)^2 + (y_0 - y)^2 \\ [A^2 + B^2] \overline{PP'}^2 &= [A^2 + B^2] \left[(x_0 - x)^2 + (y_0 - y)^2 \right] \\ &\geq [A(x - x_0) + B(y - y_0)]^2 \\ &= [-(Ax_0 + By_0 + C)]^2 = (Ax_0 + By_0 + C)^2 \\ \overline{PP'}^2 &\geq \frac{(Ax_0 + By_0 + C)^2}{A^2 + B^2} \\ \overline{PP'} &\geq \frac{|Ax_0 + By_0 + C|}{\sqrt{A^2 + B^2}}\end{aligned}$$

Chapter 23

real symmetric matrix diagonalizable

<https://ccjou.wordpress.com/2011/02/09/%E6%97%A5%E5%AF%BC%E7%9F%A5%E6%9D%A1%E5%8A%A8%E5%88%86%E6%9E%90/>

<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。

$$\begin{array}{ll}
 \left\{ \begin{array}{l} A \in \mathcal{M}_{n \times n}(\mathbb{R}) \\ A^\top = A \end{array} \right. & \begin{array}{l} \text{real matrix} \\ \text{symmetric matrix} \end{array} \\
 \downarrow & \text{real symmetric matrix} \\
 \left\{ \begin{array}{l} Ax = \lambda x \\ \lambda \in \mathbb{C} \\ \mathbf{0} \neq x \in \mathbb{C}^n \end{array} \right. & \begin{array}{l} \text{complex eigenvalue} \\ \text{complex eigenvector} \end{array} \\
 \downarrow & \\
 \left\{ \begin{array}{l} \lambda \in \mathbb{R} \\ x \in \mathbb{R}^n \end{array} \right. & \begin{array}{l} \text{real eigenvalue (1)} \\ \text{real eigenvector (2)} \end{array}
 \end{array}$$

Proof. (1)

$$\begin{aligned}
 Ax &= \lambda x \\
 \bar{A}\bar{x} &= \overline{Ax} = \overline{\lambda x} = \bar{\lambda}\bar{x} \\
 \bar{x}^\top \bar{A}^\top &= (\bar{A}\bar{x})^\top = (\bar{\lambda}\bar{x})^\top = \bar{\lambda}\bar{x}^\top \\
 \bar{x}^\top A &\stackrel{\text{symmetric}}{=} \bar{x}^\top A^\top \stackrel{\text{real}}{=} \\
 \bar{x}^\top A &= \bar{\lambda}\bar{x}^\top \\
 \lambda \bar{x}^\top x &= \bar{x}^\top (\lambda x) \stackrel{Ax=\lambda x}{=} \bar{x}^\top Ax = \bar{\lambda}\bar{x}^\top x \\
 \lambda \bar{x}^\top x &= \bar{\lambda}\bar{x}^\top x \\
 (\lambda - \bar{\lambda}) \bar{x}^\top x &= 0 \wedge \begin{cases} \bar{x}^\top x = \sum_{i=1}^n |x_i|^2 \\ x \neq \mathbf{0} \end{cases} \Rightarrow \bar{x}^\top x \neq 0 \\
 \lambda - \bar{\lambda} &= 0 \\
 \lambda = \bar{\lambda} &\Leftrightarrow \lambda \in \mathbb{R}
 \end{aligned}$$

□

Proof. (1) fast concept

$$\begin{aligned}
(\bar{A}\bar{x})^\top \bar{x} &= (\bar{x}^\top \bar{A}^\top) \bar{x} \stackrel{\text{symmetric}}{=} (\bar{x}^\top \bar{A}) \bar{x} = \bar{x}^\top (\bar{A}x) \\
(L) &= (\bar{A}\bar{x})^\top \bar{x} = \bar{x}^\top (\bar{A}x) = (R) \\
(L) &= (\bar{A}\bar{x})^\top \bar{x} \stackrel{Ax=\lambda x}{=} (\bar{\lambda}\bar{x})^\top \bar{x} = \bar{\lambda}\bar{x}^\top \bar{x} \\
(R) &= \bar{x}^\top (\bar{A}x) \stackrel{\text{real}}{=} \bar{x}^\top (Ax) \stackrel{Ax=\lambda x}{=} \bar{x}^\top (\lambda x) = \lambda \bar{x}^\top \bar{x} \\
&\quad \bar{\lambda}\bar{x}^\top \bar{x} = (\bar{A}\bar{x})^\top \bar{x} = \bar{x}^\top (\bar{A}x) = \lambda \bar{x}^\top \bar{x} \\
&\quad \bar{\lambda}\bar{x}^\top \bar{x} = \lambda \bar{x}^\top \bar{x}
\end{aligned}$$

□

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 $x \in N(A - \lambda I)$ 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。

$$\left\{
\begin{array}{ll}
\begin{cases} A \in \mathcal{M}_{n \times n}(\mathbb{R}) & \text{real matrix} \\ A^\top = A & \text{symmetric matrix} \end{cases} & \text{real symmetric matrix} \\
Ax = \lambda x & \text{23.1 } \begin{cases} \lambda \in \mathbb{R} & \text{real eigenvalue} \\ x \in \mathbb{R}^n & \text{real eigenvector} \end{cases} \\
\begin{cases} Ax_1 = \lambda_1 x_1 & (e_1) \\ Ax_2 = \lambda_2 x_2 & (e_2) \end{cases} & \lambda_1 \neq \lambda_2 \\
& \Downarrow \\
& x_1^\top x_2 = 0 \Leftrightarrow x_1 \perp x_2
\end{array}
\right.$$

Proof. (1)

$$\begin{aligned}
Ax_2 &= \lambda_2 x_2 \\
x_1^\top Ax_2 &\stackrel{x_1^\top}{=} x_1^\top \lambda_2 x_2 = \lambda_2 x_1^\top x_2 = (1) \\
Ax_1 &= \lambda_1 x_1 \\
x_1^\top A^\top &= (Ax_1)^\top = (\lambda_1 x_1)^\top = \lambda_1 x_1^\top \\
x_1^\top A^\top &= \lambda_1 x_1^\top \\
x_1^\top Ax_2 &\stackrel{\text{symmetric}}{=} x_1^\top A^\top x_2 \stackrel{x_2}{=} \lambda_1 x_1^\top x_2 = (2) \\
\lambda_2 x_1^\top x_2 &\stackrel{(1)}{=} x_1^\top Ax_2 \stackrel{(2)}{=} \lambda_1 x_1^\top x_2 \\
\lambda_2 x_1^\top x_2 &= \lambda_1 x_1^\top x_2 \\
(\lambda_2 - \lambda_1) x_1^\top x_2 &= 0 \wedge \lambda_1 \neq \lambda_2 \\
x_1^\top x_2 &= 0
\end{aligned}$$

□

Proof. (1) fast concept

$$\begin{aligned}
(Ax_1)^\top x_2 &= (x_1^\top A^\top) x_2 \stackrel{\text{symmetric}}{=} (x_1^\top A) x_2 = x_1^\top (Ax_2) \\
(L) &= (Ax_1)^\top x_2 = x_1^\top (Ax_2) = (R) \\
(L) &= (Ax_1)^\top x_2 \stackrel{(e_1)}{=} (\lambda_1 x_1)^\top x_2 = \lambda_1 x_1^\top x_2 \\
(R) &= x_1^\top (Ax_2) \stackrel{(e_2)}{=} x_1^\top (\lambda_2 x_2) = \lambda_2 x_1^\top x_2 \\
\lambda_1 x_1^\top x_2 &= (Ax_1)^\top x_2 = x_1^\top (Ax_2) = \lambda_2 x_1^\top x_2 \\
\lambda_1 x_1^\top x_2 &= \lambda_2 x_1^\top x_2
\end{aligned}$$

□

Theorem 23.3.

$$\left\{ \begin{array}{ll} A \in \mathcal{M}_{n \times n}(\mathbb{R}) & \text{real matrix} \\ A^\top = A & \text{symmetric matrix} \\ Ax_1 = \lambda x_1 & (e) \\ x_2^\top x_1 = 0 \Leftrightarrow x_2 \perp x_1 & (o) \end{array} \right. \quad \Downarrow \quad Ax_2 \perp x_1 \Leftrightarrow (Ax_2)^\top x_1 = 0$$

Proof.

$$\begin{aligned} (Ax_2)^\top x_1 &= (x_2^\top A^\top) x_1 \stackrel{\text{symmetric}}{=} (x_2^\top A) x_1 \\ &= x_2^\top (Ax_1) \stackrel{(e)}{=} x_2^\top (\lambda x_1) \\ &= \lambda x_2^\top x_1 \stackrel{(o)}{=} \lambda \cdot 0 = 0 \\ (Ax_2)^\top x_1 = 0 &\Leftrightarrow Ax_2 \perp x_1 \end{aligned}$$

□

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/1z1S2kQKXD8?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>

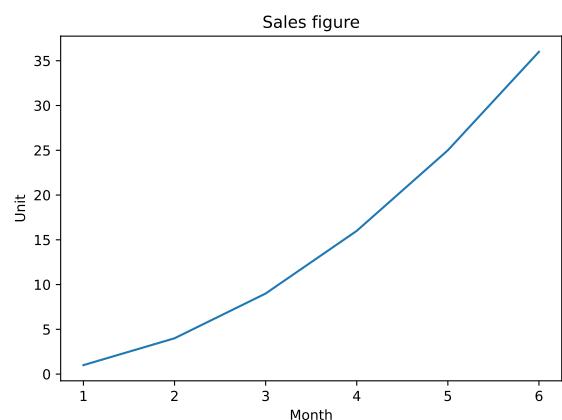
<https://tex.stackexchange.com/questions/84847/can-i-use-webp-images-in-latex>

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()`.

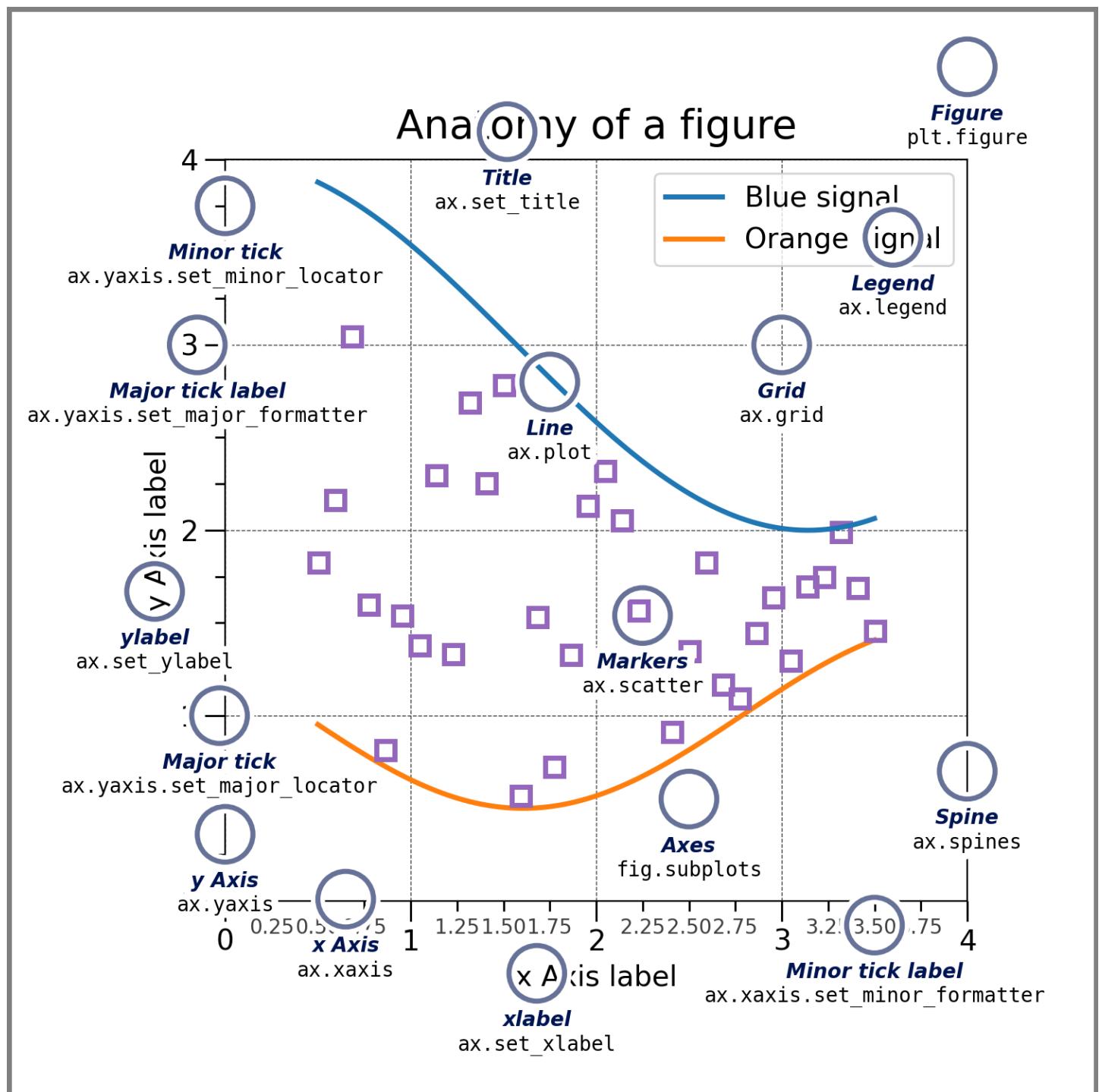


Figure 27.1: matplotlib figure anatomy

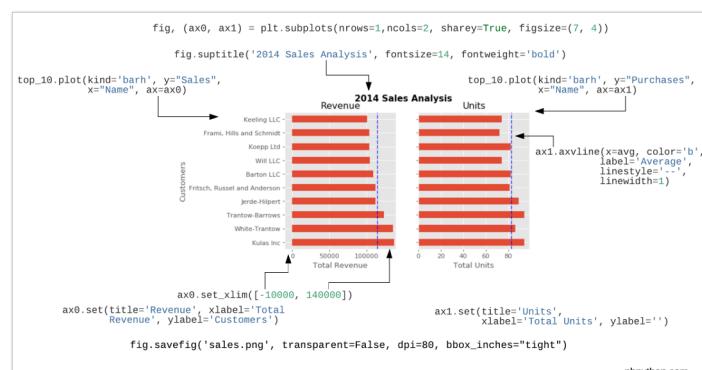


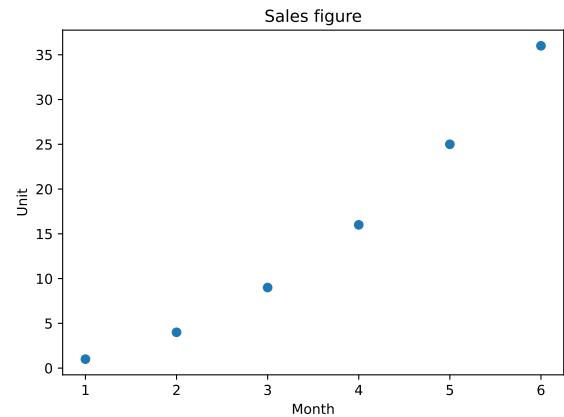
Figure 27.2: matplotlib subplot anatomy

```

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

```

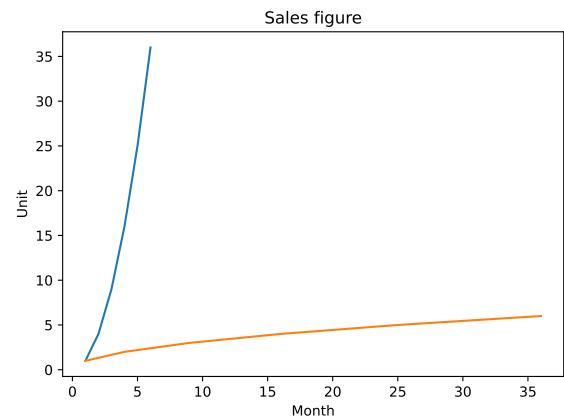


```

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)

```



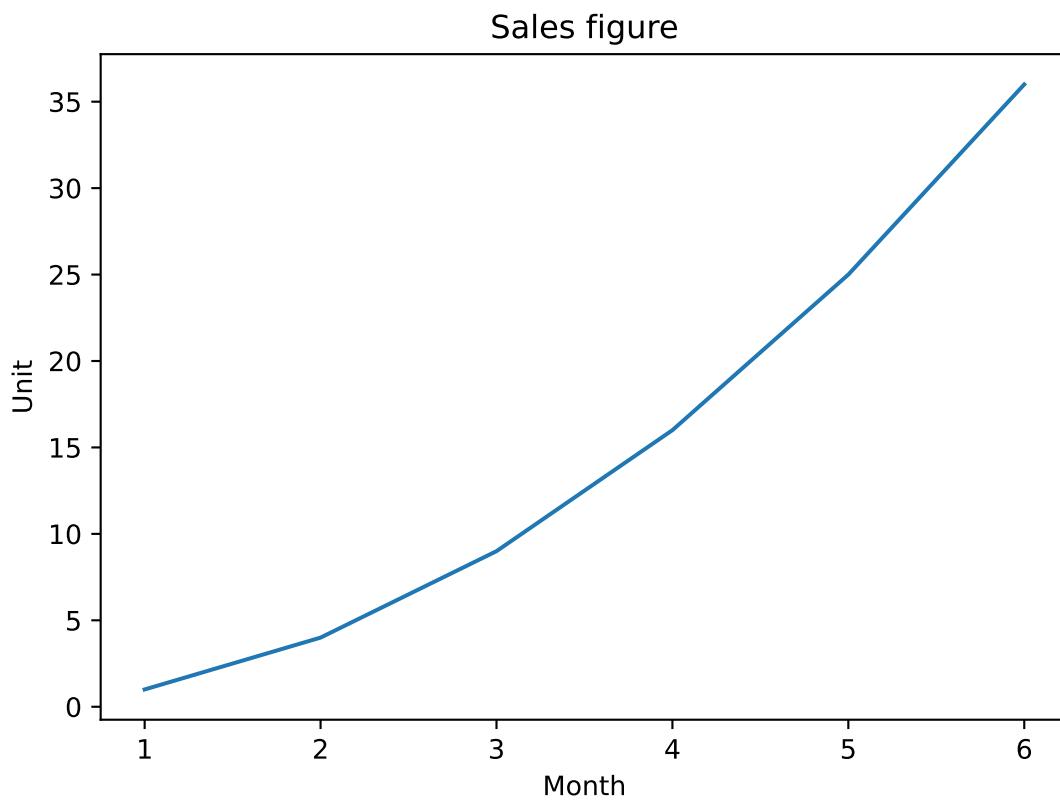
The behavior of the functional API is stateful. What's stateful? An example is when you read a text file. When you `open()` a text file to read, the library reads the next line every time you call `readline()`. It remembers where you left off, despite the fact that you do not give it the position to 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:

```

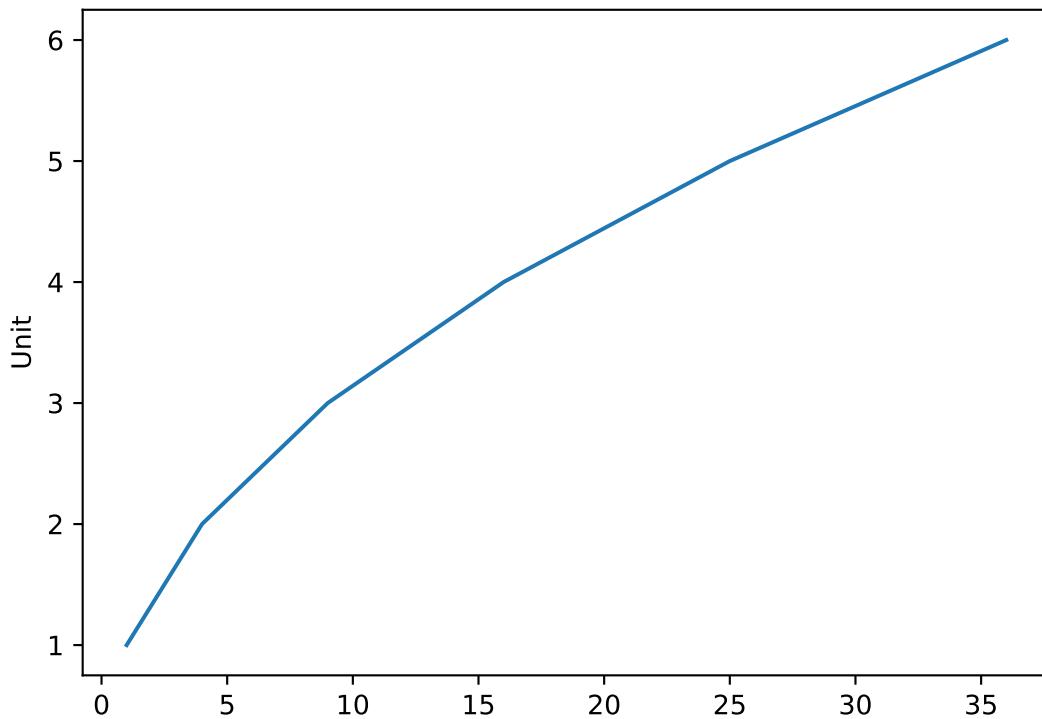
import matplotlib.pyplot as plt

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-oriented 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'>
```

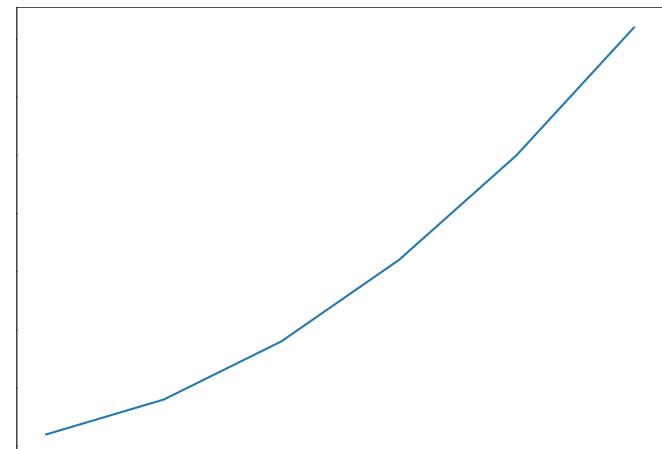
```
ax1 = fig.add_axes([0, 0, 1, 1]) # [left, bottom, width, height]
print("ax1 is of type:", type(ax1))
# ax1.plot(a1, a2)
```

```
## 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

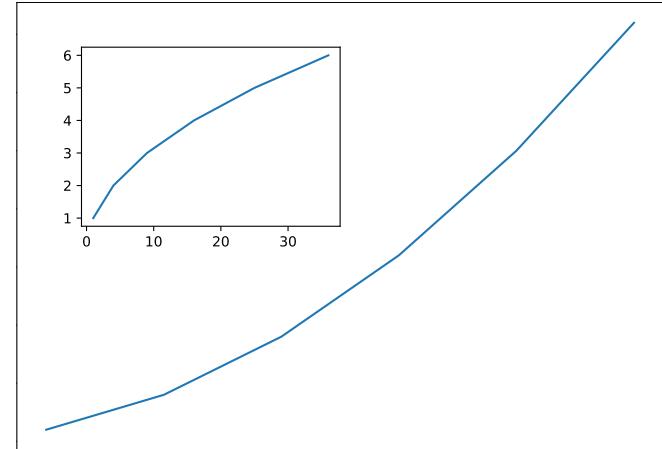
```
import matplotlib.pyplot as plt

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)
```

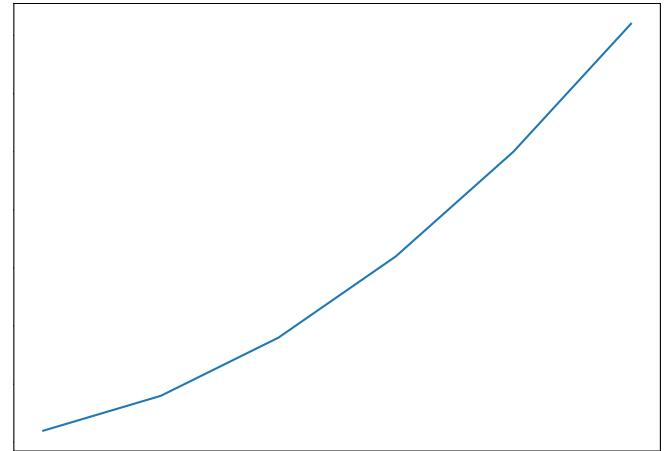


```

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])

```

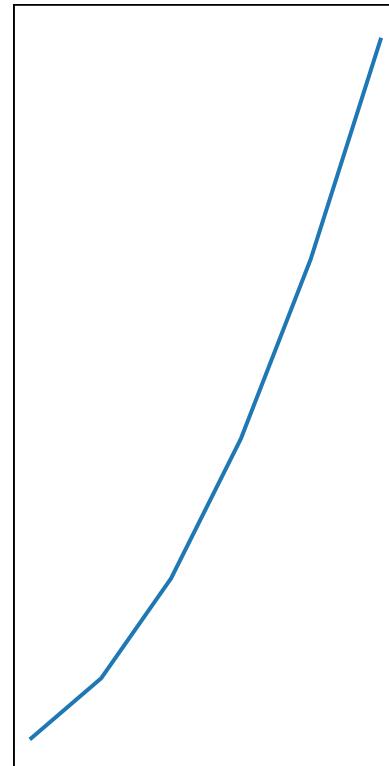
```
fig = plt.figure(figsize=(2, 4))
```

```

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)

```



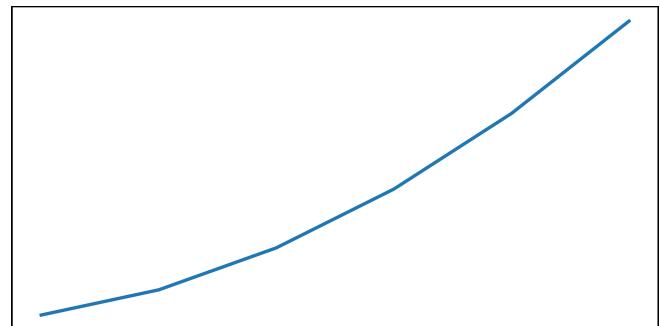
```
fig.set_size_inches((4, 2))
```

```

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: >]]

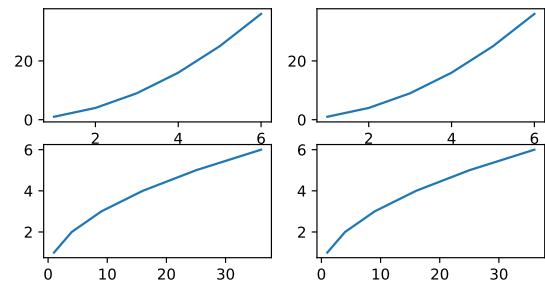
```

```

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,
# 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

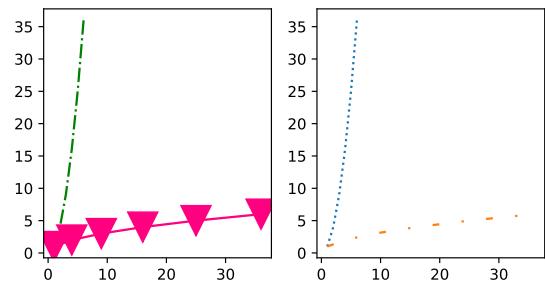
- Color
 - <https://matplotlib.org/stable/tutorials/colors/colors.html>
- line-style (ls)
 - https://matplotlib.org/stable/gallery/lines_bars_and_markers/linestyles.html
- marker
 - https://matplotlib.org/stable/api/markers_api.html
- linewidth (lw)

```

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,
# axes is a numpy array
# print(axes)
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

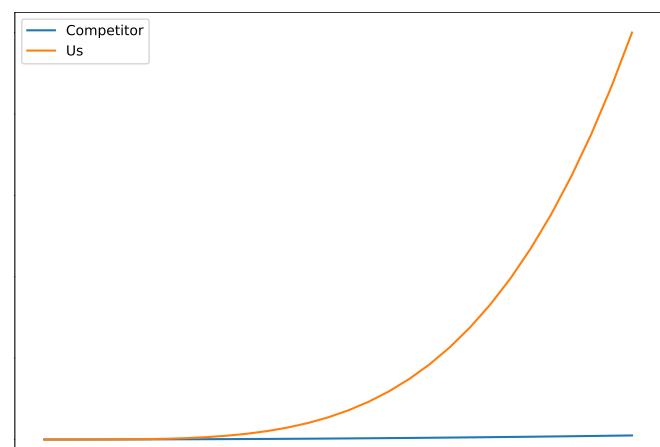
27.1.2.5 legend

```

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

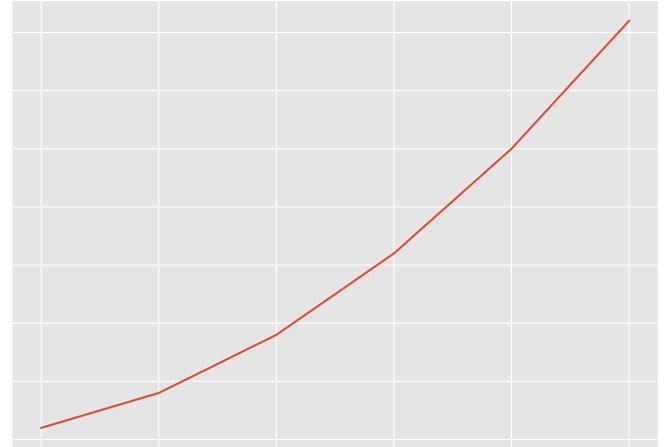
```
import matplotlib.pyplot as plt

print(plt.style.available)
```

```
## ['Solarize_Light2', '_classic_test_patch', '_mpl-gallery', '_mpl-gallery-nogrid', 'bmh', 'classic', 'dark...', 'fivethirtyeight', 'ggplot', 'seaborn', 'seaborn-bright', 'seaborn-colorblind', 'seaborn-dark', 'seaborn-dark...', 'seaborn-deep', 'seaborn-fiveth...', 'seaborn-notebook', 'seaborn-paper', 'seaborn-pastel', 'seaborn-poster', 'seaborn-ticks', 'seaborn-white', 'seaborn-white...', 'seaborn-zipkin']
```

```
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

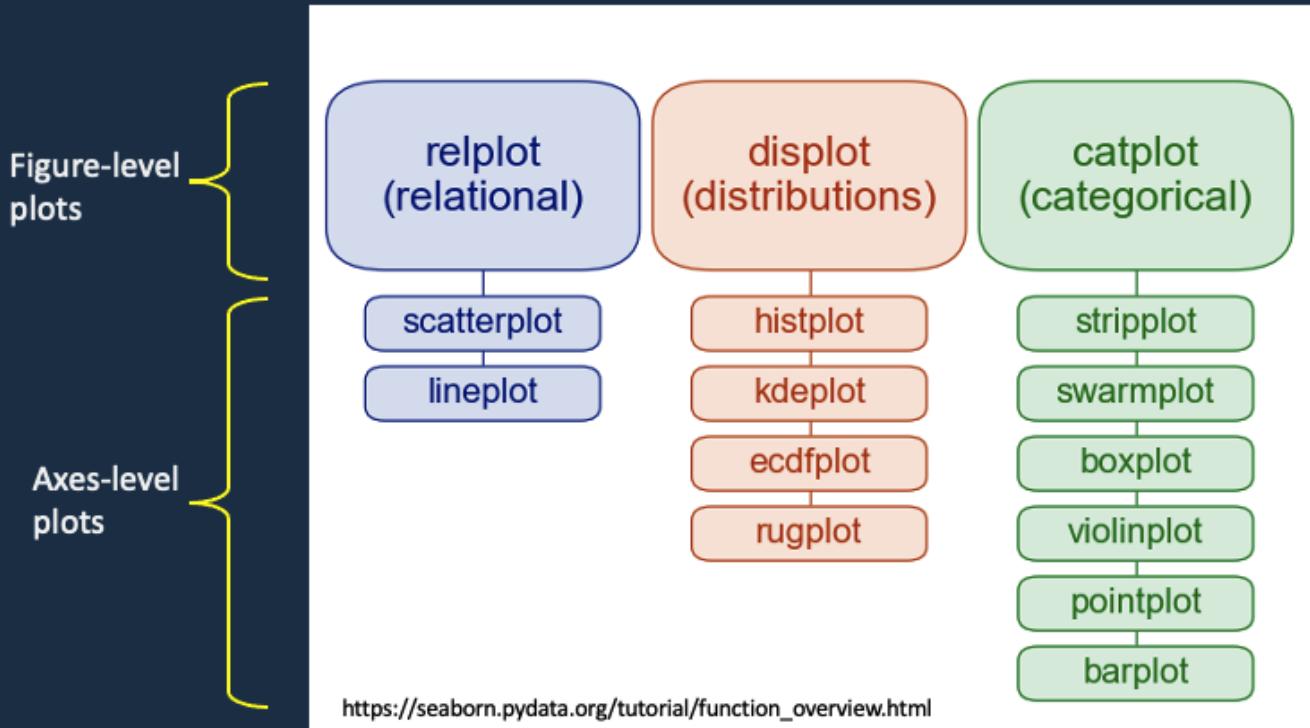
a1 = list(range(1, 7)) # [1, 2, 3, 4, 5, 6]
a2 = [1, 4, 9, 16, 25, 36]
fig = plt.figure()
ax = fig.add_axes([0, 0, 1, 1])
ax.plot(a1, a2)
fig.savefig('out.pdf', bbox_inches = 'tight')
```

27.1.3 Seaborn

- Kimberly Fessel

- visually explained
- Seaborn
- Matplotlib
- Pandas
- iPyWidgets

Three main type of Seaborn plots



Yes, that correspond to Matplotlib's concept of figure and axes

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

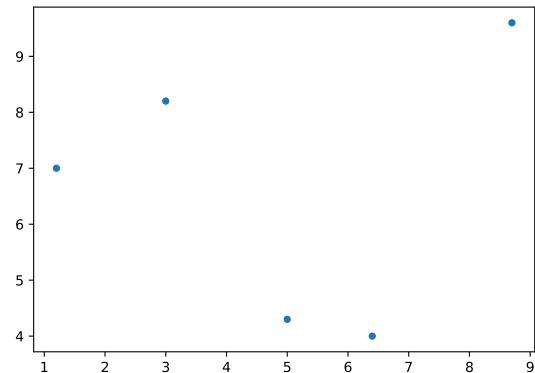
```

```

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)

```

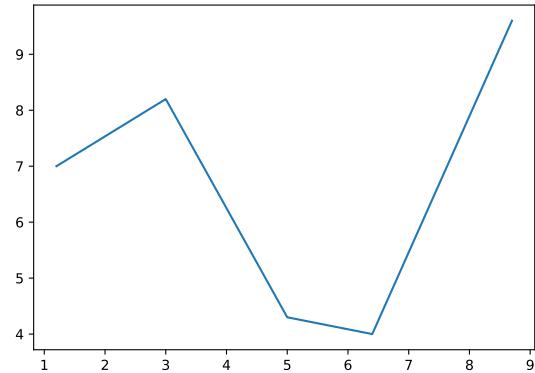


```

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

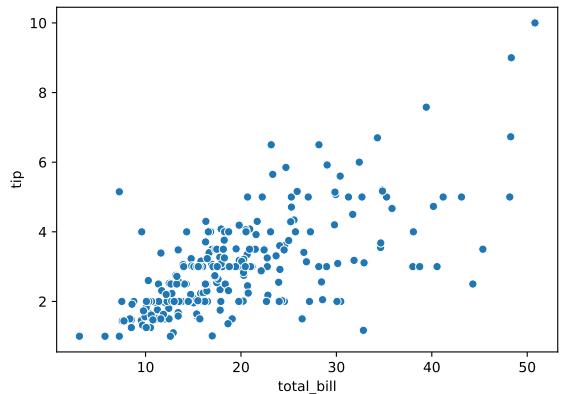
```
sns.load_dataset
```

```
import seaborn as sns
tips = sns.load_dataset("tips")
tips
```

```
##      total_bill    tip   sex smoker  day  time
## 0        16.99  1.01 Female   No Sun Dinner
## 1        10.34  1.66  Male   No Sun Dinner
## 2        21.01  3.50  Male   No Sun Dinner
## 3        23.68  3.31  Male   No Sun Dinner
## 4        24.59  3.61 Female   No Sun Dinner
## ..
## 239     29.03  5.92  Male   No Sat Dinner
## 240     27.18  2.00 Female Yes Sat Dinner
## 241     22.67  2.00  Male Yes Sat Dinner
## 242     17.82  1.75  Male   No Sat Dinner
## 243     18.78  3.00 Female   No Thur Dinner
##
## [244 rows x 7 columns]
```

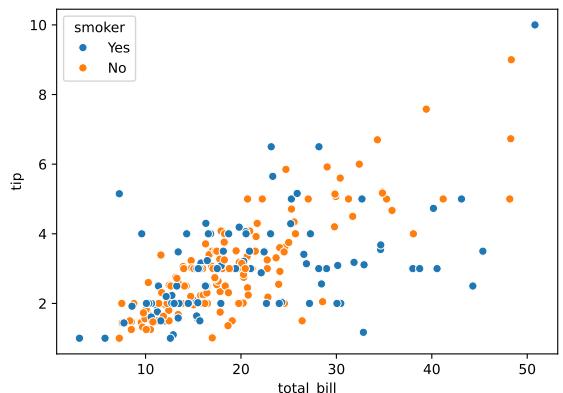
data

```
import seaborn as sns
tips = sns.load_dataset("tips")
sns.scatterplot(data=tips,
                 x='total_bill',
                 y='tip')
```



hue

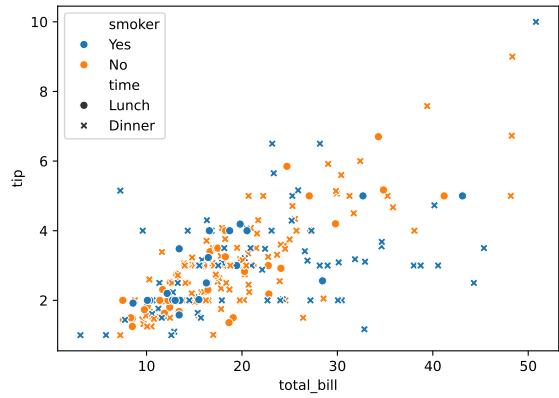
```
import seaborn as sns
tips = sns.load_dataset("tips")
sns.scatterplot(data=tips,
                 x='total_bill',
                 y='tip',
                 hue='smoker')
```



style

```
import seaborn as sns

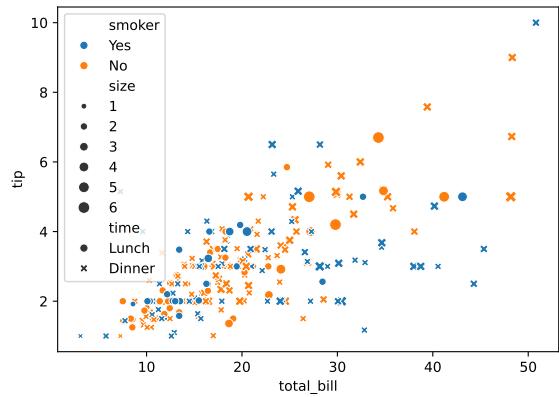
tips = sns.load_dataset("tips")
sns.scatterplot(data=tips,
                 x='total_bill',
                 y='tip',
                 hue='smoker',
                 style='time')
```



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:

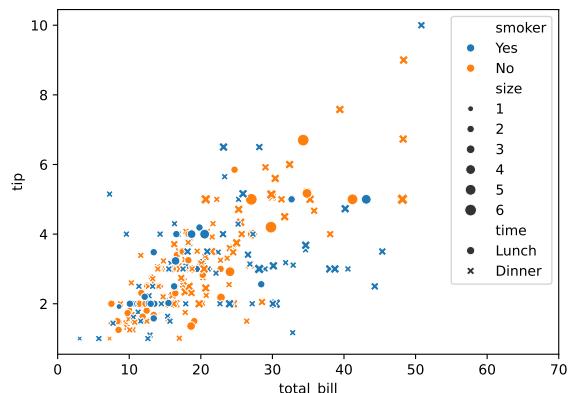
```
ax = sns.scatterplot(...)
```

<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)



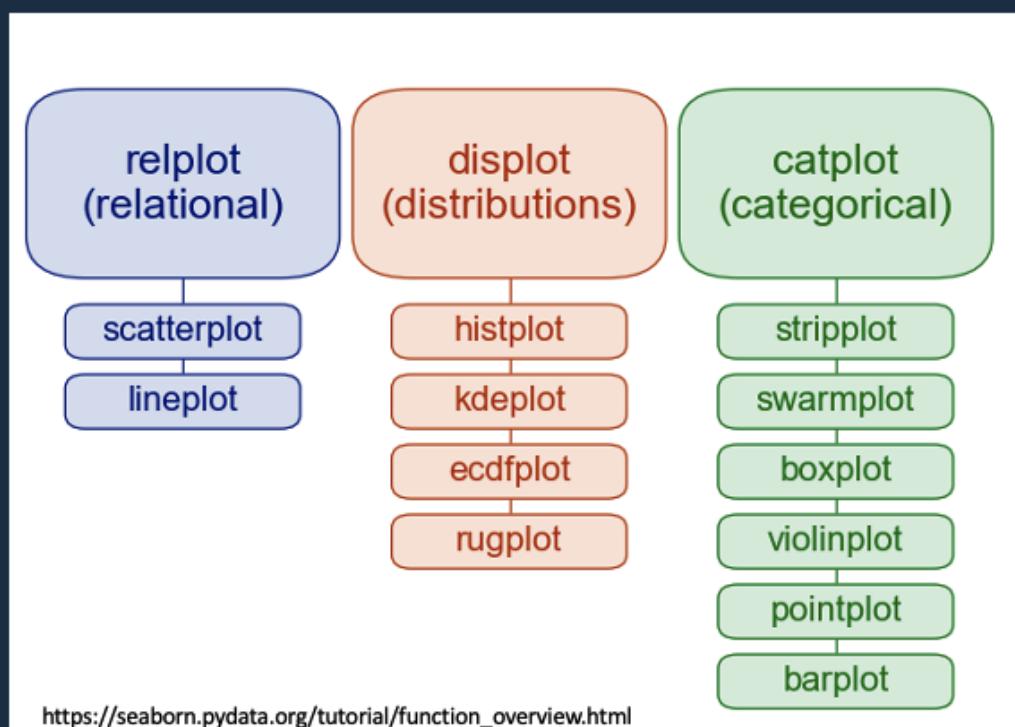
27.1.3.3 axis-level plot and figure-level plot

`sns.replot`

Three main type of Seaborn plots

Figure-level
plots

Axes-level
plots

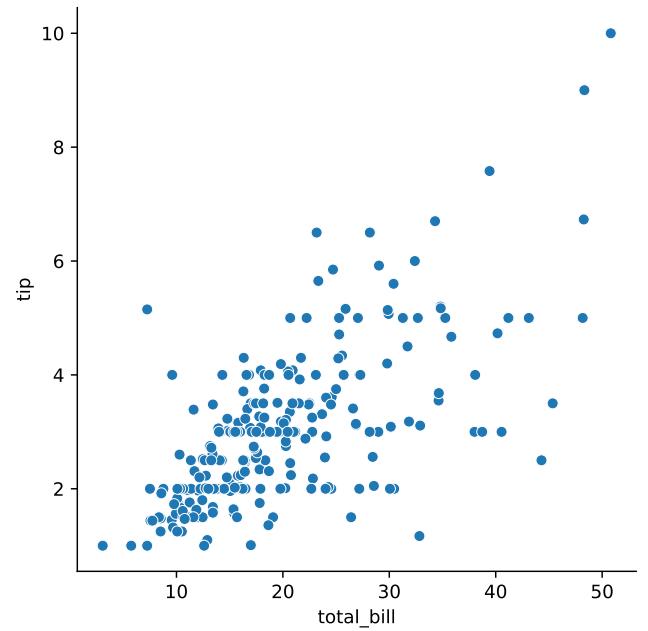


Yes, that correspond to Matplotlib's concept of figure and axes

Figure 27.4: seaborn plot type

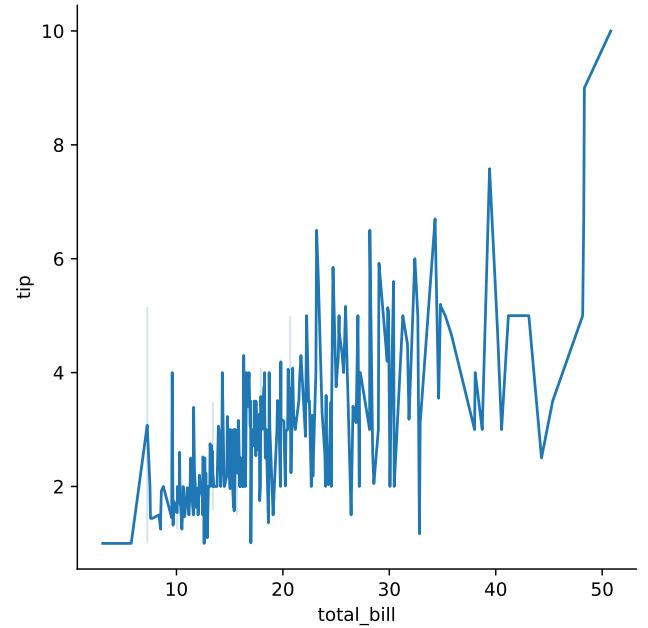
```
import seaborn as sns

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')
```



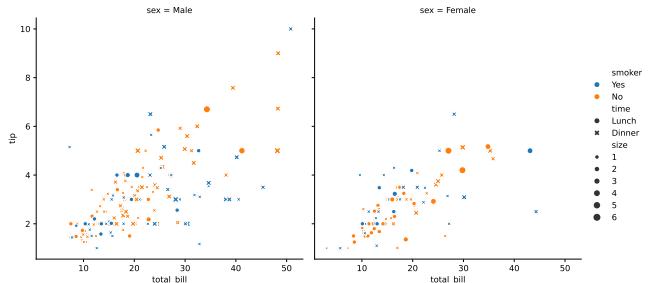
With figure-level plot, we can draw more than one plot (one axes).

Here we specify that different `sex` be on different column by specifying `col=sex`.

`col=sex`

```
import seaborn as sns

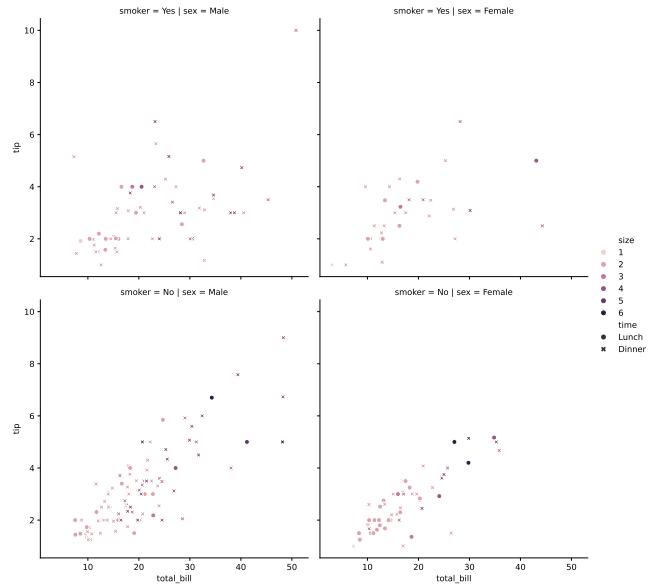
tips = sns.load_dataset("tips")
sns.relplot(kind='scatter',
            data=tips,
            x='total_bill',
            y='tip',
            hue='smoker',
            style='time',
            size='size',
            col='sex')
```



`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')
```

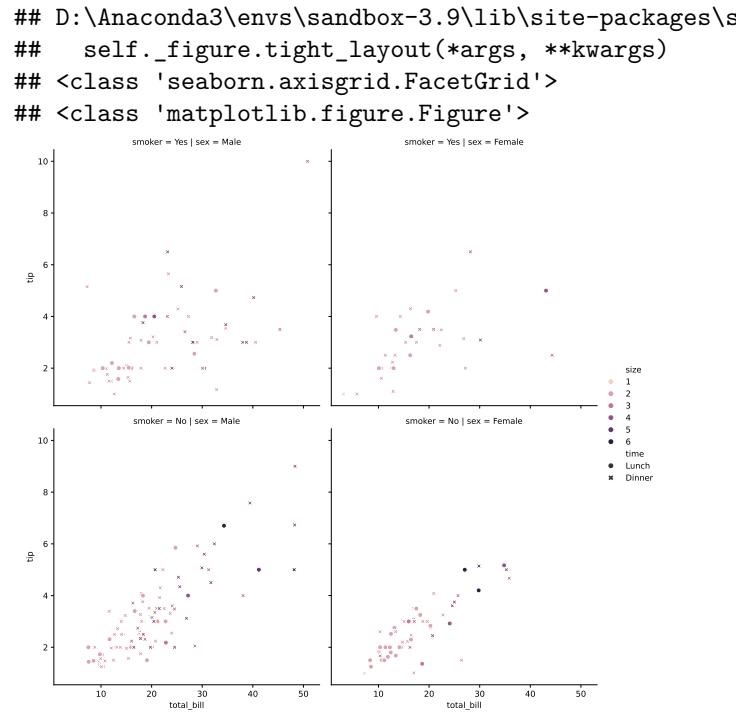


27.1.3.4 accessing figure and axes objects

Recall that Seaborn uses Matplotlib to draw the graphics. So underneath the Seaborn library, 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
```



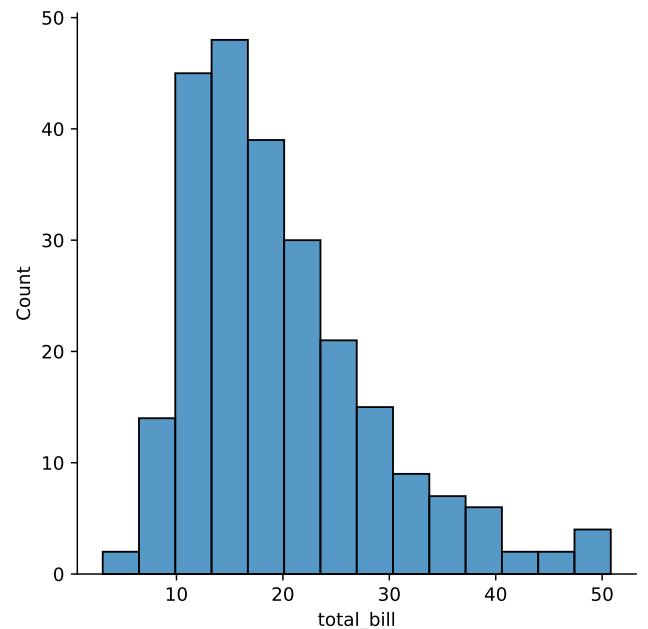
27.1.3.5 distribution plot

```
sns.displot
```

27.1.3.5.1 histogram

```
import seaborn as sns

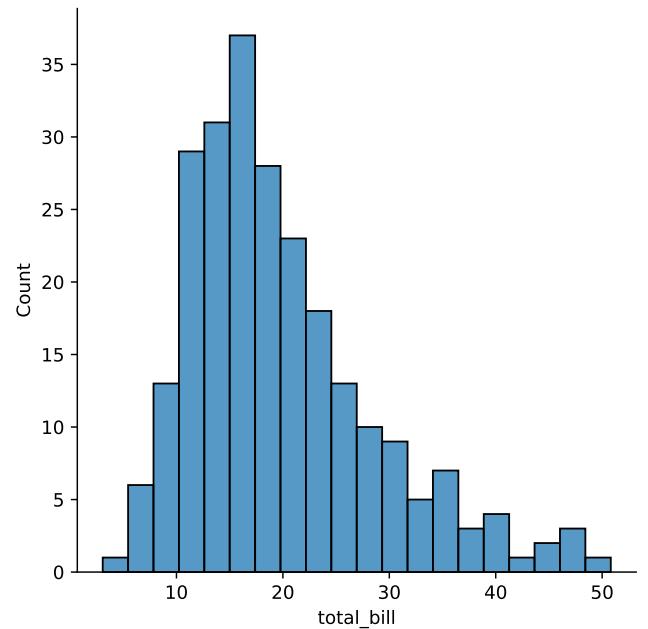
tips = sns.load_dataset("tips")
sns.displot(kind='hist',
            data=tips,
            x='total_bill')
```



bins

```
import seaborn as sns

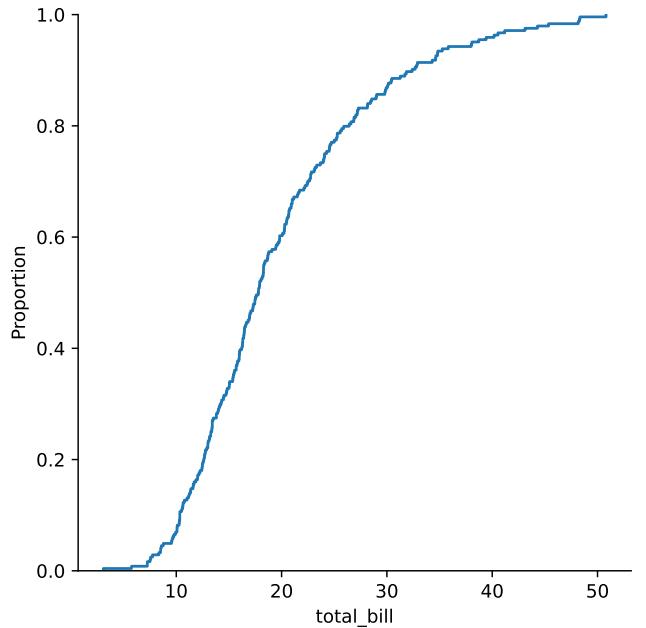
tips = sns.load_dataset("tips")
sns.displot(kind='hist', bins=20,
            data=tips,
            x='total_bill')
```



27.1.3.5.2 ECDF = empirical cumulative distrutive function

```
import seaborn as sns

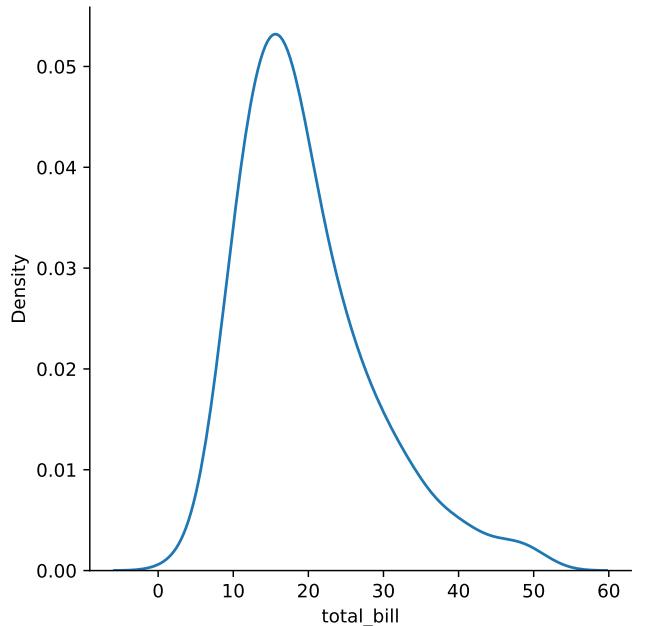
tips = sns.load_dataset("tips")
sns.displot(kind='ecdf',
            data=tips,
            x='total_bill')
```



27.1.3.5.3 KDE = kernel density estimation

```
import seaborn as sns

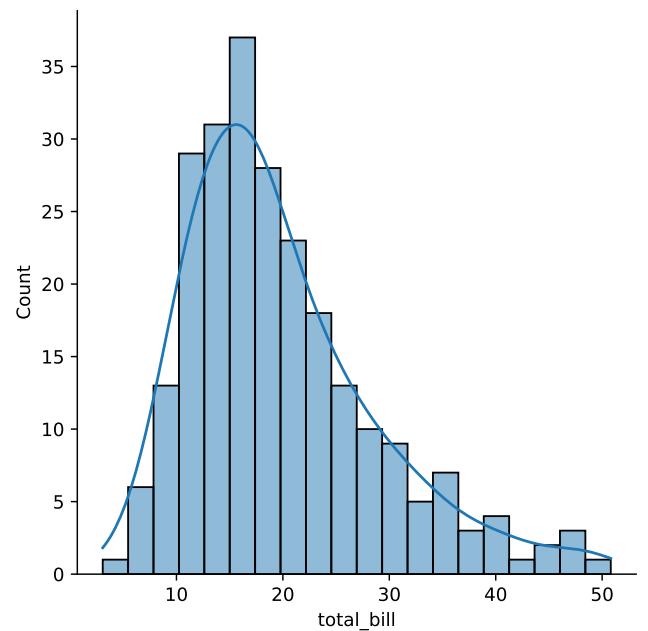
tips = sns.load_dataset("tips")
sns.displot(kind='kde',
            data=tips,
            x='total_bill')
```



histogram with KDE

```
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}, without _ in text

X ylabel={Count}, changed to ylabel={\$\text{Count}\$},: not necessary to be changed, the problem is _ in xlabel name

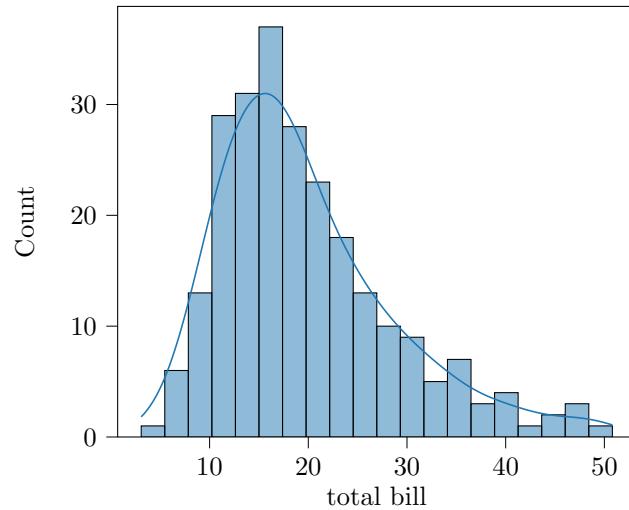
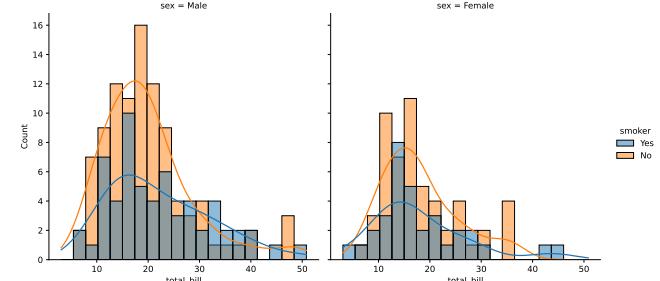


Figure 27.5: tikzplotlib

```
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'
            )
```



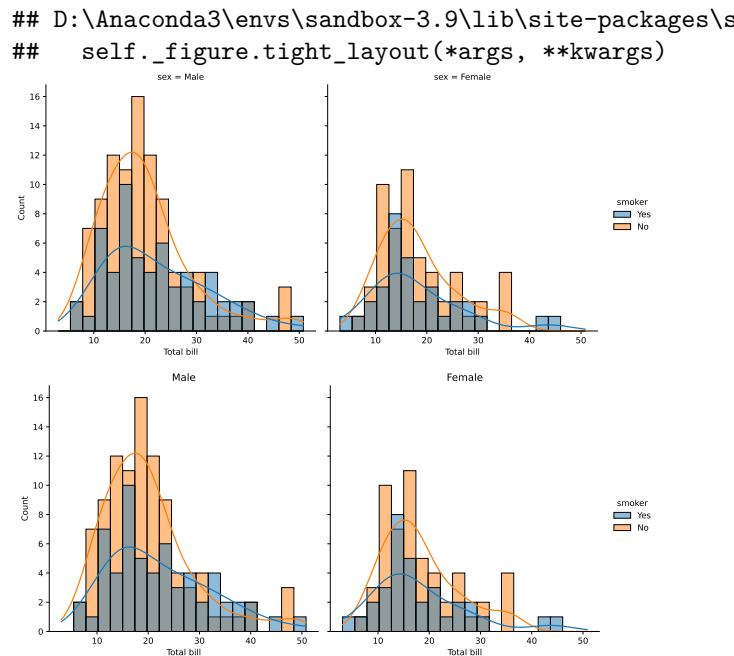
Let's customize the label and title.

```

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

```



27.1.3.6 categorical plot

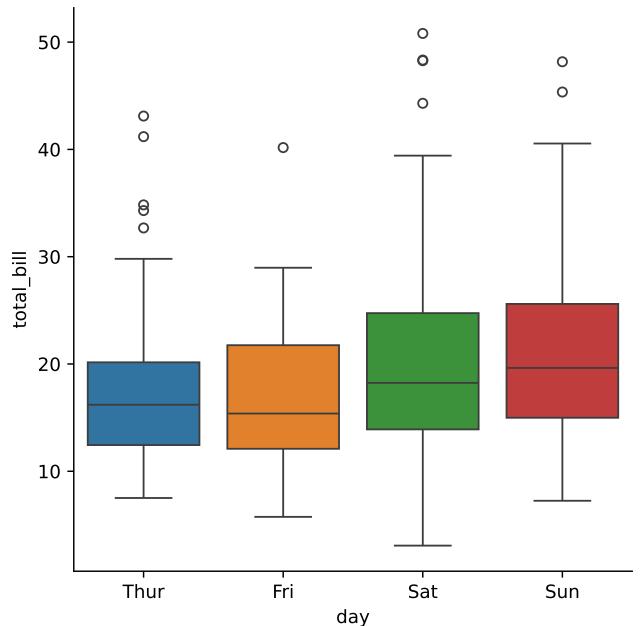
```
sns.catplot
```

```

import seaborn as sns

tips = sns.load_dataset("tips")
sns.catplot(kind='box',
            data=tips,
            x="day", hue="day",
            y="total_bill"
            )

```

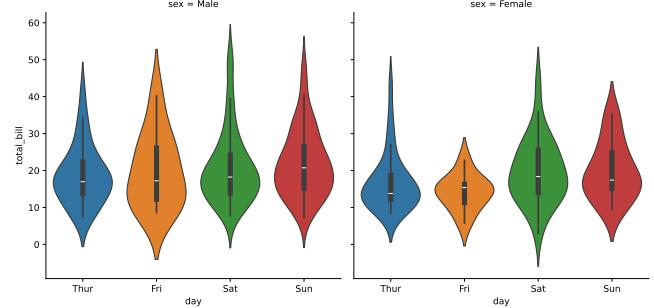


```

import seaborn as sns

tips = sns.load_dataset("tips")
sns.catplot(kind='violin',
            data=tips,
            x="day", hue="day",
            y="total_bill",
            col='sex'
            )

```



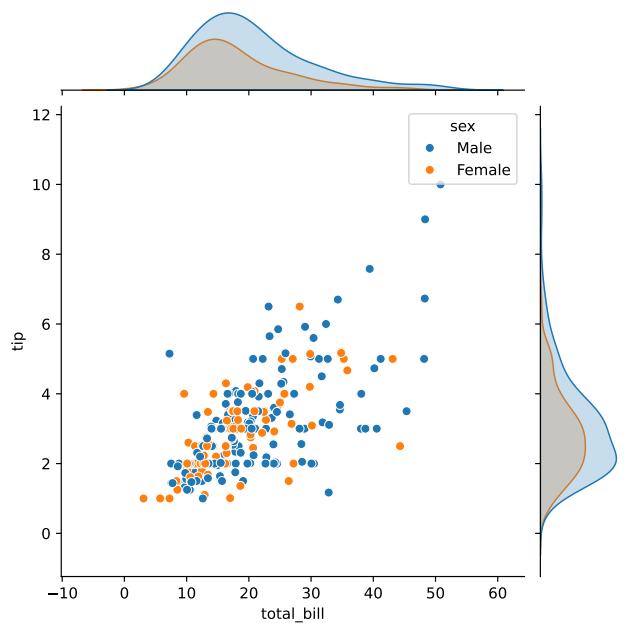
27.1.3.7 joint plot

```
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

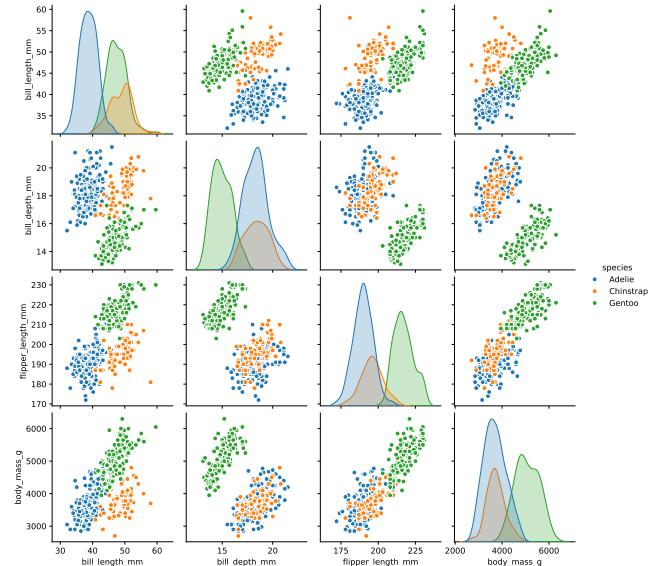


27.1.3.8 pair plot

```
sns.pairplot
```

```
import seaborn as sns

penguins = sns.load_dataset("penguins")
sns.pairplot(data=penguins,
              hue='species')
```



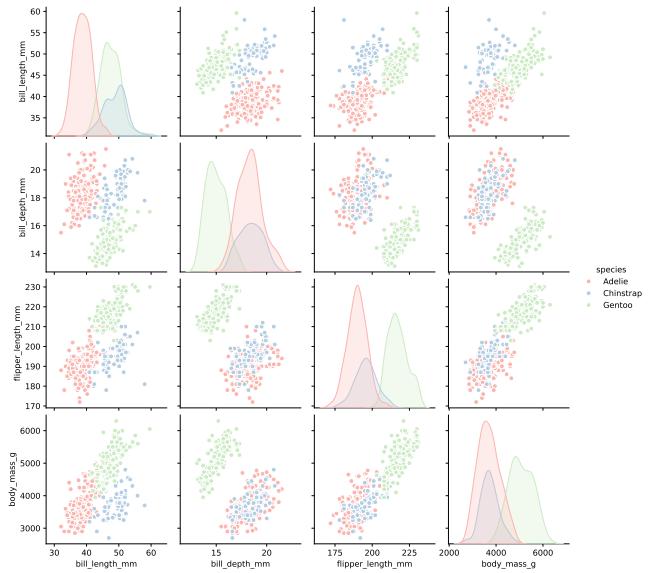
27.1.3.9 palette

Changing the color with Matplotlib's color

A list of color: <https://matplotlib.org/stable/tutorials/colors/colormaps.html>

```
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>

```
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.

```
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

survival 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

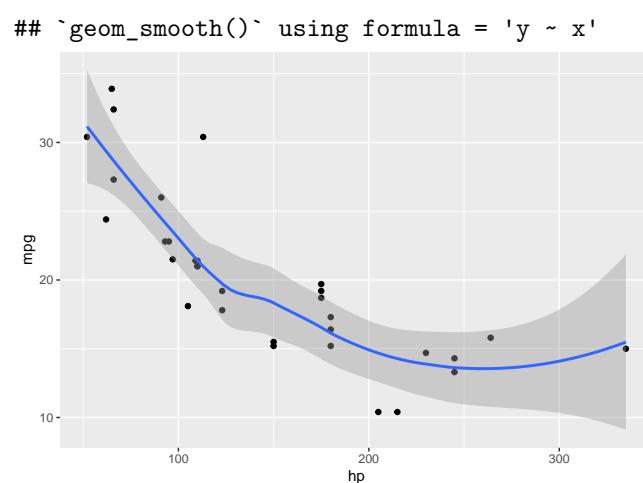
<https://bookdown.org/xiangyun/msg/system.html#chap:system>

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

```
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
```

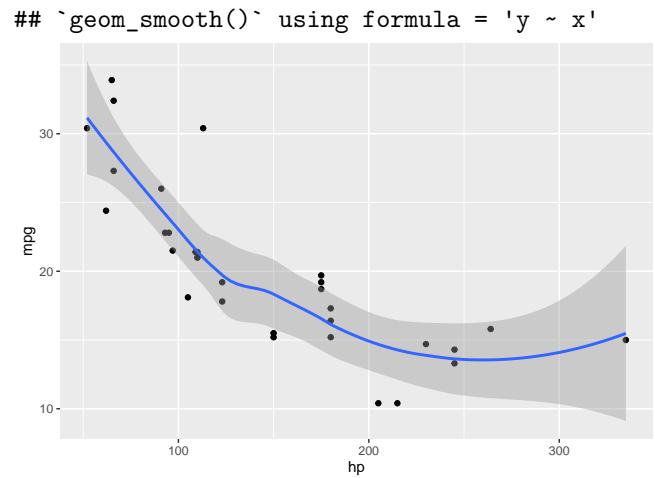


30.1 geom

<https://bookdown.org/xiangyun/msg/system.html#section-13>

```
library(ggplot2)

ggplot(aes(x = hp, y = mpg), data = mtcars) +
  geom_point() +
  geom_smooth(method = "loess")
```



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`

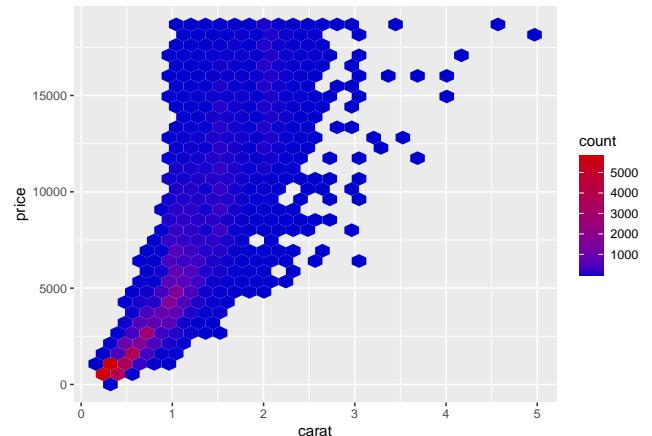
`stat_smooth`

`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.

```
# install.packages("hexbin")
```

```
library(ggplot2)

ggplot(aes(x = carat, y = price), data =
  diamonds) +
  geom_hex() +
  scale_fill_gradient(low = "blue3", high =
    "red3")
```

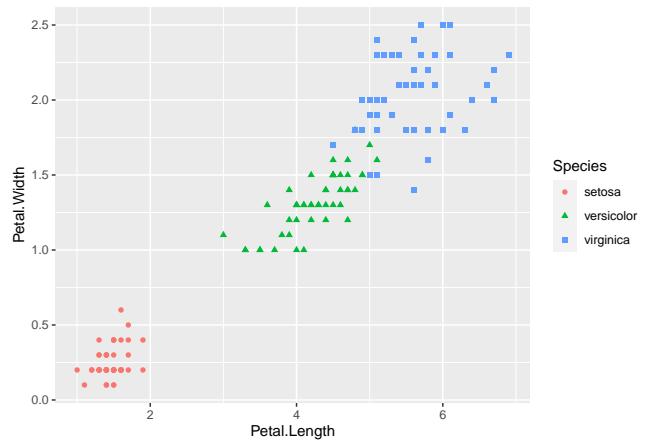


gradient color scales https://ggplot2.tidyverse.org/reference/scale_gradient.html?q=scale_fill_gradient#ref-usage

https://ggplot2.tidyverse.org/reference/scale_gradient.html?q=scale_fill_gradient#ref-examples

```
library(ggplot2)

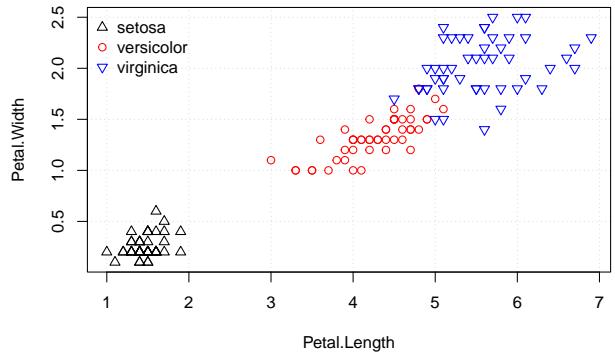
ggplot(aes(x = Petal.Length, y = Petal.Width),
       data = iris) +
  geom_point(aes(color = Species, shape =
                 Species))
```



basic plot system

<https://bookdown.org/xiangyun/msg/elements.html#sec:points>

```
# 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.Width
	## 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	2.0	2.6	

```
iris$Species
```

```
## [1] setosa      setosa      setosa      setosa      setosa
## [7] setosa      setosa      setosa      setosa      setosa
## [13] setosa     setosa      setosa      setosa      setosa
## [19] setosa     setosa      setosa      setosa      setosa
## [25] setosa     setosa      setosa      setosa      setosa
## [31] setosa     setosa      setosa      setosa      setosa
## [37] setosa     setosa      setosa      setosa      setosa
## [43] setosa     setosa      setosa      setosa      setosa
## [49] setosa     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 virginica
## [109] virginica virginica virginica virginica virginica
## [115] virginica virginica virginica virginica virginica
## [121] virginica virginica virginica virginica virginica
## [127] virginica virginica virginica virginica virginica
## [133] virginica virginica virginica virginica virginica
## [139] virginica virginica virginica virginica virginica
## [145] virginica virginica virginica virginica virginica
## Levels: setosa versicolor virginica
```

```
iris[["Species"]]
```

```
## [1] setosa      setosa      setosa      setosa      setosa
## [7] setosa      setosa      setosa      setosa      setosa
## [13] setosa     setosa      setosa      setosa      setosa
## [19] setosa     setosa      setosa      setosa      setosa
## [25] setosa     setosa      setosa      setosa      setosa
## [31] setosa     setosa      setosa      setosa      setosa
## [37] setosa     setosa      setosa      setosa      setosa
## [43] setosa     setosa      setosa      setosa      setosa
## [49] setosa     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 virginica
## [109] virginica virginica virginica virginica virginica
## [115] virginica virginica virginica virginica virginica
## [121] virginica virginica virginica virginica virginica
## [127] virginica virginica virginica virginica virginica
## [133] virginica virginica virginica virginica virginica
## [139] virginica virginica virginica virginica virginica
## [145] virginica virginica virginica virginica virginica
## Levels: setosa versicolor virginica
```

```
as.integer(iris[["Species"]])
```

```
levels(iris[["Species"]])
```

```
## [1] "setosa"      "versicolor" "virginica"
```

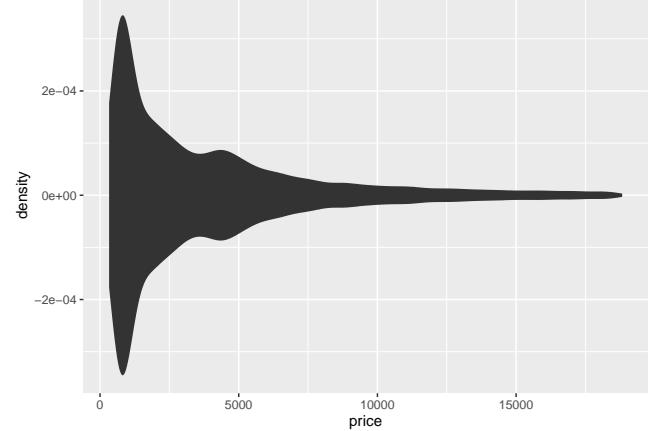
30.2 statistic

<https://bookdown.org/xiangyun/msg/system.html#section-14>

```
## 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 see
## generated.
```

```
library(ggplot2)

ggplot(diamonds, aes(x = price)) +
  stat_density(aes(ymax = ..density.., ymin =
    ↵ -..density..),
    geom = "ribbon", position = "identity"
  )
```



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

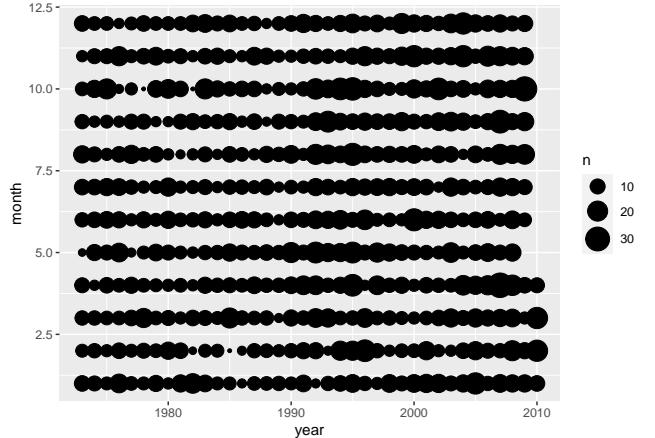
ggplot2

<https://bookdown.org/xiangyun/msg/system.html#section-15>

```
## Warning: package 'ggplot2' was built under R version 3.6.3
## Warning: The dot-dot notation (`..n..`) was deprecated in lifecycle v0.2.0.
## Please use `after_stat(n)` instead.
## This warning is displayed once every 8 hours.
## Call `lifecycle::last_lifecycle_warnings()` to see all.
## generated.
```

```
library(ggplot2)

data(quake6, package = "MSG")
ggplot(quake6, aes(x = year, y = month)) +
  stat_sum(aes(size = ..n..)) +
  scale_size(range = c(1, 8))
```



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 is 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

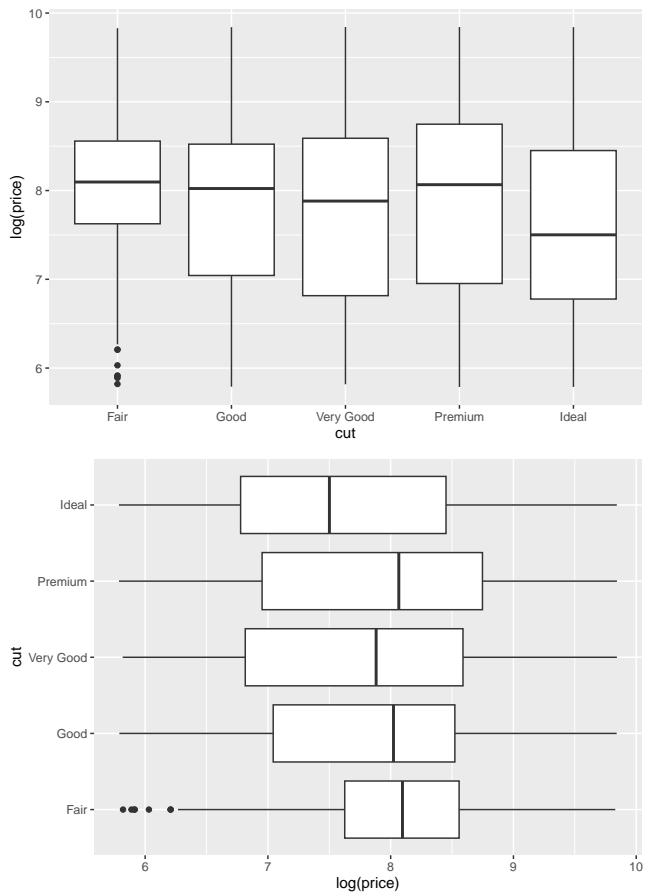
`scale_size`

30.4 coordinate system

<https://bookdown.org/xiangyun/msg/system.html#section-16>

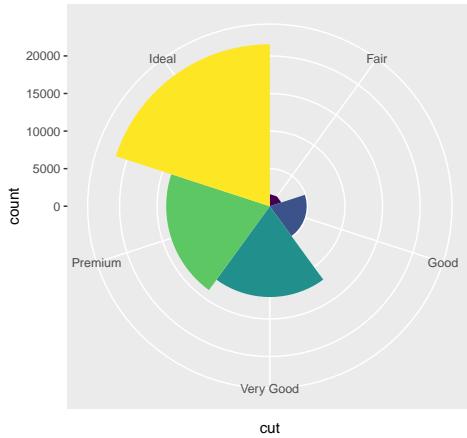
```
library(ggplot2)

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)
```

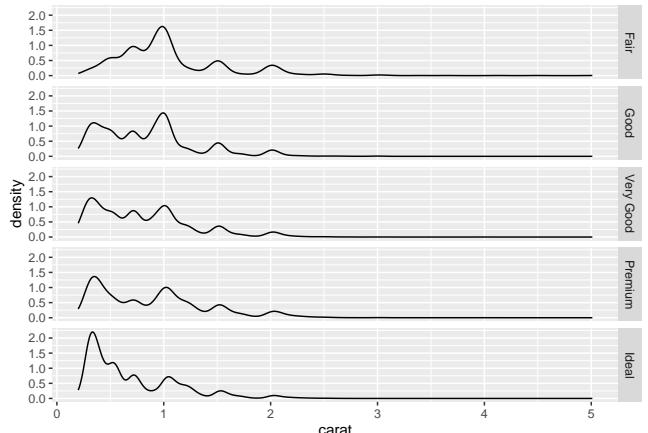


30.5 facet

<https://bookdown.org/xiangyun/msg/system.html#subsec:facet>

```
library(ggplot2)

ggplot(aes(x = carat), data = diamonds) +
  geom_density() +
  facet_grid(cut ~ .)
```

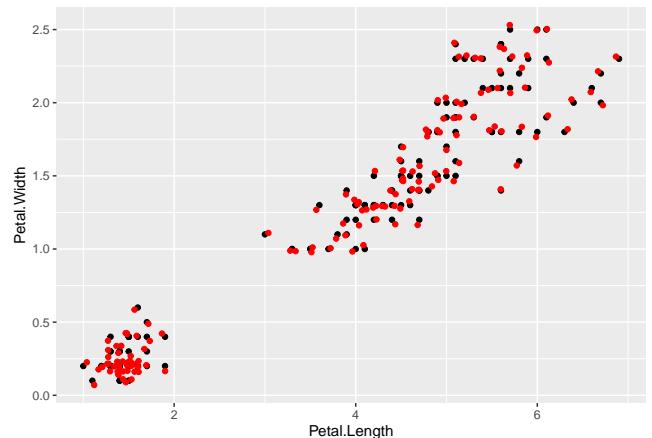


30.6 jitter

<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

$X \ O$

31.2.2 one-group pretest-posttest design

$O \ X \ O$

paired t test

⁹ p.62

O	X	O
O_t	X	O_t
O_0	X	O_1

Sources of Invalidity								
	Internal					External		
	History	Maturat	Testing	Instrumentation	Regression	Selection	Mortality	Interaction of Selection and Maturat, etc.
<i>Pre-Experimental Designs:</i>								
1. One-Shot Case Study $X \ O$	—	—			—	—		—
2. One-Group Pretest-Posttest Design $O \ X \ O$	—	—	—	?	+ +	—	—	?
3. Static-Group Comparison $\underline{X} \ O \ O$	+	?	+	+	+	—	—	—
<i>True Experimental Designs:</i>								
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

$$\begin{array}{ccc} R & X & O \\ R & & O \end{array}$$

	Sources of Invalidity								
	History	Maturation	Testing	Instrumentation	Regression	Selection	Mortality	Interaction of Selection and Maturation, etc.	
									External
<i>Quasi-Experimental Designs:</i>									
7. Time Series $O \ O \ O \ O X O \ O \ O \ O$	-	+	+	?	+	+	+	+	- ? ?
8. Equivalent Time Samples Design $X_1O \ X_2O \ X_1O \ X_2O, \text{ etc.}$	+	+	+	+	+	+	+	+	- ? -
9. Equivalent Materials Samples Design $M_aX_1O \ M_bX_2O \ M_cX_1O \ M_dX_2O, \text{ etc.}$	+	+	+	+	+	+	+	+	- ? ? -
10. Nonequivalent Control Group Design $\begin{array}{c} O \quad X \quad O \\ \hline O \quad O \end{array}$	+	+	+	+	?	+	+	-	- ? ?
11. Counterbalanced Designs $\begin{array}{cccc} X_1O & X_2O & X_1O & X_2O \\ \hline X_2O & X_1O & X_2O & X_1O \\ X_3O & X_4O & X_3O & X_4O \\ \hline X_4O & X_3O & X_4O & X_3O \end{array}$	+	+	+	+	+	+	+	-	
12. Separate-Sample Pretest-Posttest Design $\begin{array}{c} R \ O \ (X) \\ R \quad X \ O \end{array}$	-	-	+	?	+	+	-	-	+ + +
12a. $\begin{array}{c} R \ O \ (X) \\ R \quad X \ O \end{array}$	+	-	+	?	+	+	-	+	+ + +
12b. $\begin{array}{c} R \ O_1 \quad O_2 \ (X) \\ R \quad O_2 \ (X) \quad O_1 \end{array}$	-	+	+	?	+	+	-	?	+ + +
12c. $\begin{array}{c} R \ O_1 \quad X \quad O_2 \quad O_3 \\ R \quad X \quad O_3 \quad O_2 \end{array}$	-	-	+	?	+	+	+	-	+ + +

Figure 31.2: quasi-experimental designs (⁵ p.40)

	Sources of Invalidity																
	History	Maturity	Testing	Instrumentation	Regression	Selection	Mortality	Interaction of Selection and Maturation, etc.									
	Internal				External												
<i>Quasi-Experimental Designs</i>																	
<i>Continued:</i>																	
13. Separate-Sample Pretest-Posttest Control Group Design	+ + + + + + + -								+ + +								
$R \quad O \quad (X)$ $R \quad X \quad O$ $\hline R \quad O$ $R \quad O$																	
13a. $\begin{cases} R \quad O \quad (X) \\ R \quad X \quad O \\ \hline R \quad O \quad (X) \\ R \quad X \quad O \\ \hline R \quad O \\ R \quad O \\ \hline R' \quad O \\ R \quad O \\ \hline R \quad O \\ R \quad O \end{cases}$	+ + + + + + + +								+ + +								
14. Multiple Time-Series	+ + + + + + + +								- - ?								
$\begin{matrix} O & O & O & X & O & O & O \\ \hline O & O & O & O & O & O & O \end{matrix}$																	
15. Institutional Cycle Design																	
Class A X O_1																	
Class B ₁ R O_2 X O_3																	
Class B ₂ R X O_4																	
Class C O_5 X																	
$O_2 < O_1$	+ - + + ? - ?								+ ? +								
$O_5 < O_4$																	
$O_2 < O_3$	- - - ? ? + +								- ? +								
$O_2 < O_4$																	
$O_6 = O_7$	- - + ? ? + ?								+ ? ?								
$O_2 = O_{2a}$	+ -																
16. Regression Discontinuity	+ + + ? + + ? +								+ - + +								

• 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

$$\begin{array}{ccc} R & X & O_{11} \\ R & P & O_{21} \end{array}$$

$$\begin{array}{ccc} R & X_g & O_{gt} \\ R & X & O_{11} \\ R & P & O_{21} \end{array}$$

$$\begin{array}{cccc} R & X & O \\ R & X_g & O_{gt} \\ R & X & O_{11} \\ R & P & O_{21} \end{array}$$

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

$$\begin{array}{ccccc} R & O_{gt} & X_g & O_{gt} \\ R & O_{10} & X & O_{11} \\ R & O_{20} & & O_{21} \\ R & & X & O_{31} \\ R & & & O_{41} \end{array}$$

$$\begin{array}{cccc} R & O & X & O \\ R & O_{gt} & X_g & O_{gt} \\ R & O_{10} & X & O_{11} \\ R & O_{20} & & O_{21} \\ R & & X & O_{31} \\ R & & & O_{41} \end{array}$$

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**

$$\begin{aligned} Y_{ij} &= \mu_i + \varepsilon_{ij} \\ &= \mu_i + b_j + \varepsilon_{ij}^*, \begin{cases} i \in \mathbb{N} \cap [1, n_i] & \mathbb{N} \ni n_i \text{ experimental units per block} \\ j \in \mathbb{N} \cap [1, n_j] & \mathbb{N} \ni n_j \text{ blocks} \end{cases} \end{aligned}$$

where

$$\varepsilon_{ij} = b_j + \varepsilon_{ij}^*$$

i.e. the variation between groups or blocks of experimental units has been identified and isolated from ε_{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.

$$\begin{aligned} Y_{ij} - Y_{i'j} &= (\mu_i + b_j + \varepsilon_{ij}^*) \\ &\quad - (\mu_{i'} + b_j + \varepsilon_{i'j}^*) \\ &= (\mu_i - \mu_{i'}) + (\varepsilon_{ij}^* - \varepsilon_{i'j}^*) \end{aligned}$$

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

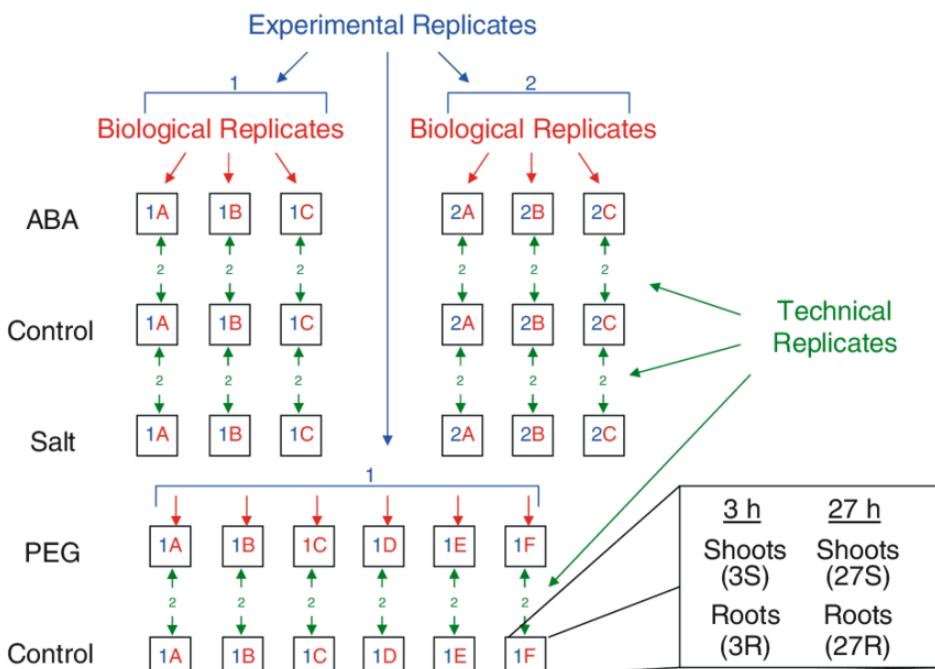


Figure 32.1: experimental, biological, technical replicates (11)

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}
```

$p = 4$ columns			
A	B	C	D
B	C	D	A
C	D	A	B
D	A	B	C

$p = |\{A, B, C, D\}| = 4$ treatments

Figure 32.2: Latin square example

$$Y_{ijk} = \mu + \alpha_i + \beta_j + \gamma_k + \varepsilon_{ijk}, \begin{cases} i \in \mathbb{N} \cap [1, p] & \mathbb{N} \ni p \text{ treatments} \\ j \in \mathbb{N} \cap [1, p] & \mathbb{N} \ni p \text{ rows} \\ k \in \mathbb{N} \cap [1, p] & \mathbb{N} \ni p \text{ columns} \end{cases}$$

$$\varepsilon_{ijk} \stackrel{\text{i.i.d.}}{\sim} n(0, \sigma^2)$$

ρ_i : i^{th} row

κ_j : j^{th} column

τ_k : k^{th} treatment

$$Y_{ijk} = \mu + \alpha_i + \beta_j + \gamma_k + \varepsilon_{ijk}, \begin{cases} i \in \mathbb{N} \cap [1, p] & \mathbb{N} \ni p \text{ treatments} \\ j \in \mathbb{N} \cap [1, p] & \mathbb{N} \ni p \text{ rows} \\ k \in \mathbb{N} \cap [1, p] & \mathbb{N} \ni p \text{ columns} \end{cases}$$

$$= \mu + \rho_i + \kappa_j + \tau_k + \varepsilon_{ijk}, \begin{cases} i \in \mathbb{N} \cap [1, p] & \mathbb{N} \ni p \text{ rows} \\ j \in \mathbb{N} \cap [1, p] & \mathbb{N} \ni p \text{ columns} \\ k \in \mathbb{N} \cap [1, p] & \mathbb{N} \ni p \text{ treatments} \end{cases}$$

32.2.5 model assumption and experimental unit, measurement/observational unit

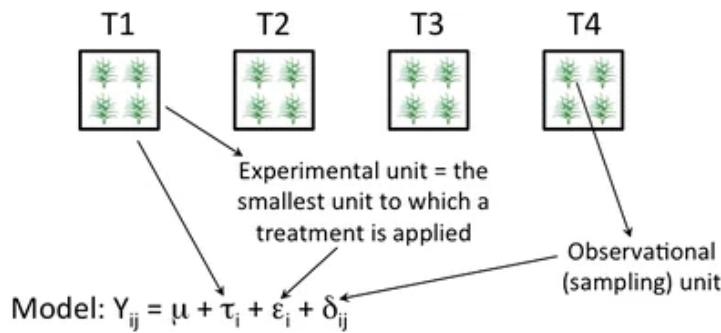
$$Y_{ijk} = \mu + \tau_i + \beta_j + \epsilon_{ij} + \Delta_{ijk}$$

32.3 experiment structure

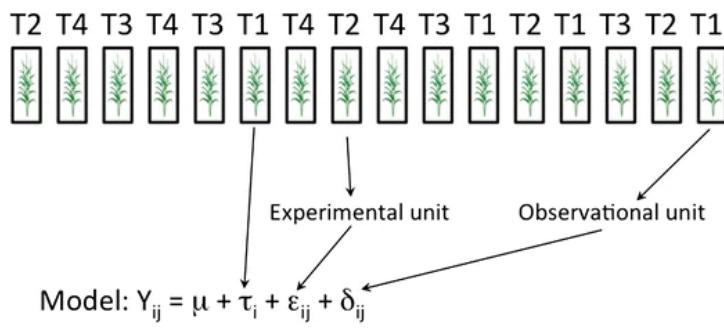
32.3.1 treatment structure

¹⁰ p.77

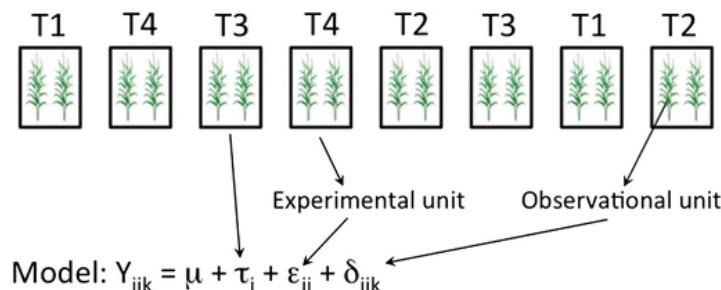
- 1-way treatment structure
- 2-way treatment structure
- factorial arrangement treatment structure



ANOVA Source of variation	df
Treatments + Experimental error	3 (fixed)
Observational error	12

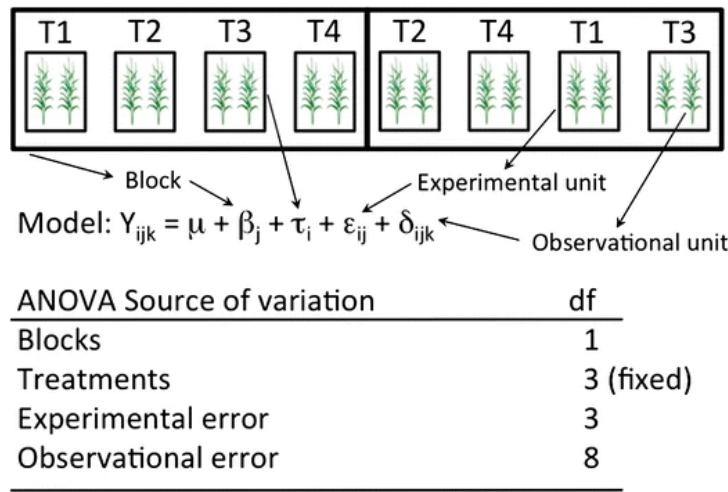
Figure 32.3: model assumption and experimental unit 1 ([12](#) fig.1)

ANOVA Source of variation	df
Treatments	3 (fixed)
Error (experimental + observational)	12

Figure 32.4: model assumption and experimental unit 2 ([12](#) fig.2)

ANOVA Source of variation	df
Treatments	3 (fixed)
Experimental error	4
Observational error	8

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 Minitab

<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

$$\begin{aligned} c &= a + bi = a + ib, \begin{cases} c \in \mathbb{C} \\ i^2 = -1 \\ a, b \in \mathbb{R} \Leftrightarrow \langle a, b \rangle \in \mathbb{R}^2 \end{cases} \\ z &= x + yi = x + iy, \begin{cases} z \in \mathbb{C} \\ i^2 = -1 \\ x, y \in \mathbb{R} \Leftrightarrow \langle x, y \rangle \in \mathbb{R}^2 \end{cases} \\ &= \sqrt{x^2 + y^2} \left(\frac{x}{\sqrt{x^2 + y^2}} + \frac{y}{\sqrt{x^2 + y^2}} i \right) = r (\cos \theta + i \sin \theta) = re^{i\theta} \end{aligned}$$

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>

$$\begin{aligned} t &= a + bi + cj = a + ib + jc, \begin{cases} t \in \mathbb{T} \\ i^2 = -1 \\ j^2 = -1 \end{cases} \\ w &= x + yi + zj = x + iy + jz, \begin{cases} w \in \mathbb{T} \\ i^2 = -1 \\ j^2 = -1 \end{cases} \\ &= \sqrt{x^2 + y^2 + z^2} \left(\frac{x}{\sqrt{x^2 + y^2 + z^2}} + \frac{y}{\sqrt{x^2 + y^2 + z^2}} i + \frac{z}{\sqrt{x^2 + y^2 + z^2}} j \right) = ? \end{aligned}$$

$$\begin{cases} A(BC) = (AB)C & (a) \text{ associativity} \\ A(B+C) = AB + AC & (d) \text{ distributivity} \end{cases}$$

$$\begin{aligned} \mathbb{T} &\ni ij = X + Yi + Zj \in \mathbb{T} \\ -j &= (i^2) j \stackrel{(a)}{=} i(ij) = i(X + Yi + Zj) \stackrel{(d)}{=} -Y + Xi + Zij \\ ij &= \frac{Y}{Z} - \frac{X}{Z}i - \frac{1}{Z}j \Rightarrow \begin{cases} X = \frac{Y}{Z} \\ Y = -\frac{X}{Z} \\ Z = -\frac{1}{Z} \end{cases} \Rightarrow Z^2 = -1 \Rightarrow Z \notin \mathbb{R} \\ -i &= i(j^2) \stackrel{(a)}{=} (ij)j = (X + Yi + Zj)j \stackrel{(d)}{=} -Z + Xj + Yij \\ ij &= \frac{Z}{Y} - \frac{1}{Y}i - \frac{X}{Y}j \Rightarrow \begin{cases} X = \frac{Z}{Y} \\ Y = -\frac{1}{Y} \\ Z = -\frac{X}{Y} \end{cases} \Rightarrow Y^2 = -1 \Rightarrow Y \notin \mathbb{R} \end{aligned}$$

34.1.3 quaternion

<https://en.wikipedia.org/wiki/Quaternion>

$$\begin{aligned} q &= a + bi + cj + dk = a + ib + jc + kd, \begin{cases} q \in \mathbb{H} \\ a, b, c, d \in \mathbb{R} \end{cases} \Leftrightarrow \langle a, b, c, d \rangle \in \mathbb{R}^4 \\ &= w = t + xi + yj + zk = t + ix + jy + kz, \begin{cases} w \in \mathbb{H} \\ t, x, y, z \in \mathbb{R} \end{cases} \Leftrightarrow \langle t, x, y, z \rangle \in \mathbb{R}^4 \\ &=? \\ q &= a + bi + cj + dk = a + ib + jc + kd, \begin{cases} q \in \mathbb{H} \\ i^2 = j^2 = k^2 = -1 = ijk \Rightarrow ij = k \\ a, b, c, d \in \mathbb{R} \end{cases} \Leftrightarrow \langle a, b, c, d \rangle \in \mathbb{R}^4 \\ &= w = t + xi + yj + zk = t + ix + jy + kz, \begin{cases} w \in \mathbb{H} \\ i^2 = j^2 = k^2 = -1 = ijk \Rightarrow ij = k \\ t, x, y, z \in \mathbb{R} \end{cases} \Leftrightarrow \langle t, x, y, z \rangle \in \mathbb{R}^4 \\ &= t + (i \ j \ k) \begin{pmatrix} x \\ y \\ z \end{pmatrix} = x_0 + (i \ j \ k) \begin{pmatrix} x \\ y \\ z \end{pmatrix} = x_0 + (e_1 \ e_2 \ e_3) \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = x_0 + x, \begin{cases} e_1 = i = i \\ e_2 = j = j \\ e_3 = k = k \end{cases} \\ &= t + \frac{ix + jy + kz}{r} r, \begin{cases} r^2 = x^2 + y^2 + z^2 \\ \|n\|^2 = \left(\frac{ix + jy + kz}{r} \right)^2 = -1 \end{cases} \Rightarrow |q|^2 = t^2 + r^2 \\ &= \sqrt{t^2 + r^2} \left(\frac{t}{\sqrt{t^2 + r^2}} + n \frac{r}{\sqrt{t^2 + r^2}} \right) = |q| \left(\cos \frac{\theta}{2} + n \sin \frac{\theta}{2} \right) = |q| e^{n \frac{\theta}{2}} \end{aligned}$$

The quaternion set is denoted \mathbb{H} for Sir R.W. **Hamilton**, because he suddenly and strikingly realized

$$\begin{cases} ij = k \\ k \in \mathbb{H} \end{cases} \Rightarrow ij \in \mathbb{H} \text{ for closure property}$$

for the sake of rigorosity, see **group theory**^[37]

$$k^2 = -1$$

$$\begin{aligned} ij &= k \\ ijk &= i(jk) \stackrel{(a)}{=} (ij)k = kk = k^2 = -1 \\ kij &= (ki)j \stackrel{(a)}{=} k(ij) = kk = k^2 = -1 \end{aligned}$$

$$\begin{aligned}
& ij = k \\
- j &= (i^2) j \stackrel{(a)}{=} i(ij) = ik \\
-i &= i(j^2) \stackrel{(a)}{=} (ij)j = kj \\
\\
-j &= (i^2) j \stackrel{(a)}{=} i(ij) = ik \\
1 &= -j^2 = j(-j) = j(ik) \stackrel{(a)}{=} (ji)k \\
k &= [1]k = [(ji)k]k \stackrel{(a)}{=} (ji)(k^2) = (ji)(-1) \\
-k &= ji \\
-i &= i(j^2) \stackrel{(a)}{=} (ij)j = kj \\
1 &= (-i)i = (kj)i \stackrel{(a)}{=} k(ji) = kji \\
1 &= kji
\end{aligned}$$

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].

$$\begin{cases} ij = k \\ ji = -k \end{cases} \Leftrightarrow ji = -k = -ij \\
\Leftrightarrow ji = -ij \\
\Leftrightarrow ij + ji = 0$$

$$\begin{cases} kij = -1 \\ kji = 1 \end{cases} \Leftrightarrow kij = -1 = -kji \\
\Leftrightarrow kij = -kji \\
\Leftrightarrow kij + kji = 0$$

$$\begin{aligned}
q &= a + bi + cj + dk = a + ib + jc + kd, \quad \begin{cases} q \in \mathbb{H} \\ i^2 = j^2 = k^2 = -1 = ijk \Rightarrow ij = k \\ a, b, c, d \in \mathbb{R} \end{cases} \Leftrightarrow \langle a, b, c, d \rangle \in \mathbb{R}^4 \\
&= w = t + xi + yj + zk = t + ix + jy + kz, \quad \begin{cases} w \in \mathbb{H} \\ i^2 = j^2 = k^2 = -1 = ijk \Rightarrow ij = k \\ t, x, y, z \in \mathbb{R} \end{cases} \Leftrightarrow \langle t, x, y, z \rangle \in \mathbb{R}^4 \\
&= t + (i \ j \ k) \begin{pmatrix} x \\ y \\ z \end{pmatrix} = x_0 + (\mathbf{i} \ \mathbf{j} \ \mathbf{k}) \begin{pmatrix} x \\ y \\ z \end{pmatrix} = x_0 + (\mathbf{e}_1 \ \mathbf{e}_2 \ \mathbf{e}_3) \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = x_0 + \mathbf{x}, \quad \begin{cases} e_1 = \mathbf{i} = i \\ e_2 = \mathbf{j} = j \\ e_3 = \mathbf{k} = k \end{cases}
\end{aligned}$$

.	1	i	j	k
1	1	i	j	k
i	i	-1	k	-j
j	j	-k	-1	i
k	k	j	-i	-1

.	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

.	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

$$\begin{aligned}
 \mathbf{x}_1 \mathbf{x}_2 &= (x_{11}i + x_{12}j + x_{13}k)(x_{21}i + x_{22}j + x_{23}k) \\
 &= x_{11}x_{21}i^2 + x_{11}x_{22}ij + x_{11}x_{23}ik \\
 &\quad + x_{12}x_{21}ji + x_{12}x_{22}j^2 + x_{12}x_{23}jk \\
 &\quad + x_{13}x_{21}ki + x_{13}x_{22}kj + x_{13}x_{23}k^2 \\
 &= -(x_{11}x_{21} + x_{12}x_{22} + x_{13}x_{23}) \\
 &\quad + (x_{12}x_{23} - x_{13}x_{22})jk + (x_{13}x_{21} - x_{11}x_{23})ki + (x_{11}x_{22} - x_{12}x_{21})ij \\
 &= -(x_{11}x_{21} + x_{12}x_{22} + x_{13}x_{23}) \\
 &\quad + (x_{12}x_{23} - x_{13}x_{22})i + (x_{13}x_{21} - x_{11}x_{23})j + (x_{11}x_{22} - x_{12}x_{21})k \\
 &= -(\mathbf{x}_1 \cdot \mathbf{x}_2) + (\mathbf{x}_1 \times \mathbf{x}_2), \begin{cases} \mathbf{x}_1 \cdot \mathbf{x}_2 = x_{11}x_{21} + x_{12}x_{22} + x_{13}x_{23} \\ \mathbf{x}_1 \times \mathbf{x}_2 = \begin{vmatrix} i & j & k \\ x_{11} & x_{12} & x_{13} \\ x_{21} & x_{22} & x_{23} \end{vmatrix} \end{cases} \\
 &= -\mathbf{x}_1 \cdot \mathbf{x}_2 + \mathbf{x}_1 \times \mathbf{x}_2
 \end{aligned}$$

product of two general quaternions / ordinary quaternions = Grassmann product

$$\begin{aligned}
q_1 q_2 &= (q_{10} + q_{11}\mathbf{i} + q_{12}\mathbf{j} + q_{13}\mathbf{k})(q_{20} + q_{21}\mathbf{i} + q_{22}\mathbf{j} + q_{23}\mathbf{k}) \\
&= (x_{10} + \mathbf{x}_1)(x_{20} + \mathbf{x}_2), \quad \begin{cases} x_{i\mu} = q_{i\mu} & \mu \in \{0\} \cup (\mathbb{N} \cap [1, 3]) \\ \mathbf{x}_i = x_{ij}\mathbf{e}_j & i, j \in \mathbb{N} \cap [1, 3], \end{cases} \quad \begin{cases} \mathbf{e}_1 = \mathbf{i} \\ \mathbf{e}_2 = \mathbf{j} \\ \mathbf{e}_3 = \mathbf{k} \end{cases} \\
&= x_{10}x_{20} + x_{10}\mathbf{x}_2 + x_{20}\mathbf{x}_1 + \mathbf{x}_1\mathbf{x}_2 \\
&= x_{10}x_{20} + x_{10}\mathbf{x}_2 + x_{20}\mathbf{x}_1 - \mathbf{x}_1 \cdot \mathbf{x}_2 + \mathbf{x}_1 \times \mathbf{x}_2 \\
&= (x_{10}x_{20} - \mathbf{x}_1 \cdot \mathbf{x}_2) + (x_{10}\mathbf{x}_2 + x_{20}\mathbf{x}_1 + \mathbf{x}_1 \times \mathbf{x}_2) \\
x_{10}x_{20} - \mathbf{x}_1 \cdot \mathbf{x}_2 &= (q_{10} \quad q_{11} \quad q_{12} \quad q_{13}) \begin{pmatrix} 1 & & & \\ & -1 & & \\ & & -1 & \\ & & & -1 \end{pmatrix} \begin{pmatrix} q_{20} \\ q_{21} \\ q_{22} \\ q_{23} \end{pmatrix} = q_1^\mu \eta_{\mu\nu} q_2^\nu \\
&= (q_{10} \quad q_{11} \quad q_{12} \quad q_{13}) \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix} \begin{pmatrix} q_{20} \\ q_{21} \\ q_{22} \\ q_{23} \end{pmatrix} = \mathbf{q}_1^\top H \mathbf{q}_2, H = [\eta_{\mu\nu}]_{4 \times 4} = \eta_{\mu\nu}
\end{aligned}$$

$$\begin{aligned}
q_1 q_2 &= (x_{10}x_{20} - \mathbf{x}_1 \cdot \mathbf{x}_2) + (x_{10}\mathbf{x}_2 + x_{20}\mathbf{x}_1 + \mathbf{x}_1 \times \mathbf{x}_2) \\
QP &= (Q_0 P_0 - \mathbf{Q} \cdot \mathbf{P}) + (Q_0 \mathbf{P} + P_0 \mathbf{Q} + \mathbf{Q} \times \mathbf{P})
\end{aligned}$$

$$ab = (a_0 b_0 - \mathbf{a} \cdot \mathbf{b}) + (a_0 \mathbf{b} + b_0 \mathbf{a} + \mathbf{a} \times \mathbf{b})$$

Minkowski metric tensor

$$\eta = H = [\eta_{\mu\nu}]_{4 \times 4} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix} = \begin{pmatrix} 1 & & & \\ & -1 & & \\ & & -1 & \\ & & & -1 \end{pmatrix} = \eta_{\mu\nu}$$

and quaternions as 4-vectors or four-vectors

$$\mathbf{q}_1^\top = (q_{10} \quad q_{11} \quad q_{12} \quad q_{13})$$

$$\mathbf{q}_2 = \begin{pmatrix} q_{20} \\ q_{21} \\ q_{22} \\ q_{23} \end{pmatrix}$$

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

$$AB + BA = 0 \Leftrightarrow AB = -BA$$

$$AB = -BA \Rightarrow AB \neq BA$$

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

$$ab = (a_0 b_0 - \mathbf{a} \cdot \mathbf{b}) + (a_0 \mathbf{b} + b_0 \mathbf{a} + \mathbf{a} \times \mathbf{b})$$

pure imaginary

$$\begin{aligned} ab &= (a_0 b_0 - \mathbf{a} \cdot \mathbf{b}) + (a_0 \mathbf{b} + b_0 \mathbf{a} + \mathbf{a} \times \mathbf{b}) \\ &\stackrel{\begin{cases} a_0 = 0 \\ b_0 = 0 \end{cases}}{=} (00 - \mathbf{a} \cdot \mathbf{b}) + (0\mathbf{b} + 0\mathbf{a} + \mathbf{a} \times \mathbf{b}) \\ &= -\mathbf{a} \cdot \mathbf{b} + \mathbf{a} \times \mathbf{b} \end{aligned}$$

pure imaginary product can get both (real & imaginary) / (scalar & vector) parts

$$ab = \mathbf{a}\mathbf{b} = -\mathbf{a} \cdot \mathbf{b} + \mathbf{a} \times \mathbf{b}, \begin{cases} a = 0 + \mathbf{a} = \mathbf{a} \\ b = 0 + \mathbf{b} = \mathbf{b} \end{cases}$$

triple product

https://en.wikipedia.org/wiki/Triple_product

pure imaginary

$$\begin{cases} a = 0 + \mathbf{a} = \mathbf{a} \\ b = 0 + \mathbf{b} = \mathbf{b} \\ c = 0 + \mathbf{c} = \mathbf{c} \end{cases}$$

$$\begin{aligned} abc &= \mathbf{a}\mathbf{b}\mathbf{c} = (\mathbf{a}\mathbf{b})\mathbf{c} = (-\mathbf{a} \cdot \mathbf{b} + \mathbf{a} \times \mathbf{b})\mathbf{c} \\ &= \mathbf{a}(\mathbf{b}\mathbf{c}) = \mathbf{a}(-\mathbf{b} \cdot \mathbf{c} + \mathbf{b} \times \mathbf{c}) \end{aligned}$$

$$\begin{aligned} abc &= (\mathbf{a}\mathbf{b})\mathbf{c} = (-\mathbf{a} \cdot \mathbf{b} + \mathbf{a} \times \mathbf{b})\mathbf{c} \\ &= -(\mathbf{a} \cdot \mathbf{b})\mathbf{c} + (\mathbf{a} \times \mathbf{b})\mathbf{c} \\ &= -(\mathbf{a} \cdot \mathbf{b})\mathbf{c} + (-(\mathbf{a} \times \mathbf{b}) \cdot \mathbf{c} + (\mathbf{a} \times \mathbf{b}) \times \mathbf{c}) \\ &= [-(\mathbf{a} \cdot \mathbf{b}) \cdot \mathbf{c}] + [(\mathbf{a} \times \mathbf{b}) \times \mathbf{c} - (\mathbf{a} \cdot \mathbf{b}) \mathbf{c}] \\ &= \mathbf{a}(\mathbf{b}\mathbf{c}) = \mathbf{a}(-\mathbf{b} \cdot \mathbf{c} + \mathbf{b} \times \mathbf{c}) \\ &= -\mathbf{a}(\mathbf{b} \cdot \mathbf{c}) + \mathbf{a}(\mathbf{b} \times \mathbf{c}) \\ &= -\mathbf{a}(\mathbf{b} \cdot \mathbf{c}) + (-\mathbf{a} \cdot (\mathbf{b} \times \mathbf{c}) + \mathbf{a} \times (\mathbf{b} \times \mathbf{c})) \\ &= [-\mathbf{a} \cdot (\mathbf{b} \times \mathbf{c})] + [\mathbf{a} \times (\mathbf{b} \times \mathbf{c}) - \mathbf{a}(\mathbf{b} \cdot \mathbf{c})] \end{aligned}$$

by comparing (real & imaginary) / (scalar & vector) parts,

$$\begin{aligned} & \begin{cases} -(\mathbf{a} \times \mathbf{b}) \cdot \mathbf{c} = -\mathbf{a} \cdot (\mathbf{b} \times \mathbf{c}) \\ (\mathbf{a} \times \mathbf{b}) \times \mathbf{c} - (\mathbf{a} \cdot \mathbf{b}) \mathbf{c} = \mathbf{a} \times (\mathbf{b} \times \mathbf{c}) - \mathbf{a} (\mathbf{b} \cdot \mathbf{c}) \end{cases} \\ \Rightarrow & \begin{cases} (\mathbf{a} \times \mathbf{b}) \cdot \mathbf{c} = \mathbf{a} \cdot (\mathbf{b} \times \mathbf{c}) & (s) \\ (\mathbf{a} \times \mathbf{b}) \times \mathbf{c} - (\mathbf{a} \cdot \mathbf{b}) \mathbf{c} = \mathbf{a} \times (\mathbf{b} \times \mathbf{c}) - \mathbf{a} (\mathbf{b} \cdot \mathbf{c}) & (v) \end{cases} \end{aligned}$$

permutation

$$\sigma = \begin{pmatrix} x_1 & x_2 & \cdots \\ \sigma(x_1) & \sigma(x_2) & \cdots \end{pmatrix}$$

$$\begin{pmatrix} a & b & c \\ a & b & c \end{pmatrix}, \begin{pmatrix} a & b & c \\ b & c & a \end{pmatrix}, \begin{pmatrix} a & b & c \\ c & a & b \end{pmatrix}, \begin{pmatrix} a & b & c \\ a & c & b \end{pmatrix}, \begin{pmatrix} a & b & c \\ b & a & c \end{pmatrix}, \begin{pmatrix} a & b & c \\ c & b & a \end{pmatrix}$$

$$(s) \Rightarrow \begin{cases} (\mathbf{a} \times \mathbf{b}) \cdot \mathbf{c} = \mathbf{a} \cdot (\mathbf{b} \times \mathbf{c}) & \begin{pmatrix} a & b & c \\ a & b & c \end{pmatrix}, s_1 \\ (\mathbf{b} \times \mathbf{c}) \cdot \mathbf{a} = \mathbf{b} \cdot (\mathbf{c} \times \mathbf{a}) & \begin{pmatrix} a & b & c \\ b & c & a \end{pmatrix}, s_2 \\ (\mathbf{c} \times \mathbf{a}) \cdot \mathbf{b} = \mathbf{c} \cdot (\mathbf{a} \times \mathbf{b}) & \begin{pmatrix} a & b & c \\ c & a & b \end{pmatrix}, s_3 \\ (\mathbf{b} \times \mathbf{a}) \cdot \mathbf{c} = \mathbf{b} \cdot (\mathbf{a} \times \mathbf{c}) & \begin{pmatrix} a & b & c \\ b & a & c \end{pmatrix}, s_4 \\ (\mathbf{a} \times \mathbf{c}) \cdot \mathbf{b} = \mathbf{a} \cdot (\mathbf{c} \times \mathbf{b}) & \begin{pmatrix} a & b & c \\ a & c & b \end{pmatrix}, s_5 \\ (\mathbf{c} \times \mathbf{b}) \cdot \mathbf{a} = \mathbf{c} \cdot (\mathbf{b} \times \mathbf{a}) & \begin{pmatrix} a & b & c \\ c & b & a \end{pmatrix}, s_6 \end{cases}$$

$$\stackrel{\text{commutative}}{\Rightarrow} (\mathbf{a} \times \mathbf{b}) \cdot \mathbf{c} \stackrel{s_1}{=} (\mathbf{b} \times \mathbf{c}) \cdot \mathbf{a} \stackrel{s_2}{=} (\mathbf{c} \times \mathbf{a}) \cdot \mathbf{b} \stackrel{s_3}{=} (\mathbf{a} \times \mathbf{b}) \cdot \mathbf{c}$$

$$\stackrel{\text{anticommutative}}{=} -(\mathbf{b} \times \mathbf{a}) \cdot \mathbf{c} \stackrel{s_6}{=} -(\mathbf{c} \times \mathbf{b}) \cdot \mathbf{a} \stackrel{s_5}{=} -(\mathbf{a} \times \mathbf{c}) \cdot \mathbf{b} \stackrel{s_4}{=} -(\mathbf{b} \times \mathbf{a}) \cdot \mathbf{c}$$

$$\Leftrightarrow \mathbf{a} \cdot (\mathbf{b} \times \mathbf{c}) = \mathbf{b} \cdot (\mathbf{c} \times \mathbf{a}) = \mathbf{c} \cdot (\mathbf{a} \times \mathbf{b})$$

$$= -\mathbf{a} \cdot (\mathbf{c} \times \mathbf{b}) = -\mathbf{b} \cdot (\mathbf{a} \times \mathbf{c}) = -\mathbf{c} \cdot (\mathbf{b} \times \mathbf{a})$$

$$\begin{aligned}
 (v) \Rightarrow & \left\{ \begin{array}{l} (\mathbf{a} \times \mathbf{b}) \times \mathbf{c} - (\mathbf{a} \cdot \mathbf{b}) \mathbf{c} = \mathbf{a} \times (\mathbf{b} \times \mathbf{c}) - \mathbf{a} (\mathbf{b} \cdot \mathbf{c}) \\ (\mathbf{b} \times \mathbf{c}) \times \mathbf{a} - (\mathbf{b} \cdot \mathbf{c}) \mathbf{a} = \mathbf{b} \times (\mathbf{c} \times \mathbf{a}) - \mathbf{b} (\mathbf{c} \cdot \mathbf{a}) \\ (\mathbf{c} \times \mathbf{a}) \times \mathbf{b} - (\mathbf{c} \cdot \mathbf{a}) \mathbf{b} = \mathbf{c} \times (\mathbf{a} \times \mathbf{b}) - \mathbf{c} (\mathbf{a} \cdot \mathbf{b}) \\ (\mathbf{b} \times \mathbf{a}) \times \mathbf{c} - (\mathbf{b} \cdot \mathbf{a}) \mathbf{c} = \mathbf{b} \times (\mathbf{a} \times \mathbf{c}) - \mathbf{b} (\mathbf{a} \cdot \mathbf{c}) \\ (\mathbf{a} \times \mathbf{c}) \times \mathbf{b} - (\mathbf{a} \cdot \mathbf{c}) \mathbf{b} = \mathbf{a} \times (\mathbf{c} \times \mathbf{b}) - \mathbf{a} (\mathbf{c} \cdot \mathbf{b}) \\ (\mathbf{c} \times \mathbf{b}) \times \mathbf{a} - (\mathbf{c} \cdot \mathbf{b}) \mathbf{a} = \mathbf{c} \times (\mathbf{b} \times \mathbf{a}) - \mathbf{c} (\mathbf{b} \cdot \mathbf{a}) \end{array} \right. \begin{array}{l} \left(\begin{matrix} a & b & c \\ a & b & c \end{matrix} \right), v_1 \\ \left(\begin{matrix} a & b & c \\ b & c & a \end{matrix} \right), v_2 \\ \left(\begin{matrix} a & b & c \\ c & a & b \end{matrix} \right), v_3 \\ \left(\begin{matrix} a & b & c \\ b & a & c \end{matrix} \right), v_4 \\ \left(\begin{matrix} a & b & c \\ a & c & b \end{matrix} \right), v_5 \\ \left(\begin{matrix} a & b & c \\ c & b & a \end{matrix} \right), v_6 \end{array} \\
 \Rightarrow & \left\{ \begin{array}{l} -Z - C = X - A \quad v_1 \\ -X - A = Y - B \quad v_2 \\ -Y - B = Z - C \quad v_3 \\ Z - C = -Y - B \quad v_4 \\ Y - B = -X - A \quad v_5 \\ X - A = -Z - C \quad v_6 \end{array} \right. , \left\{ \begin{array}{ll} \cdot \text{ and scalar-vector product} & \text{commutative} \\ \times & \text{anticommutative} \end{array} \right. ,
 \end{aligned}$$

$$\left\{ \begin{array}{l} X = \mathbf{a} \times (\mathbf{b} \times \mathbf{c}) = -(\mathbf{b} \times \mathbf{c}) \times \mathbf{a} = -\mathbf{a} \times (\mathbf{c} \times \mathbf{b}) = (\mathbf{c} \times \mathbf{b}) \times \mathbf{a} \\ Y = \mathbf{b} \times (\mathbf{c} \times \mathbf{a}) = -(\mathbf{c} \times \mathbf{a}) \times \mathbf{b} = -\mathbf{b} \times (\mathbf{a} \times \mathbf{c}) = (\mathbf{a} \times \mathbf{c}) \times \mathbf{b} \\ Z = \mathbf{c} \times (\mathbf{a} \times \mathbf{b}) = -(\mathbf{a} \times \mathbf{b}) \times \mathbf{c} = -\mathbf{c} \times (\mathbf{b} \times \mathbf{a}) = (\mathbf{b} \times \mathbf{a}) \times \mathbf{c} \\ A = \mathbf{a} (\mathbf{b} \cdot \mathbf{c}) = \mathbf{a} (\mathbf{c} \cdot \mathbf{b}) = (\mathbf{c} \cdot \mathbf{b}) \mathbf{a} = (\mathbf{b} \cdot \mathbf{c}) \mathbf{a} \\ B = \mathbf{b} (\mathbf{c} \cdot \mathbf{a}) = \mathbf{b} (\mathbf{a} \cdot \mathbf{c}) = (\mathbf{a} \cdot \mathbf{c}) \mathbf{b} = (\mathbf{c} \cdot \mathbf{a}) \mathbf{b} \\ C = \mathbf{c} (\mathbf{a} \cdot \mathbf{b}) = \mathbf{c} (\mathbf{b} \cdot \mathbf{a}) = (\mathbf{b} \cdot \mathbf{a}) \mathbf{c} = (\mathbf{a} \cdot \mathbf{b}) \mathbf{c} \end{array} \right. \\
 \Rightarrow & \left\{ \begin{array}{l} -Z - C = X - A \quad v_1 = v_6 \\ -X - A = Y - B \quad v_2 = v_5 \\ -Y - B = Z - C \quad v_3 = v_4 \end{array} \right. \Leftrightarrow \left\{ \begin{array}{l} Z + X = A - C \\ X + Y = B - A \\ Y + Z = C - B \end{array} \right. \Leftrightarrow \left\{ \begin{array}{l} X + Y = B - A \\ Y + Z = C - B \\ Z + X = A - C \end{array} \right. \\
 \Leftrightarrow & \left\{ \begin{array}{l} 2(X + Y + Z) = 0 \Rightarrow X + Y + Z = 0 \\ Y + Z = C - B \Rightarrow X = B - C \Leftrightarrow \mathbf{a} \times (\mathbf{b} \times \mathbf{c}) = \mathbf{b} (\mathbf{c} \cdot \mathbf{a}) - \mathbf{c} (\mathbf{a} \cdot \mathbf{b}) \text{ "back cab"} \\ Z + X = A - C \Rightarrow Y = C - A \Leftrightarrow \mathbf{b} \times (\mathbf{c} \times \mathbf{a}) = \mathbf{c} (\mathbf{a} \cdot \mathbf{b}) - \mathbf{a} (\mathbf{b} \cdot \mathbf{c}) \\ X + Y = B - A \Rightarrow Z = A - B \Leftrightarrow \mathbf{c} \times (\mathbf{a} \times \mathbf{b}) = \mathbf{a} (\mathbf{b} \cdot \mathbf{c}) - \mathbf{b} (\mathbf{c} \cdot \mathbf{a}) \end{array} \right.
 \end{aligned}$$

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

$$\begin{aligned}
 D &= \frac{\partial}{\partial t} + i \frac{\partial}{\partial x} + j \frac{\partial}{\partial y} + k \frac{\partial}{\partial z} = \partial_t + i \partial_x + j \partial_y + k \partial_z \\
 &= \frac{\partial}{\partial t} + i \frac{\partial}{\partial x} + j \frac{\partial}{\partial y} + k \frac{\partial}{\partial z} = \partial_t + i \partial_x + j \partial_y + k \partial_z \\
 &= \frac{\partial}{\partial x_0} + \mathbf{e}_1 \frac{\partial}{\partial x_1} + \mathbf{e}_2 \frac{\partial}{\partial x_2} + \mathbf{e}_3 \frac{\partial}{\partial x_3} = \partial_0 + \mathbf{e}_i \partial_i = \partial_0 + \nabla
 \end{aligned}$$

$$D = \partial_0 + i \partial_1 + j \partial_2 + k \partial_3 = \partial_0 + \nabla = D_0 + \mathbf{D}$$

34.1.3.5.2 nabla nabla = spatial differential operator = 3-differential operator / three-differential operator

$$\nabla = \mathbf{e}_i \partial_i = \sum_{i=1}^3 \mathbf{e}_i \partial_i = \sum_{i=1}^3 \mathbf{e}_i \frac{\partial}{\partial x_i} = \begin{pmatrix} \frac{\partial}{\partial x} & \frac{\partial}{\partial y} & \frac{\partial}{\partial z} \end{pmatrix}^\top = \begin{pmatrix} \frac{\partial}{\partial x} \\ \frac{\partial}{\partial y} \\ \frac{\partial}{\partial z} \end{pmatrix}$$

34.1.3.5.3 Laplace operator Laplace operator = Laplacian

$$\Delta = \nabla^2 = \nabla \cdot \nabla = \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2}$$

34.1.3.5.4 d'Alembert operator

$$\square = \square_c = \frac{1}{c^2} \frac{\partial^2}{\partial t^2} - \frac{\partial^2}{\partial x^2} - \frac{\partial^2}{\partial y^2} - \frac{\partial^2}{\partial z^2} = \frac{1}{c^2} \frac{\partial^2}{\partial t^2} - \Delta = \frac{1}{c^2} \frac{\partial^2}{\partial t^2} - \nabla^2$$

$$\square_1 = \square_{c=1} = \frac{\partial^2}{\partial t^2} - \frac{\partial^2}{\partial x^2} - \frac{\partial^2}{\partial y^2} - \frac{\partial^2}{\partial z^2} = \frac{\partial^2}{\partial t^2} - \Delta = \frac{\partial^2}{\partial t^2} - \nabla^2$$

34.1.3.6 electromagnetism

Maxwell

34.1.3.6.1 4-potential electromagnetic 4-potential / four-potential

$$A = A_0 + iA_1 + jA_2 + kA_3 = A_0 + \mathbf{A}$$

4-differential operator^[34.1.3.5.1]

$$\mathbf{D} = \partial_0 + i\partial_1 + j\partial_2 + k\partial_3 = \partial_0 + \nabla = \mathbf{D}_0 + \mathbf{D}$$

$$QP = (Q_0 P_0 - \mathbf{Q} \cdot \mathbf{P}) + (Q_0 \mathbf{P} + P_0 \mathbf{Q} + \mathbf{Q} \times \mathbf{P})$$

commutative bracket^[34.1.3.3.1.1]

$$\begin{aligned} [\mathbf{D}, A] &= \frac{\mathbf{D}A - AD}{2} \\ 2[\mathbf{D}, A] &= DA - AD \\ &= (\partial_0 + i\partial_1 + j\partial_2 + k\partial_3)(A_0 + iA_1 + jA_2 + kA_3) \\ &\quad - (A_0 + iA_1 + jA_2 + kA_3)(\partial_0 + i\partial_1 + j\partial_2 + k\partial_3) \\ DA &= (\mathbf{D}_0 A_0 - \mathbf{D} \cdot \mathbf{A}) + (\mathbf{D}_0 \mathbf{A} + A_0 \mathbf{D} + \mathbf{D} \times \mathbf{A}) \\ &= (\mathbf{D}_0 A_0 - \mathbf{D} \cdot \mathbf{A}) + (\mathbf{D}_0 \mathbf{A} + \mathbf{D} A_0 + \mathbf{D} \times \mathbf{A}) \\ &= \mathbf{D}_0 (A_0 + \mathbf{A}) - \mathbf{D} \cdot \mathbf{A} + \mathbf{D} A_0 + \mathbf{D} \times \mathbf{A} \\ &= \mathbf{D}_0 A - \mathbf{D} \cdot \mathbf{A} + \mathbf{D} A_0 + \mathbf{D} \times \mathbf{A} \\ &= \partial_0 A - \nabla \cdot \mathbf{A} + \nabla A_0 + \nabla \times \mathbf{A} \\ AD &= (A_0 \mathbf{D}_0 - \mathbf{A} \cdot \mathbf{D}) + (A_0 \mathbf{D} + \mathbf{D}_0 \mathbf{A} + \mathbf{A} \times \mathbf{D}) \\ &= (\mathbf{D}_0 A_0 - \mathbf{D} \cdot \mathbf{A}) + (\mathbf{D} A_0 + \mathbf{D}_0 \mathbf{A} - \mathbf{D} \times \mathbf{A}) \\ &= (\mathbf{D}_0 A_0 - \mathbf{D} \cdot \mathbf{A}) + (\mathbf{D}_0 \mathbf{A} + \mathbf{D} A_0 - \mathbf{D} \times \mathbf{A}) \\ &= \mathbf{D}_0 (A_0 + \mathbf{A}) - \mathbf{D} \cdot \mathbf{A} + \mathbf{D} A_0 - \mathbf{D} \times \mathbf{A} \\ &= \mathbf{D}_0 A - \mathbf{D} \cdot \mathbf{A} + \mathbf{D} A_0 - \mathbf{D} \times \mathbf{A} \\ &= \partial_0 A - \nabla \cdot \mathbf{A} + \nabla A_0 - \nabla \times \mathbf{A} \end{aligned}$$

$$\begin{aligned} \mathbf{D}A &= \mathbf{D}_0 A - \mathbf{D} \cdot \mathbf{A} + \mathbf{D} A_0 + \mathbf{D} \times \mathbf{A} = \partial_0 A - \nabla \cdot \mathbf{A} + \nabla A_0 + \nabla \times \mathbf{A} \\ AD &= \mathbf{D}_0 A - \mathbf{D} \cdot \mathbf{A} + \mathbf{D} A_0 - \mathbf{D} \times \mathbf{A} = \partial_0 A - \nabla \cdot \mathbf{A} + \nabla A_0 - \nabla \times \mathbf{A} \end{aligned}$$

$$\begin{aligned} \mathbf{D}\mathbf{A} - \mathbf{A}\mathbf{D} &= 2\mathbf{D} \times \mathbf{A} = 2\boldsymbol{\nabla} \times \mathbf{A} \\ [\mathbf{D}, \mathbf{A}] &= \frac{\mathbf{D}\mathbf{A} - \mathbf{A}\mathbf{D}}{2} = \mathbf{D} \times \mathbf{A} = \boldsymbol{\nabla} \times \mathbf{A} \end{aligned}$$

anticommutative bracket^[34.1.3.3.1.2]

$$\begin{aligned} \mathbf{D}\mathbf{A} + \mathbf{A}\mathbf{D} &= 2(\mathbf{D}_0\mathbf{A} - \mathbf{D} \cdot \mathbf{A} + \mathbf{D}\mathbf{A}_0) = 2(\partial_0\mathbf{A} - \boldsymbol{\nabla} \cdot \mathbf{A} + \boldsymbol{\nabla}\mathbf{A}_0) \\ \{\mathbf{D}, \mathbf{A}\} &= \frac{\mathbf{D}\mathbf{A} + \mathbf{A}\mathbf{D}}{2} = \mathbf{D}_0\mathbf{A} - \mathbf{D} \cdot \mathbf{A} + \mathbf{D}\mathbf{A}_0 = \partial_0\mathbf{A} - \boldsymbol{\nabla} \cdot \mathbf{A} + \boldsymbol{\nabla}\mathbf{A}_0 \end{aligned}$$

commutation and anticommutation on differential operator and any quaternion

$$\begin{aligned} [\mathbf{D}, \mathbf{Q}] &= \boldsymbol{\nabla} \times \mathbf{Q} \\ \{\mathbf{D}, \mathbf{Q}\} &= \partial_0\mathbf{Q} - \boldsymbol{\nabla} \cdot \mathbf{Q} + \boldsymbol{\nabla}\mathbf{Q}_0 \end{aligned}$$

or more evident

$$\begin{aligned} [\mathbf{D}, \mathbf{Q}] &= \boldsymbol{\nabla} \times \mathbf{Q} \\ \{\mathbf{D}, \mathbf{Q}\} &= \partial_0\mathbf{Q} - \boldsymbol{\nabla} \cdot \mathbf{Q} + \boldsymbol{\nabla}\mathbf{Q}_0 \\ &= \partial_0(\mathbf{Q}_0 + \mathbf{Q}) - \boldsymbol{\nabla} \cdot \mathbf{Q} + \boldsymbol{\nabla}\mathbf{Q}_0 \\ &= (\partial_0\mathbf{Q}_0 - \boldsymbol{\nabla} \cdot \mathbf{Q}) + (\partial_0\mathbf{Q} + \boldsymbol{\nabla}\mathbf{Q}_0) \\ &= \left(\frac{\partial\mathbf{Q}_0}{\partial t} - \boldsymbol{\nabla} \cdot \mathbf{Q} \right) + \left(\frac{\partial\mathbf{Q}}{\partial t} + \boldsymbol{\nabla}\mathbf{Q}_0 \right) \end{aligned}$$

34.1.3.6.2 Maxwell compromise for both quaternion and 3-vector electric potential and vector potential

$$\mathbf{A} = \mathbf{A}_0 + \mathbf{i}\mathbf{A}_1 + \mathbf{j}\mathbf{A}_2 + \mathbf{k}\mathbf{A}_3 = \mathbf{A}_0 + \mathbf{A} = \mathbf{U} + \mathbf{A}$$

electric quaternion and electric field

$$\begin{aligned} \mathbf{E} &= -\{\mathbf{D}, \mathbf{A}\} \\ &= -(\partial_0\mathbf{A} - \boldsymbol{\nabla} \cdot \mathbf{A} + \boldsymbol{\nabla}\mathbf{A}_0) \\ &= -\partial_0\mathbf{A} + \boldsymbol{\nabla} \cdot \mathbf{A} - \boldsymbol{\nabla}\mathbf{A}_0 \\ &= -\partial_t(\mathbf{U} + \mathbf{A}) + \boldsymbol{\nabla} \cdot \mathbf{A} - \boldsymbol{\nabla}\mathbf{U} \\ &= -\frac{\partial\mathbf{U}}{\partial t} + \boldsymbol{\nabla} \cdot \mathbf{A} - \boldsymbol{\nabla}\mathbf{U} - \frac{\partial\mathbf{A}}{\partial t} \\ &= \mathbf{E}_0 + \mathbf{E}, \quad \begin{cases} \mathbf{E}_0 = -\frac{\partial\mathbf{U}}{\partial t} + \boldsymbol{\nabla} \cdot \mathbf{A} \\ \mathbf{E} = -\boldsymbol{\nabla}\mathbf{U} - \frac{\partial\mathbf{A}}{\partial t} \end{cases} \quad \text{electric field 3-vector} \end{aligned}$$

magnetic field

$$\mathbf{B} = [\mathbf{D}, \mathbf{A}] = \boldsymbol{\nabla} \times \mathbf{A} = \begin{vmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ \frac{\partial}{\partial x} & \frac{\partial}{\partial y} & \frac{\partial}{\partial z} \\ \mathbf{A}_1 & \mathbf{A}_2 & \mathbf{A}_3 \end{vmatrix} = \mathbf{B}$$

Work on time? Yes.

$$q\mathbf{E} = q\mathbf{E}_0 + q\mathbf{E} = q\mathbf{E}_0 + \mathbf{F}_E$$

force equivalent on time

$$q\mathbf{E}_0$$

$$\begin{cases} E = -\{\mathbf{D}, A\} = E_0 + \mathbf{E} \\ B = +[\mathbf{D}, A] = B_0 + \mathbf{B} = 0 + \mathbf{B} = \mathbf{B} \quad B_0 = 0 \end{cases}$$

for any quaternion commuting and anticommutating with differential operator

$$\begin{aligned} [\mathbf{D}, Q] &= \nabla \times \mathbf{Q} \\ \{\mathbf{D}, Q\} &= \partial_0 Q - \nabla \cdot \mathbf{Q} + \nabla Q_0 \\ &= \partial_0 (Q_0 + \mathbf{Q}) - \nabla \cdot \mathbf{Q} + \nabla Q_0 \\ &= (\partial_0 Q_0 - \nabla \cdot \mathbf{Q}) + (\partial_0 \mathbf{Q} + \nabla Q_0) \\ &= \left(\frac{\partial Q_0}{\partial t} - \nabla \cdot \mathbf{Q} \right) + \left(\frac{\partial \mathbf{Q}}{\partial t} + \nabla Q_0 \right) \end{aligned}$$

$$\begin{cases} [\mathbf{D}, E] = \nabla \times \mathbf{E} = (0) + (\nabla \times \mathbf{E}) \\ \{\mathbf{D}, E\} = (\partial_0 E_0 - \nabla \cdot \mathbf{E}) + (\partial_0 \mathbf{E} + \nabla E_0) = \left(\frac{\partial E_0}{\partial t} - \nabla \cdot \mathbf{E} \right) + \left(\frac{\partial \mathbf{E}}{\partial t} + \nabla E_0 \right) \\ [\mathbf{D}, B] = \nabla \times \mathbf{B} = (0) + (\nabla \times \mathbf{B}) \\ \{\mathbf{D}, B\} = (\partial_0 B_0 - \nabla \cdot \mathbf{B}) + (\partial_0 \mathbf{B} + \nabla B_0) \stackrel{B_0=0}{=} -\nabla \cdot \mathbf{B} + \partial_0 \mathbf{B} = (-\nabla \cdot \mathbf{B}) + \left(\frac{\partial \mathbf{B}}{\partial t} \right) \end{cases}$$

by comparing (real & imaginary) / (scalar & vector) parts,

Maxwell equations without source terms

$$\begin{aligned} \begin{cases} [\mathbf{D}, B] = +\{\mathbf{D}, E\} \Leftrightarrow (0) + (\nabla \times \mathbf{B}) = \left(\frac{\partial E_0}{\partial t} - \nabla \cdot \mathbf{E} \right) + \left(\frac{\partial \mathbf{E}}{\partial t} + \nabla E_0 \right) \\ [\mathbf{D}, E] = -\{\mathbf{D}, B\} \Leftrightarrow (0) + (\nabla \times \mathbf{E}) = (-\nabla \cdot \mathbf{B}) + \left(\frac{\partial \mathbf{B}}{\partial t} \right) \end{cases} \\ \Leftrightarrow \begin{cases} \frac{\partial E_0}{\partial t} - \nabla \cdot \mathbf{E} = 0 & \Leftrightarrow \nabla \cdot \mathbf{E} = \frac{\partial E_0}{\partial t} \\ \frac{\partial \mathbf{E}}{\partial t} + \nabla E_0 = \nabla \times \mathbf{B} & \Leftrightarrow \nabla \times \mathbf{B} = \frac{\partial \mathbf{E}}{\partial t} + \nabla E_0 \text{ 動電生磁} \\ -\nabla \cdot \mathbf{B} = 0 & \Leftrightarrow \nabla \cdot \mathbf{B} = 0 \\ \frac{\partial \mathbf{B}}{\partial t} = \nabla \times \mathbf{E} & \Leftrightarrow \nabla \times \mathbf{E} = \frac{\partial \mathbf{B}}{\partial t} \text{ 動磁生電} \end{cases} \end{aligned}$$

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

$$J = J_0 + iJ_1 + jJ_2 + kJ_3 = J_0 + \mathbf{J} = \rho + \mathbf{J}$$

Maxwell equations with source terms

$$\begin{aligned} \begin{cases} [\mathbf{D}, B] = J + \{\mathbf{D}, E\} \Leftrightarrow (0) + (\nabla \times \mathbf{B}) = \left(\rho + \frac{\partial E_0}{\partial t} - \nabla \cdot \mathbf{E} \right) + \left(\mathbf{J} + \frac{\partial \mathbf{E}}{\partial t} + \nabla E_0 \right) \\ [\mathbf{D}, E] = 0 - \{\mathbf{D}, B\} \Leftrightarrow (0) + (\nabla \times \mathbf{E}) = (-\nabla \cdot \mathbf{B}) + \left(\frac{\partial \mathbf{B}}{\partial t} \right) \end{cases} \\ \Leftrightarrow \begin{cases} \rho + \frac{\partial E_0}{\partial t} - \nabla \cdot \mathbf{E} = 0 & \Leftrightarrow \nabla \cdot \mathbf{E} = \rho + \frac{\partial E_0}{\partial t} \\ \mathbf{J} + \frac{\partial \mathbf{E}}{\partial t} + \nabla E_0 = \nabla \times \mathbf{B} & \Leftrightarrow \nabla \times \mathbf{B} = \mathbf{J} + \frac{\partial \mathbf{E}}{\partial t} + \nabla E_0 \text{ 動電生磁} \\ -\nabla \cdot \mathbf{B} = 0 & \Leftrightarrow \nabla \cdot \mathbf{B} = 0 \\ \frac{\partial \mathbf{B}}{\partial t} = \nabla \times \mathbf{E} & \Leftrightarrow \nabla \times \mathbf{E} = \frac{\partial \mathbf{B}}{\partial t} \text{ 動磁生電} \end{cases} \end{aligned}$$

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

$$\mathbf{r} = (x, y) = \langle x, y \rangle = \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} r \cos \alpha \\ r \sin \alpha \end{pmatrix}$$

$$\mathbf{r}' = (x', y') = \langle x', y' \rangle = \begin{pmatrix} x' \\ y' \end{pmatrix} = \begin{pmatrix} r \cos(\alpha + \theta) \\ r \sin(\alpha + \theta) \end{pmatrix}$$

$$\begin{aligned} \mathbf{r}' &= \begin{pmatrix} r \cos(\alpha + \theta) \\ r \sin(\alpha + \theta) \end{pmatrix} = r \begin{pmatrix} \cos(\alpha + \theta) \\ \sin(\alpha + \theta) \end{pmatrix} \\ &= r \begin{pmatrix} \cos \alpha \cos \theta - \sin \alpha \sin \theta \\ \sin \alpha \cos \theta + \cos \alpha \sin \theta \end{pmatrix} = r \begin{pmatrix} \cos \theta \cos \alpha - \sin \theta \sin \alpha \\ \sin \theta \cos \alpha + \cos \theta \sin \alpha \end{pmatrix} \\ &= r \begin{pmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{pmatrix} \begin{pmatrix} \cos \alpha \\ \sin \alpha \end{pmatrix} = \begin{pmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{pmatrix} \begin{pmatrix} r \cos \alpha \\ r \sin \alpha \end{pmatrix} \\ &= R \mathbf{r}, \quad \begin{cases} R = \begin{pmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{pmatrix} = R(\theta) = R_\theta \\ \mathbf{r} = \begin{pmatrix} r \cos \alpha \\ r \sin \alpha \end{pmatrix}, \mathbf{r}' = \begin{pmatrix} r \cos(\alpha + \theta) \\ r \sin(\alpha + \theta) \end{pmatrix} \end{cases} \end{aligned}$$

orthonormal matrix

$$\mathbf{r}' = O \mathbf{r}$$

$$|\mathbf{r}'|^2 = |\mathbf{r}|^2$$

$$\mathbf{r}' \cdot \mathbf{r}' = \mathbf{r} \cdot \mathbf{r}$$

$$\mathbf{r}'^\top \mathbf{r}' = \mathbf{r}^\top \mathbf{r}$$

$$(O \mathbf{r})^\top (O \mathbf{r}) =$$

$$\mathbf{r}^\top O^\top O \mathbf{r} =$$

$$\mathbf{r}^\top O^\top O \mathbf{r} = \mathbf{r}^\top \mathbf{r}$$

$$O^\top O = 1 = I = I_2$$

$$\begin{aligned} R^\top R &= \begin{pmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{pmatrix}^\top \begin{pmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{pmatrix} \\ &= \begin{pmatrix} \cos \theta & \sin \theta \\ -\sin \theta & \cos \theta \end{pmatrix} \begin{pmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{pmatrix} \\ &= \begin{pmatrix} \cos^2 \theta + \sin^2 \theta & -\cos \theta \sin \theta + \sin \theta \cos \theta \\ -\sin \theta \cos \theta + \cos \theta \sin \theta & \sin^2 \theta + \cos^2 \theta \end{pmatrix} \\ &= \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = 1 = I \end{aligned}$$

$$R^\top R = 1 \Rightarrow R \in \{O | O^\top O = 1\}$$

https://en.wikipedia.org/wiki/Transformation_matrix#Affine_transformations

reflection matrix

$$\begin{cases} P_x = \begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix} & P_x^\top P_x = P_x^2 = \begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix}^2 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = 1 \Rightarrow P_x \in \{O | O^\top O = 1\} \\ P_y = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} & P_y^\top P_y = P_y^2 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}^2 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = 1 \Rightarrow P_y \in \{O | O^\top O = 1\} \end{cases}$$

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

$$\begin{aligned} O(2) &= \{1, R, P_x, P_y\} \\ &= \{I_2, R_\theta, P_x, P_y\} \subseteq \{O | O^\top O = 1\} \end{aligned}$$

$$\begin{aligned} 1 &= O^\top O \\ 1 &= \det 1 = \det I = \det(I_2) \\ &= \det(O^\top O) = (\det O^\top)(\det O) = (\det O)(\det O) = (\det O)^2 \\ 1 &= (\det O)^2 \\ \det O &= \pm 1 \end{aligned}$$

$$\det R = \det R_\theta = \begin{vmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{vmatrix} = \cos^2 \theta - (-\sin^2 \theta) = \cos^2 \theta + \sin^2 \theta = 1$$

$$\begin{cases} P_x = \begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix} & \det P_x = -1 \\ P_y = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} & \det P_y = -1 \end{cases}$$

special orthonormal group of degree 2

$$\begin{aligned} SO(2) &= \{1, R\} = \{I_2, R_\theta\} \subseteq \left\{ O \middle| \begin{cases} O^\top O = 1 \\ \det O = 1 \end{cases} \right\} \\ &\subset \{1, R, P_x, P_y\} = O(2) \subseteq \{O | O^\top O = 1\} \end{aligned}$$

34.1.4.1.2 complex

$$z = r(\cos \alpha + i \sin \alpha) = r e^{i\alpha}$$

$$z' = r(\cos(\alpha + \theta) + i \sin(\alpha + \theta)) = r e^{i(\alpha + \theta)}$$

$$\begin{aligned} z' &= z_\theta z \\ z_\theta &= \frac{z'}{z} = \frac{r' e^{i(\alpha+\theta)}}{r e^{i\alpha}} = \frac{r'}{r} e^{i\theta} = \frac{r'}{r} (\cos \theta + i \sin \theta) \\ z_\theta z &= \left[\frac{r'}{r} (\cos \theta + i \sin \theta) \right] [r(\cos \alpha + i \sin \alpha)] \\ &= r' [(\cos \theta \cos \alpha - \sin \theta \sin \alpha) + i(\sin \theta \cos \alpha + \cos \theta \sin \alpha)] \\ &= r' [\cos(\alpha + \theta) + i \sin(\alpha + \theta)] = z' \end{aligned}$$

$$\hat{z}_\theta = z_\theta \left(\frac{r'}{r} = 1 \right) = e^{i\theta} = \cos \theta + i \sin \theta$$

$$\hat{z}_\theta^* = \overline{\hat{z}_\theta} = e^{-i\theta} = \cos \theta - i \sin \theta$$

$$\hat{z}_\theta^* \hat{z}_\theta = e^{i\theta} e^{-i\theta} = e^{i\theta + (-i\theta)} = e^{i0} = e^0 = 1$$

unitary group of degree 1

$$U(1) = \{1, \hat{z}_\theta\} = \{e^{i0}, e^{i\theta}\}$$

34.1.4.1.3 $SO(2) \cong U(1)$ $\mathbb{C} \leftrightarrow M_{2 \times 2}(\mathbb{R}) = M_2(\mathbb{R})$ complex group representation^[37.4]

$$x + y\mathbf{i} \leftrightarrow \begin{pmatrix} x & -y \\ y & x \end{pmatrix} = xI + yJ$$

$$\begin{aligned} U(1) &= \{1, \hat{z}_\theta\} = \{e^{i0}, e^{i\theta}\} \\ &= \{\cos 0 + i \sin 0, \cos \theta + i \sin \theta\} \\ &\leftrightarrow \left\{ \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \cos 0 + \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \sin 0, \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \cos \theta + \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \sin \theta \right\} \\ &= \left\{ \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} 1 + \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} 0, \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \cos \theta + \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \sin \theta \right\} \\ &= \left\{ \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \begin{pmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{pmatrix} \right\} = \{I_2, R_\theta\} = \{1, R\} = SO(2) \end{aligned}$$

$$U(1) \cong SO(2) \Leftrightarrow SO(2) \cong U(1)$$

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

$$\begin{aligned} \mathbf{r}' &= \begin{pmatrix} r \cos(\alpha + \theta) \\ r \sin(\alpha + \theta) \\ z' \end{pmatrix} \stackrel{z' = z}{=} \begin{pmatrix} r \cos(\alpha + \theta) \\ r \sin(\alpha + \theta) \\ z \end{pmatrix} = \begin{pmatrix} R(\theta) & \\ & 1 \end{pmatrix} \begin{pmatrix} r \cos \alpha \\ r \sin \alpha \\ z \end{pmatrix} \\ &= \begin{pmatrix} \cos \theta & -\sin \theta & 0 \\ \sin \theta & \cos \theta & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} r \cos \alpha \\ r \sin \alpha \\ z \end{pmatrix} = R_z(\theta) \mathbf{r}, \begin{cases} R_z(\theta) = \begin{pmatrix} \cos \theta & -\sin \theta & 0 \\ \sin \theta & \cos \theta & 0 \\ 0 & 0 & 1 \end{pmatrix} \\ \mathbf{r} = \begin{pmatrix} r \cos \alpha \\ r \sin \alpha \\ z \end{pmatrix} \end{cases} \end{aligned}$$

$$\begin{aligned} \mathbf{r}' &= \begin{pmatrix} x' \\ r \cos(\alpha + \theta) \\ r \sin(\alpha + \theta) \end{pmatrix} \stackrel{x' = x}{=} \begin{pmatrix} x \\ r \cos(\alpha + \theta) \\ r \sin(\alpha + \theta) \end{pmatrix} = \begin{pmatrix} 1 & \\ & R(\theta) \end{pmatrix} \begin{pmatrix} x \\ r \cos \alpha \\ r \sin \alpha \end{pmatrix} \\ &= \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos \theta & -\sin \theta \\ 0 & \sin \theta & \cos \theta \end{pmatrix} \begin{pmatrix} x \\ r \cos \alpha \\ r \sin \alpha \end{pmatrix} = R_x(\theta) \mathbf{r}, \begin{cases} R_x(\theta) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos \theta & -\sin \theta \\ 0 & \sin \theta & \cos \theta \end{pmatrix} \\ \mathbf{r} = \begin{pmatrix} x \\ r \cos \alpha \\ r \sin \alpha \end{pmatrix} \end{cases} \end{aligned}$$

$$\begin{aligned} \mathbf{r}' &= \begin{pmatrix} r \sin(\alpha + \theta) \\ y' \\ r \cos(\alpha + \theta) \end{pmatrix} \stackrel{y' = y}{=} \begin{pmatrix} r \sin(\alpha + \theta) \\ y \\ r \cos(\alpha + \theta) \end{pmatrix} \\ &= \begin{pmatrix} \cos \theta & 0 & \sin \theta \\ 0 & 1 & 0 \\ -\sin \theta & 0 & \cos \theta \end{pmatrix} \begin{pmatrix} r \sin \alpha \\ y \\ r \cos \alpha \end{pmatrix} = R_y(\theta) \mathbf{r}, \begin{cases} R_y(\theta) = \begin{pmatrix} \cos \theta & 0 & \sin \theta \\ 0 & 1 & 0 \\ -\sin \theta & 0 & \cos \theta \end{pmatrix} \\ \mathbf{r} = \begin{pmatrix} r \sin \alpha \\ y \\ r \cos \alpha \end{pmatrix} \end{cases} \end{aligned}$$

34.1.4.2.1.2 Euler angle $z \rightarrow x \rightarrow z : \alpha \rightarrow \beta \rightarrow \gamma$

$$\begin{aligned}
& R_z(\gamma) R_x(\beta) R_z(\alpha) \\
&= \begin{pmatrix} \cos \gamma & -\sin \gamma & 0 \\ \sin \gamma & \cos \gamma & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos \beta & -\sin \beta \\ 0 & \sin \beta & \cos \beta \end{pmatrix} \begin{pmatrix} \cos \alpha & -\sin \alpha & 0 \\ \sin \alpha & \cos \alpha & 0 \\ 0 & 0 & 1 \end{pmatrix} \\
&= \begin{pmatrix} \cos \gamma & -\sin \gamma & 0 \\ \sin \gamma & \cos \gamma & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} \cos \alpha & -\sin \alpha & 0 \\ \cos \beta \sin \alpha & \cos \beta \cos \alpha & -\sin \beta \\ \sin \beta \sin \alpha & \sin \beta \cos \alpha & \cos \beta \end{pmatrix} \\
&= \begin{pmatrix} \cos \gamma \cos \alpha - \sin \gamma \cos \beta \sin \alpha & -\cos \gamma \sin \alpha - \sin \gamma \cos \beta \cos \alpha & \sin \gamma \sin \beta \\ \sin \gamma \cos \alpha + \cos \gamma \cos \beta \sin \alpha & -\sin \gamma \sin \alpha - \cos \gamma \cos \beta \cos \alpha & -\cos \gamma \sin \beta \\ \sin \beta \sin \alpha & \sin \beta \cos \alpha & \cos \beta \end{pmatrix}
\end{aligned}$$

$x \rightarrow y \rightarrow z : \alpha \rightarrow \beta \rightarrow \gamma$

$$\begin{aligned}
& R_z(\gamma) R_y(\beta) R_x(\alpha) \\
&= \begin{pmatrix} \cos \gamma & -\sin \gamma & 0 \\ \sin \gamma & \cos \gamma & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} \cos \beta & 0 & \sin \beta \\ 0 & 1 & 0 \\ -\sin \beta & 0 & \cos \beta \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos \alpha & -\sin \alpha \\ 0 & \sin \alpha & \cos \alpha \end{pmatrix} \\
&= \begin{pmatrix} \cos \gamma & -\sin \gamma & 0 \\ \sin \gamma & \cos \gamma & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} \cos \beta & \sin \beta \sin \alpha & \sin \beta \cos \alpha \\ 0 & \cos \alpha & -\sin \alpha \\ -\sin \beta & \cos \beta \sin \alpha & \cos \beta \cos \alpha \end{pmatrix} \\
&= \begin{pmatrix} \cos \gamma \cos \beta & \cos \gamma \sin \beta \sin \alpha - \sin \gamma \cos \alpha & \cos \gamma \sin \beta \cos \alpha + \sin \gamma \sin \alpha \\ \sin \gamma \cos \beta & \sin \gamma \sin \beta \sin \alpha + \cos \gamma \cos \alpha & \sin \gamma \sin \beta \cos \alpha - \cos \gamma \sin \alpha \\ -\sin \beta & \cos \beta \sin \alpha & \cos \beta \cos \alpha \end{pmatrix}
\end{aligned}$$

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

$$\begin{cases} \hat{x} = r \sin \theta \cos \phi & \stackrel{r=1}{=} \sin \theta \cos \phi \\ \hat{y} = r \sin \theta \sin \phi & \stackrel{r=1}{=} \sin \theta \sin \phi \\ \hat{z} = r \cos \theta & \stackrel{r=1}{=} \cos \theta \end{cases}$$

although I prefer θ and ϕ switched back to be compatible with 2D coordinate

$$\begin{cases} \hat{x} = r \sin \phi \cos \theta & \stackrel{r=1}{=} \sin \phi \cos \theta \\ \hat{y} = r \sin \phi \sin \theta & \stackrel{r=1}{=} \sin \phi \sin \theta \\ \hat{z} = r \cos \phi & \stackrel{r=1}{=} \cos \phi \end{cases}$$

or cos first in x, y -plane

$$\begin{cases} \hat{x} = r \cos \phi \cos \theta & \stackrel{r=1}{=} \cos \phi \cos \theta \\ \hat{y} = r \cos \phi \sin \theta & \stackrel{r=1}{=} \cos \phi \sin \theta \\ \hat{z} = r \sin \phi & \stackrel{r=1}{=} \sin \phi \end{cases}$$

still use the most convention

$$\hat{\mathbf{n}} = \begin{pmatrix} \hat{x} \\ \hat{y} \\ \hat{z} \end{pmatrix} = \begin{pmatrix} \sin \theta \cos \phi \\ \sin \theta \sin \phi \\ \cos \theta \end{pmatrix}$$

$$\begin{cases} \hat{\mathbf{n}} = \begin{pmatrix} \sin \theta \cos \phi \\ \sin \theta \sin \phi \\ \cos \theta \end{pmatrix} \\ \hat{\mathbf{u}} = \begin{pmatrix} \cos \theta \cos \phi \\ \cos \theta \sin \phi \\ -\sin \theta \end{pmatrix} \Leftarrow \cos \theta \sin \theta - \cos \theta \sin \theta = 0 \Rightarrow \hat{\mathbf{u}} \cdot \hat{\mathbf{n}} = 0 \Leftrightarrow \hat{\mathbf{u}} \perp \hat{\mathbf{n}} \\ \hat{\mathbf{v}} = \hat{\mathbf{n}} \times \hat{\mathbf{u}} = \begin{vmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ \sin \theta \cos \phi & \sin \theta \sin \phi & \cos \theta \\ \cos \theta \cos \phi & \cos \theta \sin \phi & -\sin \theta \end{vmatrix} = \begin{pmatrix} -\sin \phi \\ \cos \phi \\ 0 \end{pmatrix} \Rightarrow \begin{cases} \hat{\mathbf{v}} \perp \hat{\mathbf{n}} \\ \hat{\mathbf{v}} \perp \hat{\mathbf{u}} \end{cases} \end{cases}$$

$S = \{\hat{\mathbf{n}}, \hat{\mathbf{u}}, \hat{\mathbf{v}}\} = \{\hat{\mathbf{u}}, \hat{\mathbf{v}}, \hat{\mathbf{n}}\}$ is a basis of the spherical coordinate

$$[\mathbf{V}]_S = \begin{pmatrix} u \\ v \\ n \end{pmatrix}$$

$$\begin{aligned} \mathbf{V} &= (\hat{\mathbf{u}} \quad \hat{\mathbf{v}} \quad \hat{\mathbf{n}}) \begin{pmatrix} u \\ v \\ n \end{pmatrix} = u\hat{\mathbf{u}} + v\hat{\mathbf{v}} + n\hat{\mathbf{n}} \\ &= \begin{pmatrix} \cos \theta \cos \phi & -\sin \phi & \sin \theta \cos \phi \\ \cos \theta \sin \phi & \cos \phi & \sin \theta \sin \phi \\ -\sin \theta & 0 & \cos \theta \end{pmatrix} \begin{pmatrix} u \\ v \\ n \end{pmatrix} = S[\mathbf{V}]_S \\ \begin{pmatrix} u \\ v \\ n \end{pmatrix} &= \begin{pmatrix} \cos \theta \cos \phi & -\sin \phi & \sin \theta \cos \phi \\ \cos \theta \sin \phi & \cos \phi & \sin \theta \sin \phi \\ -\sin \theta & 0 & \cos \theta \end{pmatrix}^{-1} \mathbf{V} \\ [\mathbf{V}]_S &= S^{-1}\mathbf{V} \end{aligned}$$

$$\begin{aligned} S^{-1} &= \begin{pmatrix} \cos \theta \cos \phi & -\sin \phi & \sin \theta \cos \phi \\ \cos \theta \sin \phi & \cos \phi & \sin \theta \sin \phi \\ -\sin \theta & 0 & \cos \theta \end{pmatrix}^{-1}, S \in \{O | O^T O = 1\} \\ &= \begin{pmatrix} \cos \theta \cos \phi & -\sin \phi & \sin \theta \cos \phi \\ \cos \theta \sin \phi & \cos \phi & \sin \theta \sin \phi \\ -\sin \theta & 0 & \cos \theta \end{pmatrix}^T \\ &= \begin{pmatrix} \cos \theta \cos \phi & \cos \theta \sin \phi & -\sin \theta \\ -\sin \phi & \cos \phi & 0 \\ \sin \theta \cos \phi & \sin \theta \sin \phi & \cos \theta \end{pmatrix} = \begin{pmatrix} \hat{\mathbf{u}}^T \\ \hat{\mathbf{v}}^T \\ \hat{\mathbf{n}}^T \end{pmatrix} \\ S^{-1}S &= \begin{pmatrix} \cos \theta \cos \phi & \cos \theta \sin \phi & -\sin \theta \\ -\sin \phi & \cos \phi & 0 \\ \sin \theta \cos \phi & \sin \theta \sin \phi & \cos \theta \end{pmatrix} \begin{pmatrix} \cos \theta \cos \phi & -\sin \phi & \sin \theta \cos \phi \\ \cos \theta \sin \phi & \cos \phi & \sin \theta \sin \phi \\ -\sin \theta & 0 & \cos \theta \end{pmatrix} \\ &= \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} = 1 \end{aligned}$$

$\hat{\mathbf{n}}$ as z direction

$$\begin{aligned} [\mathbf{V}]_S &= S^{-1}\mathbf{V} \\ [\mathbf{V}']_S &= R_z(\gamma)[\mathbf{V}]_S = R_z(\gamma)S^{-1}\mathbf{V} \\ \mathbf{V}' &= S[\mathbf{V}']_S = SR_z(\gamma)[\mathbf{V}]_S = SR_z(\gamma)S^{-1}\mathbf{V} \\ \mathbf{V}' &= SR_z(\gamma)S^{-1}\mathbf{V} \end{aligned}$$

$$\begin{aligned}
& SR_z(\gamma) S^{-1} \\
&= \begin{pmatrix} \cos \theta \cos \phi & -\sin \phi & \sin \theta \cos \phi \\ \cos \theta \sin \phi & \cos \phi & \sin \theta \sin \phi \\ -\sin \theta & 0 & \cos \theta \end{pmatrix} \begin{pmatrix} \cos \gamma & -\sin \gamma & 0 \\ \sin \gamma & \cos \gamma & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} \cos \theta \cos \phi & \cos \theta \sin \phi & -\sin \theta \\ -\sin \phi & \cos \phi & 0 \\ \sin \theta \cos \phi & \sin \theta \sin \phi & \cos \theta \end{pmatrix} \\
&= \begin{pmatrix} c_1 c_2 & -s_2 & s_1 c_2 \\ c_1 s_2 & c_2 & s_1 s_2 \\ -s_1 & 0 & c_1 \end{pmatrix} \begin{pmatrix} c_3 & -s_3 & 0 \\ s_3 & c_3 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} c_1 c_2 & c_1 s_2 & -s_1 \\ -s_2 & c_2 & 0 \\ s_1 c_2 & s_1 s_2 & c_1 \end{pmatrix}, \begin{cases} c_1 = \cos \theta & s_1 = \sin \theta \\ c_2 = \cos \phi & s_2 = \sin \phi \\ c_3 = \cos \gamma & s_3 = \sin \gamma \end{cases} \\
&= \begin{pmatrix} c_1 c_2 & -s_2 & s_1 c_2 \\ c_1 s_2 & c_2 & s_1 s_2 \\ -s_1 & 0 & c_1 \end{pmatrix} \begin{pmatrix} c_3 c_1 c_2 + s_3 s_2 & c_3 c_1 s_2 - c_2 s_3 & -c_3 s_1 \\ c_1 c_2 s_3 - c_3 s_2 & c_3 c_2 + c_1 s_3 s_2 & -s_3 s_1 \\ c_2 s_1 & s_2 s_1 & c_1 \end{pmatrix} \\
&= \begin{pmatrix} c_1^2 c_2^2 c_3 + c_3 s_2^2 + c_2^2 s_1^2 & c_1 s_2 (c_1 c_2 c_3 - s_2 s_3) + c_2 (-c_1 c_2 s_3 - c_3 s_2) + c_2 s_2 s_1^2 & c_1 c_2 s_1 - s_1 (c_1 c_2 c_3 - s_2 s_3) \\ c_1 c_2 (c_1 c_3 s_2 + c_2 s_3) - s_2 (c_2 c_3 - c_1 s_2 s_3) + c_2 s_2 s_1^2 & c_1^2 c_3 s_2^2 + s_2^2 s_1^2 + c_2^2 c_3 & c_1 s_2 s_1 - s_1 (c_1 c_3 s_2 + c_2 s_1) \\ -c_1 c_2 c_3 s_1 - s_2 s_1 s_3 + c_1 c_2 s_1 & -c_1 c_3 s_2 s_1 + c_2 s_1 s_3 + c_1 s_2 s_1 & c_3 s_1^2 + c_1^2 \end{pmatrix}
\end{aligned}$$

34.1.4.2.2 quaternion <https://math.stackexchange.com/questions/328117/how-does-one-derive-this-rotation-quaternion-formula>

$$\begin{aligned}
q &= a + bi + cj + dk = a + ib + jc + kd, \begin{cases} q \in \mathbb{H} \\ i^2 = j^2 = k^2 = -1 = ijk \Rightarrow ij = k \\ a, b, c, d \in \mathbb{R} \end{cases} \Leftrightarrow \langle a, b, c, d \rangle \in \mathbb{R}^4 \\
w &= t + xi + yj + zk = t + ix + jy + kz, \begin{cases} w \in \mathbb{H} \\ i^2 = j^2 = k^2 = -1 = ijk \Rightarrow ij = k \\ t, x, y, z \in \mathbb{R} \end{cases} \Leftrightarrow \langle t, x, y, z \rangle \in \mathbb{R}^4 \\
&= t + (i \ j \ k) \begin{pmatrix} x \\ y \\ z \end{pmatrix} = x_0 + (i \ j \ k) \begin{pmatrix} x \\ y \\ z \end{pmatrix} = x_0 + (e_1 \ e_2 \ e_3) \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = x_0 + \mathbf{x}, \begin{cases} e_1 = i = i \\ e_2 = j = j \\ e_3 = k = k \end{cases}
\end{aligned}$$

$$q_1 q_2 = (x_{10} x_{20} - \mathbf{x}_1 \cdot \mathbf{x}_2) + (x_{10} \mathbf{x}_2 + x_{20} \mathbf{x}_1 + \mathbf{x}_1 \times \mathbf{x}_2)$$

$$QP = (Q_0 P_0 - \mathbf{Q} \cdot \mathbf{P}) + (Q_0 \mathbf{P} + P_0 \mathbf{Q} + \mathbf{Q} \times \mathbf{P})$$

$$q = q_0 + \mathbf{q} = q_0 + q_1 \mathbf{i} + q_2 \mathbf{j} + q_3 \mathbf{k} = q_0 + q_1 \mathbf{i} + q_2 \mathbf{j} + q_3 \mathbf{k}$$

$$\begin{aligned}
q^* &= \bar{q} = \overline{q_0 + \mathbf{q}} = q_0 - \mathbf{q} \\
&= \overline{q_0 + q_1 \mathbf{i} + q_2 \mathbf{j} + q_3 \mathbf{k}} = q_0 - q_1 \mathbf{i} - q_2 \mathbf{j} - q_3 \mathbf{k} \\
&= q_0 - q_1 \mathbf{i} - q_2 \mathbf{j} - q_3 \mathbf{k}
\end{aligned}$$

$$v = v_0 + \mathbf{v} = v_0 + v_1 \mathbf{i} + v_2 \mathbf{j} + v_3 \mathbf{k} = v_0 + v_1 \mathbf{i} + v_2 \mathbf{j} + v_3 \mathbf{k}$$

$$\begin{aligned}
qv &= (q_0 v_0 - \mathbf{q} \cdot \mathbf{v}) + (q_0 \mathbf{v} + v_0 \mathbf{q} + \mathbf{q} \times \mathbf{v}) \\
&\stackrel{v_0=0}{=} (q_0 0 - \mathbf{q} \cdot \mathbf{v}) + (q_0 \mathbf{v} + 0 \mathbf{q} + \mathbf{q} \times \mathbf{v}) \\
&= (-\mathbf{q} \cdot \mathbf{v}) + (q_0 \mathbf{v} + \mathbf{q} \times \mathbf{v}) \\
q\mathbf{v} &= (-\mathbf{q} \cdot \mathbf{v}) + (q_0 \mathbf{v} + \mathbf{q} \times \mathbf{v})
\end{aligned}$$

$$\begin{aligned}
v\bar{q} &= (v_0 q_0 - \mathbf{v} \cdot \bar{\mathbf{q}}) + (v_0 \bar{\mathbf{q}} + q_0 \mathbf{v} + \mathbf{v} \times \bar{\mathbf{q}}) \\
&\stackrel{v_0=0}{=} (0 q_0 - \mathbf{v} \cdot \bar{\mathbf{q}}) + (0 \bar{\mathbf{q}} + q_0 \mathbf{v} + \mathbf{v} \times \bar{\mathbf{q}}) \\
&= (-\mathbf{v} \cdot \bar{\mathbf{q}}) + (q_0 \mathbf{v} + \mathbf{v} \times \bar{\mathbf{q}}) \\
&= (-\mathbf{v} \cdot (-\mathbf{q})) + (q_0 \mathbf{v} + \mathbf{v} \times (-\mathbf{q})) \\
&= (\mathbf{v} \cdot \mathbf{q}) + (q_0 \mathbf{v} - \mathbf{v} \times \mathbf{q}) \\
\mathbf{v}\bar{q} &= (\mathbf{v} \cdot \mathbf{q}) + (q_0 \mathbf{v} - \mathbf{v} \times \mathbf{q})
\end{aligned}$$

<https://math.stackexchange.com/questions/41574/can-eulers-identity-be-extended-to-quaternions>

$$\begin{aligned}
q &= a + bi + cj + dk = a + ib + jc + kd \\
&= a + r \frac{ib + jc + kd}{r}, r^2 = b^2 + c^2 + d^2 \\
&= a + \theta \frac{ib + jc + kd}{\theta}, \theta^2 = b^2 + c^2 + d^2 \\
\left(\frac{ib + jc + kd}{r} \right)^2 &= \frac{-b^2 - c^2 - d^2 + bc(ij + ji) + cd(jk + kj) + db(ki + ik)}{r^2} \\
&= \frac{-b^2 - c^2 - d^2 + bc(k - k) + cd(i - i) + db(j - j)}{r^2} \\
&= \frac{-b^2 - c^2 - d^2 + bc0 + cd0 + db0}{r^2} = \frac{-b^2 - c^2 - d^2}{r^2} \\
&= \frac{-b^2 - c^2 - d^2}{b^2 + c^2 + d^2} = -1 \\
\frac{ib + jc + kd}{r} &= \pm \sqrt{-1}
\end{aligned}$$

$$\begin{aligned}
e^q &= e^{a+ib+jc+kd} \\
&= e^{a+r\sqrt{-1}} = e^a e^{r\sqrt{-1}} = e^{a+\theta\sqrt{-1}} = e^a e^{\theta\sqrt{-1}} \\
&= e^a (\cos r + \sqrt{-1} \sin r) = e^a (\cos \theta + \sqrt{-1} \sin \theta) \\
&= e^a \left(\cos r + \frac{ib + jc + kd}{r} \sin r \right) = e^a \left[\cos r + (ib + jc + kd) \frac{\sin r}{r} \right] \\
&= e^a \left(\cos \theta + \frac{ib + jc + kd}{\theta} \sin \theta \right) = e^a \left[\cos \theta + (ib + jc + kd) \frac{\sin \theta}{\theta} \right]
\end{aligned}$$

$$\begin{aligned}
q &= a + bi + cj + dk = a + ib + jc + kd \\
&= a + r \frac{ib + jc + kd}{r}, r^2 = b^2 + c^2 + d^2 \\
&= \sqrt{a^2 + r^2} \left(\frac{a}{\sqrt{a^2 + r^2}} + \frac{ib + jc + kd}{r} \frac{r}{\sqrt{a^2 + r^2}} \right) \\
&= \rho (\cos \phi + \sqrt{-1} \sin \phi), \begin{cases} \rho = \sqrt{a^2 + r^2} \\ \tan \phi = \frac{r}{a} \Leftrightarrow \phi = \arctan \frac{r}{a} \end{cases} \\
&= \rho e^{\phi\sqrt{-1}} \\
q &= \rho e^{\phi\sqrt{-1}}, \begin{cases} q = a + ib + jc + kd \\ \rho = \sqrt{a^2 + r^2} = \sqrt{a^2 + b^2 + c^2 + d^2} \\ \tan \phi = \frac{r}{a} \Leftrightarrow \phi = \arctan \frac{r}{a} = \arctan \frac{\pm\sqrt{b^2 + c^2 + d^2}}{a} \end{cases}
\end{aligned}$$

$$\begin{aligned}
\rho e^{-\phi\sqrt{-1}} &= \rho [\cos(-\phi) + \sqrt{-1} \sin(-\phi)] \\
&= \rho [\cos \phi - \sqrt{-1} \sin \phi] \\
&= \sqrt{a^2 + r^2} \left[\frac{a}{\sqrt{a^2 + r^2}} - \frac{ib + jc + kd}{r} \frac{r}{\sqrt{a^2 + r^2}} \right] \\
&= a - (ib + jc + kd) = a - ib - jc - kd = \bar{q} = q^*
\end{aligned}$$

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

$$\mathbf{v} \rightarrow \mathbf{v}'$$

$$\mathbf{v} \xrightarrow{\text{rotate about } \mathbf{u}} \mathbf{v}'$$

$$\mathbf{v} \xrightarrow{\text{rotate about } \mathbf{n}} \mathbf{v}'$$

$$\begin{cases} \mathbf{v} = \mathbf{v}_{\parallel n} + \mathbf{v}_{\perp n} = \mathbf{v}_{\parallel} + \mathbf{v}_{\perp} \Rightarrow \mathbf{v}_{\perp} = \mathbf{v} - \mathbf{v}_{\parallel} \\ \mathbf{v}' = \mathbf{v}'_{\parallel n} + \mathbf{v}'_{\perp n} = \mathbf{v}'_{\parallel} + \mathbf{v}'_{\perp} \Rightarrow \mathbf{v}'_{\perp} = \mathbf{v}' - \mathbf{v}'_{\parallel} \end{cases}$$

$$\begin{aligned} \mathbf{v}_{\parallel n} &= \mathbf{v}_{\parallel} = \text{proj}_n \mathbf{v} = \frac{\mathbf{v} \cdot \mathbf{n}}{\|\mathbf{n}\|} \hat{\mathbf{n}} \\ &= \frac{\mathbf{v} \cdot \mathbf{n}}{\|\mathbf{n}\|} \frac{\mathbf{n}}{\|\mathbf{n}\|} = \frac{\mathbf{v} \cdot \mathbf{n}}{\|\mathbf{n}\|^2} \mathbf{n} \end{aligned}$$

$$\begin{aligned} \mathbf{v}_{\parallel n} &= \mathbf{v}_{\parallel} = \text{proj}_n \mathbf{v} = \frac{\mathbf{v} \cdot \hat{\mathbf{n}}}{\|\hat{\mathbf{n}}\|} \hat{\mathbf{n}} \stackrel{\|\hat{\mathbf{n}}\|=1}{=} (\mathbf{v} \cdot \hat{\mathbf{n}}) \hat{\mathbf{n}} = \mathbf{v} \cdot \hat{\mathbf{n}} \hat{\mathbf{n}} \\ &= \frac{\mathbf{v} \cdot \hat{\mathbf{n}}}{\|\hat{\mathbf{n}}\|} \frac{\hat{\mathbf{n}}}{\|\hat{\mathbf{n}}\|} = \frac{\mathbf{v} \cdot \hat{\mathbf{n}}}{\|\hat{\mathbf{n}}\|^2} \hat{\mathbf{n}} \stackrel{\|\hat{\mathbf{n}}\|=1}{=} (\mathbf{v} \cdot \hat{\mathbf{n}}) \hat{\mathbf{n}} = \mathbf{v} \cdot \hat{\mathbf{n}} \hat{\mathbf{n}} \end{aligned}$$

$$\mathbf{n} = \hat{\mathbf{n}}$$

$$\begin{aligned} \mathbf{v}_{\parallel n} &= \mathbf{v}_{\parallel} = \text{proj}_n \mathbf{v} = \frac{\mathbf{v} \cdot \mathbf{n}}{\|\mathbf{n}\|} \hat{\mathbf{n}} \stackrel{\|\mathbf{n}\|=1}{=} (\mathbf{v} \cdot \hat{\mathbf{n}}) \hat{\mathbf{n}} = \mathbf{v} \cdot \hat{\mathbf{n}} \hat{\mathbf{n}} \\ &= \frac{\mathbf{v} \cdot \mathbf{n}}{\|\mathbf{n}\|} \frac{\mathbf{n}}{\|\mathbf{n}\|} = \frac{\mathbf{v} \cdot \mathbf{n}}{\|\mathbf{n}\|^2} \mathbf{n} \stackrel{\|\mathbf{n}\|=1}{=} (\mathbf{v} \cdot \mathbf{n}) \mathbf{n} = \mathbf{v} \cdot \mathbf{n} \mathbf{n} \end{aligned}$$

$$\mathbf{v}_{\parallel} = (\mathbf{v} \cdot \mathbf{n}) \mathbf{n}$$

$$\mathbf{v}_{\perp} = \mathbf{v} - \mathbf{v}_{\parallel} = \mathbf{v} - (\mathbf{v} \cdot \mathbf{n}) \mathbf{n}$$

$$\mathbf{v}_{\parallel} + \mathbf{v}_{\perp} = \mathbf{v} \xrightarrow{\text{rotate about } \mathbf{n}} \mathbf{v}' = \mathbf{v}'_{\parallel} + \mathbf{v}'_{\perp}$$

$$\mathbf{v}_{\parallel} \xrightarrow{\text{rotate about } \mathbf{n}} \mathbf{v}'_{\parallel}$$

$$\mathbf{v}_{\parallel} = \mathbf{v}_{\parallel n} = \mathbf{v}'_{\parallel n} = \mathbf{v}'_{\parallel}$$

$$\mathbf{v}_{\parallel} = \mathbf{v}'_{\parallel}$$

$$\mathbf{v}'_{\parallel} = \mathbf{v}_{\parallel}$$

$$\begin{cases} \mathbf{u} = \mathbf{n} \times \mathbf{v}_\perp \\ \|\mathbf{u}\| = \|\mathbf{n} \times \mathbf{v}_\perp\| = \|\mathbf{n}\| \|\mathbf{v}_\perp\| \sin \frac{\pi}{2} = \|\mathbf{v}_\perp\| \end{cases} \quad \begin{cases} \mathbf{u} \perp \mathbf{n} \\ \mathbf{u} \perp \mathbf{v}_\perp \\ \mathbf{n} \perp \mathbf{v}_\perp \\ \|\mathbf{n}\| = 1 \end{cases}$$

$$\begin{aligned} \mathbf{v}'_\perp &= \mathbf{v}'_{\parallel \mathbf{v}_\perp} + \mathbf{v}'_{\parallel \mathbf{u}} \\ &= (\cos \theta) \mathbf{v}_\perp + (\sin \theta) \mathbf{u}, \theta = \angle \mathbf{v}'_\perp \mathbf{v}_\perp \\ &= \cos \theta \mathbf{v}_\perp + \sin \theta \mathbf{u} \\ &= \cos \theta \mathbf{v}_\perp + \sin \theta (\mathbf{n} \times \mathbf{v}_\perp) \\ \mathbf{v}'_\perp &= \cos \theta \mathbf{v}_\perp + \sin \theta (\mathbf{n} \times \mathbf{v}_\perp) \end{aligned}$$

$$\mathbf{v}'_\perp = \cos \theta \mathbf{v}_\perp + \sin \theta (\mathbf{n} \times \mathbf{v}_\perp), \theta = \angle \mathbf{v}'_\perp \mathbf{v}_\perp$$

$$\mathbf{v}'_\perp = \cos \theta \mathbf{v}_\perp + \sin \theta (\mathbf{n} \times \mathbf{v}_\perp)$$

$$\begin{aligned} \mathbf{v}' &= \mathbf{v}'_\parallel + \mathbf{v}'_\perp, \begin{cases} \mathbf{v}'_\parallel = \mathbf{v}_\parallel \\ \mathbf{v}'_\perp = \cos \theta \mathbf{v}_\perp + \sin \theta (\mathbf{n} \times \mathbf{v}_\perp), \theta = \angle \mathbf{v}'_\perp \mathbf{v}_\perp \end{cases} \\ &= \mathbf{v}_\parallel + \cos \theta \mathbf{v}_\perp + \sin \theta (\mathbf{n} \times \mathbf{v}_\perp), \theta = \angle \mathbf{v}'_\perp \mathbf{v}_\perp \\ &= \mathbf{v}_\parallel + \cos \theta \mathbf{v}_\perp + \sin \theta (\mathbf{n} \times \mathbf{v}_\perp), \begin{cases} \mathbf{v}_\parallel = (\mathbf{v} \cdot \mathbf{n}) \mathbf{n} \\ \mathbf{v}_\perp = \mathbf{v} - \mathbf{v}_\parallel = \mathbf{v} - (\mathbf{v} \cdot \mathbf{n}) \mathbf{n} \end{cases} \\ &= (\mathbf{v} \cdot \mathbf{n}) \mathbf{n} + \cos \theta [\mathbf{v} - (\mathbf{v} \cdot \mathbf{n}) \mathbf{n}] + \sin \theta [\mathbf{n} \times (\mathbf{v} - \mathbf{v}_\parallel)] \\ &= (\mathbf{v} \cdot \mathbf{n}) \mathbf{n} + \cos \theta \mathbf{v} - \cos \theta (\mathbf{v} \cdot \mathbf{n}) \mathbf{n} + \sin \theta [\mathbf{n} \times \mathbf{v} - \mathbf{n} \times \mathbf{v}_\parallel] \\ &= \cos \theta \mathbf{v} + (1 - \cos \theta) (\mathbf{v} \cdot \mathbf{n}) \mathbf{n} + \sin \theta [\mathbf{n} \times \mathbf{v} - \mathbf{0}] \\ &= \cos \theta \mathbf{v} + (1 - \cos \theta) (\mathbf{v} \cdot \mathbf{n}) \mathbf{n} + \sin \theta \mathbf{n} \times \mathbf{v} \end{aligned}$$

$$\mathbf{v}' = \cos \theta \mathbf{v} + (1 - \cos \theta) (\mathbf{v} \cdot \mathbf{n}) \mathbf{n} + \sin \theta \mathbf{n} \times \mathbf{v}$$

$$\begin{aligned} \mathbf{v}' &= \cos \theta \mathbf{v} + (1 - \cos \theta) (\mathbf{v} \cdot \mathbf{n}) \mathbf{n} + \sin \theta \mathbf{n} \times \mathbf{v} \\ &= \cos \theta \mathbf{v} + \left[1 - \left(1 - 2 \sin^2 \frac{\theta}{2} \right) \right] (\mathbf{v} \cdot \mathbf{n}) \mathbf{n} + \left[2 \sin \frac{\theta}{2} \cos \frac{\theta}{2} \right] \mathbf{n} \times \mathbf{v} \\ &= \cos \theta \mathbf{v} + \left[2 \sin^2 \frac{\theta}{2} \right] (\mathbf{v} \cdot \mathbf{n}) \mathbf{n} + \left[2 \sin \frac{\theta}{2} \cos \frac{\theta}{2} \right] \mathbf{n} \times \mathbf{v} \\ &= \cos \theta \mathbf{v} + 2 \sin \frac{\theta}{2} \left[\left(\sin \frac{\theta}{2} \right) (\mathbf{v} \cdot \mathbf{n}) \mathbf{n} + \left(\cos \frac{\theta}{2} \right) \mathbf{n} \times \mathbf{v} \right] \\ &= \cos \theta \mathbf{v} + 2 \sin \frac{\theta}{2} \left[\left(\cos \frac{\theta}{2} \right) \mathbf{n} \times \mathbf{v} + \left(\sin \frac{\theta}{2} \right) (\mathbf{v} \cdot \mathbf{n}) \mathbf{n} \right] \\ &= \left[2 \cos^2 \frac{\theta}{2} - 1 \right] \mathbf{v} + 2 \sin \frac{\theta}{2} \left[\left(\cos \frac{\theta}{2} \right) \mathbf{n} \times \mathbf{v} + \left(\sin \frac{\theta}{2} \right) (\mathbf{v} \cdot \mathbf{n}) \mathbf{n} \right] \end{aligned}$$

34.2.2 quaternion operation

$$q = a + bi + cj + dk$$

$$\varrho = \alpha + \beta i + \gamma j + \delta k$$

$$\begin{aligned}
\begin{pmatrix} a \\ b \\ c \\ d \end{pmatrix} \begin{pmatrix} \alpha \\ \beta \\ \gamma \\ \delta \end{pmatrix} &\leftarrow q\varrho = (a + bi + cj + dk)(\alpha + \beta i + \gamma j + \delta k) \\
&= a(\alpha + \beta i + \gamma j + \delta k) \\
&+ bi(\alpha + \beta i + \gamma j + \delta k) \\
&+ cj(\alpha + \beta i + \gamma j + \delta k) \\
&+ dk(\alpha + \beta i + \gamma j + \delta k) \\
&= a\alpha + a\beta i + a\gamma j + a\delta k \\
&+ b\alpha i - b\beta + b\gamma k - b\delta j \\
&+ c\alpha j - c\beta k - c\gamma + c\delta i \\
&+ d\alpha k + d\beta j - d\gamma i - d\delta \\
&= (a\alpha - b\beta - c\gamma - d\delta) \\
&+ (b\alpha + a\beta - d\gamma + c\delta)i \\
&+ (c\alpha + d\beta + a\gamma - b\delta)j \\
&+ (d\alpha - c\beta + b\gamma + a\delta)k \\
\begin{pmatrix} a & -b & -c & -d \\ b & a & -d & c \\ c & d & a & -b \\ d & -c & b & a \end{pmatrix} \begin{pmatrix} \alpha \\ \beta \\ \gamma \\ \delta \end{pmatrix} &\leftarrow \\
&= (\alpha a - \beta b - \gamma c - \delta d) \\
&+ (\beta a + \alpha b + \delta c - \gamma d)i \\
&+ (\gamma a - \delta b + \alpha c + \beta d)j \\
&+ (\delta a + \gamma b - \beta c + \alpha d)k \\
\begin{pmatrix} \alpha & -\beta & -\gamma & -\delta \\ \beta & \alpha & \delta & -\gamma \\ \gamma & -\delta & \alpha & \beta \\ \delta & \gamma & -\beta & \alpha \end{pmatrix} \begin{pmatrix} a \\ b \\ c \\ d \end{pmatrix} &\leftarrow \\
qv &\leftrightarrow \begin{pmatrix} a & -b & -c & -d \\ b & a & -d & c \\ c & d & a & -b \\ d & -c & b & a \end{pmatrix} \begin{pmatrix} v_0 \\ v_1 \\ v_2 \\ v_3 \end{pmatrix} = Q_i v
\end{aligned}$$

concept like quaternion group representation^[37.5]

$$\begin{aligned}
L(q) = Q_i = &\begin{pmatrix} a & -b & -c & -d \\ b & a & -d & c \\ c & d & a & -b \\ d & -c & b & a \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} a + \begin{pmatrix} 0 & -1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & -1 \\ 0 & 0 & 1 & 0 \end{pmatrix} b \\
&+ \begin{pmatrix} 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \end{pmatrix} c + \begin{pmatrix} 0 & 0 & 0 & -1 \\ 0 & 0 & -1 & 0 \\ 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{pmatrix} d \\
&\leftrightarrow 1a + ib + jc + kd
\end{aligned}$$

$$\begin{aligned}
ij &\leftrightarrow \begin{pmatrix} 0 & -1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & -1 \\ 0 & 0 & 1 & 0 \end{pmatrix} \begin{pmatrix} 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \end{pmatrix} = \begin{pmatrix} 0 & 0 & 0 & -1 \\ 0 & 0 & -1 & 0 \\ 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{pmatrix} \leftrightarrow k \\
ji &\leftrightarrow \begin{pmatrix} 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \end{pmatrix} \begin{pmatrix} 0 & -1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & -1 \\ 0 & 0 & 1 & 0 \end{pmatrix} = -\begin{pmatrix} 0 & 0 & 0 & -1 \\ 0 & 0 & -1 & 0 \\ 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{pmatrix} \leftrightarrow -k
\end{aligned}$$

$$ki \leftrightarrow \begin{pmatrix} 0 & 0 & 0 & -1 \\ 0 & 0 & -1 & 0 \\ 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} 0 & -1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & -1 \\ 0 & 0 & 1 & 0 \end{pmatrix} = \begin{pmatrix} 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \end{pmatrix} \leftrightarrow j$$

$$Q_i = \begin{pmatrix} a & -b & -c & -d \\ b & a & -d & c \\ c & d & a & -b \\ d & -c & b & a \end{pmatrix} = \begin{pmatrix} a \\ b \\ c \\ d \end{pmatrix} \begin{pmatrix} -(b & c & d) \\ a & -d & c \\ d & a & -b \\ -c & b & a \end{pmatrix} = \begin{pmatrix} q_0 & -\mathbf{q}^\top \\ \mathbf{q} & Q_i \end{pmatrix}, \begin{cases} a = q_0 \\ b \\ c \\ d \end{cases} \mathbf{q} = \begin{pmatrix} b \\ c \\ d \\ a \end{pmatrix}, Q_i = \begin{pmatrix} a & -d & c \\ d & a & -b \\ -c & b & a \end{pmatrix}$$

Grassmann product

$$qv \leftrightarrow \begin{pmatrix} a & -b & -c & -d \\ b & a & -d & c \\ c & d & a & -b \\ d & -c & b & a \end{pmatrix} \begin{pmatrix} v_0 \\ v_1 \\ v_2 \\ v_3 \end{pmatrix} = Q_i v = \begin{pmatrix} q_0 & -\mathbf{q}^\top \\ \mathbf{q} & Q_i \end{pmatrix} \begin{pmatrix} v_0 \\ v_1 \\ v_2 \\ v_3 \end{pmatrix} = \begin{pmatrix} q_0 v_0 - \mathbf{q}^\top v \\ \mathbf{q} v_0 + Q_i v \end{pmatrix} = \begin{pmatrix} q_0 v_0 - \mathbf{q} \cdot \mathbf{v} \\ \mathbf{q} v_0 + q_0 \mathbf{v} + \mathbf{q} \times \mathbf{v} \end{pmatrix}$$

$$vq \leftrightarrow \begin{pmatrix} a & -b & -c & -d \\ b & a & d & -c \\ c & -d & a & b \\ d & c & -b & a \end{pmatrix} \begin{pmatrix} v_0 \\ v_1 \\ v_2 \\ v_3 \end{pmatrix} = Q_r v$$

$$R(q) = Q_r = \begin{pmatrix} a & -b & -c & -d \\ b & a & d & -c \\ c & -d & a & b \\ d & c & -b & a \end{pmatrix} = a \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} + b \begin{pmatrix} 0 & -1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & -1 & 0 \end{pmatrix} + c \begin{pmatrix} 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{pmatrix} + d \begin{pmatrix} 0 & 0 & 0 & -1 \\ 0 & 0 & 1 & 0 \\ 0 & -1 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{pmatrix} \leftrightarrow a1 + bi + cj + dk$$

$$ij \leftrightarrow \begin{pmatrix} 0 & -1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & -1 & 0 \end{pmatrix} \begin{pmatrix} 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{pmatrix} = - \begin{pmatrix} 0 & 0 & 0 & -1 \\ 0 & 0 & 1 & 0 \\ 0 & -1 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{pmatrix} \leftrightarrow -k$$

$$ji \leftrightarrow \begin{pmatrix} 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{pmatrix} \begin{pmatrix} 0 & -1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & -1 & 0 \end{pmatrix} = \begin{pmatrix} 0 & 0 & 0 & -1 \\ 0 & 0 & 1 & 0 \\ 0 & -1 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{pmatrix} \leftrightarrow k$$

$$ki \leftrightarrow \begin{pmatrix} 0 & 0 & 0 & -1 \\ 0 & 0 & 1 & 0 \\ 0 & -1 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} 0 & -1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & -1 & 0 \end{pmatrix} = - \begin{pmatrix} 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{pmatrix} \leftrightarrow -j$$

$$R(q) = Q_r = \begin{pmatrix} a & -b & -c & -d \\ b & a & d & -c \\ c & -d & a & b \\ d & c & -b & a \end{pmatrix} \leftrightarrow \begin{pmatrix} a+bi & -c+di \\ c+di & a-bi \end{pmatrix} = \begin{pmatrix} \alpha & -\bar{\beta} \\ \beta & \bar{\alpha} \end{pmatrix}, \begin{cases} \alpha = a+bi & \bar{\alpha} = a-bi \\ \beta = c+di & \bar{\beta} = c-di \end{cases}$$

$$Q_r = \begin{pmatrix} a & -b & -c & -d \\ b & a & d & -c \\ c & -d & a & b \\ d & c & -b & a \end{pmatrix} = \begin{pmatrix} a \\ b \\ c \\ d \end{pmatrix} \begin{pmatrix} - (b & c & d) \\ a & d & -c \\ -d & a & b \\ c & -b & a \end{pmatrix} = \begin{pmatrix} q_0 & -\mathbf{q}^\top \\ \mathbf{q} & Q_i^\top \end{pmatrix}, \begin{cases} a = q_0 \\ \mathbf{q} = \begin{pmatrix} b \\ c \\ d \end{pmatrix} \\ Q_i = \begin{pmatrix} a & -d & c \\ d & a & -b \\ -c & b & a \end{pmatrix} \end{cases}$$

$$vq \leftrightarrow \begin{pmatrix} a & -b & -c & -d \\ b & a & d & -c \\ c & -d & a & b \\ d & c & -b & a \end{pmatrix} \begin{pmatrix} v_0 \\ v_1 \\ v_2 \\ v_3 \end{pmatrix} = Q_r v = \begin{pmatrix} q_0 & -\mathbf{q}^\top \\ \mathbf{q} & Q_i^\top \end{pmatrix} \begin{pmatrix} v_0 \\ v \\ v_2 \\ v_3 \end{pmatrix} = \begin{pmatrix} v_0 q_0 - \mathbf{q}^\top \mathbf{v} \\ v_0 \mathbf{q} + Q_i^\top \mathbf{v} \end{pmatrix} = \begin{pmatrix} v_0 q_0 - \mathbf{v} \cdot \mathbf{q} \\ v_0 \mathbf{q} + q_0 \mathbf{v} + \mathbf{v} \times \mathbf{q} \end{pmatrix} = \begin{pmatrix} q_0 v_0 - \mathbf{q} \cdot \mathbf{v} \\ q v_0 + q_0 \mathbf{v} - \mathbf{q} \times \mathbf{v} \end{pmatrix}$$

$$\begin{cases} qv \leftrightarrow Q_i v = \begin{pmatrix} q_0 & -\mathbf{q}^\top \\ \mathbf{q} & Q_i \end{pmatrix} \begin{pmatrix} v_0 \\ \mathbf{v} \end{pmatrix} = \begin{pmatrix} q_0 v_0 - \mathbf{q}^\top \mathbf{v} \\ \mathbf{q} v_0 + Q_i \mathbf{v} \end{pmatrix} = \begin{pmatrix} q_0 v_0 - \mathbf{q} \cdot \mathbf{v} \\ \mathbf{q} v_0 + q_0 \mathbf{v} + \mathbf{q} \times \mathbf{v} \end{pmatrix} = \begin{pmatrix} q_0 v_0 - \mathbf{q} \cdot \mathbf{v} \\ \mathbf{q} v_0 + q_0 \mathbf{v} + \mathbf{q} \times \mathbf{v} \end{pmatrix} \\ vq \leftrightarrow Q_r v = \begin{pmatrix} q_0 & -\mathbf{q}^\top \\ \mathbf{q} & Q_i^\top \end{pmatrix} \begin{pmatrix} v_0 \\ \mathbf{v} \end{pmatrix} = \begin{pmatrix} v_0 q_0 - \mathbf{q}^\top \mathbf{v} \\ v_0 \mathbf{q} + Q_i^\top \mathbf{v} \end{pmatrix} = \begin{pmatrix} v_0 q_0 - \mathbf{v} \cdot \mathbf{q} \\ v_0 \mathbf{q} + v_0 \mathbf{q} + \mathbf{v} \times \mathbf{q} \end{pmatrix} = \begin{pmatrix} q_0 v_0 - \mathbf{q} \cdot \mathbf{v} \\ \mathbf{q} v_0 + q_0 \mathbf{v} - \mathbf{q} \times \mathbf{v} \end{pmatrix} \end{cases}$$

$$\begin{cases} qv = \begin{pmatrix} q_0 \\ \mathbf{q} \end{pmatrix} \begin{pmatrix} v_0 \\ \mathbf{v} \end{pmatrix} = L(q)v = Q_i v = \begin{pmatrix} q_0 & -\mathbf{q}^\top \\ \mathbf{q} & Q_i \end{pmatrix} \begin{pmatrix} v_0 \\ \mathbf{v} \end{pmatrix} \\ vq = \begin{pmatrix} v_0 \\ \mathbf{v} \end{pmatrix} \begin{pmatrix} q_0 \\ \mathbf{q} \end{pmatrix} = R(q)v = Q_r v = \begin{pmatrix} q_0 & -\mathbf{q}^\top \\ \mathbf{q} & Q_i^\top \end{pmatrix} \begin{pmatrix} v_0 \\ \mathbf{v} \end{pmatrix} \end{cases}$$

$$\begin{cases} \mathbf{v} \xrightarrow{\text{rotate about } \mathbf{n}} \mathbf{v}' \\ \begin{cases} \mathbf{v} = \mathbf{v}_{\parallel n} + \mathbf{v}_{\perp n} = \mathbf{v}_{\parallel} + \mathbf{v}_{\perp} \Rightarrow \mathbf{v}_{\perp} = \mathbf{v} - \mathbf{v}_{\parallel} \\ \mathbf{v}' = \mathbf{v}'_{\parallel n} + \mathbf{v}'_{\perp n} = \mathbf{v}'_{\parallel} + \mathbf{v}'_{\perp} \Rightarrow \mathbf{v}'_{\perp} = \mathbf{v}' - \mathbf{v}'_{\parallel} \end{cases} \\ \begin{cases} \mathbf{u} = \mathbf{n} \times \mathbf{v}_{\perp} \\ \|\mathbf{u}\| = \|\mathbf{n} \times \mathbf{v}_{\perp}\| = \|\mathbf{n}\| \|\mathbf{v}_{\perp}\| \sin \frac{\pi}{2} = \|\mathbf{v}_{\perp}\| \end{cases} \quad \begin{cases} \mathbf{u} \perp \mathbf{n} \\ \mathbf{u} \perp \mathbf{v}_{\perp} \\ \mathbf{n} \perp \mathbf{v}_{\perp} \\ \|\mathbf{n}\| = 1 \end{cases} \\ \begin{cases} \mathbf{v}'_{\parallel} = \mathbf{v}_{\parallel} \\ \mathbf{v}'_{\perp} = \cos \theta \mathbf{v}_{\perp} + \sin \theta (\mathbf{n} \times \mathbf{v}_{\perp}), \theta = \angle \mathbf{v}'_{\perp} \mathbf{v}_{\perp} \end{cases} \end{cases}$$

$$\begin{cases} \mathbf{v} \xrightarrow{\text{rotate about } n} \mathbf{v}' \\ \begin{cases} \mathbf{v} = \begin{pmatrix} 0 \\ \mathbf{v} \end{pmatrix} = \begin{pmatrix} 0 \\ \mathbf{v}_{\parallel n} \end{pmatrix} + \begin{pmatrix} 0 \\ \mathbf{v}_{\perp n} \end{pmatrix} = \begin{pmatrix} 0 \\ \mathbf{v}_{\parallel} \end{pmatrix} + \begin{pmatrix} 0 \\ \mathbf{v}_{\perp} \end{pmatrix} = \mathbf{v}_{\parallel} + \mathbf{v}_{\perp} \Rightarrow \mathbf{v}_{\perp} = \mathbf{v} - \mathbf{v}_{\parallel} \\ \mathbf{v}' = \begin{pmatrix} 0 \\ \mathbf{v}' \end{pmatrix} = \begin{pmatrix} 0 \\ \mathbf{v}'_{\parallel n} \end{pmatrix} + \begin{pmatrix} 0 \\ \mathbf{v}'_{\perp n} \end{pmatrix} = \begin{pmatrix} 0 \\ \mathbf{v}'_{\parallel} \end{pmatrix} + \begin{pmatrix} 0 \\ \mathbf{v}'_{\perp} \end{pmatrix} = \mathbf{v}'_{\parallel} + \mathbf{v}'_{\perp} \Rightarrow \mathbf{v}'_{\perp} = \mathbf{v}' - \mathbf{v}'_{\parallel} \end{cases} \\ \begin{cases} \mathbf{u} = \begin{pmatrix} 0 \\ \mathbf{u} \end{pmatrix} = \begin{pmatrix} 0 \\ \mathbf{n} \times \mathbf{v}_{\perp} \end{pmatrix} = \begin{pmatrix} 00 - 0 \\ n_0 + 0 \mathbf{v}_{\perp} + \mathbf{n} \times \mathbf{v}_{\perp} \end{pmatrix} = \begin{pmatrix} n_0 \mathbf{v}_{\perp 0} - \mathbf{n} \cdot \mathbf{v}_{\perp} \\ n \mathbf{v}_{\perp 0} + n_0 \mathbf{v}_{\perp} + \mathbf{n} \times \mathbf{v}_{\perp} \end{pmatrix} = n \mathbf{v}_{\perp} \end{cases} \quad \begin{cases} \mathbf{u} \perp \mathbf{n} \\ \mathbf{u} \perp \mathbf{v}_{\perp} \\ \mathbf{n} \perp \mathbf{v}_{\perp} \\ \|\mathbf{n}\| = 1 \end{cases} \\ \begin{cases} \|\mathbf{u}\| = \|\mathbf{n} \times \mathbf{v}_{\perp}\| = \|\mathbf{n}\| \|\mathbf{v}_{\perp}\| \sin \frac{\pi}{2} = \|\mathbf{v}_{\perp}\| \\ \mathbf{v}'_{\parallel} = \begin{pmatrix} 0 \\ \mathbf{v}'_{\parallel} \end{pmatrix} = \begin{pmatrix} 0 \\ \mathbf{v}_{\parallel} \end{pmatrix} = \mathbf{v}_{\parallel} \\ \mathbf{v}'_{\perp} = \begin{pmatrix} 0 \\ \mathbf{v}'_{\perp} \end{pmatrix} = \cos \theta \begin{pmatrix} 0 \\ \mathbf{v}_{\perp} \end{pmatrix} + \sin \theta \begin{pmatrix} 0 \\ \mathbf{n} \times \mathbf{v}_{\perp} \end{pmatrix} = \cos \theta \mathbf{v}_{\perp} + \sin \theta n \mathbf{v}_{\perp} = (\cos \theta + \sin \theta n) \mathbf{v}_{\perp} \end{cases} \end{cases}$$

$$\mathbf{v}'_{\perp} = (\cos \theta + \sin \theta n) \mathbf{v}_{\perp}$$

$$\mathbf{v}'_{\perp} = p \mathbf{v}_{\perp}, p = \cos \theta + \sin \theta n = \cos \theta \begin{pmatrix} 1 \\ \mathbf{0} \end{pmatrix} + \sin \theta \begin{pmatrix} 0 \\ \mathbf{n} \end{pmatrix} = \begin{pmatrix} \cos \theta \\ \sin \theta \mathbf{n} \end{pmatrix}$$

$$p = \begin{pmatrix} p_0 \\ \mathbf{p} \end{pmatrix} = \begin{pmatrix} \cos \theta \\ \sin \theta \mathbf{n} \end{pmatrix} = \cos \theta + \sin \theta n, \|\mathbf{n}\| = 1$$

$$|p| = \left\| \begin{pmatrix} \cos \theta \\ \sin \theta \mathbf{n} \end{pmatrix} \right\| = \sqrt{\begin{pmatrix} \cos \theta \\ \sin \theta \mathbf{n} \end{pmatrix}^\top \begin{pmatrix} \cos \theta \\ \sin \theta \mathbf{n} \end{pmatrix}} = \sqrt{\cos^2 \theta + \sin^2 \theta \|\mathbf{n}\|^2} = \sqrt{\cos^2 \theta + \sin^2 \theta} = 1$$

$$|p| = \left\| \begin{pmatrix} \cos \theta \\ \sin \theta \mathbf{n} \end{pmatrix} \right\| = 1$$

$$p^* = \bar{p} = \begin{pmatrix} p_0 \\ -\mathbf{p} \end{pmatrix} = \begin{pmatrix} \cos \theta \\ -\sin \theta \mathbf{n} \end{pmatrix} = \cos \theta - \sin \theta n$$

$$\begin{aligned} p^* p &= \bar{p} p = \begin{pmatrix} p_0 \\ -\mathbf{p} \end{pmatrix} \begin{pmatrix} p_0 \\ \mathbf{p} \end{pmatrix} = \begin{pmatrix} p_0 p_0 - (-\mathbf{p}) \cdot \mathbf{p} \\ -\mathbf{p} p_0 + p_0 \mathbf{p} + (-\mathbf{p}) \times \mathbf{p} \end{pmatrix} \\ &= \begin{pmatrix} p_0 p_0 + \mathbf{p} \cdot \mathbf{p} \\ \mathbf{0} + \mathbf{0} \end{pmatrix} = p_0 p_0 + \mathbf{p} \cdot \mathbf{p} = p_0^2 + \|\mathbf{p}\|^2 = |p|^2 = 1^2 = 1 \\ &= \begin{pmatrix} p_0 p_0 - \mathbf{p} \cdot (-\mathbf{p}) \\ \mathbf{p} p_0 + p_0 (-\mathbf{p}) + \mathbf{p} \times (-\mathbf{p}) \end{pmatrix} = p \bar{p} = pp^* \end{aligned}$$

$$p^* p = \bar{p} p = 1 = p \bar{p} = pp^* = |p|^2 = p_0^2 + \|\mathbf{p}\|^2 \Rightarrow p^* = \frac{1}{p} = p^{-1}$$

$$\begin{aligned} p^2 &= pp = \begin{pmatrix} p_0 p_0 - \mathbf{p} \cdot \mathbf{p} \\ \mathbf{p} p_0 + p_0 \mathbf{p} + \mathbf{p} \times \mathbf{p} \end{pmatrix} = \begin{pmatrix} \cos^2 \theta - \sin^2 \theta \|\mathbf{n}\|^2 \\ 2(\sin \theta \mathbf{n}) \cos \theta + \mathbf{0} \end{pmatrix} \\ &= \begin{pmatrix} \cos^2 \theta - \sin^2 \theta \\ 2 \sin \theta \cos \theta \mathbf{n} \end{pmatrix} = \begin{pmatrix} \cos 2\theta \\ \sin 2\theta \mathbf{n} \end{pmatrix} \\ \begin{pmatrix} \cos \theta \\ \sin \theta \mathbf{n} \end{pmatrix}^2 &= \begin{pmatrix} \cos 2\theta \\ \sin 2\theta \mathbf{n} \end{pmatrix} \\ q^2 &= \begin{pmatrix} \cos \frac{\theta}{2} \\ \sin \frac{\theta}{2} \mathbf{n} \end{pmatrix}^2 = \begin{pmatrix} \cos \theta \\ \sin \theta \mathbf{n} \end{pmatrix} = p \end{aligned}$$

$$qv_{\parallel} = v_{\parallel}q$$

$$\begin{aligned} \begin{pmatrix} \alpha \\ \beta \mathbf{n} \end{pmatrix} v_{\parallel \mathbf{n}} &= \begin{pmatrix} \alpha \\ \beta \mathbf{n} \end{pmatrix} v_{\parallel} = \begin{pmatrix} \alpha \\ \beta \mathbf{n} \end{pmatrix} \begin{pmatrix} 0 \\ \mathbf{v}_{\parallel} \end{pmatrix} \\ &= \begin{pmatrix} \alpha 0 - \beta \mathbf{n} \cdot \mathbf{v}_{\parallel} \\ \beta \mathbf{n} 0 + \alpha \mathbf{v}_{\parallel} + \beta \mathbf{n} \times \mathbf{v}_{\parallel} \end{pmatrix} = \begin{pmatrix} -\beta \mathbf{n} \cdot \mathbf{v}_{\parallel} \\ \alpha \mathbf{v}_{\parallel} + \mathbf{0} \end{pmatrix} = \begin{pmatrix} -\beta \mathbf{n} \cdot \mathbf{v}_{\parallel} \\ \alpha \mathbf{v}_{\parallel} \end{pmatrix} \\ &= \begin{pmatrix} 0\alpha - \mathbf{v}_{\parallel} \cdot \beta \mathbf{n} \\ \mathbf{v}_{\parallel} \alpha + 0\beta \mathbf{n} + \mathbf{v}_{\parallel} \times \beta \mathbf{n} \end{pmatrix} = \begin{pmatrix} -\mathbf{v}_{\parallel} \cdot \beta \mathbf{n} \\ \alpha \mathbf{v}_{\parallel} + \mathbf{0} \end{pmatrix} = \begin{pmatrix} -\beta \mathbf{n} \cdot \mathbf{v}_{\parallel} \\ \alpha \mathbf{v}_{\parallel} \end{pmatrix} \\ &= \begin{pmatrix} 0 \\ \mathbf{v}_{\parallel} \end{pmatrix} \begin{pmatrix} \alpha \\ \beta \mathbf{n} \end{pmatrix} = v_{\parallel} \begin{pmatrix} \alpha \\ \beta \mathbf{n} \end{pmatrix} = v_{\parallel \mathbf{n}} \begin{pmatrix} \alpha \\ \beta \mathbf{n} \end{pmatrix} \\ \begin{pmatrix} \alpha \\ \beta \mathbf{n} \end{pmatrix} v_{\parallel} &= v_{\parallel} \begin{pmatrix} \alpha \\ \beta \mathbf{n} \end{pmatrix} \end{aligned}$$

$$qv_{\perp} = v_{\perp}q^*$$

$$\begin{aligned} \begin{pmatrix} \alpha \\ \beta \mathbf{n} \end{pmatrix} v_{\perp \mathbf{n}} &= \begin{pmatrix} \alpha \\ \beta \mathbf{n} \end{pmatrix} v_{\perp} = \begin{pmatrix} \alpha \\ \beta \mathbf{n} \end{pmatrix} \begin{pmatrix} 0 \\ \mathbf{v}_{\perp} \end{pmatrix} \\ &= \begin{pmatrix} \alpha 0 - \beta \mathbf{n} \cdot \mathbf{v}_{\perp} \\ \beta \mathbf{n} 0 + \alpha \mathbf{v}_{\perp} + \beta \mathbf{n} \times \mathbf{v}_{\perp} \end{pmatrix} = \begin{pmatrix} 0 - 0 \\ \alpha \mathbf{v}_{\perp} + \beta \mathbf{n} \times \mathbf{v}_{\perp} \end{pmatrix} = \begin{pmatrix} 0 \\ \alpha \mathbf{v}_{\perp} + \mathbf{v}_{\perp} \times (-\beta \mathbf{n}) \end{pmatrix} \\ &= \begin{pmatrix} 0\alpha - \mathbf{v}_{\perp} \cdot (-\beta \mathbf{n}) \\ \mathbf{v}_{\perp} \alpha + 0(-\beta \mathbf{n}) + \mathbf{v}_{\perp} \times (-\beta \mathbf{n}) \end{pmatrix} = \begin{pmatrix} 0 - 0 \\ \alpha \mathbf{v}_{\perp} + \mathbf{v}_{\perp} \times (-\beta \mathbf{n}) \end{pmatrix} = \begin{pmatrix} 0 \\ \alpha \mathbf{v}_{\perp} + \beta \mathbf{n} \times \mathbf{v}_{\perp} \end{pmatrix} \\ &= \begin{pmatrix} 0 \\ \mathbf{v}_{\perp} \end{pmatrix} \begin{pmatrix} \alpha \\ -\beta \mathbf{n} \end{pmatrix} = v_{\perp} \begin{pmatrix} \alpha \\ -\beta \mathbf{n} \end{pmatrix} = v_{\perp \mathbf{n}} \begin{pmatrix} \alpha \\ -\beta \mathbf{n} \end{pmatrix} \\ \begin{pmatrix} \alpha \\ \beta \mathbf{n} \end{pmatrix} v_{\perp} &= v_{\perp} \begin{pmatrix} \alpha \\ -\beta \mathbf{n} \end{pmatrix} = v_{\perp} \begin{pmatrix} \alpha \\ \beta \mathbf{n} \end{pmatrix}^* = v_{\perp} \overline{\begin{pmatrix} \alpha \\ \beta \mathbf{n} \end{pmatrix}} \end{aligned}$$

$$\begin{aligned}
q = \cos \frac{\theta}{2} + \sin \frac{\theta}{2} n &\in \left\{ p \left| \begin{array}{l} p = \cos \theta + \sin \theta n \\ n = \begin{pmatrix} 0 \\ \mathbf{n} \end{pmatrix}, \|\mathbf{n}\| = 1 \end{array} \right. \right\} \subset \left\{ \begin{pmatrix} \alpha \\ \beta \mathbf{n} \end{pmatrix} \left| \begin{array}{l} \alpha, \beta \in \mathbb{R} \\ \|\mathbf{n}\| = 1 \end{array} \right. \right\} \\
&\Rightarrow \begin{cases} q^* q = \bar{q} q = 1 = q \bar{q} = q q^* = |q|^2 = q_0^2 + \|\mathbf{q}\|^2 \Rightarrow q^* = \frac{1}{q} = q^{-1} & (u) \text{ unit quaternion} \\ q v_{\parallel} = v_{\parallel} q \\ q v_{\perp} = v_{\perp} q^* = v_{\perp} \bar{q} \Leftrightarrow \begin{pmatrix} \cos \frac{\theta}{2} \\ \sin \frac{\theta}{2} \mathbf{n} \end{pmatrix} v_{\perp} = v_{\perp} \begin{pmatrix} \cos \frac{\theta}{2} \\ -\sin \frac{\theta}{2} \mathbf{n} \end{pmatrix} & (c) \text{ commutativity} \\ q v_{\perp} = v_{\perp} q^* = v_{\perp} \bar{q} \Leftrightarrow \begin{pmatrix} \cos \frac{\theta}{2} \\ \sin \frac{\theta}{2} \mathbf{n} \end{pmatrix} v_{\perp} = v_{\perp} \begin{pmatrix} \cos \frac{\theta}{2} \\ -\sin \frac{\theta}{2} \mathbf{n} \end{pmatrix} & (a) \text{ anticommutativity} \end{cases} \\
v' = v'_{\parallel} + v'_{\perp}, &\begin{cases} v'_{\parallel} = v_{\parallel} & q = \cos \frac{\theta}{2} + \sin \frac{\theta}{2} n, |n| = 1 \\ v'_{\perp} = p v_{\perp} = q^2 v_{\perp} & p = q^2 = \cos \theta + \sin \theta n \end{cases} \\
&= v_{\parallel} + q^2 v_{\perp} \stackrel{(u)}{=} 1 v_{\parallel} + q q v_{\perp} \\
&\stackrel{(a)}{=} q q^* v_{\parallel} + q v_{\perp} q^* \stackrel{(c)}{=} q v_{\parallel} q^* + q v_{\perp} q^* \\
&= q(v_{\parallel} + v_{\perp}) q^* = q(v) q^* = q v q^*
\end{aligned}$$

$$\begin{aligned}
v' = q v q^{-1}, &q = \cos \frac{\theta}{2} + \sin \frac{\theta}{2} n, |n| = 1 \\
&= q v q^* = q(v_{\parallel} + v_{\perp}) q^* \\
&= q v_{\parallel} q^* + q v_{\perp} q^* = q q^* v_{\parallel} + q q v_{\perp} = 1 v_{\parallel} + q^2 v_{\perp} \\
&= v_{\parallel} + p v_{\perp}, p = q^2 = \cos \theta + \sin \theta n
\end{aligned}$$

$$\begin{aligned}
v' = q v q^* = q(v q^*) &\leftrightarrow L(q) R(q^*) v = Q_l Q_r v = \begin{pmatrix} q_0 & -\mathbf{q}^T \\ \mathbf{q} & Q_i \end{pmatrix} \begin{pmatrix} q_0 & -\mathbf{q}^T \\ \mathbf{q} & Q_i^T \end{pmatrix} \begin{pmatrix} v_0 \\ \mathbf{v} \end{pmatrix} \\
&= (qv) q^* \leftrightarrow R(q^*) L(q) v = Q_r Q_l v = \begin{pmatrix} q_0 & -\mathbf{q}^T \\ \mathbf{q} & Q_i^T \end{pmatrix} \begin{pmatrix} q_0 & -\mathbf{q}^T \\ \mathbf{q} & Q_i \end{pmatrix} \begin{pmatrix} v_0 \\ \mathbf{v} \end{pmatrix}
\end{aligned}$$

$$\begin{pmatrix} q_0 & -\mathbf{q}^T \\ \mathbf{q} & Q_i \end{pmatrix} \begin{pmatrix} q_0 & -\mathbf{q}^T \\ \mathbf{q} & Q_i^T \end{pmatrix} = \begin{pmatrix} q_0^2 - \mathbf{q}^T \mathbf{q} & -q_0 \mathbf{q}^T - \mathbf{q}^T Q_i^T \\ \mathbf{q} q_0 + Q_i \mathbf{q} & -\mathbf{q} \mathbf{q}^T + Q_i Q_i^T \end{pmatrix} = L(q) R(q^*)$$

$$\begin{pmatrix} q_0 & -\mathbf{q}^T \\ \mathbf{q} & Q_i^T \end{pmatrix} \begin{pmatrix} q_0 & -\mathbf{q}^T \\ \mathbf{q} & Q_i \end{pmatrix} = \begin{pmatrix} q_0^2 - \mathbf{q}^T \mathbf{q} & -q_0 \mathbf{q}^T - \mathbf{q}^T Q_i \\ \mathbf{q} q_0 + Q_i^T \mathbf{q} & -\mathbf{q} \mathbf{q}^T + Q_i^T Q_i \end{pmatrix} = R(q^*) L(q)$$

$$\begin{aligned}
q = a + bi + cj + dk &= \cos \frac{\theta}{2} + \sin \frac{\theta}{2} n = \cos \frac{\theta}{2} + bi + cj + dk \\
\Rightarrow Q_i Q_i^T &= \begin{pmatrix} a & -d & c \\ d & a & -b \\ -c & b & a \end{pmatrix} \begin{pmatrix} a & d & -c \\ -d & a & b \\ c & -b & a \end{pmatrix} = \begin{pmatrix} a^2 + d^2 + c^2 & -cb & -db \\ -cb & d^2 + a^2 + b^2 & -dc \\ -db & -dc & c^2 + b^2 + a^2 \end{pmatrix} \\
&= Q_i^T Q_i = \begin{pmatrix} a & d & -c \\ -d & a & b \\ c & -b & a \end{pmatrix} \begin{pmatrix} a & -d & c \\ d & a & -b \\ -c & b & a \end{pmatrix} = \begin{pmatrix} a^2 + d^2 + c^2 & -cb & -db \\ -cb & d^2 + a^2 + b^2 & -dc \\ -db & -dc & c^2 + b^2 + a^2 \end{pmatrix} \\
&\mathbf{q}^T Q_i^T = (b \quad c \quad d) \begin{pmatrix} a & d & -c \\ -d & a & b \\ c & -b & a \end{pmatrix} = (ba \quad ca \quad da) \\
&= \mathbf{q}^T Q_i = (b \quad c \quad d) \begin{pmatrix} a & -d & c \\ d & a & -b \\ -c & b & a \end{pmatrix} = (ba \quad ca \quad da) \\
&Q_i \mathbf{q} = \begin{pmatrix} a & -d & c \\ d & a & -b \\ -c & b & a \end{pmatrix} \begin{pmatrix} b \\ c \\ d \end{pmatrix} = \begin{pmatrix} ab \\ ac \\ ad \end{pmatrix} \\
&= Q_i^T \mathbf{q} = \begin{pmatrix} a & d & -c \\ -d & a & b \\ c & -b & a \end{pmatrix} \begin{pmatrix} b \\ c \\ d \end{pmatrix} = \begin{pmatrix} ab \\ ac \\ ad \end{pmatrix}
\end{aligned}$$

$$\begin{aligned}
& \begin{pmatrix} q_0 & -\mathbf{q}^\top \\ \mathbf{q} & Q_i \end{pmatrix} \begin{pmatrix} q_0 & -\mathbf{q}^\top \\ \mathbf{q} & Q_i^\top \end{pmatrix} = \begin{pmatrix} q_0^2 - \mathbf{q}^\top \mathbf{q} & -q_0 \mathbf{q}^\top - \mathbf{q}^\top Q_i^\top \\ \mathbf{q} q_0 + Q_i \mathbf{q} & -\mathbf{q} \mathbf{q}^\top + Q_i Q_i^\top \end{pmatrix} = L(q) R(q^*) \\
& = \begin{pmatrix} q_0 & -\mathbf{q}^\top \\ \mathbf{q} & Q_i^\top \end{pmatrix} \begin{pmatrix} q_0 & -\mathbf{q}^\top \\ \mathbf{q} & Q_i \end{pmatrix} = \begin{pmatrix} q_0^2 - \mathbf{q}^\top \mathbf{q} & -q_0 \mathbf{q}^\top - \mathbf{q}^\top Q_i \\ \mathbf{q} q_0 + Q_i^\top \mathbf{q} & -\mathbf{q} \mathbf{q}^\top + Q_i^\top Q_i \end{pmatrix} = R(q^*) L(q) \\
& \Downarrow \\
L(q) R(q^*) &= R(q^*) L(q) \\
&\Updownarrow \\
v' &= qvq^* = q(vq^*) = (qv)q^*
\end{aligned}$$

$$a = \cos \frac{\theta}{2} \Rightarrow \theta = 2 \arccos a = 2 \cos^{-1} a$$

$$\sin \frac{\theta}{2} n = \sin \frac{2 \arccos a}{2} n = \sin (\cos^{-1} a) n$$

$$n = \begin{pmatrix} 0 \\ \mathbf{n} \end{pmatrix} = \begin{pmatrix} 0 \\ \hat{\mathbf{n}} \end{pmatrix}, \|n\| = 1, \mathbf{n} = \begin{pmatrix} \sin \gamma \cos \alpha \\ \sin \gamma \sin \alpha \\ \cos \gamma \end{pmatrix}, \begin{cases} \alpha = \angle \hat{\mathbf{n}}_{xy} \hat{\mathbf{x}} \\ \gamma = \angle \hat{\mathbf{n}} \hat{\mathbf{z}} \end{cases}$$

$$\begin{aligned}
q &= a + bi + cj + dk \\
&= \cos \frac{\theta}{2} + \sin \frac{\theta}{2} n, n = \begin{pmatrix} 0 \\ \mathbf{n} \end{pmatrix} = \begin{pmatrix} 0 \\ \hat{\mathbf{n}} \end{pmatrix} = \begin{pmatrix} 0 \\ \sin \gamma \cos \alpha \\ \sin \gamma \sin \alpha \\ \cos \gamma \end{pmatrix}, \|n\| = 1, \begin{cases} \alpha = \angle \hat{\mathbf{n}}_{xy} \hat{\mathbf{x}} \\ \gamma = \angle \hat{\mathbf{n}} \hat{\mathbf{z}} \end{cases} \\
&= \cos \frac{\theta}{2} + \sin \frac{\theta}{2} \mathbf{n}, \mathbf{n} = (\sin \gamma \cos \alpha) \mathbf{i} + (\sin \gamma \sin \alpha) \mathbf{j} + (\cos \gamma) \mathbf{k} \\
&= \cos \frac{\theta}{2} + \sin \frac{\theta}{2} [(\sin \gamma \cos \alpha) \mathbf{i} + (\sin \gamma \sin \alpha) \mathbf{j} + (\cos \gamma) \mathbf{k}] \\
&= \cos \frac{\theta}{2} + \mathbf{n} \sin \frac{\theta}{2} \\
q^{-1} &= q^* = \cos \frac{\theta}{2} - \mathbf{n} \sin \frac{\theta}{2} \\
&= \cos \frac{\theta}{2} - \sin \frac{\theta}{2} [(\sin \gamma \cos \alpha) \mathbf{i} + (\sin \gamma \sin \alpha) \mathbf{j} + (\cos \gamma) \mathbf{k}] \\
\mathbf{v}' &= v' = qvq^{-1} = q\mathbf{v}q^{-1}, \begin{cases} v = \mathbf{v} = v_1 \mathbf{i} + v_2 \mathbf{j} + v_3 \mathbf{k} \\ v' = \mathbf{v}' = v'_1 \mathbf{i} + v'_2 \mathbf{j} + v'_3 \mathbf{k} \end{cases} \\
&= \left\{ \cos \frac{\theta}{2} + \sin \frac{\theta}{2} [(\sin \gamma \cos \alpha) \mathbf{i} + (\sin \gamma \sin \alpha) \mathbf{j} + (\cos \gamma) \mathbf{k}] \right\} \mathbf{v} \\
&\quad \left\{ \cos \frac{\theta}{2} - \sin \frac{\theta}{2} [(\sin \gamma \cos \alpha) \mathbf{i} + (\sin \gamma \sin \alpha) \mathbf{j} + (\cos \gamma) \mathbf{k}] \right\} \\
&= q(\theta, \gamma, \alpha) \mathbf{v} q^{-1}(\theta, \gamma, \alpha) \\
&= \left\{ \cos \frac{\theta}{2} + \sin \frac{\theta}{2} [n_x \mathbf{i} + n_y \mathbf{j} + n_z \mathbf{k}] \right\} \mathbf{v} \left\{ \cos \frac{\theta}{2} - \sin \frac{\theta}{2} [n_x \mathbf{i} + n_y \mathbf{j} + n_z \mathbf{k}] \right\} \\
&= \left\{ \cos \frac{\theta}{2} + \sin \frac{\theta}{2} [n_x \mathbf{i} \pm \sqrt{1 - n_x^2 - n_z^2} \mathbf{j} + n_z \mathbf{k}] \right\} \mathbf{v} \\
&\quad \left\{ \cos \frac{\theta}{2} - \sin \frac{\theta}{2} [n_x \mathbf{i} \pm \sqrt{1 - n_x^2 - n_z^2} \mathbf{j} + n_z \mathbf{k}] \right\} \\
&= q(\theta, n_z, n_x) \mathbf{v} q^{-1}(\theta, n_z, n_x)
\end{aligned}$$

$$\mathbf{v}' = v' = qvq^{-1} = q\mathbf{v}q^{-1}, \begin{cases} v = \mathbf{v} = v_1\mathbf{i} + v_2\mathbf{j} + v_3\mathbf{k} \\ v' = \mathbf{v}' = v'_1\mathbf{i} + v'_2\mathbf{j} + v'_3\mathbf{k} \\ q = \cos \frac{\theta}{2} + \sin \frac{\theta}{2} [(\sin \gamma \cos \alpha) \mathbf{i} + (\sin \gamma \sin \alpha) \mathbf{j} + (\cos \gamma) \mathbf{k}] \\ q^{-1} = \cos \frac{\theta}{2} - \sin \frac{\theta}{2} [(\sin \gamma \cos \alpha) \mathbf{i} + (\sin \gamma \sin \alpha) \mathbf{j} + (\cos \gamma) \mathbf{k}] \end{cases}$$

34.2.3 matrix form

$$q \stackrel{|q|=1}{=} a + bi + cj + dk, \begin{cases} a = \cos \frac{\theta}{2} \\ b = \sin \frac{\theta}{2} \sin \gamma \cos \alpha \\ c = \sin \frac{\theta}{2} \sin \gamma \sin \alpha \\ d = \sin \frac{\theta}{2} \cos \gamma \end{cases}$$

$$L(q) = Q_l = \begin{pmatrix} a & -b & -c & -d \\ b & a & -d & c \\ c & d & a & -b \\ d & -c & b & a \end{pmatrix}$$

$$R(q) = Q_r = \begin{pmatrix} a & -b & -c & -d \\ b & a & d & -c \\ c & -d & a & b \\ d & c & -b & a \end{pmatrix}$$

$$R(q^*) = \begin{pmatrix} a & b & c & d \\ -b & a & -d & c \\ -c & d & a & -b \\ -d & -c & b & a \end{pmatrix}$$

$$\begin{aligned} v' &= qvq^* = (qv)q^* \leftrightarrow R(q^*)L(q)v \\ &= q(vq^*) \leftrightarrow L(q)R(q^*)v \\ &= \begin{pmatrix} a & -b & -c & -d \\ b & a & -d & c \\ c & d & a & -b \\ d & -c & b & a \end{pmatrix} \begin{pmatrix} a & b & c & d \\ -b & a & -d & c \\ -c & d & a & -b \\ -d & -c & b & a \end{pmatrix} v \\ &= \begin{pmatrix} a^2 + b^2 + c^2 + d^2 & 0 & 0 & 0 \\ 0 & b^2 + a^2 - d^2 - c^2 & -2ad + 2bc & 2ac + 2bd \\ 0 & 2ad + 2bc & c^2 - d^2 + a^2 - b^2 & 2cd - 2ab \\ 0 & -2ac + 2bd & 2cd + 2ab & d^2 - c^2 - b^2 + a^2 \end{pmatrix} v \\ &= \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 - 2(d^2 + c^2) & 2(bc - ad) & 2(bd + ac) \\ 0 & 2(bc + ad) & 1 - 2(d^2 + b^2) & 2(cd - ab) \\ 0 & 2(bd - ac) & 2(cd + ab) & 1 - 2(c^2 + b^2) \end{pmatrix} v \\ \begin{pmatrix} 0 \\ v' \end{pmatrix} &= \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 - 2(d^2 + c^2) & 2(bc - ad) & 2(bd + ac) \\ 0 & 2(bc + ad) & 1 - 2(d^2 + b^2) & 2(cd - ab) \\ 0 & 2(bd - ac) & 2(cd + ab) & 1 - 2(c^2 + b^2) \end{pmatrix} \begin{pmatrix} 0 \\ v \end{pmatrix} \end{aligned}$$

$$\mathbf{v}' = \begin{pmatrix} 1 - 2(d^2 + c^2) & 2(bc - ad) & 2(bd + ac) \\ 2(bc + ad) & 1 - 2(d^2 + b^2) & 2(cd - ab) \\ 2(bd - ac) & 2(cd + ab) & 1 - 2(c^2 + b^2) \end{pmatrix} \mathbf{v}$$

$$q \stackrel{|q|=1}{=} a + bi + cj + dk, \begin{cases} a = \cos \frac{\theta}{2} \\ b = \sin \frac{\theta}{2} \sin \gamma \cos \alpha \\ c = \sin \frac{\theta}{2} \sin \gamma \sin \alpha \\ d = \sin \frac{\theta}{2} \cos \gamma \end{cases}$$

34.2.4 exponential form

$$\begin{aligned} q &= a + bi + cj + dk = a + ib + jc + kd \\ &= a + r \frac{ib + jc + kd}{r}, r^2 = b^2 + c^2 + d^2 \\ &\stackrel{|q|=1}{=} \cos \frac{\theta}{2} + \sin \frac{\theta}{2} \mathbf{n}, \begin{cases} a = \cos \frac{\theta}{2} & \cos \frac{\theta}{2} = a \\ r = \sin \frac{\theta}{2} & \sin \frac{\theta}{2} = r \\ \frac{ib + jc + kd}{r} = \mathbf{n} & \mathbf{n} = \frac{ib + jc + kd}{r} \end{cases} \end{aligned}$$

$$\begin{aligned} \|\mathbf{n}\|^2 &= \left(\frac{ib + jc + kd}{r} \right)^2 = \frac{-b^2 - c^2 - d^2 + bc(ij + ji) + cd(jk + kj) + db(ki + ik)}{r^2} \\ &= \frac{-b^2 - c^2 - d^2 + bc(k - k) + cd(i - i) + db(j - j)}{r^2} \\ &= \frac{-b^2 - c^2 - d^2 + bc0 + cd0 + db0}{r^2} = \frac{-b^2 - c^2 - d^2}{r^2} \\ &= \frac{-b^2 - c^2 - d^2}{b^2 + c^2 + d^2} = -1 \\ \mathbf{n} &= \frac{ib + jc + kd}{r} = \pm \sqrt{-1} \end{aligned}$$

$$\begin{aligned} q &= a + bi + cj + dk = a + ib + jc + kd \\ &= a + r \frac{ib + jc + kd}{r}, r^2 = b^2 + c^2 + d^2 \\ &= \sqrt{a^2 + r^2} \left(\frac{a}{\sqrt{a^2 + r^2}} + \frac{ib + jc + kd}{r} \frac{r}{\sqrt{a^2 + r^2}} \right) \\ &= \rho (\cos \phi + \sqrt{-1} \sin \phi), \begin{cases} \rho = \sqrt{a^2 + r^2} \\ \tan \phi = \frac{r}{a} \Leftrightarrow \phi = \arctan \frac{r}{a} = \frac{\theta}{2} \end{cases} \\ &= \rho e^{\phi \sqrt{-1}} \stackrel{|q|=1}{=} |q| e^{\mathbf{n} \frac{\theta}{2}} = e^{\mathbf{n} \frac{\theta}{2}} \\ q &= \rho e^{\phi \sqrt{-1}} \stackrel{|q|=1}{=} e^{\mathbf{n} \frac{\theta}{2}}, \begin{cases} q = a + ib + jc + kd \\ \rho = \sqrt{a^2 + r^2} = |q| = \sqrt{a^2 + b^2 + c^2 + d^2} = 1 \\ \tan \frac{\theta}{2} = \frac{r}{a} \Leftrightarrow \frac{\theta}{2} = \arctan \frac{r}{a} = \arctan \frac{\pm \sqrt{b^2 + c^2 + d^2}}{a} \end{cases} \end{aligned}$$

$$\begin{aligned}
\mathbf{v}' &= v' = qvq^{-1} = q\mathbf{v}q^{-1}, \quad \left\{ \begin{array}{l} v = \mathbf{v} = v_1\mathbf{i} + v_2\mathbf{j} + v_3\mathbf{k} \\ v' = \mathbf{v}' = v'_1\mathbf{i} + v'_2\mathbf{j} + v'_3\mathbf{k} \end{array} \right., \quad \mathbf{n} = \begin{pmatrix} \sin \gamma \cos \alpha \\ \sin \gamma \sin \alpha \\ \cos \gamma \end{pmatrix}, \quad \left\{ \begin{array}{l} \alpha = \angle \hat{\mathbf{n}}_{xy} \hat{\mathbf{x}} \\ \gamma = \angle \hat{\mathbf{n}} \hat{\mathbf{z}} \end{array} \right. \\
&= \left\{ \cos \frac{\theta}{2} + \sin \frac{\theta}{2} [(\sin \gamma \cos \alpha) \mathbf{i} + (\sin \gamma \sin \alpha) \mathbf{j} + (\cos \gamma) \mathbf{k}] \right\} \mathbf{v} \\
&\quad \left\{ \cos \frac{\theta}{2} - \sin \frac{\theta}{2} [(\sin \gamma \cos \alpha) \mathbf{i} + (\sin \gamma \sin \alpha) \mathbf{j} + (\cos \gamma) \mathbf{k}] \right\} \\
&= q(\theta, \gamma, \alpha) \mathbf{v} q^{-1}(\theta, \gamma, \alpha) = q\left(\frac{\theta}{2}, \gamma, \alpha\right) \mathbf{v} q^{-1}\left(\frac{\theta}{2}, \gamma, \alpha\right) \\
&= \left\{ \cos \frac{\theta}{2} + \sin \frac{\theta}{2} [n_x \mathbf{i} + n_y \mathbf{j} + n_z \mathbf{k}] \right\} \mathbf{v} \left\{ \cos \frac{\theta}{2} - \sin \frac{\theta}{2} [n_x \mathbf{i} + n_y \mathbf{j} + n_z \mathbf{k}] \right\} \\
&= \left\{ \cos \frac{\theta}{2} + \sin \frac{\theta}{2} \left[n_x \mathbf{i} \pm \sqrt{1 - n_x^2 - n_z^2} \mathbf{j} + n_z \mathbf{k} \right] \right\} \mathbf{v} \\
&\quad \left\{ \cos \frac{\theta}{2} - \sin \frac{\theta}{2} \left[n_x \mathbf{i} \pm \sqrt{1 - n_x^2 - n_z^2} \mathbf{j} + n_z \mathbf{k} \right] \right\} \\
&= q(\theta, n_z, n_x) \mathbf{v} q^{-1}(\theta, n_z, n_x) = q\left(\frac{\theta}{2}, n_z, n_x\right) \mathbf{v} q^{-1}\left(\frac{\theta}{2}, n_z, n_x\right) \\
&= e^{\mathbf{n} \frac{\theta}{2}} \mathbf{v} e^{-\mathbf{n} \frac{\theta}{2}}
\end{aligned}$$

$$\mathbf{v}' = e^{\mathbf{n} \frac{\theta}{2}} \mathbf{v} e^{-\mathbf{n} \frac{\theta}{2}}$$

general quaternion exponential form

$$\begin{aligned}
q &= \rho e^{\phi \sqrt{-1}} = |q| e^{\mathbf{n} \phi}, \quad \left\{ \begin{array}{l} q = a + ib + jc + kd = a + r \frac{ib + jc + kd}{r} = a + r\mathbf{n} \\ \rho = \sqrt{a^2 + r^2} = |q| = \sqrt{a^2 + b^2 + c^2 + d^2} \\ \tan \phi = \frac{r}{a} \Leftrightarrow \phi = \arctan \frac{r}{a} = \arctan \frac{\pm \sqrt{b^2 + c^2 + d^2}}{a} \end{array} \right. \\
&= e^{\ln |q|} e^{\mathbf{n} \phi} = e^{\ln |q| + \mathbf{n} \phi}
\end{aligned}$$

34.2.5 double cover

<https://www.bilibili.com/video/BV1rj41117VW/?t=15m15s>

$$\begin{array}{ccccccccccccccc}
\theta & 0 & \rightarrow & \frac{\pi}{2} & \rightarrow & \pi & \rightarrow & \frac{3\pi}{2} & \rightarrow & 2\pi & \rightarrow & \frac{5\pi}{2} & \rightarrow & 3\pi & \rightarrow & \frac{7\pi}{2} & \rightarrow & 4\pi \\
\frac{\theta}{2} & 0 & \rightarrow & \frac{\pi}{4} & \rightarrow & \frac{\pi}{2} & \rightarrow & \frac{3\pi}{4} & \rightarrow & \pi & \rightarrow & \frac{5\pi}{4} & \rightarrow & \frac{3\pi}{2} & \rightarrow & \frac{7\pi}{4} & \rightarrow & 2\pi \\
e^{\mathbf{n} \frac{\theta}{2}} & 1 & \rightarrow & e^{\mathbf{n} \frac{\pi}{4}} & \rightarrow & \mathbf{n} & \rightarrow & e^{\mathbf{n} \frac{3\pi}{4}} & \rightarrow & -1 & \rightarrow & e^{\mathbf{n} \frac{5\pi}{4}} & \rightarrow & -\mathbf{n} & \rightarrow & e^{\mathbf{n} \frac{7\pi}{4}} & \rightarrow & 1 \\
e^{-\mathbf{n} \frac{\theta}{2}} & 1 & \rightarrow & e^{\mathbf{n} \frac{-\pi}{4}} & \rightarrow & -\mathbf{n} & \rightarrow & e^{\mathbf{n} \frac{-3\pi}{4}} & \rightarrow & -1 & \rightarrow & e^{\mathbf{n} \frac{-5\pi}{4}} & \rightarrow & \mathbf{n} & \rightarrow & e^{\mathbf{n} \frac{-7\pi}{4}} & \rightarrow & 1 \\
\mathbf{v}' &= e^{\mathbf{n} \frac{\theta}{2}} \mathbf{v} e^{-\mathbf{n} \frac{\theta}{2}} & \mathbf{v} & \rightarrow & & \rightarrow & -nvn & \rightarrow & & \rightarrow & \mathbf{v} & \rightarrow & & \rightarrow & -nvn & \rightarrow & \mathbf{v}
\end{array}$$

$$SU(2) = \left\{ 1(\cdot), e^{\mathbf{n} \frac{\theta}{2}}(\cdot) e^{-\mathbf{n} \frac{\theta}{2}} \right\}$$

3D rotation about an arbitrary axis^[34.1.4.2.1.3]

$$\mathbf{V}' = SR_z(\gamma) S^{-1} \mathbf{V}$$

$$\mathbf{v}' = SR_z(\theta) S^{-1} \mathbf{v}$$

$$SO(3) = \left\{ 1, SR_z(\theta) S^{-1} \right\} = \left\{ \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}, SR_z(\theta) S^{-1} \right\}$$

$$SU(2) \cong 2 \cdot SO(3)$$

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/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

$$\mathbf{v}' = e^{\mathbf{n} \frac{\theta}{2}} \mathbf{v} e^{-\mathbf{n} \frac{\theta}{2}}$$

$$\mathbf{v}' = e^{-\mathbf{n} \frac{\theta}{2}} \mathbf{v} e^{\mathbf{n} \frac{\theta}{2}}$$

right-hand = anticlockwise

$$[\mathbf{v}']_R = \left[e^{\mathbf{n} \frac{\theta}{2}} \right]_R [\mathbf{v}]_R \left[e^{-\mathbf{n} \frac{\theta}{2}} \right]_R$$

$$[\mathbf{v}']_L = \left[e^{-\mathbf{n} \frac{\theta}{2}} \right]_R [\mathbf{v}]_L \left[e^{\mathbf{n} \frac{\theta}{2}} \right]_R$$

$$[\mathbf{v}']_L = \left[e^{\mathbf{n} \frac{\theta}{2}} \right]_L [\mathbf{v}]_L \left[e^{-\mathbf{n} \frac{\theta}{2}} \right]_L$$

34.2.6.2 Rodrigues rotation in right-hand vs. left-hand coordinates

$$\mathbf{u} = \mathbf{n} \times \mathbf{v}_\perp$$

$$[\mathbf{u}]_R = [\mathbf{n}]_R \times_R [\mathbf{v}_\perp]_R = [\mathbf{n}]_L \times_L [\mathbf{v}_\perp]_L = [\mathbf{u}]_L$$

$$\begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \times_R \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \times_L \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}$$

$$[[\mathbf{u}]_L]_R = -[\mathbf{u}]_R$$

$$\left[\begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}_L \right]_R = - \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 0 \\ -1 \\ 0 \end{pmatrix}_R$$

$$\det O = \pm 1$$

$\det O = -1$ 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/040072a>

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_\(mathematics\)](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

$$\begin{aligned} & \forall \langle M_1, M_2 \rangle \in \mathcal{M}^2, \exists M_1 M_2 \in \mathcal{M} [M_1 M_2 = M_1 \cdot M_2] \\ \Leftrightarrow & \cdot : \mathcal{M} \times \mathcal{M} = \mathcal{M}^2 \rightarrow \mathcal{M} \\ \Leftrightarrow & \cdot : \mathcal{M} \times \mathcal{M} \rightarrow \mathcal{M} \\ \Leftrightarrow & \mathcal{M} \times \mathcal{M} \xrightarrow{\cdot} \mathcal{M} \\ \Leftrightarrow & \mathcal{M}^2 \xrightarrow{\cdot} \mathcal{M} \end{aligned}$$

matrix group

$$\begin{cases} \forall \langle M_1, M_2, M_3 \rangle \in \mathcal{M}^3 [M_1 (M_2 M_3) = (M_1 M_2) M_3] & \text{associativity} \\ \exists I = I_2 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \in \mathcal{M}, \forall M \in \mathcal{M} [IM = M] & \text{left unit element} \\ \forall M \in \mathcal{M}, \exists M^{-1} \in \mathcal{M} [M^{-1} M = I] & \text{left inverse (element)} \end{cases}$$

$\Rightarrow \mathcal{M} = (\mathcal{M}, \cdot)$ is a matrix group

37.2 group definition and basic theorem

[https://en.wikipedia.org/wiki/Group_\(mathematics\)#Elementary_consequences_of_the_group_axioms](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

$$\begin{cases} \circ : G \times G = G^2 \rightarrow G & \text{binary operation} \\ \forall \langle g_1, g_2, g_3 \rangle \in G^3 [g_1 \circ (g_2 \circ g_3) = (g_1 \circ g_2) \circ g_3] & \text{associativity} \\ \exists e \in G, \forall g \in G [e \circ g = g] & \text{left unit element} \\ \forall g \in G, \exists g^{-1} \in G [g^{-1} \circ g = e] & \text{left inverse (element)} \end{cases}$$

$\Leftrightarrow G = (G, \circ)$ is a group

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

$$\begin{aligned} G &= \{1, i, -1, -i\} \\ &= \{1^0, i^1, i^2, i^3\} \end{aligned}$$

$$\begin{aligned} \forall \langle g_1, g_2 \rangle \in G^2, \exists g_1 g_2 \in G [g_1 g_2 = g_1 \cdot g_2] \\ \Leftrightarrow \cdot : G \times G = G^2 \rightarrow G \end{aligned}$$

<https://tex.stackexchange.com/questions/627708/tikz-how-to-put-tables-within-arbitrarily-placed-nodes>

.	1	i	-1	-i
1	1	i	-1	-i
i	i	-1	-i	1
-1	-1	-i	1	i
-i	-i	1	i	-1

.	1	i	-1	-i
1	1	i	-1	-i
i	i	-1	-i	1
-1	-1	-i	1	i
-i	-i	1	i	-1

Figure 37.1: complex basis group table

37.4.2 $\mathbb{C} \rightarrow \mathcal{M}_{2 \times 2}(\mathbb{R})$

$$1 \leftrightarrow \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = I_2 = I$$

$$\begin{aligned} c_1 \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} + c_2 \begin{pmatrix} a & b \\ c & d \end{pmatrix} &= x \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} + y \begin{pmatrix} a & b \\ c & d \end{pmatrix}, \langle a, b, c, d \rangle \in \mathbb{R}^4 \\ &= xI + yJ, J = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \mathcal{M}_{2 \times 2}(\mathbb{R}), \langle x, y \rangle \in \mathbb{R}^2 \end{aligned}$$

$$J^2 = -I$$

$$J^2 = -I$$

$$\begin{pmatrix} a^2 + bc & ab + bd \\ ca + cd & cb + d^2 \end{pmatrix} = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \begin{pmatrix} a & b \\ c & d \end{pmatrix} = - \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix}$$

$$\begin{cases} a^2 + bc = -1 & b = 0 \Rightarrow a^2 = -1 \Rightarrow a \in \mathbb{R} \Rightarrow b \neq 0 \\ ab + bd = 0 & (b = 0) \vee (a = -d) \xrightarrow{b \neq 0} a = -d \\ ca + cd = 0 & \text{if } a = -d \\ cb + d^2 = -1 & a^2 = d^2 \Rightarrow (a = d) \vee (a = -d) \end{cases} \Rightarrow a = d = 0 \Rightarrow bc = -1$$

$$J = \begin{pmatrix} a & b \\ c & d \end{pmatrix} = \begin{pmatrix} a & b \\ \frac{-a^2 - 1}{b} & -a \end{pmatrix} = J(a, b), b \neq 0$$

$$J(a, b) = \begin{pmatrix} a & b \\ \frac{-a^2 - 1}{b} & -a \end{pmatrix}, b \neq 0$$

$$J(a=1, b) = \begin{pmatrix} 1 & b \\ -2 & -1 \end{pmatrix} \Rightarrow J^2(a=1, b) = \begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix} = -I$$

$$xI + yJ(a=1, b) = \begin{pmatrix} x+y & yb \\ y \cdot \frac{-2}{b} & x-y \end{pmatrix}$$

$$J(a=0, b) = \begin{pmatrix} 0 & b \\ -1 & 0 \end{pmatrix}$$

$$J(a=0, b=1) = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$$

$$\begin{aligned} J(a=0, b=-1) &= \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} = - \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} = -J(a=0, b=1) \\ \Rightarrow J^2(a=0, b=-1) &= \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} = \begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix} = -I \end{aligned}$$

$$\begin{aligned} J &= J(a=0, b=-1) = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \\ &\Rightarrow \begin{cases} 1 \leftrightarrow I = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \\ i \leftrightarrow J = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \end{cases} \\ &\Rightarrow x + yi \leftrightarrow xI + yJ \\ &= x \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} + y \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} = \begin{pmatrix} x & -y \\ y & x \end{pmatrix} \end{aligned}$$

$$x + yi \leftrightarrow \begin{pmatrix} x & -y \\ y & x \end{pmatrix} = xI + yJ$$

realizing

$$\mathbb{C} \rightarrow \mathcal{M}_{2 \times 2}(\mathbb{R}) = \mathcal{M}_2(\mathbb{R})$$

37.4.3 (determinant of complex group representation) equivalent to (squared modulus of complex number)

$$\det(xI + yJ) = \det \begin{pmatrix} x & -y \\ y & x \end{pmatrix} = \begin{vmatrix} x & -y \\ y & x \end{vmatrix} = x^2 + y^2 = |x + yi|^2$$

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]

$$\begin{aligned} & \det [(aI + bJ)(cI + dJ)] \\ &= \det \left[\begin{pmatrix} a & -b \\ b & a \end{pmatrix} \begin{pmatrix} c & -d \\ d & c \end{pmatrix} \right] \\ &= \det \begin{pmatrix} ac - bd & -ad - bc \\ ad + bc & ac - bd \end{pmatrix} = |(ac - bd) + (ad + bc)i|^2 = (ac - bd)^2 + (ad + bc)^2 \\ &= \left[\det \begin{pmatrix} a & -b \\ b & a \end{pmatrix} \right] \left[\det \begin{pmatrix} c & -d \\ d & c \end{pmatrix} \right] = |a + bi|^2 |c + di|^2 = (a^2 + b^2)(c^2 + d^2) \end{aligned}$$

$$|a + bi|^2 |c + di|^2 = (a^2 + b^2)(c^2 + d^2) = (ac - bd)^2 + (ad + bc)^2$$

$$\begin{aligned} & \det [(x_1 I + y_1 J)(x_2 I + y_2 J)] \\ &= \det \left[\begin{pmatrix} x_1 & -y_1 \\ y_1 & x_1 \end{pmatrix} \begin{pmatrix} x_2 & -y_2 \\ y_2 & x_2 \end{pmatrix} \right] \\ &= \det \begin{pmatrix} x_1 x_2 - y_1 y_2 & -x_1 y_2 - y_1 x_2 \\ x_1 y_2 + y_1 x_2 & x_1 x_2 - y_1 y_2 \end{pmatrix} = |(x_1 x_2 - y_1 y_2) + (x_1 y_2 + y_1 x_2)i|^2 = (x_1 x_2 - y_1 y_2)^2 + (x_1 y_2 + y_1 x_2)^2 \\ &= \left[\det \begin{pmatrix} x_1 & -y_1 \\ y_1 & x_1 \end{pmatrix} \right] \left[\det \begin{pmatrix} x_2 & -y_2 \\ y_2 & x_2 \end{pmatrix} \right] = |x_1 + y_1 i|^2 |x_2 + y_2 i|^2 = (x_1^2 + y_1^2)(x_2^2 + y_2^2) \end{aligned}$$

$$|x_1 + y_1 i|^2 |x_2 + y_2 i|^2 = (x_1^2 + y_1^2)(x_2^2 + y_2^2) = (x_1 x_2 - y_1 y_2)^2 + (x_1 y_2 + y_1 x_2)^2$$

37.4.4 Euler formula proved by complex group representation

$$\begin{aligned} \begin{pmatrix} x' \\ y' \end{pmatrix} &= \begin{pmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = R_\theta \begin{pmatrix} x \\ y \end{pmatrix} \\ &= \begin{pmatrix} \cos n \frac{\theta}{n} & -\sin n \frac{\theta}{n} \\ \sin n \frac{\theta}{n} & \cos n \frac{\theta}{n} \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} \cos \frac{\theta}{n} & -\sin \frac{\theta}{n} \\ \sin \frac{\theta}{n} & \cos \frac{\theta}{n} \end{pmatrix}^n \begin{pmatrix} x \\ y \end{pmatrix} = R_{\frac{\theta}{n}}^n \begin{pmatrix} x \\ y \end{pmatrix} \end{aligned}$$

$$\begin{aligned} \lim_{n \rightarrow \infty} R_{\frac{\theta}{n}} &= \lim_{n \rightarrow \infty} \begin{pmatrix} \cos \frac{\theta}{n} & -\sin \frac{\theta}{n} \\ \sin \frac{\theta}{n} & \cos \frac{\theta}{n} \end{pmatrix} = \begin{pmatrix} 1 & -\frac{\theta}{n} \\ \frac{\theta}{n} & 1 \end{pmatrix} \\ &= \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} + \frac{\theta}{n} \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \\ &= \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} + \frac{\theta}{n} \begin{pmatrix} \cos \frac{\pi}{2} & -\sin \frac{\pi}{2} \\ \sin \frac{\pi}{2} & \cos \frac{\pi}{2} \end{pmatrix} \\ &= I + \frac{\theta}{n} R_{\frac{\pi}{2}} \end{aligned}$$

$$\lim_{n \rightarrow \infty} R_{\frac{\theta}{n}} = \lim_{n \rightarrow \infty} \begin{pmatrix} \cos \frac{\theta}{n} & -\sin \frac{\theta}{n} \\ \sin \frac{\theta}{n} & \cos \frac{\theta}{n} \end{pmatrix} = \begin{pmatrix} 1 & -\frac{\theta}{n} \\ \frac{\theta}{n} & 1 \end{pmatrix} = I + \frac{\theta}{n} R_{\frac{\pi}{2}}$$

$$\begin{aligned} \begin{pmatrix} x' \\ y' \end{pmatrix} &= \lim_{n \rightarrow \infty} \begin{pmatrix} x' \\ y' \end{pmatrix} = \lim_{n \rightarrow \infty} R_{\frac{\theta}{n}}^n \begin{pmatrix} x \\ y \end{pmatrix} = \left[\lim_{n \rightarrow \infty} R_{\frac{\theta}{n}}^n \right] \left[\lim_{n \rightarrow \infty} \begin{pmatrix} x \\ y \end{pmatrix} \right] = \left[\lim_{n \rightarrow \infty} R_{\frac{\theta}{n}}^n \right] \begin{pmatrix} x \\ y \end{pmatrix} \\ &= \lim_{n \rightarrow \infty} \left[\lim_{n \rightarrow \infty} R_{\frac{\theta}{n}} \right]^n \begin{pmatrix} x \\ y \end{pmatrix} = \lim_{n \rightarrow \infty} \left[I + \frac{\theta}{n} R_{\frac{\pi}{2}} \right]^n \begin{pmatrix} x \\ y \end{pmatrix} \\ &= \lim_{n \rightarrow \infty} \left[I + \frac{\theta}{n} \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \right]^n \begin{pmatrix} x \\ y \end{pmatrix} = \lim_{n \rightarrow \infty} \left[I + \frac{\theta J}{n} \right]^n \begin{pmatrix} x \\ y \end{pmatrix}, J = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \\ &= e^{J\theta} \begin{pmatrix} x \\ y \end{pmatrix} \end{aligned}$$

$$\begin{aligned} \begin{cases} \begin{pmatrix} x' \\ y' \end{pmatrix} = \begin{pmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} \\ \begin{pmatrix} x' \\ y' \end{pmatrix} = e^{J\theta} \begin{pmatrix} x \\ y \end{pmatrix} \end{cases} \\ \Rightarrow e^{J\theta} = \begin{pmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{pmatrix} \stackrel{x+y \leftrightarrow \begin{pmatrix} x & -y \\ y & x \end{pmatrix}}{\Rightarrow} \stackrel{xI+yJ}{e^{i\theta} = \cos \theta + i \sin \theta} \end{aligned}$$

□

37.5 quaternion group representation

https://groupprops.subwiki.org/wiki/Linear_representation_theory_of_quaternion_group#Two-dimensional_irreducible_representation_over_a_splitting_field

$$\begin{aligned} q &= a + bi + cj + dk = a + ib + jc + kd, \begin{cases} q \in \mathbb{H} \\ a, b, c, d \in \mathbb{R} \Leftrightarrow \langle a, b, c, d \rangle \in \mathbb{R}^4 \end{cases} \\ &= w = t + xi + yj + zk = t + ix + jy + kz, \begin{cases} w \in \mathbb{H} \\ t, x, y, z \in \mathbb{R} \Leftrightarrow \langle t, x, y, z \rangle \in \mathbb{R}^4 \end{cases} \\ &= a1 + bi + cj + dk = t1 + xi + yj + zk = x_0 1 + e_i x_i \end{aligned}$$

37.5.1 $\mathbb{H} \rightarrow \mathcal{M}_{2 \times 2}(\mathbb{C})$

$$1 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = 1_2 = 1$$

$$\begin{aligned} e &= \begin{pmatrix} a & b \\ c & d \end{pmatrix} = \begin{pmatrix} a & b \\ \frac{-a^2 - 1}{b} & -a \end{pmatrix} \\ &= \begin{cases} \begin{pmatrix} a & b \\ \frac{-a^2 - 1}{b} & -a \end{pmatrix} = \begin{pmatrix} 0 & b \\ -1 & 0 \end{pmatrix} = \begin{pmatrix} 0 & \beta \\ -\beta & 0 \end{pmatrix} \Rightarrow e_2 = J = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} = j & a = 0 \\ \begin{pmatrix} a & b \\ \frac{-a^2 - 1}{b} & -a \end{pmatrix} = \begin{pmatrix} \alpha & \beta \\ \beta & -\alpha \end{pmatrix} & a \neq 0 \end{cases} \\ &\quad \left(\begin{pmatrix} \alpha^2 + \beta^2 & 0 \\ 0 & \beta^2 + \alpha^2 \end{pmatrix} = \begin{pmatrix} \alpha & \beta \\ \beta & -\alpha \end{pmatrix} \begin{pmatrix} \alpha & \beta \\ \beta & -\alpha \end{pmatrix} = e^2 = -1 = - \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix} \right) \\ &\quad \downarrow \\ &\quad \alpha^2 + \beta^2 = -1 \Leftrightarrow \beta^2 + \alpha^2 = -1 \end{aligned}$$

$$\alpha^2 + \beta^2 = -1 \Rightarrow \langle \alpha, \beta \rangle \notin \mathbb{R}^2 \Rightarrow \langle \alpha, \beta \rangle \in \mathbb{R}^2 \Rightarrow \alpha^2 + \beta^2 \geq 0$$

quaternion group has no irreducible two-dimensional representation over the reals ¹

$$\langle \alpha, \beta \rangle \in \mathbb{C}^2 - \mathbb{R}^2$$

$$\alpha^2 + \beta^2 = -1 = \beta^2 + \alpha^2$$

$$\begin{pmatrix} \beta & \alpha \\ \alpha & -\beta \end{pmatrix} \begin{pmatrix} \beta & \alpha \\ \alpha & -\beta \end{pmatrix} = \begin{pmatrix} \beta^2 + \alpha^2 & 0 \\ 0 & \alpha^2 + \beta^2 \end{pmatrix} = \begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix} = - \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = -1$$

¹https://groupprops.subwiki.org/wiki/Linear_representation_theory_of_quaternion_group#Four-dimensional_irreducible_representation_over_a_non-splitting_field

$$1 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, e_1 = \begin{pmatrix} \alpha & \beta \\ \beta & -\alpha \end{pmatrix}, e_2 = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$$

$$1 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, i = \begin{pmatrix} \alpha & \beta \\ \beta & -\alpha \end{pmatrix}, j = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$$

$$ij = \begin{pmatrix} \alpha & \beta \\ \beta & -\alpha \end{pmatrix} \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} = \begin{pmatrix} \beta & -\alpha \\ -\alpha & -\beta \end{pmatrix} = k$$

$$1 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, e_1 = \begin{pmatrix} \alpha & \beta \\ \beta & -\alpha \end{pmatrix}, e_2 = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}, e_3 = \begin{pmatrix} \beta & -\alpha \\ -\alpha & -\beta \end{pmatrix}$$

$$1 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, i = \begin{pmatrix} \alpha & \beta \\ \beta & -\alpha \end{pmatrix}, j = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}, k = \begin{pmatrix} \beta & -\alpha \\ -\alpha & -\beta \end{pmatrix}$$

$$jk = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} \beta & -\alpha \\ -\alpha & -\beta \end{pmatrix} = \begin{pmatrix} \alpha & \beta \\ \beta & -\alpha \end{pmatrix} = i$$

$$\alpha^2 + \beta^2 = -1 \quad \begin{cases} \alpha = \sqrt{-1} \\ \beta = 0 \end{cases} \quad \begin{cases} \alpha = \sqrt{-2} \\ \beta = 1 \end{cases} \quad \beta = \alpha^2, n \in \{1, 2, 4, 5\}$$

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	$\begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix}$	$\begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix}$	$\begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix}$	$\begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix}$
i	$\begin{pmatrix} \alpha & \beta \\ \beta & -\alpha \end{pmatrix}$	$\begin{pmatrix} \sqrt{-1} & 0 \\ 0 & -\sqrt{-1} \end{pmatrix}$	$\begin{pmatrix} \sqrt{-2} & 1 \\ 1 & -\sqrt{-2} \end{pmatrix}$	$\begin{pmatrix} e^{\pi \frac{n}{3}\sqrt{-1}} & e^{\pi \frac{2n}{3}\sqrt{-1}} \\ e^{\pi \frac{2n}{3}\sqrt{-1}} & -e^{\pi \frac{n}{3}\sqrt{-1}} \end{pmatrix}$
-i	$\begin{pmatrix} -\alpha & -\beta \\ -\beta & \alpha \end{pmatrix}$	$\begin{pmatrix} -\sqrt{-1} & 0 \\ 0 & \sqrt{-1} \end{pmatrix}$	$\begin{pmatrix} -\sqrt{-2} & -1 \\ -1 & \sqrt{-2} \end{pmatrix}$	$\begin{pmatrix} -e^{\pi \frac{n}{3}\sqrt{-1}} & -e^{\pi \frac{2n}{3}\sqrt{-1}} \\ -e^{\pi \frac{2n}{3}\sqrt{-1}} & e^{\pi \frac{n}{3}\sqrt{-1}} \end{pmatrix}$
j	$\begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$	$\begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$	$\begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$	$\begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$
-j	$\begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$	$\begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$	$\begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$	$\begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$
k	$\begin{pmatrix} \beta & -\alpha \\ -\alpha & -\beta \end{pmatrix}$	$\begin{pmatrix} 0 & -\sqrt{-1} \\ -\sqrt{-1} & 0 \end{pmatrix}$	$\begin{pmatrix} 1 & -\sqrt{-2} \\ -\sqrt{-2} & -1 \end{pmatrix}$	$\begin{pmatrix} e^{\pi \frac{2n}{3}\sqrt{-1}} & -e^{\pi \frac{n}{3}\sqrt{-1}} \\ -e^{\pi \frac{n}{3}\sqrt{-1}} & -e^{\pi \frac{2n}{3}\sqrt{-1}} \end{pmatrix}$
-k	$\begin{pmatrix} -\beta & \alpha \\ \alpha & \beta \end{pmatrix}$	$\begin{pmatrix} 0 & \sqrt{-1} \\ \sqrt{-1} & 0 \end{pmatrix}$	$\begin{pmatrix} -1 & \sqrt{-2} \\ \sqrt{-2} & 1 \end{pmatrix}$	$\begin{pmatrix} -e^{\pi \frac{2n}{3}\sqrt{-1}} & e^{\pi \frac{n}{3}\sqrt{-1}} \\ e^{\pi \frac{n}{3}\sqrt{-1}} & e^{\pi \frac{2n}{3}\sqrt{-1}} \end{pmatrix}$

$$-1 = \alpha^2 + \beta^2$$

$$\stackrel{\beta=\alpha^2}{=} \alpha^2 + \alpha^4$$

$$\alpha^4 + \alpha^2 + 1 = 0, \alpha^4 + \alpha^2 + 1 = (\alpha^2 + \alpha + 1)(\alpha^2 - \alpha + 1)$$

$$(\alpha^2 - 1)(\alpha^4 + \alpha^2 + 1) = 0$$

$$\alpha^6 - 1 = 0$$

$$\alpha^6 - 1 = e^{2\pi k\sqrt{-1}}, k \in \mathbb{Z}$$

$$\alpha = e^{2\pi \frac{n}{6}\sqrt{-1}}, n \in \{0, 1, 2, 3, 4, 5\} - \{0, 3\}$$

$$= e^{\pi \frac{n}{3}\sqrt{-1}}, n \in \{1, 2, 4, 5\}$$

$$\begin{array}{llll}
 \alpha^2 + \beta^2 = -1 & \begin{cases} \alpha = i \\ \beta = 0 \end{cases} & \begin{cases} \alpha = \sqrt{2}i \\ \beta = 1 \end{cases} & \omega = e^{i\pi\frac{n}{3}}, n \in \{1, 2, 4, 5\} \\
 1 & \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} & \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} & \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \\
 -1 & \begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix} & \begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix} & \begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix} \\
 i & \begin{pmatrix} \alpha & \beta \\ \beta & -\alpha \end{pmatrix} & \begin{pmatrix} i & 0 \\ 0 & -i \end{pmatrix} & \begin{pmatrix} \sqrt{2}i & 1 \\ 1 & -\sqrt{2}i \end{pmatrix} \\
 -i & \begin{pmatrix} -\alpha & -\beta \\ -\beta & \alpha \end{pmatrix} & \begin{pmatrix} -i & 0 \\ 0 & i \end{pmatrix} & \begin{pmatrix} -\sqrt{2}i & -1 \\ -1 & \sqrt{2}i \end{pmatrix} \\
 j & \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} & \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} & \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \\
 -j & \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} & \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} & \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \\
 k & \begin{pmatrix} \beta & -\alpha \\ -\alpha & -\beta \end{pmatrix} & \begin{pmatrix} 0 & -i \\ -i & 0 \end{pmatrix} & \begin{pmatrix} 1 & -\sqrt{2}i \\ -\sqrt{2}i & -1 \end{pmatrix} \\
 -k & \begin{pmatrix} -\beta & \alpha \\ \alpha & \beta \end{pmatrix} & \begin{pmatrix} 0 & i \\ i & 0 \end{pmatrix} & \begin{pmatrix} -1 & \sqrt{2}i \\ \sqrt{2}i & 1 \end{pmatrix}
 \end{array}$$

realizing

$$\mathbb{H} \rightarrow \mathcal{M}_{2 \times 2}(\mathbb{C}) = \mathcal{M}_2(\mathbb{C})$$

37.5.1.1 Euler identity

cf. Lagrange identity^[37.4.3.1]

$$\begin{aligned}
 & \det(a + bi + cj + dk) \\
 &= \det \left[a \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} + b \begin{pmatrix} i & 0 \\ 0 & -i \end{pmatrix} + c \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} + d \begin{pmatrix} 0 & -i \\ -i & 0 \end{pmatrix} \right] \\
 &= \det \begin{pmatrix} a + bi & -c - di \\ c - di & a - bi \end{pmatrix} = \begin{vmatrix} a + bi & -c - di \\ c - di & a - bi \end{vmatrix} \\
 &= (a^2 + b^2) + (c^2 + d^2) = a^2 + b^2 + c^2 + d^2
 \end{aligned}$$

$$\det(a + bi + cj + dk) = \det \begin{pmatrix} a + bi & -c - di \\ c - di & a - bi \end{pmatrix} = a^2 + b^2 + c^2 + d^2$$

$$\begin{aligned}
& \det [(q_{10} + q_{11}\mathbf{i} + q_{12}\mathbf{j} + q_{13}\mathbf{k})(q_{20} + q_{21}\mathbf{i} + q_{22}\mathbf{j} + q_{23}\mathbf{k})] = \det [(a + b\mathbf{i} + c\mathbf{j} + d\mathbf{k})(\alpha + \beta\mathbf{i} + \gamma\mathbf{j} + \delta\mathbf{k})] \\
&= \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. \\
&\quad \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. \\
&\quad \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} = (a^2 + b^2 + c^2 + d^2)(\alpha^2 + \beta^2 + \gamma^2 + \delta^2) \\
&= \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 \{(a\alpha - b\beta - c\gamma - d\delta) + (a\beta + b\alpha + c\delta - d\gamma)\mathbf{i} + (a\gamma - b\delta + c\alpha + d\beta)\mathbf{j} + (a\delta + b\gamma - c\beta + d\alpha)\mathbf{k}\} \\
&= (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 \\
&\quad + (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 \\
&\quad (a^2 + b^2 + c^2 + d^2)(\alpha^2 + \beta^2 + \gamma^2 + \delta^2) \\
&\quad = (a\alpha - b\beta - c\gamma - d\delta)^2 + (a\beta + b\alpha + c\delta - d\gamma)^2 \\
&\quad \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.

$$\forall \langle m_1, m_2 \rangle \in (\mathbb{N} \cup \{0\})^2, \exists \langle k_1, k_2, k_3, k_4 \rangle \in (\mathbb{N} \cup \{0\})^4 [m_1 m_2 = k_1^2 + k_2^2 + k_3^2 + k_4^2]$$

Proof:

$$\text{Let } \begin{cases} m_1 = a^2 + b^2 + c^2 + d^2 & \langle a, b, c, d \rangle \in (\mathbb{N} \cup \{0\})^4 \\ m_2 = \alpha^2 + \beta^2 + \gamma^2 + \delta^2 & \langle \alpha, \beta, \gamma, \delta \rangle \in (\mathbb{N} \cup \{0\})^4 \end{cases} \xrightarrow{\text{closure property}} \begin{pmatrix} m_1 \\ m_2 \end{pmatrix} = \begin{pmatrix} a^2 + b^2 + c^2 + d^2 \\ \alpha^2 + \beta^2 + \gamma^2 + \delta^2 \end{pmatrix} \in (\mathbb{N} \cup \{0\})^2,$$

$$\begin{aligned}
m_1 m_2 &= (a^2 + b^2 + c^2 + d^2)(\alpha^2 + \beta^2 + \gamma^2 + \delta^2) \\
&\stackrel{\text{Euler identity}}{=} (a\alpha - b\beta - c\gamma - d\delta)^2 + (a\beta + b\alpha + c\delta - d\gamma)^2 \\
&\quad + (a\gamma - b\delta + c\alpha + d\beta)^2 + (a\delta + b\gamma - c\beta + d\alpha)^2 \\
&= |a\alpha - b\beta - c\gamma - d\delta|^2 \\
&\quad + |a\beta + b\alpha + c\delta - d\gamma|^2 \\
&\quad + |a\gamma - b\delta + c\alpha + d\beta|^2 \\
&\quad + |a\delta + b\gamma - c\beta + d\alpha|^2 \\
&= k_1^2 + k_2^2 + k_3^2 + k_4^2, \begin{pmatrix} k_1 \\ k_2 \\ k_3 \\ k_4 \end{pmatrix} = \begin{pmatrix} |a\alpha - b\beta - c\gamma - d\delta| \\ |a\beta + b\alpha + c\delta - d\gamma| \\ |a\gamma - b\delta + c\alpha + d\beta| \\ |a\delta + b\gamma - c\beta + d\alpha| \end{pmatrix} \in (\mathbb{N} \cup \{0\})^4
\end{aligned}$$

$$\begin{aligned} & \therefore \begin{cases} \langle a, b, c, d \rangle \in (\mathbb{N} \cup \{0\})^4 \\ \langle \alpha, \beta, \gamma, \delta \rangle \in (\mathbb{N} \cup \{0\})^4 \end{cases} \\ \xrightarrow{\text{closure property}} & \begin{pmatrix} |a\alpha - b\beta - c\gamma - d\delta| \\ |a\beta + b\alpha + c\delta - d\gamma| \\ |a\gamma - b\delta + c\alpha + d\beta| \\ |a\delta + b\gamma - c\beta + d\alpha| \end{pmatrix} \in (\mathbb{N} \cup \{0\})^4 \end{aligned}$$

□

37.5.2 $\mathbb{H} \rightarrow \mathcal{M}_{4 \times 4}(\mathbb{R})$

https://groupprops.subwiki.org/wiki/Linear_representation_theory_of_quaternion_group#Four-dimensional_irreducible_representation_over_a_non-splitting_field

$$\mathbb{H} \rightarrow \mathcal{M}_2(\mathbb{C}) \xrightarrow{\mathbb{C} \rightarrow \mathcal{M}_2(\mathbb{R})} \mathcal{M}_{4 \times 4}(\mathbb{R}) = \mathcal{M}_4(\mathbb{R})$$

$$\mathbb{C} \rightarrow \mathcal{M}_2(\mathbb{R}) \leftarrow \begin{cases} 1 \rightarrow \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \\ i \rightarrow \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \end{cases}$$

$$\begin{aligned} 1 & \rightarrow \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \rightarrow \begin{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} & \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix} \\ \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix} & \begin{pmatrix} 0 & 1 \\ 0 & 1 \end{pmatrix} \end{pmatrix} \rightarrow \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \\ -1 & \rightarrow \begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix} \rightarrow \begin{pmatrix} \begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix} & \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix} \\ \begin{pmatrix} 0 & 0 \\ -1 & 0 \end{pmatrix} & \begin{pmatrix} 0 & 1 \\ 0 & -1 \end{pmatrix} \end{pmatrix} \rightarrow \begin{pmatrix} -1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix} \\ i & \rightarrow \begin{pmatrix} i & 0 \\ 0 & -i \end{pmatrix} \rightarrow \begin{pmatrix} \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} & \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix} \\ \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} & \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \end{pmatrix} \rightarrow \begin{pmatrix} 0 & -1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & -1 & 0 \end{pmatrix} \\ -i & \rightarrow \begin{pmatrix} -i & 0 \\ 0 & i \end{pmatrix} \rightarrow \begin{pmatrix} \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} & \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix} \\ \begin{pmatrix} 0 & 0 \\ 0 & -1 \end{pmatrix} & \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix} \end{pmatrix} \rightarrow \begin{pmatrix} 0 & 1 & 0 & 0 \\ -1 & 0 & 0 & 0 \\ 0 & 0 & 0 & -1 \\ 0 & 0 & 1 & 0 \end{pmatrix} \\ j & \rightarrow \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \rightarrow \begin{pmatrix} \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix} & \begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix} \\ \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} & \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix} \end{pmatrix} \rightarrow \begin{pmatrix} 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{pmatrix} \\ -j & \rightarrow \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \rightarrow \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} \rightarrow \begin{pmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ -1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \end{pmatrix} \\ k & \rightarrow \begin{pmatrix} 0 & -i \\ -i & 0 \end{pmatrix} \rightarrow \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} \rightarrow \begin{pmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & -1 & 0 \\ 0 & 1 & 0 & 0 \\ -1 & 0 & 0 & 0 \end{pmatrix} \\ -k & \rightarrow \begin{pmatrix} 0 & i \\ i & 0 \end{pmatrix} \rightarrow \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} \rightarrow \begin{pmatrix} 0 & 0 & 0 & -1 \\ 0 & 0 & 1 & 0 \\ 0 & -1 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{pmatrix} \end{aligned}$$

some examinations

$$ij \rightarrow \begin{pmatrix} 0 & -1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & -1 & 0 \end{pmatrix} \begin{pmatrix} 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{pmatrix} = \begin{pmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & -1 & 0 \\ 0 & 1 & 0 & 0 \\ -1 & 0 & 0 & 0 \end{pmatrix} \leftarrow k$$

$$ji \rightarrow \begin{pmatrix} 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{pmatrix} \begin{pmatrix} 0 & -1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & -1 & 0 \end{pmatrix} = \begin{pmatrix} 0 & 0 & 0 & -1 \\ 0 & 0 & 1 & 0 \\ 0 & -1 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{pmatrix} \leftarrow -k$$

$$i^2 = ii \rightarrow \begin{pmatrix} 0 & -1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & -1 & 0 \end{pmatrix} \begin{pmatrix} 0 & -1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & -1 & 0 \end{pmatrix} = \begin{pmatrix} -1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix} \leftarrow -1$$

	$\begin{cases} \alpha = i \\ \beta = 0 \end{cases}$	$\begin{cases} 1 \rightarrow \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \\ i \rightarrow \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \end{cases}$	$\begin{cases} \alpha = \sqrt{2}i \\ \beta = 1 \end{cases}$
1	$\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$	$\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$	$\begin{cases} 1 \rightarrow \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \\ i \rightarrow \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \end{cases}$
-1	$\begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix}$	$\begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix}$	$\begin{pmatrix} -1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix}$
i	$\begin{pmatrix} i & 0 \\ 0 & -i \end{pmatrix}$	$\begin{pmatrix} 0 & -1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & -1 & 0 \end{pmatrix} \quad \begin{pmatrix} \sqrt{2}i & 1 \\ 1 & -\sqrt{2}i \end{pmatrix}$	$\begin{pmatrix} 0 & -\sqrt{2} & 1 & 0 \\ \sqrt{2} & 0 & 0 & 1 \\ 1 & 0 & 0 & \sqrt{2} \\ 0 & 1 & -\sqrt{2} & 0 \end{pmatrix}$
$-i$	$\begin{pmatrix} -i & 0 \\ 0 & i \end{pmatrix}$	$\begin{pmatrix} 0 & 1 & 0 & 0 \\ -1 & 0 & 0 & 0 \\ 0 & 0 & 0 & -1 \\ 0 & 0 & 1 & 0 \end{pmatrix} \quad \begin{pmatrix} -\sqrt{2}i & -1 \\ -1 & \sqrt{2}i \end{pmatrix}$	$\begin{pmatrix} 0 & -1 & -1 & 0 \\ 1 & 0 & 0 & -1 \\ -1 & 0 & 0 & 1 \\ 0 & -1 & -1 & 0 \end{pmatrix}$
j	$\begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$	$\begin{pmatrix} 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{pmatrix} \quad \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$	$\begin{pmatrix} 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{pmatrix}$
$-j$	$\begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$	$\begin{pmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ -1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \end{pmatrix} \quad \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$	$\begin{pmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ -1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \end{pmatrix}$
k	$\begin{pmatrix} 0 & -i \\ -i & 0 \end{pmatrix}$	$\begin{pmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & -1 & 0 \\ 0 & 1 & 0 & 0 \\ -1 & 0 & 0 & 0 \end{pmatrix} \quad \begin{pmatrix} 1 & -\sqrt{2}i \\ -\sqrt{2}i & -1 \end{pmatrix}$	$\begin{pmatrix} 1 & 0 & 0 & \sqrt{2} \\ 0 & 1 & -\sqrt{2} & 0 \\ 0 & \sqrt{2} & -1 & 0 \\ -\sqrt{2} & 0 & 0 & -1 \end{pmatrix}$
$-k$	$\begin{pmatrix} 0 & i \\ i & 0 \end{pmatrix}$	$\begin{pmatrix} 0 & 0 & 0 & -1 \\ 0 & 0 & 1 & 0 \\ 0 & -1 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{pmatrix} \quad \begin{pmatrix} -1 & \sqrt{2}i \\ \sqrt{2}i & 1 \end{pmatrix}$	$\begin{pmatrix} -1 & 0 & 0 & -\sqrt{2} \\ 0 & -1 & \sqrt{2} & 0 \\ 0 & \sqrt{2} & 1 & 0 \\ \sqrt{2} & 0 & 0 & 1 \end{pmatrix}$

some examinations

$$ij \rightarrow \begin{pmatrix} 0 & -\sqrt{2} & 1 & 0 \\ \sqrt{2} & 0 & 0 & 1 \\ 1 & 0 & 0 & \sqrt{2} \\ 0 & 1 & -\sqrt{2} & 0 \end{pmatrix} \begin{pmatrix} 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 & \sqrt{2} \\ 0 & 1 & -\sqrt{2} & 0 \\ 0 & \sqrt{2} & -1 & 0 \\ -\sqrt{2} & 0 & 0 & -1 \end{pmatrix} \leftarrow k$$

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/%E7%BA%BF%E6%8D%80%E6%95%99%E5%8A%A8%E4%BD%9C/>

39.1 linear equations

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n &= b_1 \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n &= b_2 \\ &\vdots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n &= b_m \end{aligned}$$

$$\begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{pmatrix}$$

$$A\mathbf{x} = \mathbf{b}$$

$$(a_{i1} \quad a_{i2} \quad \cdots \quad a_{in}) \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} = (b_i) = b_i$$

$$(a_{i1} \quad \cdots \quad a_{ij} \quad \cdots \quad a_{in}) \begin{pmatrix} x_1 \\ \vdots \\ x_j \\ \vdots \\ x_n \end{pmatrix} = (b_i) = b_i$$

$$(\cdots \quad a_{ij} \quad \cdots) \begin{pmatrix} \vdots \\ x_j \\ \vdots \end{pmatrix} = b_i$$

$$\mathbf{a}_i^\top \mathbf{x} = \mathbf{a}_i \cdot \mathbf{x} = b_i$$

$$a_{ij}x_j = \mathbf{a}_i^\top \mathbf{x} = \mathbf{a}_i \cdot \mathbf{x} = b_i$$

$$\mathbf{a}_i \cdot \mathbf{x} = \mathbf{a}_i^\top \mathbf{x} = a_{ij}x_j = \cdots + a_{ij}x_j + \cdots$$

if finite,

$$\mathbf{a}_i \cdot \mathbf{x} = \mathbf{a}_i^\top \mathbf{x} = a_{ij}x_j = a_{i1}x_1 + \cdots + a_{ij}x_j + \cdots + a_{in}x_n$$

39.2 matrix multiplication

$$\begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix} \begin{pmatrix} x_{11} & x_{12} & \cdots & x_{1p} \\ x_{21} & x_{22} & \cdots & x_{2p} \\ \vdots & \vdots & \ddots & \vdots \\ x_{n1} & x_{n2} & \cdots & x_{np} \end{pmatrix} = \begin{pmatrix} b_{11} & b_{12} & \cdots & b_{1p} \\ b_{21} & b_{22} & \cdots & b_{2p} \\ \vdots & \vdots & \ddots & \vdots \\ b_{m1} & b_{m2} & \cdots & b_{mp} \end{pmatrix}$$

$$AX = B$$

$$(a_{i1} \quad a_{i2} \quad \cdots \quad a_{in}) \begin{pmatrix} x_{1k} \\ x_{2k} \\ \vdots \\ x_{nk} \end{pmatrix} = (b_{ik}) = b_{ik}$$

$$(a_{i1} \quad \cdots \quad a_{ij} \quad \cdots \quad a_{in}) \begin{pmatrix} x_{1k} \\ \vdots \\ x_{jk} \\ \vdots \\ x_{nk} \end{pmatrix} = (b_{ik}) = b_{ik}$$

$$(\cdots \quad a_{ij} \quad \cdots) \begin{pmatrix} \vdots \\ x_{jk} \\ \vdots \end{pmatrix} = b_{ik}$$

$$a_{ij}x_j = \mathbf{a}_i^\top \mathbf{x} = \mathbf{a}_i \cdot \mathbf{x} = b_i$$

$$\mathbf{a}_i \cdot \mathbf{x} = \mathbf{a}_i^\top \mathbf{x} = a_{ij}x_j$$

\iddots in MathJax

<https://math.meta.stackexchange.com/questions/23273/mathjax-and-iddots-udots-or-reflectbox>

$$b_{ik} = a_{ij}x_{jk} = \mathbf{a}_i \cdot \mathbf{x}_k = \mathbf{a}_i^\top \mathbf{x}_k$$

$$\text{row}(A) \text{col}(X) = b_{\text{row}, \text{col}}$$

$$\mathbf{a}_i \cdot \mathbf{x}_k = \mathbf{a}_i^\top \mathbf{x}_k = a_{ij}x_{jk}$$

$$\mathbf{a}_i \cdot \mathbf{x}_j = \mathbf{a}_i^\top \mathbf{x}_j = a_{ik}x_{kj}$$

$$\mathbf{a}_i \cdot \mathbf{x}_j = \mathbf{a}_i^\top \mathbf{x}_j = a_{ik}x_{kj} = \cdots + a_{ik}x_{kj} + \cdots$$

if finite,

$$\mathbf{a}_i \cdot \mathbf{x}_j = \mathbf{a}_i^\top \mathbf{x}_j = a_{ik}x_{kj} = a_{i1}x_{1j} + \cdots + a_{ik}x_{kj} + \cdots + a_{in}x_{nj}$$

39.3 functional

(inner product or dot product) or linear equations

$$\mathbf{a}_i \cdot \mathbf{x} = \mathbf{a}_i^\top \mathbf{x} = a_{ij}x_j = \cdots + a_{ij}x_j + \cdots$$

if finite,

$$\mathbf{a}_i \cdot \mathbf{x} = \mathbf{a}_i^\top \mathbf{x} = a_{ij}x_j = a_{i1}x_1 + \cdots + a_{ij}x_j + \cdots + a_{in}x_n$$

actually, several rows

$$\begin{array}{ccccccc} \vdots & & \vdots & & \vdots & & \vdots \\ \mathbf{a}_i \cdot \mathbf{x} & = & \mathbf{a}_i^\top \mathbf{x} & = & a_{ij}x_j & = & \cdots + a_{ij}x_j + \cdots \\ \vdots & & \vdots & & \vdots & & \vdots \end{array}$$

if finite,

$$\begin{array}{ccccccc} \vdots & & \vdots & & \vdots & & \vdots \\ \mathbf{a}_i \cdot \mathbf{x} & = & \mathbf{a}_i^\top \mathbf{x} & = & a_{ij}x_j & = & a_{i1}x_1 + \cdots + a_{ij}x_j + \cdots + a_{in}x_n \\ \vdots & & \vdots & & \vdots & & \vdots \end{array}$$

in functional aspect,

$$\begin{array}{ccccccc} \vdots & & \vdots & & \vdots & & \vdots \\ f_i(\mathbf{x}) & = & \mathbf{a}_i \cdot \mathbf{x} & = & \mathbf{a}_i^\top \mathbf{x} & = & a_{ij}x_j = \cdots + a_{ij}x_j + \cdots \\ \vdots & & \vdots & & \vdots & & \vdots \end{array}$$

if finite,

$$\begin{array}{ccccccc} \vdots & & \vdots & & \vdots & & \vdots \\ f_i(\mathbf{x}) & = & \mathbf{a}_i \cdot \mathbf{x} & = & \mathbf{a}_i^\top \mathbf{x} & = & a_{ij}x_j = a_{i1}x_1 + \cdots + a_{ij}x_j + \cdots + a_{in}x_n \\ \vdots & & \vdots & & \vdots & & \vdots \end{array}$$

$$\begin{array}{ccccccc} \vdots & & \vdots & & \vdots & & \vdots \\ f_i(\cdots, x_j, \cdots) & = & f_i(x_j) & = & f_i(\mathbf{x}) & = & \mathbf{a}_i \cdot \mathbf{x} = \mathbf{a}_i^\top \mathbf{x} = a_{ij}x_j = \cdots + a_{ij}x_j + \cdots \\ \vdots & & \vdots & & \vdots & & \vdots \end{array}$$

if finite,

$$\begin{array}{ccccccc} \vdots & & \vdots & & \vdots & & \vdots \\ f_i(x_1, \cdots, x_j, \cdots, x_n) & = & f_i(x_j) & = & f_i(\mathbf{x}) & = & \mathbf{a}_i \cdot \mathbf{x} = \mathbf{a}_i^\top \mathbf{x} = a_{ij}x_j = a_{i1}x_1 + \cdots + a_{ij}x_j + \cdots + a_{in}x_n \\ \vdots & & \vdots & & \vdots & & \vdots \end{array}$$

or simply

$$f_i(\cdots, x_j, \cdots) = f_i(x_j) = f_i(\mathbf{x}) = \mathbf{a}_i \cdot \mathbf{x}_j = \mathbf{a}_i^\top \mathbf{x}_j = a_{ik}x_{kj} = \cdots + a_{ij}x_j + \cdots$$

if scalar with complex as the field,

$$f_i : \mathbb{C}^\infty \rightarrow \mathbb{C}$$

if scalar with a field,

$$f_i : \mathbb{F}^\infty \rightarrow \mathbb{F}$$

or more abstract notation,

$$f_i : F^\infty \rightarrow F$$

if scalar with real as the field,

$$f_i : \mathbb{R}^\infty \rightarrow \mathbb{R}$$

if finite,

$$f_i(x_1, \dots, x_j, \dots, x_n) = f_i(x_j) = f_i(\mathbf{x}) = \mathbf{a}_i \cdot \mathbf{x}_j = \mathbf{a}_i^\top \mathbf{x}_j = a_{ij}x_j = a_{i1}x_1 + \dots + a_{ij}x_j + \dots + a_{in}x_n$$

$$\begin{aligned} f_i(x_1, x_2, \dots, x_n) &= f_i(x_1, \dots, x_j, \dots, x_n) = f_i(x_j) = f_i(\mathbf{x}) \\ &= \mathbf{a}_i \cdot \mathbf{x}_j = \mathbf{a}_i^\top \mathbf{x}_j = a_{ij}x_j = a_{i1}x_1 + \dots + a_{ij}x_j + \dots + a_{in}x_n \end{aligned}$$

if scalar with complex as the field,

$$f_i : \mathbb{C}^n \rightarrow \mathbb{C}$$

if scalar with a field,

$$f_i : \mathbb{F}^n \rightarrow \mathbb{F}$$

or more abstract notation,

$$f_i : F^n \rightarrow F$$

if scalar with real as the field,

$$f_i : \mathbb{R}^n \rightarrow \mathbb{R}$$

functionals are a set of fuctions mapping n -dimensional vectors to scalars

$$f_i : F^n \rightarrow F$$

$$\begin{array}{ccccccccc} \vdots & & \vdots & & \vdots & & \vdots & & \vdots \\ f_i(\dots, x_j, \dots) & = & f_i(x_j) & = & f_i(\mathbf{x}) & = & \mathbf{a}_i \cdot \mathbf{x} & = & \mathbf{a}_i^\top \mathbf{x} = a_{ij}x_j = \dots + a_{ij}x_j + \dots \\ \vdots & & \vdots & & \vdots & & \vdots & & \vdots \end{array}$$

$$f = \{f_i | f_i : F^\infty \rightarrow F\} = \left\{ \begin{array}{c} f_i(\dots, x_j, \dots) = f_i(x_j) = f_i(\mathbf{x}) = \mathbf{a}_i \cdot \mathbf{x} = \mathbf{a}_i^\top \mathbf{x} = a_{ij}x_j, \\ \vdots \end{array} \right\}$$

if finite,

$$\begin{array}{ccccccccc} \vdots & & \vdots & & \vdots & & \vdots & & \vdots \\ f_i(x_1, \dots, x_j, \dots, x_n) & = & f_i(x_j) & = & f_i(\mathbf{x}) & = & \mathbf{a}_i \cdot \mathbf{x} & = & \mathbf{a}_i^\top \mathbf{x} = a_{ij}x_j = a_{i1}x_1 + \dots + a_{ij}x_j + \dots + a_{in}x_n \\ \vdots & & \vdots & & \vdots & & \vdots & & \vdots \end{array}$$

$$f = \{f_i | f_i : F^n \rightarrow F\} = \left\{ \begin{array}{l} \vdots \\ f_i(x_1, x_2, \dots, x_n) = f_i(\mathbf{x}) = a_{i1}x_1 + \dots + a_{ij}x_j + \dots + a_{in}x_n, \\ \vdots \end{array} \right\}$$

linear functionals are a set of functions mapping vectors to scalars linearly

$$f = \{f_i | f_i : F^n \rightarrow F\} = \left\{ \begin{array}{l} \vdots \\ f_i(x_1, x_2, \dots, x_n) = f_i(\mathbf{x}) = a_{i1}x_1 + \dots + a_{ij}x_j + \dots + a_{in}x_n, \\ \vdots \end{array} \right\}$$

39.4 definition of linear functional

functionals generalized to general vector space

$$f = \{f_i | f_i : V \rightarrow F\}$$

linear functionals generalized to general vector space is a linear transformation

$$f = \left\{ f_i \left| \begin{array}{l} f_i : V \rightarrow F \\ \forall \langle \mathbf{x}, \mathbf{y} \rangle \in V^2 [f_i(\mathbf{x} + \mathbf{y}) = f_i(\mathbf{x}) + f_i(\mathbf{y})] \\ \forall \mathbf{x} \in V, c \in F [f_i(c\mathbf{x}) = cf_i(\mathbf{x})] \end{array} \right. \right\}$$

$$f = \left\{ f_i \left| \begin{array}{l} f_i : V \rightarrow F \\ \forall \mathbf{x}, \mathbf{y} \in V [f_i(\mathbf{x} + \mathbf{y}) = f_i(\mathbf{x}) + f_i(\mathbf{y})] \\ \forall \mathbf{x} \in V, c \in F [f_i(c\mathbf{x}) = cf_i(\mathbf{x})] \end{array} \right. \right\}$$

$$\text{if } \begin{cases} F = \mathbb{C} \\ V = \mathbb{C}^n \end{cases},$$

$$f = \left\{ f_i \left| \begin{array}{l} f_i : \mathbb{C}^n \rightarrow \mathbb{C} \\ \forall \langle \mathbf{x}, \mathbf{y} \rangle \in (\mathbb{C}^n)^2 [f_i(\mathbf{x} + \mathbf{y}) = f_i(\mathbf{x}) + f_i(\mathbf{y})] \\ \forall \mathbf{x} \in \mathbb{C}^n, c \in \mathbb{C} [f_i(c\mathbf{x}) = cf_i(\mathbf{x})] \end{array} \right. \right\}$$

$$f = \left\{ f_i \left| \begin{array}{l} f_i : \mathbb{C}^n \rightarrow \mathbb{C} \\ \forall \mathbf{x}, \mathbf{y} \in \mathbb{C}^n [f_i(\mathbf{x} + \mathbf{y}) = f_i(\mathbf{x}) + f_i(\mathbf{y})] \\ \forall \mathbf{x} \in \mathbb{C}^n, c \in \mathbb{C} [f_i(c\mathbf{x}) = cf_i(\mathbf{x})] \end{array} \right. \right\}$$

then

$$f_i(\mathbf{x}) = a_{i1}x_1 + \dots + a_{ij}x_j + \dots + a_{in}x_n$$

satisfying

$$\begin{aligned} f_i(\mathbf{x} + \mathbf{y}) &= a_{i1}(x + y)_1 + \dots + a_{ij}(x + y)_j + \dots + a_{in}(x + y)_n \\ &= a_{i1}(x_1 + y_1) + \dots + a_{ij}(x_j + y_j) + \dots + a_{in}(x_n + y_n) \\ &= (a_{i1}x_1 + a_{i1}y_1) + \dots + (a_{ij}x_j + a_{ij}y_j) + \dots + (a_{in}x_n + a_{in}y_n) \\ &= (a_{i1}x_1 + \dots + a_{ij}x_j + \dots + a_{in}x_n) + (a_{i1}y_1 + \dots + a_{ij}y_j + \dots + a_{in}y_n) \\ &= f_i(\mathbf{x}) + f_i(\mathbf{y}) \end{aligned}$$

$$\begin{aligned} f_i(c\mathbf{x}) &= a_{i1}(cx)_1 + \dots + a_{ij}(cx)_j + \dots + a_{in}(cx)_n \\ &= a_{i1}(cx_1) + \dots + a_{ij}(cx_j) + \dots + a_{in}(cx_n) \\ &= c(a_{i1}x_1) + \dots + c(a_{ij}x_j) + \dots + c(a_{in}x_n) \\ &= c(a_{i1}x_1 + \dots + a_{ij}x_j + \dots + a_{in}x_n) \\ &= cf_i(\mathbf{x}) \end{aligned}$$

different functional has different a_{ij}

let

$$a_{ij} = f_i(\mathbf{e}_j), \mathbf{e}_j = \left\langle \underbrace{0, \dots, 0}_{j-1}, 1, 0, \dots, 0 \right\rangle = (0 \ \dots \ 0 \ 1 \ 0 \ \dots \ 0)^\top = \begin{pmatrix} 0 \\ \vdots \\ 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}_{n \times 1} = \begin{pmatrix} 0 \\ \vdots \\ 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}_{n \times 1}$$

if f_i is a linear functional, then

$$\begin{aligned} f_i(\mathbf{x}) &= f_i(x_1 \mathbf{e}_1 + \dots + x_j \mathbf{e}_j + \dots + x_n \mathbf{e}_n) \\ &= f_i(x_1 \mathbf{e}_1) + \dots + f_i(x_j \mathbf{e}_j) + \dots + f_i(x_n \mathbf{e}_n) \\ &= x_1 f_i(\mathbf{e}_1) + \dots + x_j f_i(\mathbf{e}_j) + \dots + x_n f_i(\mathbf{e}_n) \\ &= x_1 a_{i1} + \dots + x_j a_{ij} + \dots + x_n a_{in} \\ &= a_{i1} x_1 + \dots + a_{ij} x_j + \dots + a_{in} x_n \\ &= f_i(\mathbf{x}) \end{aligned}$$

39.5 set of all linear transformations is a vector space

<https://ccjou.wordpress.com/2011/04/08/%E7%BA%BF%E6%8D%A2%E5%85%83%E5%95%9C/>

vector space^[40]

<https://math.stackexchange.com/questions/2381942/the-set-of-all-linear-maps-tv-w-is-a-vector-space>

$$T : V \rightarrow W \Leftrightarrow \forall \mathbf{v} \in V, \exists! \mathbf{w} \in W [\mathbf{w} = T(\mathbf{v})]$$

$$\begin{cases} V, W \text{ are vector spaces} \\ T : V \rightarrow W \\ \begin{cases} \forall \mathbf{u}, \mathbf{v} \in V [T(\mathbf{u} + \mathbf{v}) = T(\mathbf{u}) + T(\mathbf{v})] \\ \forall \mathbf{v} \in V, c \in F [T(c\mathbf{v}) = cT(\mathbf{v})] \end{cases} \quad \text{linearity} \end{cases}$$

$\Leftrightarrow T$ is a linear transformation

$$\begin{cases} V, W \text{ are vector spaces, both over } F \\ T, U : V \rightarrow W \\ \begin{cases} T : V \rightarrow W \\ U : V \rightarrow W \end{cases} \\ T, U \text{ are both linear transformations} \\ \mathbf{v} \in V \\ c \in F \end{cases}$$

There is still linearity over linear transformations

(va)

$$\begin{aligned} (T + U)(\mathbf{u} + \mathbf{v}) &= T(\mathbf{u} + \mathbf{v}) + U(\mathbf{u} + \mathbf{v}) \\ &= [T(\mathbf{u}) + T(\mathbf{v})] + [U(\mathbf{u}) + U(\mathbf{v})] \\ &= [T(\mathbf{u}) + U(\mathbf{u})] + [T(\mathbf{v}) + U(\mathbf{v})] \\ &= (T + U)(\mathbf{u}) + (T + U)(\mathbf{v}) \end{aligned}$$

(sm)

$$\begin{aligned}
 (T + U)(cv) &= T(cv) + U(cv) \\
 &= cT(v) + cU(v) \\
 &= c[T(v) + U(v)] \\
 &= c(T + U)(v)
 \end{aligned}$$

so we can define

$$\begin{cases} (T + U)(v) = T(v) + U(v) & \text{linear transformation addition} \\ (cT)(v) = cT(v) & \text{scalar linear transformation multiplication} \end{cases}$$

the set of all linear transformations is a vector space

\mathcal{T} is the set of all linear transformations

$$\left\{
 \begin{array}{ll}
 F & (f) F \text{ is a field} \\
 \mathcal{T} \neq \emptyset & (ne) \text{ nonempty set} \\
 + : \mathcal{T} \times \mathcal{T} \xrightarrow{\perp} \mathcal{T} \Leftrightarrow \forall T, U \in \mathcal{T}, \exists S \in \mathcal{T} [S = T + U] & (va) \text{ vector addition} \\
 \cdot : F \times \mathcal{T} \xrightarrow{\perp} \mathcal{T} \Leftrightarrow \forall c \in F, \forall T \in \mathcal{T}, \exists U \in \mathcal{T} [U = cT = c \cdot T] & (sm) \text{ scalar multiplication} \\
 \left\{
 \begin{array}{ll}
 \forall S, T, U \in \mathcal{T} [S + (T + U) = (S + T) + U] & (a) \\
 \forall T, U \in \mathcal{T} [T + U = U + T] & (c) \\
 \exists O \in \mathcal{T}, \forall T \in \mathcal{T} [O + T = T] & (e) \\
 \forall T \in \mathcal{T}, \exists! -T \in \mathcal{T} [(-T) + T = O] & (i) \\
 \forall b, c \in F, T \in \mathcal{T} [b(cT) = (bc)T] & (a) \\
 \exists! 1 \in F, \forall T \in \mathcal{T} [1T = T] & (e) \\
 \forall c \in F, T, U \in \mathcal{T} [c(T + U) = cT + cU] & (dv) \\
 \forall b, c \in F, T \in \mathcal{T} [(b + c)T = bT + cT] & (ds)
 \end{array}
 \right. & \text{(va) vector addition axioms} \\
 \Leftrightarrow \mathcal{T} = \mathcal{T}(F, +, \cdot) = (\mathcal{T}, F, +, \cdot) \text{ is a vector space over the field } F & \\
 \Leftrightarrow \mathcal{T} \text{ is a vector space} &
 \end{array}
 \right.$$

Selected proofs of 8 vector space axioms due to some trivial field and vector space properties:

(va) (a)

$$\begin{aligned}
 (S + (T + U))(v) &= S(v) + (T + U)(v) \\
 &= S(v) + T(v) + U(v) \\
 &= (S + T)(v) + U(v) \\
 &= ((S + T) + U)(v)
 \end{aligned}$$

(va) (c)

$$\begin{aligned}
 (T + U)(v) &= T(v) + U(v) \\
 &= U(v) + T(v) \\
 &= (U + T)(v)
 \end{aligned}$$

(va) (e)

$$O(v) = 0w \in W$$

$$\begin{aligned}
 (O + T)(v) &= O(v) + T(v) \\
 &= 0w + T(v) \\
 &= T(v)
 \end{aligned}$$

$$O_1(v) - O_2(v) = 0w - 0w = 0w \Rightarrow O_1(v) = O_2(v)$$

$(sm)(dv)$

$$\begin{aligned}(c(T+U))(\mathbf{v}) &= c(T+U)(\mathbf{v}) \\ &= c[T(\mathbf{v}) + U(\mathbf{v})] \\ &= cT(\mathbf{v}) + cU(\mathbf{v}) \\ &= (cT + cU)(\mathbf{v})\end{aligned}$$

The set of all linear tranformations \mathcal{T} is a vector space.

□

39.6 definition of dual space

$$\begin{aligned}V^* &= L(V, F) \\ &= \left\{ f_i \middle| \begin{array}{l} f_i : V \rightarrow F \\ \left\{ \begin{array}{l} \forall \mathbf{x}, \mathbf{y} \in V [f_i(\mathbf{x} + \mathbf{y}) = f_i(\mathbf{x}) + f_i(\mathbf{y})] \\ \forall \mathbf{x} \in V, c \in F [f_i(c\mathbf{x}) = cf_i(\mathbf{x})] \end{array} \right. \\ \text{(L) linearity} \end{array} \right\} \quad \text{functional mapping vector to field scalar}\end{aligned}$$

$\Leftrightarrow V^*$ is a dual space, a set of linear functionals f_i mapping vectors in the vector space V to scalars in the field F

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/%E9%BE%94%E6%98%87/>

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

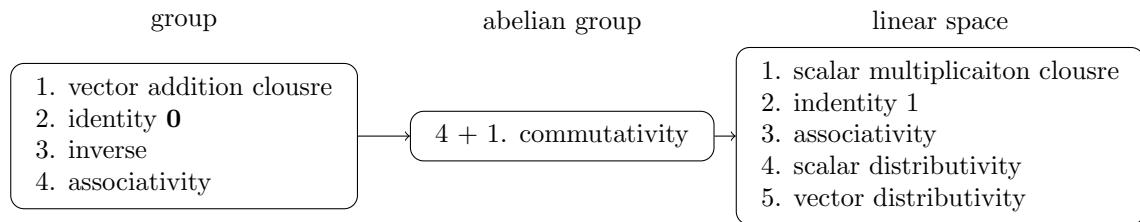


Figure 40.1: vector space construction

$$\left\{ \begin{array}{l} F \text{ is a field} \\ V \neq \emptyset \\ + : V \times V = V^2 \rightrightarrows V \Leftrightarrow \forall \mathbf{u}, \mathbf{v} \in V, \exists! \mathbf{w} \in V [\mathbf{w} = \mathbf{u} + \mathbf{v}] \\ \cdot : F \times V \rightarrow V \Leftrightarrow \forall s \in F, \forall \mathbf{v} \in V, \exists! \mathbf{u} \in V [\mathbf{u} = s\mathbf{v} = s \cdot \mathbf{v}] \\ \left\{ \begin{array}{ll} \exists! \mathbf{0} \in V, \forall \mathbf{v} \in V [\mathbf{0} + \mathbf{v} = \mathbf{v}] & (e) \text{ identity} \\ \forall \mathbf{v} \in V, \exists! -\mathbf{v} \in V [(-\mathbf{v}) + \mathbf{v} = \mathbf{0}] & (i) \text{ inverse} \\ \forall \mathbf{u}, \mathbf{v}, \mathbf{w} \in V [\mathbf{u} + (\mathbf{v} + \mathbf{w}) = (\mathbf{u} + \mathbf{v}) + \mathbf{w}] & (a) \text{ associativity} \\ \forall \mathbf{u}, \mathbf{v} \in V [\mathbf{u} + \mathbf{v} = \mathbf{v} + \mathbf{u}] & (c) \text{ commutativity} \end{array} \right. \\ \left\{ \begin{array}{ll} \exists! 1 \in F, \forall \mathbf{v} \in V [1\mathbf{v} = \mathbf{v}] & (e) \text{ identity} \\ \forall s, t \in F, \mathbf{v} \in V [s(t\mathbf{v}) = (st)\mathbf{v}] & (a) \text{ associativity} \\ \forall s, t \in F, \mathbf{v} \in V [(s+t)\mathbf{v} = s\mathbf{v} + t\mathbf{v}] & (ds) \text{ scalar distributivity} \\ \forall s \in F, \mathbf{u}, \mathbf{v} \in V [s(\mathbf{u} + \mathbf{v}) = su + sv] & (dv) \text{ vector distributivity} \end{array} \right. \end{array} \right. \Rightarrow V = V(F, +, \cdot) = (V, F, +, \cdot) \text{ is a vector space over the field } F \\ \Leftrightarrow V \text{ is a vector space}$$

40.2.1 commutative group structure of vector space

(va) axioms = vector addition axioms

$$\begin{aligned}
& V = (V, +) \text{ is a commutative group} \Leftrightarrow V = (V, +) \text{ is an abelian group} \\
\Leftrightarrow & \left\{ \begin{array}{ll} V = (V, +) = (V, +_V) \text{ is a group} & (g) \text{ group} \\ \forall \mathbf{u}, \mathbf{v} \in V [\mathbf{u} + \mathbf{v} = \mathbf{v} + \mathbf{u}] & (c) \text{ commutativity} \end{array} \right. \\
\Leftrightarrow & \left\{ \begin{array}{lll} + : V \times V = V^2 \xrightarrow{\dagger} V \Leftrightarrow \forall \mathbf{u}, \mathbf{v} \in V, \exists! \mathbf{w} \in V [\mathbf{w} = \mathbf{u} + \mathbf{v}] & (cl) \text{ closure} \\ \exists! \mathbf{0} \in V, \forall \mathbf{v} \in V [\mathbf{0} + \mathbf{v} = \mathbf{v}] & (e) \text{ identity} \\ \forall \mathbf{v} \in V, \exists! -\mathbf{v} \in V [(-\mathbf{v}) + \mathbf{v} = \mathbf{0}] & (i) \text{ inverse} \\ \forall \mathbf{u}, \mathbf{v}, \mathbf{w} \in V [\mathbf{u} + (\mathbf{v} + \mathbf{w}) = (\mathbf{u} + \mathbf{v}) + \mathbf{w}] & (a) \text{ associativity} \\ \forall \mathbf{u}, \mathbf{v} \in V [\mathbf{u} + \mathbf{v} = \mathbf{v} + \mathbf{u}] & (c) \end{array} \right. \quad (g)
\end{aligned}$$

V is a vector space

$\Leftrightarrow V = V(F, +, \cdot) = (V, F, +, \cdot)$ is a vector space over the field F

$$\begin{aligned}
& \Leftrightarrow \left\{ \begin{array}{ll} 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 \xrightarrow{\dagger} V \Leftrightarrow \forall s \in F, \forall \mathbf{v} \in V, \exists! \mathbf{u} \in V [\mathbf{u} = s\mathbf{v} = s \cdot \mathbf{v}] & (sm) \text{ scalar multiplication} \\ \left\{ \begin{array}{ll} \exists! 1 \in F, \forall \mathbf{v} \in V [1\mathbf{v} = \mathbf{v}] & (e) \text{ identity} \\ \forall s, t \in F, \mathbf{v} \in V [s(t\mathbf{v}) = (st)\mathbf{v}] & (a) \text{ associativity} \\ \forall s, t \in F, \mathbf{v} \in V [(s+t)\mathbf{v} = s\mathbf{v} + t\mathbf{v}] & (ds) \text{ scalar distributivity} \\ \forall s \in F, \mathbf{u}, \mathbf{v} \in V [s(\mathbf{u} + \mathbf{v}) = s\mathbf{u} + s\mathbf{v}] & (dv) \text{ vector distributivity} \end{array} \right. & (sm) \text{ axioms} \end{array} \right. \\
& \Leftrightarrow \left\{ \begin{array}{ll} F \text{ is a field} & (f) \text{ field} \\ V \neq \emptyset & (ne) \text{ nonempty set} \\ \left\{ \begin{array}{ll} V = (V, +) = (V, +_V) \text{ is a group} & (g) \text{ group} \\ \forall \mathbf{u}, \mathbf{v} \in V [\mathbf{u} + \mathbf{v} = \mathbf{v} + \mathbf{u}] & (c) \text{ commutativity} \end{array} \right. & (va) \text{ vector addition} \\ \cdot = \cdot_{F \times V} : F \times V \xrightarrow{\dagger} V \Leftrightarrow \forall s \in F, \forall \mathbf{v} \in V, \exists! \mathbf{u} \in V [\mathbf{u} = s\mathbf{v} = s \cdot \mathbf{v}] & (sm) \text{ scalar multiplication} \\ \left\{ \begin{array}{ll} \exists! 1 \in F, \forall \mathbf{v} \in V [1\mathbf{v} = \mathbf{v}] & (e) \text{ identity} \\ \forall s, t \in F, \mathbf{v} \in V [s(t\mathbf{v}) = (st)\mathbf{v}] & (a) \text{ associativity} \\ \forall s, t \in F, \mathbf{v} \in V [(s+t)\mathbf{v} = s\mathbf{v} + t\mathbf{v}] & (ds) \text{ scalar distributivity} \\ \forall s \in F, \mathbf{u}, \mathbf{v} \in V [s(\mathbf{u} + \mathbf{v}) = s\mathbf{u} + s\mathbf{v}] & (dv) \text{ vector distributivity} \end{array} \right. & (sm) \text{ axioms} \end{array} \right. \\
& \Leftrightarrow \left\{ \begin{array}{ll} F = F(+_F, \cdot_F) = (F, +_F, \cdot_F) = (F, +, \cdot) \text{ is a field} & (f) \\ V \neq \emptyset & (ne) \\ \left\{ \begin{array}{ll} + : V \times V = V^2 \xrightarrow{\dagger} V \Leftrightarrow \forall \mathbf{u}, \mathbf{v} \in V, \exists! \mathbf{w} \in V [\mathbf{w} = \mathbf{u} + \mathbf{v}] & (cl) \text{ closure} \\ \exists! \mathbf{0} \in V, \forall \mathbf{v} \in V [\mathbf{0} + \mathbf{v} = \mathbf{v}] & (e) \text{ identity} \\ \forall \mathbf{v} \in V, \exists! -\mathbf{v} \in V [(-\mathbf{v}) + \mathbf{v} = \mathbf{0}] & (i) \text{ inverse} \\ \forall \mathbf{u}, \mathbf{v}, \mathbf{w} \in V [\mathbf{u} + (\mathbf{v} + \mathbf{w}) = (\mathbf{u} + \mathbf{v}) + \mathbf{w}] & (a) \text{ associativity} \\ \forall \mathbf{u}, \mathbf{v} \in V [\mathbf{u} + \mathbf{v} = \mathbf{v} + \mathbf{u}] & (c) \end{array} \right. & (va) \\ \cdot : F \times V \xrightarrow{\dagger} V \Leftrightarrow \forall s \in F, \forall \mathbf{v} \in V, \exists! \mathbf{u} \in V [\mathbf{u} = s\mathbf{v} = s \cdot \mathbf{v}] & (cl) \text{ closure} \\ \exists! 1 \in F, \forall \mathbf{v} \in V [1\mathbf{v} = \mathbf{v}] & (e) \text{ identity} \\ \forall s, t \in F, \mathbf{v} \in V [s(t\mathbf{v}) = s \cdot_{F \times V} (t \cdot_{F \times V} \mathbf{v}) = (s \cdot_F t) \cdot_{F \times V} \mathbf{v} = (st)\mathbf{v}] & (a) \text{ associativity} \quad (sm) \\ \forall s, t \in F, \mathbf{v} \in V [(s+t)\mathbf{v} = (s+_F t)\mathbf{v} = s\mathbf{v} +_V t\mathbf{v}] & (ds) \text{ scalar distributivity} \\ \forall s \in F, \mathbf{u}, \mathbf{v} \in V [s(\mathbf{u} + \mathbf{v}) = s\mathbf{u} + s\mathbf{v}] & (dv) \text{ vector distributivity} \end{array} \right. \end{aligned}$$

40.2.2 scalar distributivity

(sm) (ds)

$$\forall s, t \in F, \mathbf{v} \in V [(s+t)\mathbf{v} = s\mathbf{v} + t\mathbf{v}]$$

$$\forall s, t \in F, \mathbf{v} \in V [(s+_F t)\mathbf{v} = s\mathbf{v} +_V t\mathbf{v}]$$

$$\forall s, t \in F, \mathbf{v} \in V [(s +_F t)\mathbf{v} = s\mathbf{v} +_V t\mathbf{v}]$$

40.3 linearity

$$\begin{aligned} & \begin{cases} f(x+y) = f(x) + f(y) & \text{additivity} \\ f(\lambda x) = \lambda f(x) & \text{homogeneity} \end{cases} \\ \Leftrightarrow & f(\lambda x + y) = \lambda f(x) + f(y) \\ \Leftrightarrow & f \text{ is linear} \end{aligned}$$

40.3.1 linear structure of vector space

$$\forall s \in F, \mathbf{u}, \mathbf{v} \in V [\mathbf{u} + s\mathbf{v} \in V]$$

$$\forall s \in F, \forall \mathbf{u}, \mathbf{v} \in V [\mathbf{u} + s\mathbf{v} \in V]$$

$$\forall s \in F, \langle \mathbf{u}, \mathbf{v} \rangle \in V^2 [\mathbf{u} + s\mathbf{v} \in V]$$

$$\begin{aligned} & \begin{cases} \forall \mathbf{u}, \mathbf{v} \in V [\mathbf{u} + \mathbf{v} \in V] & \text{vector addition closure} \\ \forall s \in F, \mathbf{v} \in V [s\mathbf{v} \in V] & \text{scalar multiplication closure} \end{cases} \\ \Leftrightarrow & \begin{cases} \forall \mathbf{u}, \mathbf{v} \in V [\mathbf{u} + \mathbf{v} \in V] & (a) \text{ additivity} \\ \forall s \in F, \mathbf{v} \in V [s\mathbf{v} \in V] & (h) \text{ homogeneity} \end{cases} \\ \Leftrightarrow & \forall s \in F, \mathbf{u}, \mathbf{v} \in V [\mathbf{u} + s\mathbf{v} \in V] \quad (l) \text{ linearity} \end{aligned}$$

40.3.2 linear transformation or linear map

$$\begin{aligned} & \begin{cases} V, W \text{ are vector spaces} \\ T : V \rightarrow W \\ \begin{cases} \forall \mathbf{u}, \mathbf{v} \in V [T(\mathbf{u} + \mathbf{v}) = T(\mathbf{u}) + T(\mathbf{v})] & (a) \text{ additivity} \\ \forall \mathbf{v} \in V, c \in F [T(c\mathbf{v}) = cT(\mathbf{v})] & (h) \text{ homogeneity} \end{cases} \end{cases} \quad (L) \\ \Leftrightarrow & \begin{cases} V, W \text{ are vector spaces} \\ T : V \rightarrow W \\ \forall \mathbf{u}, \mathbf{v} \in V, c \in F [T(\mathbf{u} + c\mathbf{v}) = T(\mathbf{u}) + cT(\mathbf{v})] \quad (l) \text{ linearity} \end{cases} \\ \Leftrightarrow & T \text{ is a linear map from } V \text{ to } W \\ \Leftrightarrow & T \text{ is a linear transformation} \end{aligned}$$

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

$$\begin{cases} \mathbf{e}_1 = \mathbf{a} & \mathbf{a} \times \mathbf{b} \neq \mathbf{0} \\ \mathbf{e}_2 = \mathbf{b} & \mathbf{b} \times \mathbf{c} \neq \mathbf{0} \\ \mathbf{e}_3 = \mathbf{c} & \mathbf{c} \times \mathbf{a} \neq \mathbf{0} \end{cases} \Rightarrow \begin{cases} \mathbf{e}'_1 = \frac{\mathbf{b} \times \mathbf{c}}{\Omega} \\ \mathbf{e}'_2 = \frac{\mathbf{c} \times \mathbf{a}}{\Omega} \\ \mathbf{e}'_3 = \frac{\mathbf{a} \times \mathbf{b}}{\Omega} \end{cases},$$

$$\Omega = \mathbf{a} \cdot (\mathbf{b} \times \mathbf{c}) = \mathbf{b} \cdot (\mathbf{c} \times \mathbf{a}) = \mathbf{c} \cdot (\mathbf{a} \times \mathbf{b})$$

reciprocal space as dual space and contravariant vector

$$\begin{aligned} \text{span} \{ \mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3 \} &= \text{span} \{ \mathbf{a}, \mathbf{b}, \mathbf{c} \} = V \\ &= \mathbb{R}^3 = \text{span} \{ \mathbf{e}'_1, \mathbf{e}'_2, \mathbf{e}'_3 \} = \text{span} \left\{ \frac{\mathbf{b} \times \mathbf{c}}{\Omega}, \frac{\mathbf{c} \times \mathbf{a}}{\Omega}, \frac{\mathbf{a} \times \mathbf{b}}{\Omega} \right\} \\ &= \text{span} \{ \mathbf{e}^1, \mathbf{e}^2, \mathbf{e}^3 \} = \text{span} \{ \mathbf{e}^* \}_{* \in \{1, 2, 3\}} = V^* \end{aligned}$$

40.4.1.1 Kronecker delta

$$\begin{pmatrix} \mathbf{e}_1 \cdot \mathbf{e}'_1 & \mathbf{e}_1 \cdot \mathbf{e}'_2 & \mathbf{e}_1 \cdot \mathbf{e}'_3 \\ \mathbf{e}_2 \cdot \mathbf{e}'_1 & \mathbf{e}_2 \cdot \mathbf{e}'_2 & \mathbf{e}_2 \cdot \mathbf{e}'_3 \\ \mathbf{e}_3 \cdot \mathbf{e}'_1 & \mathbf{e}_3 \cdot \mathbf{e}'_2 & \mathbf{e}_3 \cdot \mathbf{e}'_3 \end{pmatrix} = [\delta_{ij}] = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} = \begin{pmatrix} \mathbf{e}^1 \cdot \mathbf{e}_1 & \mathbf{e}^1 \cdot \mathbf{e}_2 & \mathbf{e}^1 \cdot \mathbf{e}_3 \\ \mathbf{e}^2 \cdot \mathbf{e}_1 & \mathbf{e}^2 \cdot \mathbf{e}_2 & \mathbf{e}^2 \cdot \mathbf{e}_3 \\ \mathbf{e}^3 \cdot \mathbf{e}_1 & \mathbf{e}^3 \cdot \mathbf{e}_2 & \mathbf{e}^3 \cdot \mathbf{e}_3 \end{pmatrix}$$

Kronecker delta

$$\mathbf{e}_i \cdot \mathbf{e}'_j = \delta_{ij} = \begin{cases} 1 & i = j \\ 0 & i \neq j \end{cases}$$

Kronecker delta tensor = Kronecker tensor

$$\mathbf{e}^i(\mathbf{e}_j) = \mathbf{e}^i \cdot \mathbf{e}_j = \delta_j^i = \delta^i_j = \begin{cases} 1 & i = j \\ 0 & i \neq j \end{cases}$$

$$\mathbf{v} = v_a \mathbf{a} + v_b \mathbf{b} + v_c \mathbf{c} = v_1 \mathbf{e}_1 + v_2 \mathbf{e}_2 + v_3 \mathbf{e}_3$$

$$\mathbf{e}^1 \cdot \mathbf{v} = \mathbf{v} \cdot \mathbf{e}'_1 = v_1 \mathbf{e}_1 \cdot \mathbf{e}'_1 + v_2 \mathbf{e}_2 \cdot \mathbf{e}'_1 + v_3 \mathbf{e}_3 \cdot \mathbf{e}'_1 = v_1$$

$$\mathbf{e}^2 \cdot \mathbf{v} = \mathbf{v} \cdot \mathbf{e}'_2 = v_1 \mathbf{e}_1 \cdot \mathbf{e}'_2 + v_2 \mathbf{e}_2 \cdot \mathbf{e}'_2 + v_3 \mathbf{e}_3 \cdot \mathbf{e}'_2 = v_2$$

$$\mathbf{e}^3 \cdot \mathbf{v} = \mathbf{v} \cdot \mathbf{e}'_3 = v_1 \mathbf{e}_1 \cdot \mathbf{e}'_3 + v_2 \mathbf{e}_2 \cdot \mathbf{e}'_3 + v_3 \mathbf{e}_3 \cdot \mathbf{e}'_3 = v_3$$

$$\begin{aligned} \mathbf{v} &= v_1 \mathbf{e}_1 + v_2 \mathbf{e}_2 + v_3 \mathbf{e}_3 \\ &= (\mathbf{v} \cdot \mathbf{e}'_1) \mathbf{e}_1 + (\mathbf{v} \cdot \mathbf{e}'_2) \mathbf{e}_2 + (\mathbf{v} \cdot \mathbf{e}'_3) \mathbf{e}_3 \\ &= (\mathbf{e}^1 \cdot \mathbf{v}) \mathbf{e}_1 + (\mathbf{e}^2 \cdot \mathbf{v}) \mathbf{e}_2 + (\mathbf{e}^3 \cdot \mathbf{v}) \mathbf{e}_3 \\ &= \mathbf{e}^1(\mathbf{v}) \mathbf{e}_1 + \mathbf{e}^2(\mathbf{v}) \mathbf{e}_2 + \mathbf{e}^3(\mathbf{v}) \mathbf{e}_3 \end{aligned}$$

$$\begin{cases} \mathbf{e}^1(\mathbf{v}) = \mathbf{e}^1 \cdot \mathbf{v} = \mathbf{v} \cdot \mathbf{e}'_1 = v_1 \\ \mathbf{e}^2(\mathbf{v}) = \mathbf{e}^2 \cdot \mathbf{v} = \mathbf{v} \cdot \mathbf{e}'_2 = v_2 \\ \mathbf{e}^3(\mathbf{v}) = \mathbf{e}^3 \cdot \mathbf{v} = \mathbf{v} \cdot \mathbf{e}'_3 = v_3 \end{cases}$$

$$\mathbf{e}^i(\mathbf{e}_j) = \mathbf{e}^i \cdot \mathbf{e}_j = \delta_j^i = \delta_{j,i} = \begin{cases} 1 & i=j \\ 0 & i \neq j \end{cases}$$

reciprocal space is a dual space of its original vector space

$$\begin{aligned} V &= \text{span}\{\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3\} = \{v_1 \mathbf{e}_1 + v_2 \mathbf{e}_2 + v_3 \mathbf{e}_3\} \\ &= \left\{ \sum_{j=1}^3 v_j \mathbf{e}_j \right\} = \left\{ v_j \mathbf{e}_j \middle| \begin{array}{l} v_j \in F \\ \mathbf{e}_j \in F^3 \end{array} \right\} = \{\mathbf{v} | \mathbf{v} \in V\} \\ V^* &= \text{span}\{\mathbf{e}^1, \mathbf{e}^2, \mathbf{e}^3\} = \{v^{*1} \mathbf{e}^1 + v^{*2} \mathbf{e}^2 + v^{*3} \mathbf{e}^3\} \\ &= \left\{ \sum_{i=1}^3 v^{*i} \mathbf{e}^i \right\} = \left\{ v^{*i} \mathbf{e}^i \middle| \begin{array}{l} v^{*i} \in F \\ \mathbf{e}^i \in F^3 \end{array} \right\} = \{\mathbf{v}^* | \mathbf{v}^* \in V^*\} \\ \mathbf{v}^*(\mathbf{v}) &= (v^{*i} \mathbf{e}^i)(\mathbf{v}), \mathbf{v} \in V \\ &= (v^{*1} \mathbf{e}^1 + v^{*2} \mathbf{e}^2 + v^{*3} \mathbf{e}^3)(\mathbf{v}) \\ &= v^{*1} \mathbf{e}^1(\mathbf{v}) + v^{*2} \mathbf{e}^2(\mathbf{v}) + v^{*3} \mathbf{e}^3(\mathbf{v}) \\ &= v^{*1} v_1 + v^{*2} v_2 + v^{*3} v_3 \in F \end{aligned}$$

element in dual space is a functional or mapping from its original vector space to the field

$$\mathbf{v}^* : V \rightarrow F$$

$$V \xrightarrow{\mathbf{v}^*} F$$

$$V^* = \{\mathbf{v}^* | \mathbf{v}^* : V \rightarrow F\}$$

$$\begin{array}{ccccccc} & & V & = & \{ & \mathbf{e}_1 & \mathbf{e}_2 & \mathbf{e}_3 & \mathbf{v} & \cdots \} \\ \mathbf{e}^1 & : & \downarrow & & \downarrow & & \downarrow & & \downarrow \\ \{ & \mathbf{e}^2 & \} & F & \supseteq & \{ & 1 & 0 & 0 & v_1 & \cdots \} \\ \mathbf{e}^3 & & V & = & \{ & \mathbf{e}_1 & \mathbf{e}_2 & \mathbf{e}_3 & \mathbf{v} & \cdots \} \\ \mathbf{v}^* & : & \downarrow & & \downarrow & & \downarrow & & \downarrow \\ & & F & \supseteq & \{ & v^{*1} & v^{*2} & v^{*3} & v^{*i} v_i & \cdots \} \end{array}$$

$$\begin{aligned} V^* &= \{\mathbf{v}^* | \mathbf{v}^* \in V^*\} = \{\mathbf{v}^* | \mathbf{v}^* : V \rightarrow F\} \\ &= \left\{ \mathbf{v}^* \middle| V \xrightarrow{\mathbf{v}^*} F \right\} \\ &= \{\omega | \omega : V \rightarrow F\} \\ &= \left\{ \omega^i \mathbf{e}^i \middle| \begin{array}{l} \omega^i \in F \\ \mathbf{e}^i \in F^3 \end{array} \right\} \end{aligned}$$

By defining vector addition and scalar multiplication on the dual space

$$\begin{cases} + : V^* \times V^* \rightarrow V^* \Leftrightarrow \forall \omega_1, \omega_2 \in V^*, \exists! (\omega_1 + \omega_2) \in V^* [(\omega_1 + \omega_2)(\mathbf{v}) = \omega_1(\mathbf{v}) + \omega_2(\mathbf{v})] \\ \cdot : F \times V^* \rightarrow V^* \Leftrightarrow \forall k \in F, \forall \omega \in V^*, \exists! (k\omega) \in V^* [(k\omega)(\mathbf{v}) = k \cdot \omega(\mathbf{v})] \\ \forall \omega \in V^*, \exists! \mathbf{0} \in V^* [(\omega + \mathbf{0})(\mathbf{v}) = \omega(\mathbf{v}) + \mathbf{0}(\mathbf{v}) = \omega(\mathbf{v})] \end{cases}$$

the dual space also becomes a vector space.

40.4.1.2 double dual concept

double dual space = second dual space

$$\begin{aligned} V^{**} &= (V^*)^* \\ &= \{\omega^* | \omega^* : V^* \rightarrow F\} \\ &= \{\omega^* | \omega^* \in V^{**}\} \end{aligned}$$

$$\begin{aligned} V^{**} &= (V^*)^* = \text{span} \{e^\mu\}_{\mu \in \{1, 2, 3\}}^* \\ &= \text{span} \{e^1, e^2, e^3\}^* \\ &= \text{span} \{e^{1*}, e^{2*}, e^{3*}\} \\ &= \text{span} \{e^{\nu*}\}_{\nu \in \{1, 2, 3\}} \end{aligned}$$

$$\begin{aligned} \omega^*(\omega) &= (\omega^{*\nu} e^{\nu*})(\omega), \omega \in V^* \\ &= (\omega^{*1} e^{1*} + \omega^{*2} e^{2*} + \omega^{*3} e^{3*})(\omega) \\ &= \omega^{*1} e^{1*}(\omega) + \omega^{*2} e^{2*}(\omega) + \omega^{*3} e^{3*}(\omega) \\ &= \omega^{*1} \omega_1 + \omega^{*2} \omega_2 + \omega^{*3} \omega_3 \in F \end{aligned}$$

$$V^{**} = \{\omega^* | \omega^* : V^* \rightarrow F\}$$

$$\begin{array}{ccccccc} & V^* & = & \{ & e^1 & e^2 & e^3 & \omega & \dots \} \\ \begin{matrix} e^{1*} \\ e^{2*} \\ e^{3*} \\ \omega^* \end{matrix} & : & \downarrow & & \downarrow & \downarrow & \downarrow \\ V^{**} = \{ & F & \supseteq & \{ & 1 & 0 & 0 & \omega^1 & \dots \} \\ & V^* & = & \{ & e_1 & e_2 & e_3 & v & \dots \} \\ & \vdots & & & \downarrow & \downarrow & \downarrow & & \downarrow \\ & F & \supseteq & \{ & \omega^{1*} & \omega^{2*} & \omega^{3*} & \omega^{\mu*} \omega^\mu & \dots \} \end{array}$$

$$\begin{cases} e^{1*}(\omega) = e^{1*} \cdot \omega = \omega \cdot e^{1*} = \omega(e^{1*}) \\ e^{2*}(\omega) = e^{2*} \cdot \omega = \omega \cdot e^{2*} = \omega(e^{2*}) \\ e^{3*}(\omega) = e^{3*} \cdot \omega = \omega \cdot e^{3*} = \omega(e^{3*}) \end{cases}$$

$$\omega^*(\omega) = \omega^* \cdot \omega = \omega \cdot \omega^* = \omega(\omega^*)$$

i.e. f acts on x equivalent to x acts on f

$$x(f) = x \cdot f = f \cdot x = f(x)$$

$$e^\mu(e_\nu) = e^\mu \cdot e_\nu = \delta_\nu^\mu = \delta_\nu^\mu = \begin{cases} 1 & \mu = \nu \\ 0 & \mu \neq \nu \end{cases}$$

$$\begin{aligned} e^{\nu*}(e^\mu) &\stackrel{\text{def.}}{=} e_\nu \cdot e^\mu = e^\mu \cdot e_\nu = e^\mu(e_\nu) \\ &\quad \Downarrow \\ V^{**} &= \text{span} \{e^{\nu*}\}_{\nu \in \{1, 2, 3\}} \cong \text{span} \{e_\nu\}_{\nu \in \{1, 2, 3\}} = V \\ V^{**} &\cong V \\ &\quad \Downarrow \begin{cases} V^{**} \cong V & V, V^{**} \text{ are isomorphism} \\ & \text{independent of choice of bases} \end{cases} \\ &\quad V, V^{**} \text{ are naturally isomorphism} \end{aligned}$$

$$V^{**} = \{\omega^* | \omega^* : V^* \rightarrow F\} \cong V = \{v | v : V^* \rightarrow F\}$$

$$\begin{array}{ccccccc}
& V^* & = \{ & e^1 & e^2 & e^3 & \omega & \dots \} \\
\begin{matrix} e^{1*} \\ e^{2*} \\ e^{3*} \\ \omega^* \end{matrix} : & \downarrow & \downarrow & \downarrow & \downarrow & \downarrow & \dots \\
V^{**} = \{ & F & \supseteq \{ & 1 & 0 & 0 & \omega^1 & \dots \} \\
& V^* & = \{ & e^1 & e^2 & e^3 & \omega & \dots \} \\
& \downarrow & \downarrow & \downarrow & \downarrow & \downarrow & \\
& \vdots & F & \supseteq \{ & \omega^{1*} & \omega^{2*} & \omega^{3*} & \omega^{\mu*}\omega^{\mu} & \dots \} \\
& & V^* & = \{ & e^1 & e^2 & e^3 & v^* & \dots \} \\
& e_1 & : & \downarrow & \downarrow & \downarrow & \downarrow & \downarrow \\
\cong V = \{ & e_2 & \} & F & \supseteq \{ & 1 & 0 & 0 & v^{*1} & \dots \} \\
& e_3 & \} & V^* & = \{ & e^1 & e^2 & e^3 & v^* & \dots \} \\
& v & : & \downarrow & \downarrow & \downarrow & \downarrow & \downarrow \\
& \vdots & F & \supseteq \{ & v_1 & v_2 & v_3 & v_{\mu}v^{*\mu} & \dots \}
\end{array}$$

$$V \cong V^{**}$$

$$V \cong V^{**} = \{\omega^* | \omega^* : V^* \rightarrow F\}$$

$$V = \{v | v : V^* \rightarrow F\}$$

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

$$f = \{f_i\} = \{f_1, f_2, \dots\} = \{f, g, \dots\}$$

$$\mathbf{v} : f \rightarrow F$$

$$\mathbf{v}(af + bg) = a\mathbf{v}(f) + b\mathbf{v}(g)$$

$$\mathbf{v}(fg) = f|_P \mathbf{v}(g) + \mathbf{v}(f)g|_P$$

$$\frac{\mathrm{d}}{\mathrm{d}x} [f(x)g(x)]|_{x=x_0} = f(x_0) \frac{\mathrm{d}}{\mathrm{d}x} g(x)|_{x=x_0} + \frac{\mathrm{d}}{\mathrm{d}x} f(x)|_{x=x_0} g(x_0)$$

$$V = \{\mathbf{v} | \mathbf{v} : f \rightarrow F\}$$

$$\begin{aligned} f &= f(\mathbf{x}) \\ &= f(x_1, \dots, x_j, \dots, x_n) \\ &= f(x^1, \dots, x^j, \dots, x^n) \end{aligned}$$

$$\mathbf{x} = \langle x^1, \dots, x^j, \dots, x^n \rangle$$

$$\mathbf{x}(t) = \langle x^1(t), \dots, x^j(t), \dots, x^n(t) \rangle$$

$$\begin{aligned} \frac{\mathrm{d}f}{\mathrm{d}t} &= \frac{\mathrm{d}x^1}{\mathrm{d}t} \frac{\partial f}{\partial x^1} + \dots + \frac{\mathrm{d}x^j}{\mathrm{d}t} \frac{\partial f}{\partial x^j} + \dots + \frac{\mathrm{d}x^n}{\mathrm{d}t} \frac{\partial f}{\partial x^n} \\ &= \dots + \frac{\mathrm{d}x^j}{\mathrm{d}t} \frac{\partial f}{\partial x^j} + \dots = \frac{\mathrm{d}x^j}{\mathrm{d}t} \frac{\partial f}{\partial x^j} = \frac{\mathrm{d}x^j}{\mathrm{d}t} \partial_j f \end{aligned}$$

$$\begin{aligned} V &= \text{span}\{\mathbf{e}_1, \dots, \mathbf{e}_j, \dots, \mathbf{e}_n\} \\ &= \text{span}\left\{\frac{\partial}{\partial x^1}|_P, \dots, \frac{\partial}{\partial x^j}|_P, \dots, \frac{\partial}{\partial x^n}|_P\right\} \\ &= \text{span}\{\boldsymbol{\partial}_1, \dots, \boldsymbol{\partial}_j, \dots, \boldsymbol{\partial}_n\} \\ &= \left\{\boldsymbol{\partial}_t \middle| \boldsymbol{\partial}_t = a_j \mathbf{e}_j = a_j \boldsymbol{\partial}_j = a_j \frac{\partial}{\partial x^j}|_P\right\} \\ &= \left\{\frac{\partial}{\partial t}|_P \middle| \frac{\partial}{\partial t}|_P = a_1 \frac{\partial}{\partial x^1}|_P + \dots + a_j \frac{\partial}{\partial x^j}|_P + \dots + a_n \frac{\partial}{\partial x^n}|_P\right\} \end{aligned}$$

41.2 dual space of span of partials

$$V^* = \{\omega_f | \omega_f : V \rightarrow F\}$$

$$\omega_f(e_j) = \omega_f(\partial_j) = \omega_f\left(\frac{\partial}{\partial x^j}|_P\right) = \frac{\partial f}{\partial x^j}|_P \in F$$

$$\begin{aligned}\omega_{fg}(\partial_j) &= \frac{\partial fg}{\partial x^j}|_P = f|_P \frac{\partial g}{\partial x^j}|_P + \frac{\partial f}{\partial x^j}|_P g|_P \\ &= f|_P \omega_g(\partial_j) + \omega_f(\partial_j) g|_P\end{aligned}$$

$$\omega_{x^i}(\partial_j) = \omega_{x^i}\left(\frac{\partial}{\partial x^j}|_P\right) = \frac{\partial x^i}{\partial x^j}|_P = \delta_{ij} = \begin{cases} 1 & i=j \\ 0 & i \neq j \end{cases}$$

$$\begin{aligned}V^* = \{\omega_f | \omega_f : V \rightarrow F\} &= \left\{ \omega_f \left| \begin{array}{l} \omega_f(e_j) = \omega_f(\partial_j) = \omega_f\left(\frac{\partial}{\partial x^j}|_P\right) = \frac{\partial f}{\partial x^j}|_P \in F \\ \omega_{fg}(\partial_j) = f|_P \omega_g(\partial_j) + \omega_f(\partial_j) g|_P \\ \omega_{x^i}(\partial_j) = \omega_{x^i}\left(\frac{\partial}{\partial x^j}|_P\right) = \frac{\partial x^i}{\partial x^j}|_P = \delta_{ij} = \begin{cases} 1 & i=j \\ 0 & i \neq j \end{cases} \end{array} \right. \right\} \\ &= \{df|df : V \rightarrow F\} = \left\{ df \left| \begin{array}{l} df(e_j) = df(\partial_j) = df\left(\frac{\partial}{\partial x^j}|_P\right) = \frac{\partial f}{\partial x^j}|_P \in F \\ dfg(\partial_j) = f|_P (dg) + (df) g|_P \\ dx^i(\partial_j) = dx^i\left(\frac{\partial}{\partial x^j}|_P\right) = \frac{\partial x^i}{\partial x^j}|_P = \delta_{ij} = \begin{cases} 1 & i=j \\ 0 & i \neq j \end{cases} \end{array} \right. \right\}\end{aligned}$$

$$dx^i\left(\frac{\partial}{\partial x^j}|_P\right) = \delta_{ij} = e^i \cdot e_j \Rightarrow \begin{cases} e^i = dx^i \\ e_j = \frac{\partial}{\partial x^j}|_P \end{cases}$$

$$\begin{aligned}V^* = \{df|df : V \rightarrow F\} &= \left\{ df \left| \begin{array}{l} df(e_j) = df(\partial_j) = df\left(\frac{\partial}{\partial x^j}|_P\right) = \frac{\partial f}{\partial x^j}|_P \in F \\ dfg(\partial_j) = f|_P (dg) + (df) g|_P \\ dx^i(\partial_j) = dx^i\left(\frac{\partial}{\partial x^j}|_P\right) = \frac{\partial x^i}{\partial x^j}|_P = \delta_{ij} = \begin{cases} 1 & i=j \\ 0 & i \neq j \end{cases} \end{array} \right. \right\} \\ &= \text{span}\{dx^1, \dots, dx^i, \dots, dx^n\} = \text{span}\{e^1, \dots, e^i, \dots, e^n\}\end{aligned}$$

41.3 directional derivative

$$\begin{aligned}df(v) &= df(v_j e_j) = v_j df(e_j) \\ &= v_j df(\partial_j) = v_j \frac{\partial f}{\partial x^j}|_P \\ &= v_1 \frac{\partial f}{\partial x^1}|_P + v_2 \frac{\partial f}{\partial x^2}|_P + \dots + v_n \frac{\partial f}{\partial x^n}|_P \\ &= (v_1 \quad \dots \quad v_j \quad \dots \quad v_n) \nabla f\end{aligned}$$

$$\widehat{PQ} = C(t) - C(0) = Q - P$$

$$v = \frac{\partial}{\partial t}|_P$$

$$\begin{aligned}df(sv) &= df\left(s \frac{\partial}{\partial t}|_P\right) = s \frac{\partial f}{\partial t}|_P \\ &= sv(f) = s \cdot \lim_{t \rightarrow 0} \frac{f(C(t)) - f(C(0))}{t} \\ &\approx s \cdot \frac{f(Q) - f(P)}{s} = f(Q) - f(P) = \Delta f\end{aligned}$$

41.4 coefficient of linear combination for vector space and dual space

$$\begin{aligned}
V &= \{v|v : f \rightarrow F\} \\
&= \text{span}\{e_1, \dots, e_j, \dots, e_n\} \\
&= \text{span}\left\{\frac{\partial}{\partial x^1}|_P, \dots, \frac{\partial}{\partial x^j}|_P, \dots, \frac{\partial}{\partial x^n}|_P\right\} = \text{span}\{\partial_1, \dots, \partial_j, \dots, \partial_n\} \\
&= \left\{\partial_t \middle| \partial_t = a_j e_j = a_j \partial_j = a_j \frac{\partial}{\partial x^j}|_P\right\} \\
&= \left\{\frac{\partial}{\partial t}|_P \middle| \frac{\partial}{\partial t}|_P = a_1 \frac{\partial}{\partial x^1}|_P + \dots + a_j \frac{\partial}{\partial x^j}|_P + \dots + a_n \frac{\partial}{\partial x^n}|_P\right\} \\
V^* &= \{df|df : V \rightarrow F\} \\
&= \text{span}\{e^1, \dots, e^i, \dots, e^n\} \\
&= \text{span}\{dx^1, \dots, dx^i, \dots, dx^n\} \\
&= \{df|df = b^i e^i = b^i dx^i\} \\
&= \{df|df = b^1 dx^1 + \dots + b^i dx^i + \dots + b^n dx^n\}
\end{aligned}$$

or more simply to be comparison

$$\begin{array}{ll}
V &= \text{span}\{e_j = \partial_j\} = \{v = \partial_t|_P = a_j e_j = a_j \partial_j|_P : f \rightarrow F\} \\
V^* &= \text{span}\{e^i = dx^i\} = \{\omega = df = b^i e^i = b^i dx^i : V \rightarrow F\}
\end{array}$$

$$\begin{aligned}
&\left\{ dx^i(\partial_j) = dx^i\left(\frac{\partial}{\partial x^j}|_P\right) = \frac{\partial x^i}{\partial x^j}|_P = \delta_{ij} = \begin{cases} 1 & i=j \\ 0 & i \neq j \end{cases} \right. \\
&\left. \partial_t = a_j e_j = a_j \partial_j \Leftrightarrow \frac{\partial}{\partial t}|_P = a_1 \frac{\partial}{\partial x^1}|_P + \dots + a_j \frac{\partial}{\partial x^j}|_P + \dots + a_n \frac{\partial}{\partial x^n}|_P \right. \\
&\Rightarrow \left\{ \begin{array}{l} dx^i(\partial_t) = dx^i\left(\frac{\partial}{\partial t}|_P\right) = \frac{\partial x^i}{\partial t}|_P \\ dx^i(\partial_t) = dx^i(a_j \partial_j) = a_j dx^i(\partial_j) = a_j \delta_{ij} = a_i \end{array} \right. \Rightarrow a_i = dx^i(\partial_t) = \frac{\partial x^i}{\partial t}|_P \\
&\Rightarrow a_i = \frac{\partial x^i}{\partial t}|_P \Rightarrow a_j = \frac{\partial x^j}{\partial t}|_P = \partial_t x^j|_P \\
&\Rightarrow \frac{\partial}{\partial t}|_P = a_i \frac{\partial}{\partial x^i}|_P = \frac{\partial x^i}{\partial t}|_P \frac{\partial}{\partial x^i}|_P = \frac{\partial x^i}{\partial t} \frac{\partial}{\partial x^i}|_P \Rightarrow \frac{\partial}{\partial t} = \frac{\partial x^i}{\partial t} \frac{\partial}{\partial x^i} \\
&\Rightarrow \partial_t|_P = \frac{\partial x^j}{\partial t} \partial_j|_P \Leftrightarrow \partial_t|_P = \partial_t x^j \partial_j|_P
\end{aligned}$$

$$\begin{aligned}
df &= b^i e^i = b^i dx^i \\
\frac{\partial f}{\partial x^j} &= df(\partial_j) = df(e_j) = b^i e^i \cdot e_j = b^i \delta_{ij} = b^j \\
b^j &= \frac{\partial f}{\partial x^j} \\
b^i &= \frac{\partial f}{\partial x^i} = \partial_i f \\
df &= b^i e^i = b^i dx^i = \frac{\partial f}{\partial x^i} dx^i \\
df &= \frac{\partial f}{\partial x^i} dx^i \\
df &= \partial_i f dx^i
\end{aligned}$$

$$\begin{array}{ll}
V &= \text{span}\{e_j = \partial_j\} = \{v = \partial_t|_P = a_j e_j = a_j \partial_j|_P : f \rightarrow F\} \\
V^* &= \text{span}\{e^i = dx^i\} = \{\omega = df = b^i e^i = b^i dx^i : V \rightarrow F\}
\end{array}$$

41.5 change of basis / change of coordinate

$$\frac{\partial}{\partial t} = \frac{\partial x^i}{\partial t} \frac{\partial}{\partial x^i} \stackrel{t=x'^j}{\Rightarrow} \frac{\partial}{\partial x'^j} = \frac{\partial x^i}{\partial x'^j} \frac{\partial}{\partial x^i} = \frac{\partial x^1}{\partial x'^j} \frac{\partial}{\partial x^1} + \frac{\partial x^2}{\partial x'^j} \frac{\partial}{\partial x^2} + \frac{\partial x^3}{\partial x'^j} \frac{\partial}{\partial x^3}$$

$$\begin{aligned} df &= \frac{\partial f}{\partial x^i} dx^i \\ f = x'^j &\Downarrow \\ dx'^j &= \frac{\partial x'^j}{\partial x^i} dx^i \end{aligned}$$

$$\begin{cases} \frac{\partial}{\partial x'^j} = \frac{\partial x^i}{\partial x'^j} \frac{\partial}{\partial x^i} = \sum_i \frac{\partial x^i}{\partial x'^j} \frac{\partial}{\partial x^i} \\ dx'^j = \frac{\partial x'^j}{\partial x^i} dx^i = \sum_i \frac{\partial x'^j}{\partial x^i} dx^i \end{cases}$$

Chapter 42

determinant

42.1 induction of determinant axioms

42.2 determinant axioms

determinant axioms

$$\begin{aligned} & \left\{ \begin{array}{l} \det(\mathbf{u}, \mathbf{v}) = \det(\mathbf{u} + s\mathbf{v}, \mathbf{v}) \\ \det(\mathbf{u}, \mathbf{v}) = \det(\mathbf{u}, \mathbf{v} + s\mathbf{u}) \end{array} \right. && \text{translation invariance} \\ & \left\{ \begin{array}{l} \det(s\mathbf{u}, \mathbf{v}) = s \det(\mathbf{u}, \mathbf{v}) \\ \det(\mathbf{u}, s\mathbf{v}) = s \det(\mathbf{u}, \mathbf{v}) \end{array} \right. && \text{scaling} \\ & \left\{ \begin{array}{l} \det(\mathbf{u}_1 + \mathbf{u}_2, \mathbf{v}) = \det(\mathbf{u}_1, \mathbf{v}) + \det(\mathbf{u}_2, \mathbf{v}) \\ \det(\mathbf{u}, \mathbf{v}_1 + \mathbf{v}_2) = \det(\mathbf{u}, \mathbf{v}_1) + \det(\mathbf{u}, \mathbf{v}_2) \end{array} \right. && \text{decomposition} \\ \Leftrightarrow & \left\{ \begin{array}{l} \det(\mathbf{u}, \mathbf{v}) = \det(\mathbf{u} + s\mathbf{v}, \mathbf{v}) \\ \det(\mathbf{u}, \mathbf{v}) = \det(\mathbf{u}, \mathbf{v} + s\mathbf{u}) \\ \det(\mathbf{u}_1 + s\mathbf{u}_2, \mathbf{v}) = \det(\mathbf{u}_1, \mathbf{v}) + s \det(\mathbf{u}_2, \mathbf{v}) \\ \det(\mathbf{u}, \mathbf{v}_1 + s\mathbf{v}_2) = \det(\mathbf{u}, \mathbf{v}_1) + s \det(\mathbf{u}, \mathbf{v}_2) \end{array} \right. && \begin{array}{l} \text{translation invariance} \\ \text{linearity} \end{array} \end{aligned}$$

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

quantum state^[45.2.3]

Taylor vs. Fourier^[@ref(taylor-vs.-fourier)]

$$f(x) = a_0x^0 + a_1x^1 + a_2x^2 + \dots = \sum_{k=0}^{\infty} a_k x^k$$

$$f(x) = x_0x^0 + x_1x^1 + x_2x^2 + \dots = \sum_{k=0}^{\infty} x_k x^k$$

Def: 45.3

$$\langle f|g \rangle = \int_a^b \overline{f(x)} g(x) dx \stackrel{f,g:\mathbb{R} \rightarrow \mathbb{R}}{=} \int_a^b f(x) g(x) dx$$

Dirac bracket^[45.5]

$$\langle x^2|x \rangle \stackrel{x^2,x:\mathbb{R} \rightarrow \mathbb{R}}{=} \int_a^b x^2 x dx = \int_a^b x^3 dx = \left[\frac{x^4}{4} \right]_a^b \neq 0$$

$$x^0 \not\perp x^1, x^1 \not\perp x^2, \dots$$

$$\langle x_m|x^n \rangle = \int_a^b x_m x^n dx = \delta_{mn}$$

$$\langle 1|x^n \rangle = \int_a^b x_0 x^n dx = \delta_{0n} \Rightarrow x_0 = \delta(x) = \delta(x-0)$$

$$\langle x_m|x^n \rangle = \int_a^b x_m x^n dx = \delta_{mn} \Rightarrow x_m = \frac{(-1)^m}{m!} \delta^{(m)}(x)$$

$$|f\rangle = 1|f\rangle = \left(\sum_i |\hat{f}_i\rangle \langle \hat{f}_i| \right) |f\rangle = \sum_i |\hat{f}_i\rangle \langle \hat{f}_i| f \rangle$$

$$\begin{aligned}
|f\rangle &= 1 |f\rangle = \left(\sum_i |\hat{f}_i\rangle \langle \hat{f}_i| \right) |f\rangle = \sum_i |\hat{f}_i\rangle \langle \hat{f}_i| f \rangle \\
&= 1 |f\rangle = \left(\sum_n |x^n\rangle \langle x^n| \right) |f\rangle = \sum_n |x^n\rangle \langle x^n| f \rangle = \sum_n \langle x^n| f \rangle |x^n\rangle \\
\langle x^n| |f\rangle &= \langle x^n| f \rangle = \int_a^b x_n f(x) dx = \int_a^b \frac{(-1)^n}{n!} \delta^{(n)}(x) f(x) dx = \frac{f^{(n)}(0)}{n!} \\
|f\rangle &= \sum_n \langle x^n| f \rangle |x^n\rangle = \sum_n \frac{f^{(n)}(0)}{n!} |x^n\rangle \\
|f\rangle &= \sum_n \frac{f^{(n)}(0)}{n!} |x^n\rangle \\
&\Downarrow \\
f(x) &= \sum_n \frac{f^{(n)}(0)}{n!} x^n
\end{aligned}$$

43.2 beta function

$$\begin{aligned}
\binom{n}{k} &= C_k^n = \frac{n!}{(n-k)!k!} \\
&= \frac{n(n-1)\cdots(n-k+1)}{k!}, \quad \begin{cases} n \in \mathbb{N} \\ k \in (\{0\} \cup \mathbb{N}) \end{cases} \\
\binom{r}{k} &= \begin{cases} \frac{r(r-1)\cdots(r-k+1)}{k!} & k \geq 0, k \in \mathbb{Z} \\ 0 & k < 0, k \in \mathbb{Z} \end{cases}
\end{aligned}$$

$$\sum_{k=0}^n \binom{r}{k} (\cdot)$$

$$\sum_{k=-\infty}^n \binom{r}{k} (\cdot) = (0 + 0 + \dots) + \sum_{k=0}^n \binom{r}{k} (\cdot)$$

$$\sum_{k=-\infty}^{\infty} \binom{r}{k} (\cdot)$$

$$n! = \Gamma(n+1) = \int_0^\infty x^{(n+1)-1} e^{-x} dx$$

$$\Gamma(z) = \int_0^\infty x^{z-1} e^{-x} dx$$

$$\Gamma(z+1) = z\Gamma(z)$$

$$\Gamma(z)\Gamma(1-z) = \frac{\pi}{\sin(\pi z)}$$

$$\begin{aligned}
\Gamma(z)\Gamma(1-z) &= \frac{\pi}{\sin(\pi z)} \\
[\Gamma(z)\Gamma(1-z)]_{z=-n} &= \left[\frac{\pi}{\sin(\pi z)} \right]_{z=-n}, \quad n \in \mathbb{N}
\end{aligned}$$

$$\Gamma(-n)n! = \Gamma(n+1) = \Gamma(-n)\Gamma(1-(-n)) = \frac{\pi}{\sin(\pi(-n))} = \frac{\pi}{-\sin(n\pi)}$$

$$\Gamma(-n) = \frac{-\pi}{n! \sin(n\pi)} = \frac{-\pi}{n! 0} \rightarrow -\infty, \quad n \in \mathbb{N}$$

$$\binom{n}{k} = C_k^n = \frac{n!}{(n-k)!k!} = \frac{\Gamma(n+1)}{\Gamma(n-k+1)\Gamma(k+1)}$$

$$\binom{n}{k} = \frac{\Gamma(n+1)}{\Gamma(n-k+1)\Gamma(k+1)}$$

$$\binom{n}{k} = \frac{\Gamma(n+1)}{\Gamma(n-k+1)\Gamma(k+1)} \stackrel{k \leq 0}{=} \begin{cases} \frac{\Gamma(n+1)}{\Gamma(n+1)\Gamma(1)} = \frac{\Gamma(n+1)}{\Gamma(n+1)1} = 1 & k=0 \\ \frac{\Gamma(n+1)}{\Gamma(n-k+1)(-\infty)} = 0 & k \leq -1, k \in \mathbb{Z} \end{cases}$$

beta function = β function

Definition 43.1. beta function = β function

$$B(p, q) = \int_0^1 x^{p-1} (1-x)^{q-1} dx = \frac{\Gamma(p)\Gamma(q)}{\Gamma(p+q)}$$

$$\begin{aligned} \binom{n}{k} &= \frac{\Gamma(n+1)}{\Gamma(n-k+1)\Gamma(k+1)} \\ \left[\binom{n}{k} \right]_{\substack{n=a+b \\ k=a}} &= \left[\frac{\Gamma(n+1)}{\Gamma(n-k+1)\Gamma(k+1)} \right]_{\substack{n=a+b \\ k=a}} \\ \binom{a+b}{a} &= \frac{\Gamma(a+b+1)}{\Gamma(a+b-a+1)\Gamma(a+1)} \\ &= \frac{\Gamma(a+b+1)}{\Gamma(b+1)\Gamma(a+1)} \end{aligned}$$

$$\begin{aligned} B(p, q) &= \frac{\Gamma(p)\Gamma(q)}{\Gamma(p+q)} \\ [B(p, q)]_{\substack{p=a+1 \\ q=b+1}} &= \left[\frac{\Gamma(p)\Gamma(q)}{\Gamma(p+q)} \right]_{\substack{p=a+1 \\ q=b+1}} \\ B(a+1, b+1) &= \frac{\Gamma(a+1)\Gamma(b+1)}{\Gamma(a+1+b+1)} \\ &= \frac{\Gamma(a+1)\Gamma(b+1)}{\Gamma([a+b+1]+1)} = \frac{\Gamma(a+1)\Gamma(b+1)}{[a+b+1]\Gamma(a+b+1)} \end{aligned}$$

$$\begin{aligned} \binom{a+b}{a} &= \frac{\Gamma(a+b+1)}{\Gamma(b+1)\Gamma(a+1)} = \frac{1}{\frac{\Gamma(b+1)\Gamma(a+1)}{\Gamma(a+b+1)}} \\ &= \frac{1}{[a+b+1]\frac{\Gamma(b+1)\Gamma(a+1)}{[a+b+1]\Gamma(a+b+1)}} \\ &= \frac{1}{[a+b+1]B(a+1, b+1)} \end{aligned}$$

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 technique

分式微分不是難而是煩

44.1.1 principle

Theorem 44.1. *Feynman method of differentiation / derivative*

$$\begin{aligned} f(t) &= k [u(t)]^a [v(t)]^b [w(t)]^c \dots \\ f'(t) &= f(t) \left[a \frac{u'(t)}{u(t)} + b \frac{v'(t)}{v(t)} + c \frac{w'(t)}{w(t)} + \dots \right] \end{aligned}$$

Proof:

$$f(t) = k [u(t)]^a [v(t)]^b [w(t)]^c \dots$$

$$f = k u^a v^b w^c \dots = k \cdot u^a \cdot v^b \cdot w^c \dots$$

$$\begin{aligned} f &= k u^a v^b w^c \dots = k \cdot u^a \cdot v^b \cdot w^c \dots \\ \ln f &= \ln(k u^a v^b w^c \dots) = \ln k + \ln u^a + \ln v^b + \ln w^c + \dots \\ &= \ln k + a \ln u + b \ln v + c \ln w + \dots \end{aligned}$$

$$\begin{aligned} \frac{d}{dt} \ln f &= \frac{d}{dt} (\ln k + a \ln u + b \ln v + c \ln w + \dots) \\ \frac{d}{dt} \frac{f}{f} &= 0 + \frac{d}{dt} (a \ln u) + \frac{d}{dt} (b \ln v) + \frac{d}{dt} (c \ln w) + \dots \\ &= a \frac{d}{dt} \ln u + b \frac{d}{dt} \ln v + c \frac{d}{dt} \ln w + \dots \\ &= a \frac{d}{dt} u + b \frac{d}{dt} v + c \frac{d}{dt} w + \dots \\ \frac{f'}{f} &= a \frac{u'}{u} + b \frac{v'}{v} + c \frac{w'}{w} + \dots \\ f' &= f \left(a \frac{u'}{u} + b \frac{v'}{v} + c \frac{w'}{w} + \dots \right) \\ f'(t) &= f(t) \left[a \frac{u'(t)}{u(t)} + b \frac{v'(t)}{v(t)} + c \frac{w'(t)}{w(t)} + \dots \right] \end{aligned}$$

□

44.1.2 examples

$$f(x) = x^x$$

$$(x^x)' = x^x + x^x \ln x$$

$$\begin{aligned} f(x) &= x^x \\ \ln f(x) &= x \ln x \\ \frac{d}{dx} \ln f(x) &= \frac{d}{dx} [x \ln x] \\ \frac{f'(x)}{f(x)} &= [x \ln x]' \\ f'(x) &= f(x) [x \ln x]' = x^x [(x)' \ln x + x (\ln x)'] \\ &= x^x \left[1 \ln x + x \frac{1}{x} \right] = x^x [\ln x + 1] = x^x [1 + \ln x] \\ &= x^x + x^x \ln x \end{aligned}$$

□

3D delta function^[45.2.2]

$$\Delta \left(\frac{1}{r} \right) = \nabla^2 \left(\frac{1}{\sqrt{\mathbf{r} \cdot \mathbf{r}}} \right) = \nabla \cdot \nabla \left(\frac{1}{\sqrt{\mathbf{r} \cdot \mathbf{r}}} \right) = \nabla \cdot \left(\frac{\partial}{\partial x}, \frac{\partial}{\partial y}, \frac{\partial}{\partial z} \right) \left(\frac{1}{\sqrt{x^2 + y^2 + z^2}} \right)$$

$$\frac{\partial}{\partial x} \frac{1}{\sqrt{x^2 + y^2 + z^2}}$$

$$\begin{aligned} \frac{\partial}{\partial x} \frac{1}{\sqrt{x^2 + y^2 + z^2}} &= \frac{\partial}{\partial x} \left[(x^2 + y^2 + z^2)^{-\frac{1}{2}} \right] \\ &= (x^2 + y^2 + z^2)^{-\frac{1}{2}} \left[\frac{-1}{2} \frac{2x}{x^2 + y^2 + z^2} \right] \\ &= \frac{-x}{(x^2 + y^2 + z^2)^{\frac{3}{2}}} \end{aligned}$$

$$\nabla \left(\frac{1}{r} \right) = \nabla \left(\frac{1}{\sqrt{\mathbf{r} \cdot \mathbf{r}}} \right) = \left(\frac{\partial}{\partial x}, \frac{\partial}{\partial y}, \frac{\partial}{\partial z} \right) \left(\frac{1}{\sqrt{x^2 + y^2 + z^2}} \right) = \left(\frac{-x}{(x^2 + y^2 + z^2)^{\frac{3}{2}}}, \frac{-y}{(x^2 + y^2 + z^2)^{\frac{3}{2}}}, \frac{-z}{(x^2 + y^2 + z^2)^{\frac{3}{2}}} \right)$$

$$\begin{aligned} \nabla \cdot \nabla \left(\frac{1}{r} \right) &= \nabla \cdot \left(\frac{-x}{(x^2 + y^2 + z^2)^{\frac{3}{2}}}, \frac{-y}{(x^2 + y^2 + z^2)^{\frac{3}{2}}}, \frac{-z}{(x^2 + y^2 + z^2)^{\frac{3}{2}}} \right) \\ &= \frac{\partial}{\partial x} \frac{-x}{(x^2 + y^2 + z^2 + \epsilon^2)^{\frac{3}{2}}} + \frac{\partial}{\partial y} \frac{-y}{(x^2 + y^2 + z^2 + \epsilon^2)^{\frac{3}{2}}} + \frac{\partial}{\partial z} \frac{-z}{(x^2 + y^2 + z^2 + \epsilon^2)^{\frac{3}{2}}} \end{aligned}$$

$$\frac{\partial}{\partial x} \frac{-x}{(x^2 + y^2 + z^2 + \epsilon^2)^{\frac{3}{2}}}$$

$$\begin{aligned} \frac{\partial}{\partial x} \frac{-x}{(x^2 + y^2 + z^2)^{\frac{3}{2}}} &= \frac{-x}{(x^2 + y^2 + z^2)^{\frac{3}{2}}} \left[1 \cdot \frac{1}{x} + \frac{-3}{2} \frac{2x}{x^2 + y^2 + z^2} \right] \\ &= \frac{-1}{(x^2 + y^2 + z^2)^{\frac{3}{2}}} + \frac{3x^2}{(x^2 + y^2 + z^2)^{\frac{5}{2}}} \\ &= \frac{-(x^2 + y^2 + z^2) + 3x^2}{(x^2 + y^2 + z^2)^{\frac{5}{2}}} = \frac{2x^2 - y^2 - z^2}{(x^2 + y^2 + z^2)^{\frac{5}{2}}} \end{aligned}$$

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

$$\Delta \left(\frac{1}{r} \right) = \nabla \cdot \nabla \left(\frac{1}{r} \right) = 0$$

□

$$\begin{aligned}& \frac{6(1+2t^2)(t^3-t)^2}{\sqrt{t+5t^2}(4t)^{\frac{3}{2}}} + \frac{\sqrt{1+2t}}{t+\sqrt{1+t^2}} \\ & \quad \frac{d}{dt} \left[\frac{6(1+2t^2)(t^3-t)^2}{\sqrt{t+5t^2}(4t)^{\frac{3}{2}}} + \frac{\sqrt{1+2t}}{t+\sqrt{1+t^2}} \right] \\ & \quad \frac{d}{dt} \left[\frac{6(1+2t^2)(t^3-t)^2}{\sqrt{t+5t^2}(4t)^{\frac{3}{2}}} \right] \\ & \quad \frac{d}{dt} \left[\frac{6(1+2t^2)(t^3-t)^2}{\sqrt{t+5t^2}(4t)^{\frac{3}{2}}} \right] \\ & = \frac{6(1+2t^2)(t^3-t)^2}{\sqrt{t+5t^2}(4t)^{\frac{3}{2}}} \cdot [\\ & = \frac{6[1+2t^2](t^3-t)^2}{\sqrt{t+5t^2}(4t)^{\frac{3}{2}}} \cdot [1 \cdot \frac{1}{1+2t^2} \cdot 4t, [1+2t^2]] \rightarrow \begin{cases} \text{exponential :} & 1 \\ \text{linear to denominator :} & \frac{1}{1+2t^2} \\ \text{differentiation} & 4t \end{cases} \\ & = \frac{6(1+2t^2)[(t^3-t)^2]}{\sqrt{t+5t^2}(4t)^{\frac{3}{2}}} \cdot [\frac{4t}{1+2t^2} + 2 \cdot \frac{1}{t^3-t} \cdot (3t^2-1) \\ & , [(t^3-t)^2] \rightarrow \begin{cases} \text{exponential :} & 2 \\ \text{linear to denominator :} & \frac{1}{t^3-t} \\ \text{differentiation} & 3t^2-1 \end{cases} \\ & = \frac{6(1+2t^2)(t^3-t)^2}{[\sqrt{t+5t^2}](4t)^{\frac{3}{2}}} \cdot [\frac{4t}{1+2t^2} + \frac{6t^2-2}{t^3-t} + \frac{-1}{2} \cdot \frac{1}{t+5t^2} \cdot (1+10t) \\ & , [\sqrt{t+5t^2}] \rightarrow \begin{cases} \text{exponential :} & -\frac{1}{2} \\ \text{linear to denominator :} & \frac{1}{t+5t^2} \\ \text{differentiation} & 1+10t \end{cases} \\ & = \frac{6(1+2t^2)(t^3-t)^2}{\sqrt{t+5t^2}[(4t)^{\frac{3}{2}}]} \cdot \left[\frac{4t}{1+2t^2} + \frac{6t^2-2}{t^3-t} - \frac{1+10t}{2t+10t^2} + \frac{-3}{2} \cdot \frac{1}{4t} \cdot 4 \right] \\ & , [(4t)^{\frac{3}{2}}] \rightarrow \begin{cases} \text{exponential :} & -\frac{3}{2} \\ \text{linear to denominator :} & \frac{1}{4t} \\ \text{differentiation} & 4 \end{cases} \\ & = \frac{6(1+2t^2)(t^3-t)^2}{\sqrt{t+5t^2}(4t)^{\frac{3}{2}}} \left[\frac{4t}{1+2t^2} + \frac{6t^2-2}{t^3-t} - \frac{1+10t}{2t+10t^2} - \frac{3}{2t} \right]\end{aligned}$$

$$\begin{aligned} & \frac{d}{dt} \left[\frac{\sqrt{1+2t}}{t+\sqrt{1+t^2}} \right] \\ &= \frac{\sqrt{1+2t}}{t+\sqrt{1+t^2}} \left[\frac{1}{2} \frac{2}{1+2t} + (-1) \frac{1+\left[\frac{1}{2} \frac{2t}{\sqrt{1+t^2}}\right]}{t+\sqrt{1+t^2}} \right] \end{aligned}$$

□

44.2 Feynman method of integration / integral technique

44.2.1 principle

<https://www.bilibili.com/video/BV1Lj411L79X/?t=2m38s>

<https://www.youtube.com/watch?v=GW86SShcYbM>

$$I = \int f(x) dx$$

$$I(t) = \int f(x, t) dx$$

$$\begin{aligned} I &= \int f(x) dx \\ I(t) &= \int f(x, t) dx \\ \frac{d}{dt} I(t) &= I'(t) = \frac{d}{dt} \int f(x, t) dx \\ &\stackrel{\text{Leibniz integral rule}}{=} f(x, b(x), t) \frac{db(x)}{dx} - f(x, a(x), t) \frac{da(x)}{dx} + \int_{a(x)}^{b(x)} \frac{\partial}{\partial t} f(x, t) dx \\ &\stackrel{\dots}{=} \int \frac{\partial}{\partial t} f(x, t) dx = \int \frac{\partial f(x, t)}{\partial t} dx = \int f_t(x, t) dx \\ I(t) &= \int I'(t) dt = \int I'(t) dt + C = \left[\int I'(t) dt \right](t) \\ &\Downarrow \text{if } f(x, t=0) = f(x, 0) = f(x) \\ I &\stackrel{f(x,0)=f(x)}{=} \left[\int I'(t) dt \right](0) \end{aligned}$$

44.2.2 Dirichlet integral

as a example by Feynman method of integration / integral technique

https://en.wikipedia.org/wiki/Dirichlet_integral

<https://www.bilibili.com/video/BV1Lj411L79X/?t=4m38s>

$$\int_0^\infty \frac{\sin x}{x} dx$$

<https://www.youtube.com/watch?v=ZZccexuOpb4k>

<https://blog.csdn.net/zhuoqingjoking97298/article/details/127950915>

$$\int_{-\infty}^\infty \frac{\sin x}{x} dx$$

$$\int_{-\infty}^\infty \frac{\sin x}{x} dx$$

$f(x) = \frac{\sin x}{x}$ is an even function $\Leftrightarrow f(-x) = \frac{\sin(-x)}{(-x)} = \frac{-\sin(x)}{-x} = \frac{\sin x}{x} = f(x)$

$$\begin{aligned}
 \int_{-\infty}^{\infty} \frac{\sin x}{x} dx &= \int_{-\infty}^0 \frac{\sin x}{x} dx + \int_0^{\infty} \frac{\sin x}{x} dx \\
 &= \int_{x=-\infty}^{x=0} \frac{\sin x}{x} dx + \int_0^{\infty} \frac{\sin x}{x} dx \\
 &= \int_{x=-\infty}^{x=0} \frac{\sin x}{x} dx (, x' = -x \Leftrightarrow x = -x') + \int_0^{\infty} \frac{\sin x}{x} dx \\
 &= \int_{(-x')=-\infty}^{(-x')=0} \frac{\sin(-x')}{(-x')} d(-x') + \int_0^{\infty} \frac{\sin x}{x} dx \\
 &= \int_{x'=-(\infty)}^{x'=-0} \frac{-\sin x'}{-x'} (-dx') + \int_0^{\infty} \frac{\sin x}{x} dx \\
 &= \int_{x'=\infty}^{x'=0} \frac{\sin x'}{x'} (-dx') + \int_0^{\infty} \frac{\sin x}{x} dx \\
 &= - \int_{x'=\infty}^{x'=0} \frac{\sin x'}{x'} dx' + \int_0^{\infty} \frac{\sin x}{x} dx \\
 &= \int_{x'=0}^{x'=\infty} \frac{\sin x'}{x'} dx' + \int_0^{\infty} \frac{\sin x}{x} dx \\
 &= \int_0^{\infty} \frac{\sin x}{x} dx + \int_0^{\infty} \frac{\sin x}{x} dx = 2 \int_0^{\infty} \frac{\sin x}{x} dx
 \end{aligned}$$

$$\int_{-\infty}^{\infty} \frac{\sin x}{x} dx = 2 \int_0^{\infty} \frac{\sin x}{x} dx$$

$$\int_0^{\infty} \frac{\sin x}{x} dx$$

$$\begin{aligned}
I &= \int_0^\infty \frac{\sin x}{x} dx = \int_0^\infty f(x) dx, f(x) = \frac{\sin x}{x} \\
\int_0^\infty e^{-tx} \frac{\sin x}{x} dx &= \int_0^\infty e^{-tx} f(x) dx, \text{ Laplacian transform of } f(x) \\
&= \int_0^\infty f(x, t) dx, \begin{cases} f(x, t) = e^{-tx} f(x) \\ f(x, 0) = e^{-0x} f(x) = 1 \cdot f(x) = f(x) \end{cases} \Downarrow \\
I(t) &= \int_0^\infty f(x, t) dx = \int_0^\infty e^{-tx} f(x) dx = \int_0^\infty e^{-tx} \frac{\sin x}{x} dx \Rightarrow I(0) = I \\
\frac{d}{dt} I(t) &= I'(t) = \frac{d}{dt} \int_0^\infty f(x, t) dx = \int_0^\infty \frac{\partial f(x, t)}{\partial t} dx = \int_0^\infty \frac{\partial e^{-tx}}{\partial t} \frac{\sin x}{x} dx \\
&= \int_0^\infty \frac{\sin x}{x} \left[\frac{\partial e^{-tx}}{\partial t} \right] dx = \int_0^\infty \frac{\sin x}{x} [-xe^{-tx}] dx = \int_0^\infty \sin x [-e^{-tx}] dx \\
&= - \int_0^\infty e^{-tx} \sin x dx, \text{ Laplacian transform of } -\sin x \\
&= \int_0^\infty e^{-tx} (-\sin x) dx = \int_{x=0}^{x=\infty} e^{-tx} d\cos x, I'(t) = \int_0^\infty e^{-tx} (-\sin x) dx \\
&= [e^{-tx} \cos x]_{x=0}^{x=\infty} - \int_{x=0}^{x=\infty} \cos x de^{-tx} \\
&= [e^{-t\infty} \cos \infty] - [e^{-t0} \cos 0] - \int_{x=0}^{x=\infty} \cos x (-te^{-tx} dx) \\
&= [0 \cos \infty] - [1 \cdot 1] + \int_0^\infty te^{-tx} \cos x dx \\
&= 0 - 1 + \int_0^\infty te^{-tx} \cos x dx = -1 + \int_{x=0}^{x=\infty} te^{-tx} d\sin x \\
&= -1 + [te^{-tx} \sin x]_{x=0}^{x=\infty} - \int_{x=0}^{x=\infty} \sin x d(te^{-tx}) \\
&= -1 + [0 - 0] - \int_0^\infty \sin x (te^{-tx} (-t) dx) = -1 - \int_0^\infty t^2 e^{-tx} (-\sin x) dx \\
&= -1 - t^2 \int_0^\infty e^{-tx} (-\sin x) dx = -1 - t^2 I'(t) \\
I'(t) &= -1 - t^2 I'(t) \\
I'(t) &= \frac{-1}{1+t^2} \\
I(t) &= \int \frac{-1}{1+t^2} dt \stackrel{t=\tan\theta}{=} \int \frac{-1}{1+\tan^2\theta} d\tan\theta = - \int \frac{1}{\sec^2\theta} \sec^2\theta d\theta = - \int d\theta = -\theta + C \\
&= -\arctan t + C = -\tan^{-1} t + C \\
\int_0^\infty e^{-tx} \frac{\sin x}{x} dx &= I(t) = -\arctan t + C \\
\Downarrow \Rightarrow I &= \int_0^\infty \frac{\sin x}{x} dx = \int_0^\infty e^{-0x} \frac{\sin x}{x} dx = I(0) = -\arctan 0 + C = -0 + C \Rightarrow C = I \\
\Rightarrow 0 &= \int_0^\infty 0 \frac{\sin x}{x} dx = \int_0^\infty e^{-\infty x} \frac{\sin x}{x} dx = I(\infty) = -\arctan \infty + C = \frac{-\pi}{2} + C \\
0 &= \frac{-\pi}{2} + C \Rightarrow C = \frac{\pi}{2} \\
I &= I(0) = C = \frac{\pi}{2}, I = \int_0^\infty \frac{\sin x}{x} dx \\
\int_0^\infty \frac{\sin x}{x} dx &= \frac{\pi}{2}
\end{aligned}$$

$$\int_0^\infty \frac{\sin x}{x} dx = \frac{\pi}{2}$$

$$\int_{-\infty}^{\infty} \frac{\sin x}{x} dx = 2 \int_0^{\infty} \frac{\sin x}{x} dx = 2 \cdot \frac{\pi}{2} = \pi$$

$$\int_{-\infty}^{\infty} \frac{\sin x}{x} dx = \pi$$

□

44.2.2.1 residue method

<https://www.bilibili.com/video/BV1Lj411L79X/?t=7m26s>

44.2.2.2 sinc function

https://en.wikipedia.org/wiki/Sinc_function

https://en.wikipedia.org/wiki/Anti-aliasing_filter

https://en.wikipedia.org/wiki/Borwein_integral

<https://blog.udn.com/paraquat/22455342>

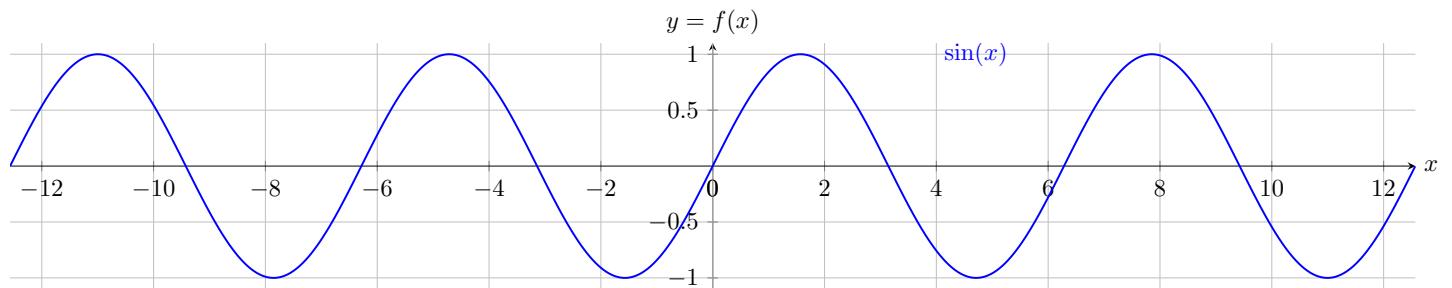


Figure 44.1: $\sin(x)$

<https://tex.stackexchange.com/questions/235006/declaring-sinc-in-tikz>

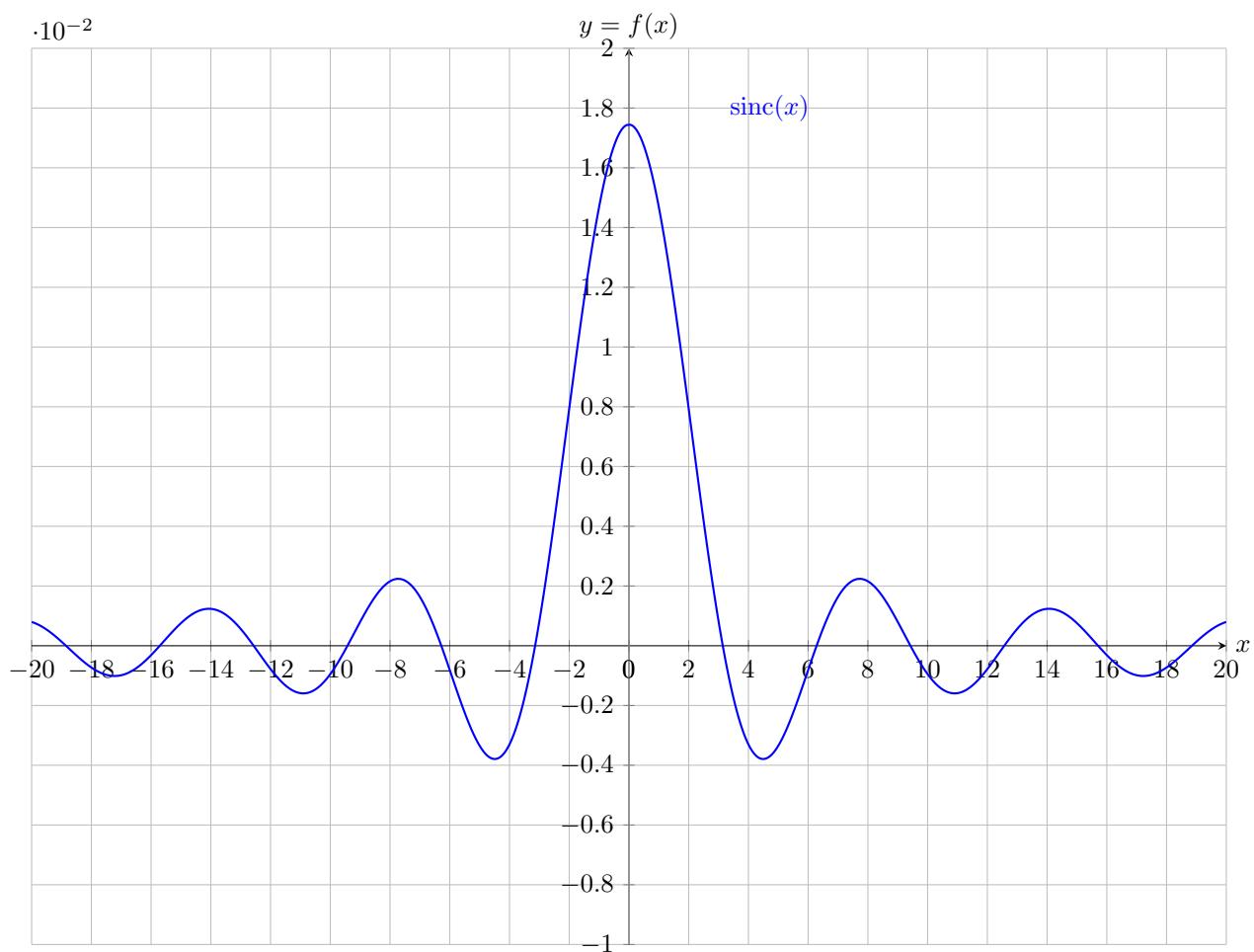
44.2.3 Gaussian integral

<https://www.youtube.com/watch?v=jP-6j6mEpRg>

$$\int_{-\infty}^{\infty} e^{-x^2} dx$$

$$\int_{-\infty}^{\infty} e^{-x^2} dx$$

$f(x) = e^{-x^2}$ is an even function $\Leftrightarrow f(-x) = e^{-(x)^2} = e^{-x^2} = f(x)$

Figure 44.2: $\text{sinc}(x)$

$$\begin{aligned}
\int_{-\infty}^{\infty} e^{-x^2} dx &= \int_{-\infty}^0 e^{-x^2} dx + \int_0^{\infty} e^{-x^2} dx \\
&= \int_{x=-\infty}^{x=0} e^{-x^2} dx + \int_0^{\infty} e^{-x^2} dx \\
&= \int_{x=-\infty}^{x=0} e^{-x^2} dx, (x' = -x \Leftrightarrow x = -x') + \int_0^{\infty} e^{-x^2} dx \\
&= \int_{(-x')=-\infty}^{(-x')=0} e^{-(-x')^2} d(-x') + \int_0^{\infty} e^{-x^2} dx \\
&= \int_{x'=-(-\infty)}^{x'=0} e^{-(x')^2} (-dx') + \int_0^{\infty} e^{-x^2} dx \\
&= \int_{x'=\infty}^{x'=0} e^{-(x')^2} dx' + \int_0^{\infty} e^{-x^2} dx \\
&= - \int_{x'=\infty}^{x'=0} e^{-(x')^2} dx' + \int_0^{\infty} e^{-x^2} dx \\
&= \int_{x'=0}^{x'=\infty} e^{-(x')^2} dx + \int_0^{\infty} e^{-x^2} dx \\
&= \int_0^{\infty} e^{-x^2} dx + \int_0^{\infty} e^{-x^2} dx = 2 \int_0^{\infty} e^{-x^2} dx
\end{aligned}$$

$$\int_{-\infty}^{\infty} e^{-x^2} dx = 2 \int_0^{\infty} e^{-x^2} dx$$

$$\int_0^{\infty} e^{-x^2} dx$$

$$I(t) = \left[\int_0^t e^{-x^2} dx \right]^2$$

$$\begin{aligned}
I'(t) &= \frac{dI(t)}{dt} = \frac{d \left[\int_0^t e^{-x^2} dx \right]^2}{dt} = \frac{d \left[\int_0^t e^{-x^2} dx \right]}{dt} \frac{d \left[\int_0^t e^{-x^2} dx \right]}{dt} = 2 \left[\int_0^t e^{-x^2} dx \right] \left[\frac{d}{dt} \int_0^t e^{-x^2} dx \right] \\
&= 2 \left[\int_0^t e^{-x^2} dx \right] \left[\frac{d}{dt} \int_0^t e^{-x^2} dx \right] \stackrel{\text{FToC}}{=} 2 \left[\int_0^t e^{-x^2} dx \right] [e^{-t^2}] \\
&= 2 \int_0^t e^{-x^2} e^{-t^2} dx = 2 \int_0^t e^{-(x^2+t^2)} dx = 2 \int_{x=0}^{x=t} e^{-(x^2+t^2)} dx, x' = \frac{x}{t} \Leftrightarrow x = tx' \\
&= 2 \int_{tx'=0}^{tx'=t} e^{-((tx')^2+t^2)} dt x' = 2 \int_{x'=0}^{x'=1} e^{-t^2((x')^2+1)} t dx' = 2 \int_0^1 t e^{-t^2(x^2+1)} dx \\
&= - \int_0^1 (-2t) e^{-t^2(x^2+1)} dx = - \int_0^1 \frac{\partial}{\partial t} \frac{e^{-t^2(x^2+1)}}{x^2+1} dx = - \frac{d}{dt} \int_0^1 \frac{e^{-t^2(x^2+1)}}{x^2+1} dx \\
&\quad \left[\int_0^t e^{-x^2} dx \right]^2 = I(t) = - \int_0^1 \frac{e^{-t^2(x^2+1)}}{x^2+1} dx + C \\
0 &= \left[\int_0^0 e^{-x^2} dx \right]^2 = I(0) = \int_0^1 \frac{1}{x^2+1} dx + C \stackrel{x=\tan\theta}{=} [-\arctan x]_0^1 + C = \frac{-\pi}{4} + C \Rightarrow C = \frac{\pi}{4} \\
&\quad \left[\int_0^t e^{-x^2} dx \right]^2 = I(t) = - \int_0^1 \frac{e^{-t^2(x^2+1)}}{x^2+1} dx + C = - \int_0^1 \frac{e^{-t^2(x^2+1)}}{x^2+1} dx + \frac{\pi}{4} \\
&\quad \left[\int_0^{\infty} e^{-x^2} dx \right]^2 = \lim_{t \rightarrow \infty} I(t) = I(\infty) = - \int_0^1 \frac{e^{-\infty^2(x^2+1)}}{x^2+1} dx + \frac{\pi}{4} = - \int_0^1 \frac{0}{x^2+1} dx + \frac{\pi}{4} = \frac{\pi}{4} \\
&\quad \left[\int_0^{\infty} e^{-x^2} dx \right]^2 = \frac{\pi}{4} \\
&\quad \int_0^{\infty} e^{-x^2} dx = \frac{\sqrt{\pi}}{2}
\end{aligned}$$

$$\int_0^\infty e^{-x^2} dx = \frac{\sqrt{\pi}}{2}$$

$$\int_{-\infty}^\infty e^{-x^2} dx = 2 \int_0^\infty e^{-x^2} dx = 2 \cdot \frac{\sqrt{\pi}}{2} = \sqrt{\pi}$$

$$\int_{-\infty}^\infty e^{-x^2} dx = \sqrt{\pi}$$

□

44.2.4 other examples

<https://www.youtube.com/watch?v=GW86SShcYbM>

$$\int_0^1 \ln x dx$$

$$I = \int_0^1 \ln x dx = \int_0^1 \ln(x) dx, f(x) = \ln x$$

$$I(t) = \int_0^1 \ln(tx) dx = \int_0^1 f(tx) dx = \int_0^1 f(x,t) dx, \begin{cases} f(x,t) = f(tx) = \ln(tx) \\ f(x,1) = f(x) = \ln x \end{cases}$$

$$I(t) = \int_0^1 f(x,t) dx = \int_0^1 f(tx) dx = \int_0^1 \ln(tx) dx \Rightarrow I(1) = I$$

$$\begin{aligned} \frac{d}{dt} I(t) &= I'(t) = \frac{d}{dt} \int_0^1 f(x,t) dx = \int_0^1 \frac{\partial f(x,t)}{\partial t} dx = \int_0^1 \frac{\partial \ln tx}{\partial t} dx \\ &= \int_0^1 \frac{1}{tx} x dx = \int_0^1 \frac{1}{t} dx = \frac{1}{t} \int_0^1 dx = \frac{1}{t} [x]_0^1 = \frac{1}{t} \\ I(t) &= \int I'(t) dt = \int \frac{1}{t} dt = \ln|t| + C \\ \Downarrow \Rightarrow I &= \int_0^1 \ln x dx = \int_0^1 \ln 1 dx = I(1) = \ln|1| + C = 0 + C \Rightarrow C = I \end{aligned}$$

no more known boundary condition

$$\begin{aligned} \Im(t) &= \int_0^1 x^t dx = \left[\frac{x^{t+1}}{t+1} \right]_{x=0}^1 = \frac{1}{t+1} \\ \frac{-1}{(t+1)^2} &= \frac{d}{dt} \frac{1}{t+1} = \frac{d}{dt} \Im(t) = \Im'(t) = \frac{d}{dt} \int_0^1 x^t dx = \int_0^1 \frac{\partial x^t}{\partial t} dx = \int_0^1 x^t \ln x dx \\ -1 &= \left[\frac{-1}{(t+1)^2} \right]_{t=0} = \Im'(0) = \int_0^1 x^0 \ln x dx = \int_0^1 1 \ln x dx = \int_0^1 \ln x dx \\ \int_0^1 \ln x dx &= -1 \end{aligned}$$

<https://www.bilibili.com/video/BV1Lj411L79X>

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

```
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 ??

$$\int_a^b f'(x) dx = f(b) - f(a) \quad (45.1)$$

$$\begin{aligned} & \left\{ \begin{array}{ll} f : [a, b] \rightarrow \mathbb{R} & (0) \\ f \text{ continuous on } [a, b] & (1) \\ f \text{ differentiable on } (a, b) & (2) \end{array} \right. \\ & \left\{ \begin{array}{ll} f' : [a, b] \rightarrow \mathbb{R} & (4) \\ f' \text{ continuous on } [a, b] & (5) \end{array} \right. \stackrel{a < b \ (3)}{\Rightarrow} \exists c \in (a, b) \left(f'(c) = \frac{\int_a^b f'(x) dx}{b-a} \right) \quad \text{MVTi1 ??} \\ & \exists c \in (a, b) \{ f'(c)(b-a) = f(b) - f(a) \} \quad 45.1 \\ \Rightarrow & \uparrow \downarrow \\ & \int_a^b f'(x) dx = f(b) - f(a) \quad 45.1 \end{aligned}$$

(↓):

$$\begin{aligned} & \left\{ \begin{array}{ll} f : [a, b] \rightarrow \mathbb{R} & (0) \\ f \text{ continuous on } [a, b] & (1) \stackrel{??}{\Rightarrow} \exists c \in (a, b) \{ f'(c)(b-a) = f(b) - f(a) \} \quad \text{MVTd1 ??} \\ f \text{ differentiable on } (a, b) & (2) \end{array} \right. \\ & [a, b] = [x_0, x_n] = [x_0, x_1] \cup [x_0, x_2] \cup \dots \cup [x_{n-1}, x_n] \quad 45.2 \\ & \stackrel{??}{\Rightarrow} \int_a^b f'(x) dx = f(b) - f(a) \end{aligned}$$

(↑):

$$\begin{aligned} & \left\{ \begin{array}{ll} f : [a, b] \rightarrow \mathbb{R} & (0) \\ f \text{ continuous on } [a, b] & (1) \stackrel{??}{\Rightarrow} \int_a^b f'(x) dx = f(b) - f(a) \quad \text{N-LF 45.1} \\ f \text{ differentiable on } (a, b) & (2) \end{array} \right. \\ & \left\{ \begin{array}{ll} f' : [a, b] \rightarrow \mathbb{R} & (4) \\ f' \text{ continuous on } [a, b] & (5) \end{array} \right. \stackrel{a < b \ (3)}{\Rightarrow} \exists c \in (a, b) \left(f'(c) = \frac{\int_a^b f'(x) dx}{b-a} \right) \quad \text{MVTi1 ??} \\ & \Rightarrow \exists c \in (a, b) \{ f'(c)(b-a) = f(b) - f(a) \} \end{aligned}$$

Proof: (\Downarrow)

重疊端點分割 $[a, b]$ 成 n 部分聯集 ($n \in \mathbb{N}$)

$$[a, b] = [x_0, x_n] = [x_0, x_1] \cup [x_0, x_2] \cup \cdots \cup [x_{n-1}, x_n] \quad (45.2)$$

$$\begin{aligned} & \left\{ \begin{array}{ll} f : [a, b] \rightarrow \mathbb{R} & (0) \\ f \text{ continuous on } [a, b] & (1) \stackrel{??}{\Rightarrow} \exists c \in (a, b) \{f'(c)(b-a) = f(b) - f(a)\} \text{ MVTd1 ??} \\ f \text{ differentiable on } (a, b) & (2) \end{array} \right. \\ & [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 \\ \Rightarrow & \left\{ \begin{array}{ll} (f : [x_0, x_n] \rightarrow \mathbb{R}) = (f : [x_0, x_1] \cup [x_0, x_2] \cup \cdots \cup [x_{n-1}, x_n] \rightarrow \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{array} \right. \\ \Rightarrow & \left\{ \begin{array}{ll} f : [x_0, x_1] \rightarrow \mathbb{R}, f : [x_0, x_2] \rightarrow \mathbb{R}, \dots, f : [x_{n-1}, x_n] \rightarrow \mathbb{R} & \Leftarrow (0) \\ f \text{ continuous on } [x_0, x_1], [x_0, x_2], \dots, [x_{n-1}, x_n] & \Leftarrow (1) \\ f \text{ differentiable on } (x_0, x_1), [x_0, x_2], \dots, [x_{n-1}, x_n] \Rightarrow f \text{ differentiable on } (x_0, x_1), (x_0, x_2), \dots, (x_{n-1}, x_n) & \Leftarrow (2) \end{array} \right. \\ \stackrel{??}{\Rightarrow} & \left\{ \begin{array}{ll} \exists c_1 \in (x_0, x_1) \{f'(c_1)(x_1 - x_0) = f(x_1) - f(x_0)\} \\ \exists c_2 \in (x_1, x_2) \{f'(c_2)(x_2 - x_1) = f(x_2) - f(x_1)\} \\ \vdots \\ \exists c_k \in (x_{k-1}, x_k) \{f'(c_k)(x_k - x_{k-1}) = f(x_k) - f(x_{k-1})\} \quad \forall k \in \mathbb{N} \cap [1, n] \\ \vdots \\ \exists c_n \in (x_{n-1}, x_n) \{f'(c_n)(x_n - x_{n-1}) = f(x_n) - f(x_{n-1})\} \end{array} \right. \\ \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}) \\ & = [f(x_1) - f(x_0)] + [f(x_2) - f(x_1)] + \cdots + [f(x_n) - f(x_{n-1})] \\ & = f(x_n) - f(x_0) = f(b) - f(a) \\ \Rightarrow & \sum_{k=1}^n f'(c_k)(x_k - x_{k-1}) = f(b) - f(a) \stackrel{\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 \rightarrow \infty} \lim_{\Delta x_k = \frac{x_n - x_0}{n}} \sum_{k=1}^n f'(c_k) \Delta x_k = \lim_{n \rightarrow \infty} \lim_{\Delta x_k = \frac{b-a}{n} \rightarrow 0} f(b) - f(a) = f(b) - f(a) \\ \Rightarrow & \int_a^b f'(x) dx = \int_{x_0}^{x_n} f'(x) dx = f(b) - f(a) \\ \Rightarrow & \int_a^b f'(x) dx = f(b) - f(a) \end{aligned}$$

$$\int_a^b f'(x) dx = f(b) - f(a)$$

得到 牛頓-萊布尼茨公式 Newton-Leibniz formula = N-LF [eq:N-LF]

Proof: (\Updownarrow)

□

$$\begin{aligned}
& \left\{ \begin{array}{ll} f : [a, b] \rightarrow \mathbb{R} & (0) \\ f \text{ continuous on } [a, b] & (1) \stackrel{??}{\Rightarrow} \int_a^b f'(x) dx = f(b) - f(a) \\ f \text{ differentiable on } (a, b) & (2) \end{array} \right. \quad \text{N-LF 45.1} \\
& \Rightarrow \left\{ \begin{array}{ll} f : [a, b] \rightarrow \mathbb{R} & (0) \\ f \text{ continuous on } [a, b] & (1) \stackrel{??}{\Rightarrow} \int_a^b f'(x) dx = f(b) - f(a) \\ f \text{ differentiable on } (a, b) & (2) \end{array} \right. \quad \text{N-LF 45.1} \\
& \Rightarrow \left\{ \begin{array}{ll} f : [a, b] \rightarrow \mathbb{R} & a < b (3) \\ f \text{ continuous on } [a, b] & (1) \Rightarrow \exists c \in (a, b) \left(f(c) = \frac{\int_a^b f(x) dx}{b-a} \right) \quad \text{MVTi1 ??} \\ f \text{ differentiable on } (a, b) & (2) \end{array} \right. \quad \text{MVTi1 ??} \\
& \Rightarrow \left\{ \begin{array}{ll} f : [a, b] \rightarrow \mathbb{R} & (0) \\ f \text{ continuous on } [a, b] & (1) \stackrel{??}{\Rightarrow} \int_a^b f'(x) dx = f(b) - f(a) \\ f \text{ differentiable on } (a, b) & (2) \end{array} \right. \quad \text{N-LF 45.1} \\
& \Rightarrow \left\{ \begin{array}{ll} \text{if } f' : [a, b] \rightarrow \mathbb{R} & (4) \\ \text{if } f' \text{ continuous on } [a, b] & (5) \Rightarrow \exists c \in (a, b) \left\{ f'(c)(b-a) = \int_a^b f'(x) dx \right\} \quad \text{MVTi1 ??} \\ f \text{ continuous on } [a, b] & (1) \end{array} \right. \\
& \Rightarrow \exists c \in (a, b) \left\{ f'(c)(b-a) = \int_a^b f'(x) dx = f(b) - f(a) \right\} \\
& \Rightarrow \exists c \in (a, b) \{ f'(c)(b-a) = f(b) - f(a) \}
\end{aligned}$$

$$\exists c \in (a, b) [f'(c)(b-a) = f(b) - f(a)]$$

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]

□

$$\begin{aligned}
& \int_a^b f'(x) dx = f(b) - f(a) \\
& \int_{x_0}^x f'(t) dt = f(x) - f(x_0) \\
& f(x) = f(x_0) + \int_{x_0}^x f'(t) dt
\end{aligned}$$

$$f(x) = f(x_0) + \int_{x_0}^x f'(t) dt$$

Lemma 45.2. univariable product rule

$$d(uv) = (du)v + udv = vdu + udv$$

$$\frac{d(uv)}{dt} = v \frac{du}{dt} + u \frac{dv}{dt}$$

Lemma 45.3. integration by parts

$$\begin{aligned} \int \frac{d(uv)}{dt} dt &= \int v \frac{du}{dt} + u \frac{dv}{dt} dt \\ \int d(uv) &= \int v \frac{du}{dt} dt + \int u \frac{dv}{dt} dt \\ uv &= \int v du + \int u dv \\ \int u \frac{dv}{dt} dt &= uv - \int v \frac{du}{dt} dt \text{ "switching derivatives"} \end{aligned} \tag{45.3}$$

$$\int u dv = uv - \int v du \text{ "switching differentials"} \tag{45.4}$$

Theorem 45.1. *univariable Taylor theorem with the remainder in integral form*

$$f(x) = \sum_{k=0}^n \frac{f^{(k)}(a)}{k!} (x-a)^k + \int_a^x f^{(n+1)}(t) \frac{(t-x)^n}{n!} dt$$

$$f(x) = \sum_{k=0}^n \frac{f^{(k)}(a)}{k!} (x-a)^k + \int_a^x f^{(n+1)}(t) \frac{(x-t)^n}{n!} dt$$

Proof:

$$\stackrel{45.1}{=} \frac{f(x) - f(a)}{\int_a^x f'(t) dt} = \int_a^x f'(t) \frac{d(t-x)}{dt} dt \stackrel{45.3}{=} [f'(t)(t-x)]_{t=a}^x - \int_a^x (t-x) \frac{df'(t)}{dt} dt \quad (45.5)$$

$$(45.6)$$

$$= [-f'(a)(a-x)] - \int_a^x f''(t)(t-x) dt = f'(a)(x-a) - \int_a^x f''(t) \frac{d\frac{(t-x)^2}{2}}{dt} dt \quad (45.7)$$

$$= f'(a)(x-a) - \left(\left[f''(t) \frac{(t-x)^2}{2} \right]_{t=a}^x - \int_a^x \frac{(t-x)^2}{2} \frac{df''(t)}{dt} dt \right) \quad (45.8)$$

$$= f'(a)(x-a) - \left(\left[-f''(a) \frac{(x-a)^2}{2} \right] - \int_a^x f'''(t) \frac{(t-x)^2}{2} dt \right) \quad (45.9)$$

$$= f'(a)(x-a) + f''(a) \frac{(x-a)^2}{2} + \int_a^x f'''(t) \frac{(t-x)^2}{2} dt \quad (45.10)$$

$$= f'(a)(x-a) + f''(a) \frac{(x-a)^2}{2} + \int_a^x f'''(t) \frac{d\frac{(t-x)^3}{2 \cdot 3}}{dt} dt \quad (45.11)$$

$$= f'(a)(x-a) + f''(a) \frac{(x-a)^2}{2} + \left(\left[f'''(t) \frac{(t-x)^3}{2 \cdot 3} \right]_{t=a}^x - \int_a^x \frac{(t-x)^3}{2 \cdot 3} \frac{df'''(t)}{dt} dt \right) \quad (45.12)$$

$$= f'(a)(x-a) + f''(a) \frac{(x-a)^2}{2} + \left(\left[f'''(a) \frac{(x-a)^3}{2 \cdot 3} \right] - \int_a^x f^{(4)}(t) \frac{(t-x)^3}{2 \cdot 3} dt \right) \quad (45.13)$$

$$= f'(a)(x-a) + f''(a) \frac{(x-a)^2}{2} + f'''(t) \frac{(x-a)^3}{2 \cdot 3} - \int_a^x f^{(4)}(t) \frac{(t-x)^3}{2 \cdot 3} dt \quad (45.14)$$

$$= f'(a)(x-a) + f''(a) \frac{(x-a)^2}{2} + f'''(t) \frac{(x-a)^3}{2 \cdot 3} - \int_a^x f^{(4)}(t) \frac{d\frac{(t-x)^4}{2 \cdot 3 \cdot 4}}{dt} dt \quad (45.15)$$

$$= f'(a)(x-a) + f''(a) \frac{(x-a)^2}{2} + f'''(t) \frac{(x-a)^3}{2 \cdot 3} \quad (45.16)$$

$$- \left(\left[f^{(4)}(t) \frac{(t-x)^4}{2 \cdot 3 \cdot 4} \right]_{t=a}^x - \int_a^x \frac{(t-x)^4}{2 \cdot 3 \cdot 4} \frac{df^{(4)}(t)}{dt} dt \right) \quad (45.17)$$

$$= f'(t)(x-a) + f''(a) \frac{(x-a)^2}{2} + f'''(t) \frac{(x-a)^3}{2 \cdot 3} \quad (45.18)$$

$$- \left(\left[-f^{(4)}(a) \frac{(x-a)^4}{2 \cdot 3 \cdot 4} \right] - \int_a^x f^{(5)}(t) \frac{(t-x)^4}{2 \cdot 3 \cdot 4} dt \right) \quad (45.19)$$

$$= f'(t)(x-a) + f''(a) \frac{(x-a)^2}{2} + f'''(a) \frac{(x-a)^3}{2 \cdot 3} + f^{(4)}(a) \frac{(x-a)^4}{2 \cdot 3 \cdot 4} \quad (45.20)$$

$$\vdots + \int_a^x f^{(5)}(t) \frac{(t-x)^4}{2 \cdot 3 \cdot 4} dt \quad (45.21)$$

$$= \sum_{k=1}^n \frac{f^{(k)}(a)}{k!} (x-a)^k + \int_a^x f^{(n+1)}(t) \frac{(t-x)^n}{n!} dt \text{ remainder in integral form} \quad (45.22)$$

□

$$\begin{aligned}
& f(x) - f(a) \\
& \stackrel{45.1}{=} \int_a^x f'(t) dt = \int_a^x f'(t) \frac{d(t-x)}{dt} dt \stackrel{45.3}{=} [f'(t)(t-x)]_{t=a}^x - \int_a^x (t-x) \frac{df'(t)}{dt} dt \\
& = [-f'(a)(a-x)] - \int_a^x f''(t)(t-x) dt = f'(a)(x-a) + \int_a^x f''(t)(x-t) dt \\
& = f'(a)(x-a) - \int_a^x f''(t)(t-x) dt = \sum_{k=1}^1 \frac{f^{(k)}(a)}{k!} (x-a)^k + \int_a^x f''(t)(x-t) dt \\
& = f'(a)(x-a) - \int_a^x f''(t) \frac{d\frac{(t-x)^2}{2}}{dt} dt \\
& = 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{df''(t)}{dt} dt \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} dt \right) \\
& = f'(a)(x-a) + f''(a) \frac{(x-a)^2}{2} + \int_a^x f'''(t) \frac{(t-x)^2}{2} dt \\
& = \sum_{k=1}^2 \frac{f^{(k)}(a)}{k!} (x-a)^k + \int_a^x f'''(t) \frac{(t-x)^2}{2} dt = \sum_{k=1}^2 \frac{f^{(k)}(a)}{k!} (x-a)^k + \int_a^x f'''(t) \frac{(x-t)^2}{2!} dt \\
& = f'(a)(x-a) + f''(a) \frac{(x-a)^2}{2} + \int_a^x f'''(t) \frac{d\frac{(t-x)^3}{2 \cdot 3}}{dt} dt \\
& = 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{df'''(t)}{dt} dt \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} dt \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} dt \\
& = \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} dt = \sum_{k=1}^3 \frac{f^{(k)}(a)}{k!} (x-a)^k + \int_a^x f^{(4)}(t) \frac{(x-t)^3}{3!} dt \\
& \stackrel{\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!} dt \\
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!} dt \\
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!} dt \\
& = \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!} dt \\
& = \sum_{k=0}^n \frac{f^{(k)}(a)}{k!} (x-a)^k + \int_a^x f^{(n+1)}(t) \frac{(x-t)^n}{n!} dt
\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 ??

$$\text{let } f(x) = \sum_{k=0}^n \frac{f^{(k)}(a)}{k!} (x-a)^k + r_n(x) \quad (45.23)$$

$$r_n(x) = \int_a^x f^{(n+1)}(t) \frac{(t-x)^n}{n!} dt \quad (45.24)$$

$$\text{let } n \in \{2k-1 | k \in \mathbb{N}\} \quad (45.25)$$

$$\text{let } f(t) \stackrel{??}{\in} [m, M] \subseteq f((a, x)) \text{ when } a < x \quad (45.26)$$

$$\int_a^x m \frac{(t-x)^n}{n!} dt \leq r_n(x) \leq \int_a^x M \frac{(t-x)^n}{n!} dt \quad (45.27)$$

$$m \int_a^x \frac{(t-x)^n}{n!} dt = M \int_a^x \frac{(t-x)^n}{n!} dt \quad (45.28)$$

$$m \left[\frac{(t-x)^{n+1}}{(n+1)!} \right]_{t=a}^x = M \left[\frac{(t-x)^{n+1}}{(n+1)!} \right]_{t=a}^x \quad (45.29)$$

$$m \frac{(x-a)^{n+1}}{(n+1)!} = M \frac{(x-a)^{n+1}}{(n+1)!} \quad (45.30)$$

$$\Downarrow \quad ?? \quad (45.31)$$

$$r_n(x) \stackrel{\exists \xi \in (a, x)}{=} f^{(n+1)}(\xi) \frac{(x-a)^{n+1}}{(n+1)!} \quad (45.32)$$

$$f(x) = \sum_{k=0}^n \frac{f^{(k)}(a)}{k!} (x-a)^k + f^{(n+1)}(\xi) \frac{(x-a)^{n+1}}{(n+1)!} \quad (45.33)$$

$$= \sum_{k=0}^n \frac{f^{(k)}(a)}{k!} (x-a)^k + \frac{f^{(n+1)}(\xi)}{(n+1)!} (x-a)^{n+1} \quad (45.34)$$

□

Theorem 45.3. univariable Taylor theorem with the remainder in big O form

$$f(a+h) = \sum_{k=0}^n \frac{f^{(k)}(a)}{k!} h^k + O(h^{n+1})$$

Proof:

$$\text{if } |f^{(n+1)}(t)| \leq K \forall t \in (a, x) \quad (45.35)$$

$$|r_n(x)| = \left| \frac{f^{(n+1)}(\xi)}{(n+1)!} (x-a)^{n+1} \right| \leq \frac{K}{(n+1)!} |x-a|^{n+1} \quad (45.36)$$

$$r_n(x) \in O((x-a)^{n+1}) \quad (45.37)$$

$$\text{let } R_n(h) = r_n(a+h) \quad (45.38)$$

$$R_n(h) \in O(h^{n+1}) \quad (45.39)$$

$$f(a+h) = \sum_{k=0}^n \frac{f^{(k)}(a)}{k!} h^k + \frac{f^{(n+1)}(\xi)}{(n+1)!} h^{n+1} \quad (45.40)$$

$$= \sum_{k=0}^n \frac{f^{(k)}(a)}{k!} h^k + O(h^{n+1}) \quad (45.41)$$

□

$$f(x) = \sum_{k=0}^n \frac{f^{(k)}(a)}{k!} (x-a)^k + \int_a^x f^{(n+1)}(t) \frac{(x-t)^n}{n!} dt$$

$$f(x) = \sin(x)$$

$$\sin(x) = \sum_{k=0}^n \frac{\sin^{(k)}(a)}{k!} (x-a)^k + \int_a^x \sin^{(n+1)}(t) \frac{(x-t)^n}{n!} dt$$

$$a = 0$$

$$\sin(x) = \sum_{k=0}^n \frac{\sin^{(k)}(0)}{k!} (x-0)^k + \int_0^x \sin^{(n+1)}(t) \frac{(x-t)^n}{n!} dt$$

$$\begin{aligned} \sin(x) &= \sum_{k=0}^n \frac{\sin^{(k)}(0)}{k!} (x-0)^k + \int_0^x \sin^{(n+1)}(t) \frac{(x-t)^n}{n!} dt \\ &= \sum_{k=0}^n \frac{\sin^{(k)}(0)}{k!} x^k + \int_0^x \sin^{(n+1)}(t) \frac{(x-t)^n}{n!} dt \\ &= \frac{\sin^{(0)}(0)}{0!} x^0 + \frac{\sin^{(1)}(0)}{1!} x^1 + \frac{\sin^{(2)}(0)}{2!} x^2 + \frac{\sin^{(3)}(0)}{3!} x^3 \\ &\quad + \int_0^x \sin^{(3+1)}(t) \frac{(x-t)^3}{3!} dt \\ &= \frac{0}{0!} x^0 + \frac{\cos(0)}{1!} x^1 + \frac{-\sin(0)}{2!} x^2 + \frac{-\cos(0)}{3!} x^3 \\ &\quad + \int_0^x \sin^{(4)}(t) \frac{(x-t)^3}{3!} dt \\ &= 0 + \frac{1}{1} x + 0 - \frac{1}{6} x^3 + \int_0^x \sin(t) \frac{(x-t)^3}{6} dt \\ &= x - \frac{1}{6} x^3 + \frac{1}{6} \int_0^x (x-t)^3 \sin(t) dt \\ \sin(x) &= x - \frac{1}{6} x^3 + \frac{1}{6} \int_0^x (x-t)^3 \sin(t) dt \end{aligned}$$

$$\begin{aligned} \sin(x) &= \sum_{k=0}^n \frac{\sin^{(k)}(0)}{k!} (x-0)^k + \int_0^x \sin^{(n+1)}(t) \frac{(x-t)^n}{n!} dt \\ &= \sum_{k=0}^n \frac{\sin^{(k)}(0)}{k!} x^k + \int_0^x \sin^{(n+1)}(t) \frac{(x-t)^n}{n!} dt \\ &= \frac{\sin^{(0)}(0)}{0!} x^0 + \frac{\sin^{(1)}(0)}{1!} x^1 + \frac{\sin^{(2)}(0)}{2!} x^2 + \frac{\sin^{(3)}(0)}{3!} x^3 + \frac{\sin^{(4)}(0)}{4!} x^4 \\ &\quad + \int_0^x \sin^{(4+1)}(t) \frac{(x-t)^4}{4!} dt \\ &= \frac{0}{0!} x^0 + \frac{\cos(0)}{1!} x^1 + \frac{-\sin(0)}{2!} x^2 + \frac{-\cos(0)}{3!} x^3 + \frac{\sin(0)}{4!} x^4 \\ &\quad + \int_0^x \sin^{(5)}(t) \frac{(x-t)^4}{4!} dt \\ &= 0 + \frac{1}{1} x + 0 - \frac{1}{6} x^3 + 0 + \int_0^x \cos(t) \frac{(x-t)^4}{24} dt \\ &= x - \frac{1}{6} x^3 + \frac{1}{24} \int_0^x (x-t)^4 \cos(t) dt \\ \sin(x) &= x - \frac{1}{6} x^3 + \frac{1}{24} \int_0^x (x-t)^4 \cos(t) dt \end{aligned}$$

$$\begin{aligned}
\sin(x) &= x - \frac{1}{6}x^3 + \frac{1}{6} \int_0^x (x-t)^3 \sin(t) dt \\
&= x - \frac{1}{6}x^3 + \frac{1}{24} \int_0^x (x-t)^4 \cos(t) dt \\
&\quad \Downarrow \\
\frac{1}{6} \int_0^x (x-t)^3 \sin(t) dt &= \frac{1}{24} \int_0^x (x-t)^4 \cos(t) dt \\
&\quad \Downarrow \\
\int_0^x (x-t)^4 \cos(t) dt &= 4 \int_0^x (x-t)^3 \sin(t) dt
\end{aligned}$$

$$f(x) = \sum_{k=0}^n \frac{f^{(k)}(a)}{k!} (x-a)^k + \int_a^x f^{(n+1)}(t) \frac{(x-t)^n}{n!} dt$$

$$f(x) = e^x = \exp(x)$$

$$e^x = \sum_{k=0}^n \frac{\exp^{(k)}(a)}{k!} (x-a)^k + \int_a^x \exp^{(n+1)}(t) \frac{(x-t)^n}{n!} dt$$

$$a = 0$$

$$e^x = \sum_{k=0}^n \frac{\exp^{(k)}(0)}{k!} (x-0)^k + \int_0^x \exp^{(n+1)}(t) \frac{(x-t)^n}{n!} dt$$

$$\begin{aligned}
e^x &= \sum_{k=0}^n \frac{\exp^{(k)}(0)}{k!} (x-0)^k + \int_0^x \exp^{(n+1)}(t) \frac{(x-t)^n}{n!} dt \\
&= \sum_{k=0}^n \frac{\exp^{(k)}(0)}{k!} x^k + \int_0^x \exp^{(n+1)}(t) \frac{(x-t)^n}{n!} dt \\
&= \sum_{k=0}^n \frac{\exp(0)}{k!} x^k + \int_0^x \exp(t) \frac{(x-t)^n}{n!} dt \\
&= \sum_{k=0}^n \frac{e^0}{k!} x^k + \int_0^x e^t \frac{(x-t)^n}{n!} dt = \sum_{k=0}^n \frac{1}{k!} x^k + \int_0^x e^t \frac{(x-t)^n}{n!} dt \\
&= \sum_{k=0}^n \frac{x^k}{k!} + \frac{1}{n!} \int_0^x (x-t)^n e^t dt \\
e^x &= \sum_{k=0}^n \frac{x^k}{k!} + \frac{1}{n!} \int_0^x (x-t)^n e^t dt
\end{aligned}$$

$$e^x = \sum_{k=0}^n \frac{x^k}{k!} + \frac{1}{n!} \int_0^x (x-t)^n e^t dt$$

$$\lim_{n \rightarrow \infty} \frac{1 + n + \frac{n^2}{2} + \cdots + \frac{n^n}{n!}}{e^n}$$

$$\begin{aligned}
& \lim_{n \rightarrow \infty} \frac{1+n+\frac{n^2}{2}+\cdots+\frac{n^n}{n!}}{e^n}, \wedge e^x = \sum_{k=0}^n \frac{x^k}{k!} + \frac{1}{n!} \int_0^x (x-t)^n e^t dt \\
&= \lim_{n \rightarrow \infty} \frac{1+n+\frac{n^2}{2}+\cdots+\frac{n^n}{n!}}{\left[\sum_{k=0}^n \frac{x^k}{k!} + \frac{1}{n!} \int_0^x (x-t)^n e^t dt \right]_{x=n}} \\
&= \lim_{n \rightarrow \infty} \frac{\sum_{k=0}^n \frac{n^k}{k!}}{\sum_{k=0}^n \frac{n^k}{k!} + \frac{1}{n!} \int_0^n (n-t)^n e^t dt} = \lim_{n \rightarrow \infty} \frac{1}{1 + \frac{\frac{1}{n!} \int_0^n (n-t)^n e^t dt}{\sum_{k=0}^n \frac{n^k}{k!}}}
\end{aligned}$$

$$\begin{aligned}
& \lim_{n \rightarrow \infty} \frac{1+n+\frac{n^2}{2}+\cdots+\frac{n^n}{n!}}{e^n}, \wedge e^x = \sum_{k=0}^n \frac{x^k}{k!} + \frac{1}{n!} \int_0^x (x-t)^n e^t dt \\
&= \lim_{n \rightarrow \infty} \frac{\sum_{k=0}^n \frac{n^k}{k!}}{e^n}, \wedge e^n = \sum_{k=0}^n \frac{n^k}{k!} + \frac{1}{n!} \int_0^n (n-t)^n e^t dt \Rightarrow \sum_{k=0}^n \frac{n^k}{k!} = e^n - \frac{1}{n!} \int_0^n (n-t)^n e^t dt \\
&= \lim_{n \rightarrow \infty} \frac{e^n - \frac{1}{n!} \int_0^n (n-t)^n e^t dt}{e^n} = \lim_{n \rightarrow \infty} \left[1 - \frac{1}{n!} \int_0^n (n-t)^n e^{t-n} dt \right], \text{ if } \lim_{n \rightarrow \infty} \frac{1}{n!} \int_0^n (t-n)^n e^{t-n} dt \in \mathbb{R} \\
&= \lim_{n \rightarrow \infty} 1 - \lim_{n \rightarrow \infty} \frac{1}{n!} \int_0^n (n-t)^n e^{t-n} dt = 1 - \lim_{n \rightarrow \infty} \frac{1}{n!} \int_0^n (n-t)^n e^{t-n} dt
\end{aligned}$$

$$\lim_{n \rightarrow \infty} \frac{1}{n!} \int_0^n (n-t)^n e^{t-n} dt$$

$$\begin{aligned}
& \lim_{n \rightarrow \infty} \frac{1}{n!} \int_0^n (n-t)^n e^{t-n} dt, x = n-t \\
&= \lim_{n \rightarrow \infty} \frac{1}{n!} \int_{t=0}^{t=n} (n-t)^n e^{t-n} dt, t = n-x \\
&= \lim_{n \rightarrow \infty} \frac{1}{n!} \int_{n-x=0}^{n-x=n} (n-(n-x))^n e^{(n-x)-n} d(n-x) \\
&= \lim_{n \rightarrow \infty} \frac{1}{n!} \left(- \int_{x=n}^{x=0} x^n e^{-x} dx \right) = \lim_{n \rightarrow \infty} \frac{1}{n!} \left(\int_{x=0}^{x=n} x^n e^{-x} dx \right) \\
&= \lim_{n \rightarrow \infty} \frac{1}{n!} \int_0^n x^n e^{-x} dx
\end{aligned}$$

$$\int x^n e^{-x} dx$$

difficult process

$$\begin{aligned}
\int x^n e^{-x} dx &= - \int x^n e^{-x} dx \\
&= - \left[x^n e^{-x} - \int e^{-x} dx^n \right] \\
&= - \left[x^n e^{-x} - \int e^{-x} n x^{n-1} dx \right] \\
&= - x^n e^{-x} + n \int x^{n-1} e^{-x} dx
\end{aligned}$$

$$\int x^n e^{-x} dx = -x^n e^{-x} + n \int x^{n-1} e^{-x} dx$$

$$\begin{aligned}
\int_0^n x^n e^{-x} dx &= [-x^n e^{-x}]_{x=0}^n + n \int_0^n x^{n-1} e^{-x} dx \\
&= [-n^n e^{-n} - (-0^n e^{-0})] + n \int_0^n x^{n-1} e^{-x} dx \\
&= -n^n e^{-n} + n \int_0^n x^{n-1} e^{-x} dx
\end{aligned}$$

$$\int_0^n x^n e^{-x} dx = -n^n e^{-n} + n \int_0^n x^{n-1} e^{-x} dx$$

more formal process with gamma function

<https://www.wolframalpha.com/input/?i2d=true&i=Divide%5BIntegrate%5BPower%5Bx%2Cn%5D%5BPower%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%5D%5BPower%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>

$$f(x) = \sum_{k=0}^n \frac{f^{(k)}(x_0)}{k!} (x - x_0)^k + \frac{f^{(n+1)}(\xi)}{(n+1)!} (x - x_0)^{n+1}$$

$$f(x_0 + \epsilon) = \sum_{k=0}^n \frac{f^{(k)}(x_0)}{k!} \epsilon^k + O(\epsilon^{n+1}), \epsilon = x - x_0$$

$$f(x + \epsilon) = \sum_{n=0}^{\infty} \frac{f^{(n)}(x)}{n!} \epsilon^n$$

$$\begin{aligned}
f(x + \epsilon) &= \sum_{n=0}^{\infty} \frac{f^{(n)}(x)}{n!} \epsilon^n = \sum_{n=0}^{\infty} \frac{\frac{d^n}{dx^n} f(x)}{n!} \epsilon^n = \sum_{n=0}^{\infty} \frac{\left(\frac{d}{dx}\right)^n f(x)}{n!} \epsilon^n \\
&= \sum_{n=0}^{\infty} \frac{\left(\epsilon \frac{d}{dx}\right)^n f(x)}{n!} = \left(\sum_{n=0}^{\infty} \frac{\left(\epsilon \frac{d}{dx}\right)^n}{n!} \right) f(x) = e^{\epsilon \frac{d}{dx}} f(x)
\end{aligned}$$

univariable Taylor operator

$$f(x + \epsilon) = e^{\epsilon \frac{d}{dx}} f(x)$$

$$e^{\epsilon \frac{d}{dx}} = \sum_{n=0}^{\infty} \frac{\left(\epsilon \frac{d}{dx}\right)^n}{n!} = 1 + \epsilon \frac{d}{dx} + \frac{1}{2} \epsilon^2 \frac{d^2}{dx^2} + \dots$$

$$f(x) = mx + b,$$

$$\begin{aligned} e^{\epsilon \frac{d}{dx}} f(x) &= \left(1 + \epsilon \frac{d}{dx} + \frac{1}{2} \epsilon^2 \frac{d^2}{dx^2} + \dots\right) (mx + b) \\ &= 1(mx + b) + \epsilon \frac{d}{dx}(mx + b) + \frac{1}{2} \epsilon^2 \frac{d^2}{dx^2}(mx + b) + \dots \\ &= (mx + b) + (\epsilon m) + \left(\frac{1}{2} \epsilon^2 0\right) + (0 + 0 + \dots) \\ &= (mx + b) + (\epsilon m) + (0 + 0 + \dots) \\ &= (mx + b) + (\epsilon m) + 0 \\ &= mx + b + m\epsilon = m(x + \epsilon) + b = f(x + \epsilon) \end{aligned}$$

<https://www.youtube.com/watch?v=HQsZG8Yxb7w&t=15m34s>

multivariable ...

$$\begin{aligned} f(x + \epsilon) &= e^{\epsilon \frac{d}{dx}} f(x) \\ f(x + \epsilon_x, y + \epsilon_y, z + \epsilon_z) &= f(x, y + \epsilon_y, z + \epsilon_z) \\ &\quad + \epsilon \frac{\partial}{\partial x} f(x, y + \epsilon_y, z + \epsilon_z) \\ &\quad + \frac{1}{2} \epsilon^2 \frac{\partial^2}{\partial x^2} f(x, y + \epsilon_y, z + \epsilon_z) \\ &\quad + \dots \\ f(x + \epsilon) &= e^{\epsilon \frac{d}{dx}} f(x) \\ f(x + \epsilon_x, y + \epsilon_y, z + \epsilon_z) &= f(\mathbf{r} + \boldsymbol{\epsilon}) = e^{\boldsymbol{\epsilon} \cdot \nabla} f(\mathbf{r}), \quad \begin{cases} \mathbf{r} = (x, y, z) \\ \boldsymbol{\epsilon} = (\epsilon_x, \epsilon_y, \epsilon_z) \\ \nabla = \left(\frac{\partial}{\partial x}, \frac{\partial}{\partial y}, \frac{\partial}{\partial z} \right) \end{cases} \\ \boldsymbol{\epsilon} \cdot \nabla &= (\epsilon_x, \epsilon_y, \epsilon_z) \cdot \left(\frac{\partial}{\partial x}, \frac{\partial}{\partial y}, \frac{\partial}{\partial z} \right) = \epsilon_x \frac{\partial}{\partial x} + \epsilon_y \frac{\partial}{\partial y} + \epsilon_z \frac{\partial}{\partial z} \end{aligned}$$

multivariable Taylor operator

$$f(\mathbf{r} + \boldsymbol{\epsilon}) = e^{\boldsymbol{\epsilon} \cdot \nabla} f(\mathbf{r})$$

<https://www.youtube.com/watch?v=HQsZG8Yxb7w&t=17m36s>

physics

45.1.1.1 making complicated equation simple

linearize

simple pendulum

$\sin(x)$ to x linearization

potential energy

$$f(x_0 + \epsilon) = \sum_{k=0}^n \frac{f^{(k)}(x_0)}{k!} \epsilon^k + O(\epsilon^{n+1})$$

$$U(x_0 + \epsilon) = \sum_{k=0}^n \frac{U^{(k)}(x_0)}{k!} \epsilon^k + O(\epsilon^{n+1}) = U(x_0) + U'(x_0)x + \frac{1}{2}U''(x_0)x^2 + \dots$$

$$F = \frac{-dU}{dx} = -U'(x_0) - U''(x_0)x - \dots \stackrel{\text{if } U'(x_0)=0}{=} -U''(x_0)x = -kx, k = U''(x_0)$$

45.1.1.2 non-relativistic limit

<https://www.youtube.com/watch?v=HQsZG8Yxb7w&t=22m32s>

$$\begin{aligned}
 E &= \sqrt{(mc^2)^2 + (pc)^2} = \sqrt{m^2c^4 + p^2c^2} \\
 &= mc^2 \sqrt{1 + \left(\frac{pc}{mc^2}\right)^2} = mc^2 \sqrt{1 + \left(\frac{p}{mc}\right)^2} = mc^2 \left(1 + \left(\frac{p}{mc}\right)^2\right)^{\frac{1}{2}} \\
 &= mc^2 \left[\frac{1}{0!} + \frac{1}{1!} \left(\frac{1}{2}\right) \left(\left(\frac{p}{mc}\right)^2\right)^1 + \frac{1}{2!} \left(\frac{1}{2}\right) \left(\frac{1}{2} - 1\right) \left(\left(\frac{p}{mc}\right)^2\right)^2 + \dots \right] \\
 &= mc^2 \left[1 + \frac{1}{2} \left(\frac{p}{mc}\right)^2 - \frac{1}{8} \left(\frac{p}{mc}\right)^4 + \dots \right] = mc^2 \left[1 + \frac{1}{2} \frac{p^2}{m^2c^2} - \frac{1}{8} \frac{p^4}{m^4c^4} + \dots \right] \\
 &= mc^2 + \frac{1}{2} \frac{p^2}{m} - \frac{1}{8} \frac{p^4}{m^3c^2} + \dots = mc^2 + \frac{p^2}{2m} - \frac{p^4}{8m^3c^2} + \dots \\
 &= E_0 + \frac{p^2}{2m} - \frac{p^4}{8m^3c^2} + \dots , \quad \begin{cases} E_0 = mc^2 \\ \frac{p^2}{2m} = \frac{(mv)^2}{2m} = \frac{mv^2}{2} = \frac{1}{2}mv^2 \end{cases}
 \end{aligned}$$

slappy 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

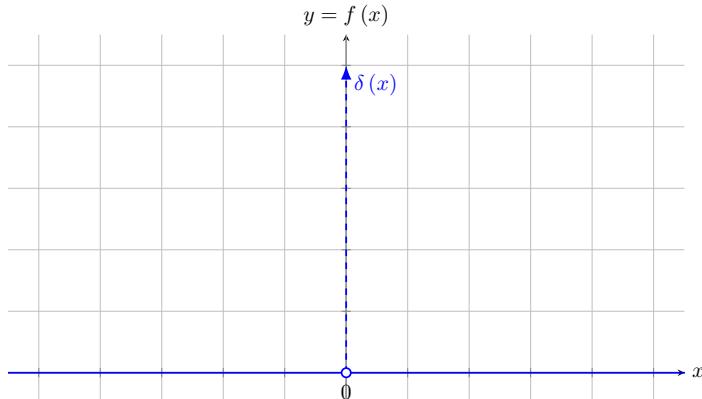
<https://www.youtube.com/watch?v=HQsZG8Yxb7w&t=29m13s>

45.2 Dirac delta function

Dirac function = Dirac delta function

https://tikz.net/delta_function/

<https://tikz.dev/tikz-arrows>

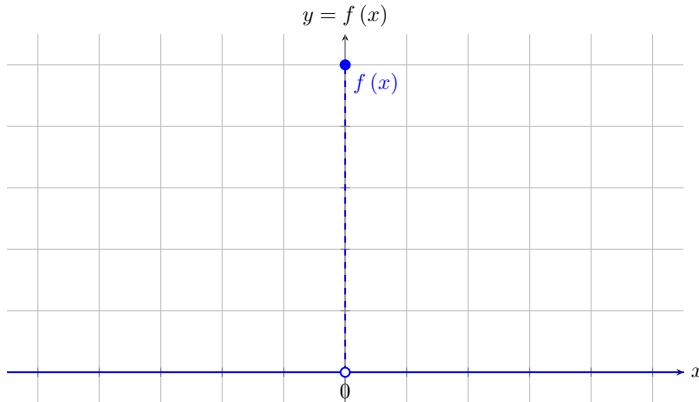


$$\delta(x) = \begin{cases} 0 & x \neq 0 \\ \infty & x = 0 \end{cases}, \int_{-\infty}^{\infty} \delta(x) dx = 1$$

Figure 45.1: Dirac function = Dirac delta function

$$\text{supp}(f) = \left\{ x \middle| \begin{cases} x \in \mathcal{D} \\ f(x) \neq 0 \end{cases} \right\}$$

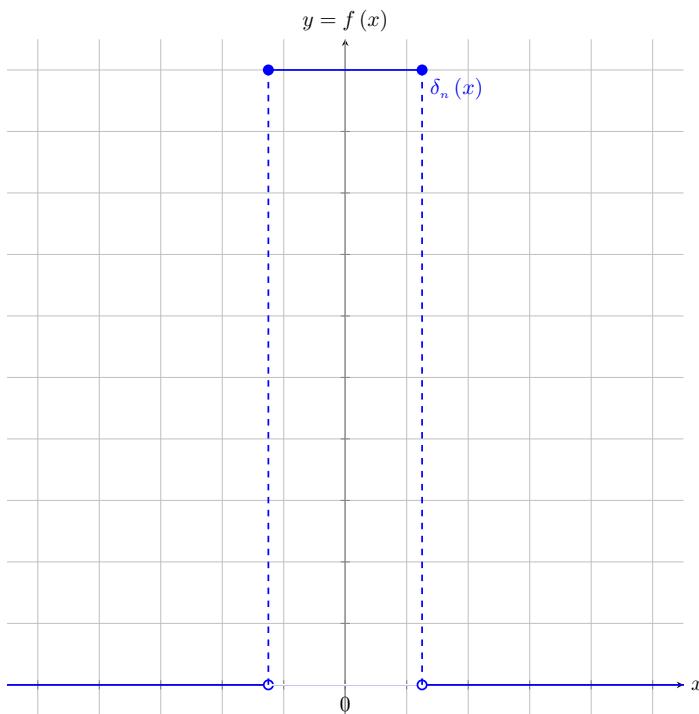
$$\text{supp}(\delta) = \left\{ x \middle| \begin{cases} x \in \mathcal{D} = \mathbb{R} \\ \delta(x) \neq 0 \end{cases} \right\} = \{0\}$$



$$f(x) = \begin{cases} 0 & x \neq 0 \\ 1 & x = 0 \end{cases}$$

Figure 45.2: $f(x) = 1$ if $x = 0$ else 0

<https://tex.stackexchange.com/questions/45275/tikz-get-values-for-predefined-dash-patterns>



$$\delta_n(x) = \begin{cases} 0 & |x| > \frac{1}{2^n} \\ n & |x| \leq \frac{1}{2^n} \end{cases}, \forall n \in \mathbb{N}$$

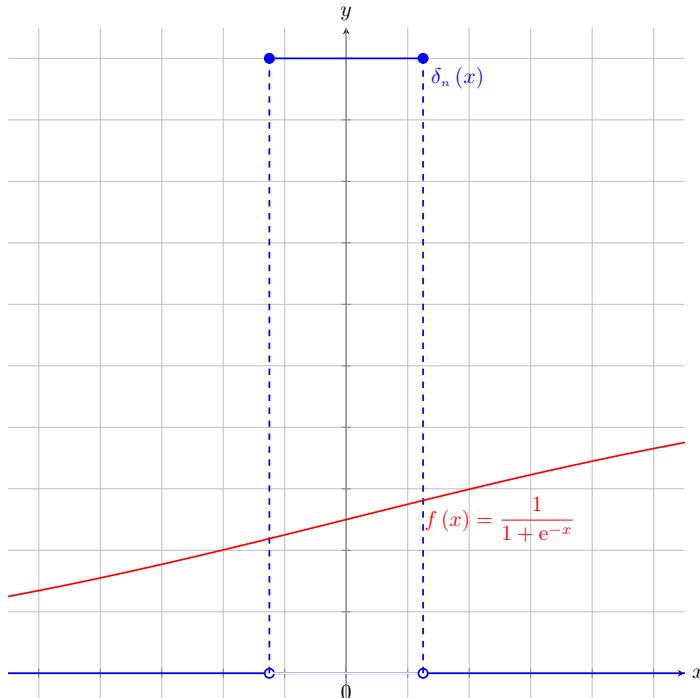
$$\delta(x) = \{\delta_n(x) | n \in \mathbb{N}\} = \{\delta_1(x), \delta_2(x), \dots\} = \lim_{n \rightarrow \infty} \delta_n(x)$$

Figure 45.3: $\delta_n(x)$

$$\delta_n(x) = \begin{cases} 0 & |x| > \frac{1}{2^n} \Leftrightarrow \begin{cases} x > \frac{1}{2^n} \\ x < -\frac{1}{2^n} \end{cases} \\ n & |x| \leq \frac{1}{2^n} \Leftrightarrow -\frac{1}{2^n} \leq x \leq \frac{1}{2^n} \Leftrightarrow x \in \left[-\frac{1}{2^n}, \frac{1}{2^n}\right] \end{cases}, \forall n \in \mathbb{N}$$

$$\int_{-\infty}^{\infty} \delta_n(x) dx = \int_{-\frac{1}{2^n}}^{\frac{1}{2^n}} n dx = n \int_{-\frac{1}{2^n}}^{\frac{1}{2^n}} dx = n [x]_{-\frac{1}{2^n}}^{\frac{1}{2^n}} = n \left[\frac{1}{2^n} - \left(-\frac{1}{2^n} \right) \right] = n \cdot \frac{1}{n} = 1$$

$$\begin{aligned}
\int_{-\infty}^{\infty} \delta(x) f(x) dx &= \lim_{n \rightarrow \infty} \int_{-\infty}^{\infty} \delta_n(x) f(x) dx \\
&= \lim_{n \rightarrow \infty} \int_{-\frac{1}{2n}}^{\frac{1}{2n}} \delta_n(x) f(x) dx \\
&= \lim_{n \rightarrow \infty} \int_{-\frac{1}{2n}}^{\frac{1}{2n}} n f(x) dx, \begin{cases} f(x) \in [m, M] \subseteq f\left(\left[\frac{-1}{2n}, \frac{1}{2n}\right]\right) \\ \Downarrow \\ m \leq f(x) \leq M \end{cases} \\
\int_{-\frac{1}{2n}}^{\frac{1}{2n}} n m dx &\leq \int_{-\frac{1}{2n}}^{\frac{1}{2n}} n f(x) dx \leq \int_{-\frac{1}{2n}}^{\frac{1}{2n}} n M dx \\
m = m \cdot 1 &= m \int_{-\frac{1}{2n}}^{\frac{1}{2n}} n dx \leq \int_{-\frac{1}{2n}}^{\frac{1}{2n}} n f(x) dx \leq M \int_{-\frac{1}{2n}}^{\frac{1}{2n}} n dx = M \cdot 1 = M \\
&\Downarrow ?? \\
\int_{-\frac{1}{2n}}^{\frac{1}{2n}} n f(x) dx &\stackrel{\exists \xi_n \in (\frac{-1}{2n}, \frac{1}{2n})}{=} f(\xi_n) \int_{-\frac{1}{2n}}^{\frac{1}{2n}} n dx = f(\xi_n) \cdot 1 = f(\xi_n) \\
\int_{-\infty}^{\infty} \delta(x) f(x) dx &= \lim_{n \rightarrow \infty} \int_{-\frac{1}{2n}}^{\frac{1}{2n}} n f(x) dx = \lim_{n \rightarrow \infty} f(\xi_n), \xi_n \in \left(\frac{-1}{2n}, \frac{1}{2n}\right) \\
&= f(0) \\
\int_{-\infty}^{\infty} \delta(x) f(x) dx &= f(0) \\
\int_{-\infty}^{\infty} \delta(x) f(x) dx &= f(0)
\end{aligned}$$



$$\begin{aligned}
\int_{-\infty}^{\infty} \delta(x) f(x) dx &= \lim_{n \rightarrow \infty} \int_{-\frac{1}{2n}}^{\frac{1}{2n}} \delta_n(x) f(x) dx \\
&= \lim_{n \rightarrow \infty} \int_{-\frac{1}{2n}}^{\frac{1}{2n}} n f(x) dx \\
&= \lim_{n \rightarrow \infty} f(\xi_n), \xi_n \in \left(\frac{-1}{2n}, \frac{1}{2n}\right)
\end{aligned}$$

$$\begin{aligned}
\int_{-\infty}^{\infty} \delta(x) f(x) dx &= f(0) \\
\int_{-\infty}^{\infty} \delta(x-0) f(x) dx &= f(0)
\end{aligned}$$

Figure 45.4: $\int_{-\infty}^{\infty} \delta(x) f(x) dx = f(0)$

$$\int_{-\infty}^{\infty} \delta(x) f(x) dx = f(0)$$

$$\int_{-\infty}^{\infty} \delta(x-0) f(x) dx = f(0)$$

$$\int_{-\infty}^{\infty} \delta(x - x_0) f(x) dx = f(x_0)$$

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

<https://www.youtube.com/watch?v=nDa3cqFk80o>

$$\begin{aligned} & \left\{ \left. \{f_n(x) | n \in \mathbb{N}\} \right| \begin{cases} f_n : \mathbb{R} \rightarrow \mathbb{R} \\ \int_{-\infty}^{\infty} f_n(x) dx = 1 \end{cases} \right\} \\ &= \left\{ \left. \begin{array}{l} \frac{n}{2} e^{-n|x|} \\ \frac{1}{\pi} \frac{n}{n^2 x^2 + 1} \\ \frac{n}{\sqrt{\pi}} e^{-n^2 x^2} \\ \frac{1}{\pi} \frac{\sin(nx)}{x} \end{array} \right| n \in \mathbb{N} \right\}, \dots \\ & \left\{ \left. \delta(x) \right| \begin{cases} \delta : \mathbb{R} \rightarrow \mathbb{R} \\ \int_{-\infty}^{\infty} \delta(x) \cdot 1 dx = 1 \end{cases} \right\} \\ & \left\{ \left. \delta(x) \right| \begin{cases} \delta : \mathbb{R} \rightarrow \mathbb{R} \\ \int_{-\infty}^{\infty} \delta(x-0) \cdot f(x) dx = f(0) \end{cases} \right\} \end{aligned}$$

In measure theory, we can define the distance of two functions by

$$d(f, g) = \sqrt{\int_{-\infty}^{\infty} [f(x) - g(x)]^2 dx}$$

for real distance of two square delta function approximations,

$$\begin{aligned} d(\delta_m, \delta_n) &= \sqrt{\int_{-\infty}^{\infty} [\delta_m(x) - \delta_n(x)]^2 dx} \\ &= \sqrt{\int_{-\infty}^{\infty} [\delta_m(x)]^2 - 2\delta_m(x)\delta_n(x) + [\delta_n(x)]^2 dx} \\ &= \sqrt{\int_{-\infty}^{\infty} [\delta_m(x)]^2 dx - 2 \int_{-\infty}^{\infty} \delta_m(x)\delta_n(x) dx + \int_{-\infty}^{\infty} [\delta_n(x)]^2 dx} \\ &= \sqrt{\int_{-\infty}^{\infty} \delta_m(x)\delta_m(x) dx - 2 \int_{-\infty}^{\infty} \delta_m(x)\delta_n(x) dx + \int_{-\infty}^{\infty} \delta_n(x)\delta_n(x) dx} \\ &= \sqrt{\delta_m(\xi_m) - 2 \int_{-\infty}^{\infty} \delta_m(x)\delta_n(x) dx + \delta_n(\xi_n)}, \begin{cases} \xi_m \in \left(\frac{-1}{2m}, \frac{1}{2m}\right) \\ \xi_n \in \left(\frac{-1}{2n}, \frac{1}{2n}\right) \end{cases} \\ &= \sqrt{m - 2 \int_{-\infty}^{\infty} \delta_m(x)\delta_n(x) dx + n}, \text{ if } m > n \\ &= \sqrt{m - 2\delta_n(\xi_n) + n}, \xi_n \in \left(\frac{-1}{2n}, \frac{1}{2n}\right) \\ &= \sqrt{m - 2n + n} = \sqrt{m - n} \in \mathbb{R} \\ d(\delta_m, \delta_n) &\stackrel{m \geq n}{=} \sqrt{m - n} \in \mathbb{R} \end{aligned}$$

$\langle d(\delta_m, \delta_n) \rangle_{n \in \mathbb{N}} = \langle \sqrt{m - n} \rangle_{n \in \mathbb{N}}$ is not a Cauchy series, not even mentioned convergence

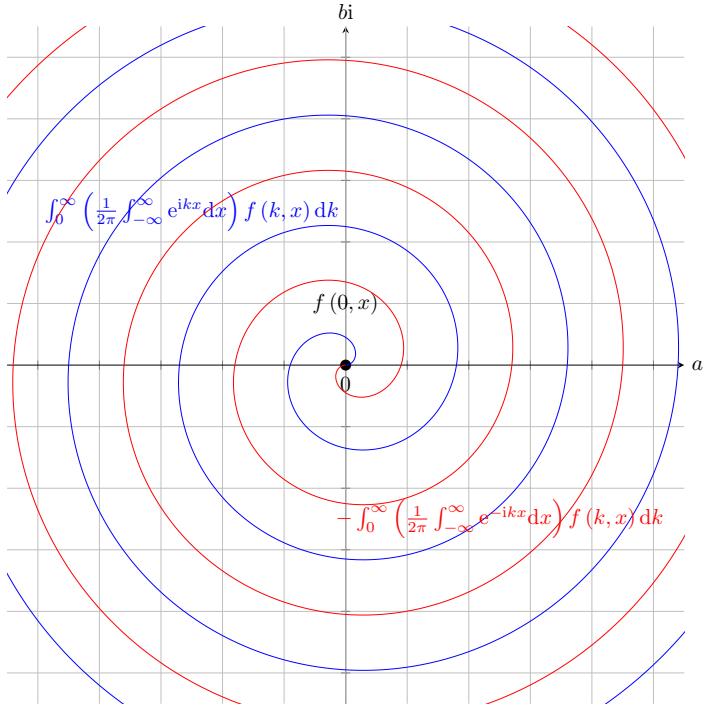
Def: 45.2

45.2.1 complex delta function

$$\delta(k) = \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{ikx} dx$$

$$\begin{aligned}
 & \int_{-\infty}^{\infty} \left(\frac{1}{2\pi} \int_{-\infty}^{\infty} e^{ikx} dx \right) f(k, x) dk \\
 &= \lim_{k' \rightarrow 0} \left[\int_{-\infty}^{k'-} \left(\frac{1}{2\pi} \int_{-\infty}^{\infty} e^{ikx} dx \right) f(k, x) dk + \int_{k'-}^{k'+} \left(\frac{1}{2\pi} \int_{-\infty}^{\infty} e^{ikx} dx \right) f(k, x) dk + \int_{k'+}^{\infty} \left(\frac{1}{2\pi} \int_{-\infty}^{\infty} e^{ikx} dx \right) f(k, x) dk \right] \\
 &\approx \lim_{k' \rightarrow 0} \left[\int_{-\infty}^0 \left(\frac{1}{2\pi} \int_{-\infty}^{\infty} e^{ikx} dx \right) f(k, x) dk + \int_{k'-}^{k'+} \left(\frac{1}{2\pi} \int_{-\infty}^{\infty} e^{i0x} dx \right) f(0, x) dk + \int_0^{\infty} \left(\frac{1}{2\pi} \int_{-\infty}^{\infty} e^{ikx} dx \right) f(k, x) dk \right] \\
 &\approx \lim_{k' \rightarrow 0} \left[\int_{-\infty}^0 \left(\frac{1}{2\pi} \int_{-\infty}^{\infty} e^{ikx} dx \right) f(k, x) dk + \int_{k'-}^{k'+} \left(\frac{1}{2\pi} \int_{-\infty}^{\infty} 1 dx \right) f(0, x) dk + \int_0^{\infty} \left(\frac{1}{2\pi} \int_{-\infty}^{\infty} e^{ikx} dx \right) f(k, x) dk \right] \\
 &\approx \lim_{k' \rightarrow 0} \left[- \int_0^{\infty} \left(\frac{1}{2\pi} \int_{-\infty}^{\infty} e^{-ikx} dx \right) f(k, x) dk + \int_{k'-}^{k'+} \left(\frac{1}{2\pi} \int_{-\infty}^{\infty} 1 dx \right) f(0, x) dk + \int_0^{\infty} \left(\frac{1}{2\pi} \int_{-\infty}^{\infty} e^{ikx} dx \right) f(k, x) dk \right] \\
 &\approx \lim_{k' \rightarrow 0} \left[- \int_0^{\infty} \left(\frac{1}{2\pi} \int_{-\infty}^{\infty} e^{-ikx} dx \right) f(x) dk + \int_{k'-}^{k'+} \left(\frac{1}{2\pi} \infty \right) f(0, x) dk + \int_0^{\infty} \left(\frac{1}{2\pi} \int_{-\infty}^{\infty} e^{ikx} dx \right) f(k, x) dk \right] \\
 &\approx \lim_{k' \rightarrow 0} \left[\int_{k'-}^{k'+} \frac{1}{dk} f(0, x) dk + \int_0^{\infty} \left(\frac{1}{2\pi} e^{ikx} - e^{-ikx} \right) f(k, x) dk \right] \\
 &\approx \lim_{k' \rightarrow 0} \left[f(0, x) + \int_0^{\infty} \left(\frac{1}{2\pi} 0 \right) f(k, x) dk \right] \approx f(0, x) \\
 f(0, x) &= \int_{-\infty}^{\infty} \left(\frac{1}{2\pi} \int_{-\infty}^{\infty} e^{ikx} dx \right) f(k, x) dk = \int_{-\infty}^{\infty} \delta(k) f(k, x) dk = \int_{-\infty}^{\infty} \delta(k-0) f(k, x) dk
 \end{aligned}$$

<https://tex.stackexchange.com/questions/150138/how-can-i-create-a-polar-plot-on-a-cartesian-grid>



$$\begin{aligned}
 f(0, x) &= \int_{-\infty}^{\infty} \left(\frac{1}{2\pi} \int_{-\infty}^{\infty} e^{ikx} dx \right) f(k, x) dk \\
 &= \int_{-\infty}^{\infty} \delta(k-0) f(k, x) dk
 \end{aligned}$$

Figure 45.5: $\int_{-\infty}^{\infty} \left(\frac{1}{2\pi} \int_{-\infty}^{\infty} e^{ikx} dx \right) f(k, x) dk$

According to Fourier transform and inverse transform, **Fourier analysis**^[50]

$$f(t) = \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{it\cdot\omega} \left(\int_{-\infty}^{\infty} e^{-i\omega\cdot s} f(s) ds \right) d\omega$$

$$f(x) = \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{ixk} \left(\int_{-\infty}^{\infty} e^{-ikx'} f(x') dx' \right) dk$$

or

$$f(t) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{it\cdot\omega} \left(\frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-i\omega\cdot s} f(s) ds \right) d\omega$$

$$f(t) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{ixk} \left(\frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-ikx'} f(x') dx' \right) dk$$

complex function if analytic always with better properties than real function

$$\begin{aligned} & \left\{ \{ \delta(k) | k \in \mathbb{R} \} \middle| \left\{ \begin{array}{l} \delta : \mathbb{R} \rightarrow \mathbb{C} \\ \int_{-\infty}^{\infty} \delta(k-0) f(k, x) dk = f(0, x) \end{array} \right\} \right\} \\ &= \left\{ \left\{ K \int_{-\infty}^{\infty} e^{ikx} (\cdot) dk \middle| k \in \mathbb{R} \right\}, \dots \right\} \end{aligned}$$

$$\delta(k) = \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{ikx} (\cdot) dx$$

$$\delta(x) = \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{ikx} (\cdot) dk$$

$$\delta(x') = \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{ikx'} (\cdot) dk$$

$$\psi(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{ikx} \psi(k) dk$$

$$\psi(k) = \frac{1}{\sqrt{2\pi}} \int e^{-ikx} \psi(x) dx$$

$$\begin{aligned} & \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{ikx} \psi(k) dk \\ &= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{ikx} \left(\frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-ikx'} \psi(x') dx' \right) dk \\ &= \frac{1}{2\pi} \int_{-\infty}^{\infty} \left(\int_{-\infty}^{\infty} e^{ikx} e^{-ikx'} \psi(x') dx' \right) dk \\ &\stackrel{\text{Fubini}}{=} \frac{1}{2\pi} \int_{-\infty}^{\infty} \left(\int_{-\infty}^{\infty} e^{ik(x-x')} \psi(x') dk \right) dx' \\ &= \int_{-\infty}^{\infty} \left(\frac{1}{2\pi} \int_{-\infty}^{\infty} e^{ik(x-x')} \psi(x') dk \right) dx' \\ &= \int_{-\infty}^{\infty} \psi(x') \left(\frac{1}{2\pi} \int_{-\infty}^{\infty} e^{ik(x-x')} dk \right) dx' \\ &= \int_{-\infty}^{\infty} \psi(x') \delta(x - x') dx' = \int_{-\infty}^{\infty} \delta(x - x') \psi(x') dx' = \psi(x) \end{aligned}$$

$$\psi(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{ikx} \psi(k) dk$$

$$\int_{-\infty}^{\infty} |\psi(k)|^2 dk = \int_{-\infty}^{\infty} |\psi(x)|^2 dx$$

$$\begin{cases} \psi(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{ikx} \psi(k) dk \\ \psi(k) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-ikx} \psi(x) dx \end{cases}$$

$$\begin{aligned} \psi^*(k) &= \overline{\psi(k)} = \overline{\frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-ikx} \psi(x) dx} \\ &= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \overline{e^{-ikx} \psi(x)} dx \\ &= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \overline{e^{-ikx} \psi(x)} dx \\ &= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{ikx} \psi^*(x) dx \end{aligned}$$

$$\begin{aligned} \int_{-\infty}^{\infty} |\psi(k)|^2 dk &= \int_{-\infty}^{\infty} \psi^*(k) \psi(k) dk \\ &= \int_{-\infty}^{\infty} \left(\frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{ikx} \psi^*(x) dx \right) \left(\frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-ikx} \psi(x) dx \right) dk \\ &= \frac{1}{2\pi} \int_{-\infty}^{\infty} \left(\int_{-\infty}^{\infty} e^{ikx} \psi^*(x) dx \right) \left(\int_{-\infty}^{\infty} e^{-ikx} \psi(x) dx \right) dk \\ &= \frac{1}{2\pi} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \psi^*(x) \int_{-\infty}^{\infty} e^{ikx} e^{-ikx'} \psi(x') dx' dx dk \\ &= \frac{1}{2\pi} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \psi^*(x) \int_{-\infty}^{\infty} e^{ik(x-x')} \psi(x') dx' dx dk \\ &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \psi^*(x) \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{ik(x-x')} \psi(x') dx' dx dk \\ &\stackrel{\text{Fubini}}{=} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \psi^*(x) \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{ik(x-x')} \psi(x') dk dx' dx \\ &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \psi^*(x) \psi(x') \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{ik(x-x')} dk dx' dx \\ &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \psi^*(x) \psi(x') \delta(x-x') dx' dx \\ &= \int_{-\infty}^{\infty} \psi^*(x) \int_{-\infty}^{\infty} \psi(x') \delta(x-x') dx' dx \\ &= \int_{-\infty}^{\infty} \psi^*(x) \psi(x) dx = \int_{-\infty}^{\infty} |\psi(x)|^2 dx \end{aligned}$$

$$\int_{-\infty}^{\infty} |\psi(k)|^2 dk = \int_{-\infty}^{\infty} |\psi(x)|^2 dx$$

convolution

$$\psi(x) = \int_{-\infty}^{\infty} \psi_1(y) \psi_2(x-y) dy \Rightarrow \psi(k) = \sqrt{2\pi} \psi_1(k) \psi_2(k)$$

$$\psi(x) = \int_{-\infty}^{\infty} \psi_1(y) \psi_2(x-y) dy$$

$$\frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{ik_1 y} \psi_1(k_1) dk_1$$

$$\psi_2(x-y) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{ik_2(x-y)} \psi_2(k_2) dk_2$$

$$\begin{cases} \psi(x) = & \int_{-\infty}^{\infty} \psi_1(y) \psi_2(x-y) dy \\ \psi_1(y) = & \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{ik_1 y} \psi_1(k_1) dk_1 \\ \psi_2(x-y) = & \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{ik_2(x-y)} \psi_2(k_2) dk_2 \end{cases}$$

$$\begin{aligned} \psi(k) &= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-ikx} \psi(x) dx = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-ikx} \int_{-\infty}^{\infty} \psi_1(y) \psi_2(x-y) dy dx \\ &= \frac{1}{(\sqrt{2\pi})^3} \int_{-\infty}^{\infty} e^{-ikx} \int_{-\infty}^{\infty} \left(\int_{-\infty}^{\infty} e^{ik_1 y} \psi_1(k_1) dk_1 \right) \left(\int_{-\infty}^{\infty} e^{ik_2(x-y)} \psi_2(k_2) dk_2 \right) dy dx \\ &= \frac{1}{(\sqrt{2\pi})^3} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} e^{ik_1 y} e^{-ik_2 y} \psi_1(k_1) e^{-ikx} e^{ik_2 x} \psi_2(k_2) dk_1 dk_2 dy dx \\ &= \frac{1}{(\sqrt{2\pi})^3} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} e^{i(k_1-k_2)y} \psi_1(k_1) e^{i(k_2-k)x} \psi_2(k_2) dk_1 dk_2 dy dx \\ &\stackrel{\text{Fubini}}{=} \frac{1}{(\sqrt{2\pi})^3} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} e^{i(k_1-k_2)y} \psi_1(k_1) e^{i(k_2-k)x} \psi_2(k_2) dy dx dk_1 dk_2 \\ &= \frac{1}{(\sqrt{2\pi})^3} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \left(\int_{-\infty}^{\infty} e^{i(k_1-k_2)y} \psi_1(k_1) dy \right) \left(\int_{-\infty}^{\infty} e^{i(k_2-k)x} \psi_2(k_2) dx \right) dk_1 dk_2 \\ &\quad \frac{1}{(\sqrt{2\pi})^3} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \left(\int_{-\infty}^{\infty} e^{i(k_1-k_2)y} \psi_1(k_1) dy \right) \left(\int_{-\infty}^{\infty} e^{i(k_2-k)x} \psi_2(k_2) dx \right) dk_1 dk_2 \\ &= \sqrt{2\pi} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \left[\frac{1}{2\pi} \int_{-\infty}^{\infty} e^{i(k_1-k_2)y} \psi_1(k_1) dy \right] \left[\frac{1}{2\pi} \int_{-\infty}^{\infty} e^{i(k_2-k)x} \psi_2(k_2) dx \right] dk_1 dk_2 \\ &= \sqrt{2\pi} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \left[\psi_1(k_1) \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{i(k_1-k_2)y} dy \right] \left[\psi_2(k_2) \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{i(k_2-k)x} dx \right] dk_1 dk_2 \\ &= \sqrt{2\pi} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} [\psi_1(k_1) \delta(k_1 - k_2)] [\psi_2(k_2) \delta(k_2 - k)] dk_1 dk_2 \\ &\stackrel{\text{Fubini}}{=} \sqrt{2\pi} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} [\psi_1(k_1) \delta(k_1 - k_2)] [\psi_2(k_2) \delta(k_2 - k)] dk_2 dk_1 \\ &= \sqrt{2\pi} \int_{-\infty}^{\infty} [\psi_1(k_1)] \left[\int_{-\infty}^{\infty} \delta(k_1 - k_2) \psi_2(k_2) \delta(k_2 - k) dk_2 \right] dk_1 \\ &= \sqrt{2\pi} \int_{-\infty}^{\infty} [\psi_1(k_1)] [\delta(k_1 - k) \psi_2(k)] dk_1 = \sqrt{2\pi} \psi_2(k) \int_{-\infty}^{\infty} \psi_1(k_1) \delta(k_1 - k) dk_1 \\ &= \sqrt{2\pi} \psi_2(k) \psi_1(k) = \sqrt{2\pi} \psi_1(k) \psi_2(k) \end{aligned}$$

$$\begin{cases} \psi(x) = & \int_{-\infty}^{\infty} \psi_1(y) \psi_2(x-y) dy \\ \psi_1(y) = & \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{ik_1 y} \psi_1(k_1) dk_1 \\ \psi_2(x-y) = & \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{ik_2(x-y)} \psi_2(k_2) dk_2 \end{cases}$$

\Downarrow

$$\psi(k) = \sqrt{2\pi} \psi_1(k) \psi_2(k)$$

45.2.2 3D delta function

<https://www.youtube.com/watch?v=Y8y965ZAmQE>

$$\delta(\mathbf{r} - \mathbf{r}') = \delta(x - x') \delta(y - y') \delta(z - z')$$

Poisson equation

[Feynman method of differentiation / derivative]^[44,1]

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

$$\Delta \left(\frac{1}{r} \right) = -4\pi\delta(\mathbf{r})$$

<https://math.stackexchange.com/questions/3774483/derivatives-of-frac1r-and-dirac-delta-function>

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

$$\Delta \left(\frac{1}{r} \right) = \nabla^2 \left(\frac{1}{r} \right) = \nabla^2 \frac{1}{|\mathbf{r}|} = \nabla^2 \frac{1}{|\mathbf{r} - \mathbf{0}|} = \Delta \frac{1}{|\mathbf{r} - \mathbf{0}|} = \Delta(\mathbf{r} - \mathbf{0}) = \Delta(\mathbf{r})$$

$$\Delta \frac{1}{r(\epsilon)} = \Delta_\epsilon(r) = \Delta_\epsilon(\mathbf{r}) = \Delta_\epsilon(\mathbf{r} - \mathbf{0})$$

$$\lim_{\epsilon \rightarrow 0} \Delta_\epsilon(\mathbf{r}) = \lim_{\epsilon \rightarrow 0} \Delta \frac{1}{r(\epsilon)} = \Delta \left(\frac{1}{r} \right) = \Delta(\mathbf{r})$$

$$K_3 \int_{\mathbb{R}^3} f(\mathbf{r}) \Delta(\mathbf{r}) d\mathbf{r} = \lim_{\epsilon \rightarrow 0} K_3 \int_{\mathbb{R}^3} f(\mathbf{r}) \Delta_\epsilon(\mathbf{r}) d\mathbf{r} = \lim_{\epsilon \rightarrow 0} K_3 \int_{\mathbb{R}^3} f(\mathbf{r}) \Delta_\epsilon(\mathbf{r} - \mathbf{0}) d\mathbf{r} = f(\mathbf{0})$$

$$K_1 \int_{\mathbb{R}^1} f(x) \delta(x) dx = \lim_{n \rightarrow \infty} K_1 \int_{\mathbb{R}^1} f(x) \delta_n(x) dx = K_1 \int_{\mathbb{R}^1} f(x) \delta_n(x - 0) dx = f(0)$$

$$\begin{aligned}
\int_{\mathbb{R}^3} f(\mathbf{r}) \Delta_\epsilon(\mathbf{r}) d\mathbf{r} &= \int_{\mathbb{R}^3} f(\mathbf{r}) \frac{-3\epsilon^2}{(r^2 + \epsilon^2)^{\frac{5}{2}}} d\mathbf{r}, \quad \begin{cases} \mathbf{r} = r\mathbf{n}_s \\ \mathbf{n}_s = \begin{pmatrix} \sin \phi \cos \theta \\ \sin \phi \sin \theta \\ \cos \phi \end{pmatrix} \end{cases} \\
&= \int_0^\infty \int_{\mathbb{S}^2} f(r\mathbf{n}_s) \frac{-3\epsilon^2}{(r^2 + \epsilon^2)^{\frac{5}{2}}} r^2 \sigma(d\mathbf{n}_s) dr, \quad \begin{cases} \mathbb{S}^2 = \text{unit sphere centered at the origin} \\ \sigma(d\mathbf{n}_s) \text{ is the surface measure of } \mathbb{S}^2 \\ r^2 \sigma(d\mathbf{n}_s) = r^2 \sin \phi d\phi d\theta \\ r = \epsilon s \end{cases} \\
&= \int_0^\infty \int_{\mathbb{S}^2} f(\epsilon s \mathbf{n}_s) \frac{-3\epsilon^2}{((\epsilon s)^2 + \epsilon^2)^{\frac{5}{2}}} (\epsilon s)^2 \sigma(d\mathbf{n}_s) d\epsilon s \\
&= \int_0^\infty \int_{\mathbb{S}^2} f(\epsilon s \mathbf{n}_s) \frac{-3\epsilon^5}{\epsilon^5 (s^2 + 1)^{\frac{5}{2}}} s^2 \sigma(d\mathbf{n}_s) ds = - \int_0^\infty \int_{\mathbb{S}^2} f(\epsilon s \mathbf{n}_s) \frac{3s^2}{(s^2 + 1)^{\frac{5}{2}}} \sigma(d\mathbf{n}_s) ds \\
&= - \int_0^\infty \int_0^{2\pi} \int_0^{2\pi} f(\epsilon s \mathbf{n}_s) \frac{3s^2}{(s^2 + 1)^{\frac{5}{2}}} \sin \phi d\phi d\theta ds = - \int_0^\infty 4\pi f(\epsilon s \mathbf{n}_s) \frac{3s^2}{(s^2 + 1)^{\frac{5}{2}}} ds \\
&\quad \lim_{\epsilon \rightarrow 0} \int_{\mathbb{R}^3} f(\mathbf{r}) \Delta_\epsilon(\mathbf{r}) d\mathbf{r} \\
&= \lim_{\epsilon \rightarrow 0} \left[- \int_0^\infty 4\pi f(\epsilon s \mathbf{n}_s) \frac{3s^2}{(s^2 + 1)^{\frac{5}{2}}} ds \right] \\
&= - 4\pi f(\mathbf{0}) \int_0^\infty \frac{3s^2}{(s^2 + 1)^{\frac{5}{2}}} ds
\end{aligned}$$

<https://www.integral-calculator.com/>

$$\begin{aligned}
&\int \frac{3s^2}{(s^2 + 1)^{\frac{5}{2}}} ds, s = \tan u \\
&= \int \frac{3 \tan^2 u}{(\tan^2 u + 1)^{\frac{5}{2}}} d\tan u = \int \frac{3 \tan^2 u}{(\tan^2 u + 1)^{\frac{5}{2}}} \sec^2 u du \\
&= \int \frac{3 \tan^2 u}{(\sec^2 u)^{\frac{5}{2}}} \sec^2 u du = \int \frac{3 \tan^2 u}{\sec^5 u} \sec^2 u du = \int \frac{3 \tan^2 u}{\sec^3 u} du \\
&= \int \frac{3 \left(\frac{\sin u}{\cos u} \right)^2}{\left(\frac{1}{\cos u} \right)^3} du = 3 \int \sin^2 u \cos u du = 3 \int \sin^2 u d\sin u = 3 \frac{\sin^3 u}{3} + C \\
&= \sin^3 u + C = \left(\frac{s}{\sqrt{s^2 + 1}} \right)^3 + C = \frac{s^3}{(s^2 + 1)^{\frac{3}{2}}} + C \stackrel{s \geq 0}{=} \frac{1}{(1 + s^{-2})^{\frac{3}{2}}} + C \\
&\int_0^\infty \frac{3s^2}{(s^2 + 1)^{\frac{5}{2}}} ds = \left[\frac{s^3}{(s^2 + 1)^{\frac{3}{2}}} \right]_0^\infty = \left[\frac{1}{(1 + s^{-2})^{\frac{3}{2}}} \right]_{s=\infty} - \left[\frac{s^3}{(s^2 + 1)^{\frac{3}{2}}} \right]_{s=0} = 1 - 0 = 1 \\
&\lim_{\epsilon \rightarrow 0} \int_{\mathbb{R}^3} f(\mathbf{r}) \Delta_\epsilon(\mathbf{r}) d\mathbf{r} \\
&= \lim_{\epsilon \rightarrow 0} \left[- \int_0^\infty 4\pi f(\epsilon s \mathbf{n}_s) \frac{3s^2}{(s^2 + 1)^{\frac{5}{2}}} ds \right] \\
&= - 4\pi f(\mathbf{0}) \int_0^\infty \frac{3s^2}{(s^2 + 1)^{\frac{5}{2}}} ds = - 4\pi f(\mathbf{0}) 1 = - 4\pi f(\mathbf{0})
\end{aligned}$$

$$\int_{\mathbb{R}^3} f(\mathbf{r}) \Delta(\mathbf{r}) d\mathbf{r} = \lim_{\epsilon \rightarrow 0} \int_{\mathbb{R}^3} f(\mathbf{r}) \Delta_\epsilon(\mathbf{r}) d\mathbf{r} = - 4\pi f(\mathbf{0}) = - 4\pi \int_{\mathbb{R}^3} f(\mathbf{r}) \delta(\mathbf{r}) d\mathbf{r}$$

$$\Delta \left(\frac{1}{r} \right) = \Delta(\mathbf{r}) = - 4\pi \delta(\mathbf{r})$$

Big Delta reciprocal r is minus 4 pi delta r.

□

$$\Delta\left(\frac{1}{r}\right) = -4\pi\delta(\mathbf{r})$$

<https://math.stackexchange.com/questions/368155/where-does-the-relation-nabla21-r-4-pi-delta3-bf-r-between-laplacian>

$$\begin{aligned} & \left\{ \{\delta(k)|k \in \mathbb{R}\} \middle| \left\{ \begin{array}{l} \delta : \mathbb{R} \rightarrow \mathbb{C} \\ \int_{-\infty}^{\infty} \delta(k-0) f(k, x) dk = f(0, x) \end{array} \right\} \right\} \\ &= \left\{ \left\{ K \int_{-\infty}^{\infty} e^{ikx} (\cdot) dk \middle| k \in \mathbb{R} \right\}, \dots \right\} \end{aligned}$$

[34.1.3]

$$\begin{aligned} & \left\{ \{\delta(\mathbf{k})|\mathbf{k} \in \mathbb{R}^3\} \middle| \left\{ \begin{array}{l} \delta : \mathbb{R}^3 \rightarrow \mathbb{H} \\ \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \delta(\mathbf{k}-\mathbf{0}) f(\mathbf{k}, \mathbf{r}) d^3\mathbf{r} = f(\mathbf{0}, \mathbf{r}) \end{array} \right\} \right\} \\ &= \left\{ \left\{ K \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} e^{i\mathbf{k}\cdot\mathbf{r}} (\cdot) d^3\mathbf{k} \middle| \mathbf{k} \in \mathbb{R}^3 \right\}, \dots \right\} \end{aligned}$$

Coulomb law

$$F_e = k \frac{Qq}{r^2}$$

$$\mathbf{F}_e = k \frac{Qq}{r^2} \hat{\mathbf{r}}$$

electric field from point electric charge

$$\begin{aligned} \mathbf{E}(\mathbf{r}) = \mathbf{E}(r) &= \mathbf{E} = k \frac{q}{r^2} \hat{\mathbf{r}} = k_e \frac{q}{|\mathbf{r}|^2} \frac{\mathbf{r}}{|\mathbf{r}|} = k_e q \frac{\mathbf{r}}{|\mathbf{r}|^3} = k_e q \frac{\mathbf{r}}{(\mathbf{r}^2)^{\frac{3}{2}}} \\ &= k_e q \frac{\mathbf{r}}{(\mathbf{r}^2)^{\frac{3}{2}}} = k_e q \left(-\nabla \frac{1}{\sqrt{\mathbf{r} \cdot \mathbf{r}}} \right) = -k_e q \nabla \frac{1}{\|\mathbf{r}\|} \\ \mathbf{E}(\mathbf{r}) &= -k_e q \nabla \frac{1}{\|\mathbf{r}\|} \end{aligned}$$

electric field from tiny point electric charge

$$\begin{aligned} d\mathbf{E}(\mathbf{r}) = d\mathbf{E}(r) &= d\mathbf{E} = k \frac{dq}{r^2} \hat{\mathbf{r}} = k \frac{dq(r)}{r^2} \hat{\mathbf{r}} = k_e dq(\mathbf{r}) \frac{\mathbf{r}}{(\mathbf{r}^2)^{\frac{3}{2}}} \\ &= k_e dq(\mathbf{r}) \frac{\mathbf{r}}{(\mathbf{r}^2)^{\frac{3}{2}}} = k_e dq(\mathbf{r}) \nabla \frac{1}{\|\mathbf{r}\|} \\ d\mathbf{E}(\mathbf{r}) &= -k_e dq(\mathbf{r}) \nabla \frac{1}{\|\mathbf{r}\|} \end{aligned}$$

$$dq = \rho dV \Leftrightarrow \rho = \frac{dq}{dV}$$

$$dq(\mathbf{r}) = \rho(\mathbf{r}) dV \Leftrightarrow \rho(\mathbf{r}) = \frac{dq(\mathbf{r})}{dV}$$

$$q = \int_q dq(\mathbf{r}) = \int_V \rho(\mathbf{r}) dV = \iiint_V \rho(\mathbf{r}) d^3\mathbf{r}$$

$$\begin{aligned}\mathbf{E} &= \iiint_V \mathbf{E}(\mathbf{r}) d^3\mathbf{r} = \iiint_V \left(-k_e \rho(\mathbf{r}) \nabla \frac{1}{\|\mathbf{r}\|} \right) d^3\mathbf{r} \\ &= -k_e \iiint_V \left(\rho(\mathbf{r}) \nabla \frac{1}{\|\mathbf{r}\|} \right) d^3\mathbf{r} \\ &= -k_e \iiint_V \left(\rho(\mathbf{r}) \nabla \frac{1}{\|\mathbf{r} - \mathbf{0}\|} \right) d^3\mathbf{r} = \mathbf{E}(\mathbf{0})\end{aligned}$$

$$\begin{aligned}\Downarrow &\begin{cases} Q \text{ at } \mathbf{r} \\ dq \text{ or } \rho \text{ at } \mathbf{r}', \text{ totally } V \end{cases} \\ \mathbf{E}(\mathbf{r}) &= \iiint_V \left(-k_e \rho(\mathbf{r}') \nabla \frac{1}{\|\mathbf{r} - \mathbf{r}'\|} \right) d^3\mathbf{r}' = \iiint_V \left(-k_e \rho(\mathbf{r}') \nabla_r \frac{1}{\|\mathbf{r} - \mathbf{r}'\|} \right) d^3\mathbf{r}' \\ \nabla_r \cdot \mathbf{E}(\mathbf{r}) &= \nabla_r \cdot \iiint_V \left(-k_e \rho(\mathbf{r}') \nabla_r \frac{1}{\|\mathbf{r} - \mathbf{r}'\|} \right) d^3\mathbf{r}' = \iiint_V \left(-k_e \rho(\mathbf{r}') \nabla_r \cdot \nabla_r \frac{1}{\|\mathbf{r} - \mathbf{r}'\|} \right) d^3\mathbf{r}' \\ &= \iiint_V \left(-k_e \rho(\mathbf{r}') \nabla_r^2 \frac{1}{\|\mathbf{r} - \mathbf{r}'\|} \right) d^3\mathbf{r}' = \iiint_V \left(-k_e \rho(\mathbf{r}') \Delta_r \frac{1}{\|\mathbf{r} - \mathbf{r}'\|} \right) d^3\mathbf{r}' \\ &= \iiint_V f(\mathbf{r}') \Delta(\mathbf{r} - \mathbf{r}') d^3\mathbf{r}', f(\mathbf{r}') = -k_e \rho(\mathbf{r}') \\ &= \iiint_V f(\mathbf{r}') \delta(\mathbf{r} - \mathbf{r}') d^3\mathbf{r}', f(\mathbf{r}') = -k_e \rho(\mathbf{r}') \\ &= -4\pi f(\mathbf{r}' = \mathbf{r}) = -4\pi [-k_e \rho(\mathbf{r}' = \mathbf{r})] \\ &= -4\pi f(\mathbf{r}) = -4\pi [-k_e \rho(\mathbf{r})] = 4\pi k_e \rho(\mathbf{r}) \\ \nabla \cdot \mathbf{E}(\mathbf{r}) &= 4\pi k_e \rho(\mathbf{r})\end{aligned}$$

$$\nabla \cdot \mathbf{E}(\mathbf{r}) = 4\pi k_e \rho(\mathbf{r})$$

<https://www.bilibili.com/video/BV1qu411578c/?t=7m26s>

https://en.wikipedia.org/wiki/Helmholtz_equation

Helmholtz equation

$$(\Delta + k^2) \frac{e^{\pm ik \cdot r}}{r} = -4\pi \delta(\mathbf{r})$$

$$(\nabla_r^2 + k^2) \frac{e^{\pm ik \cdot r}}{\|\mathbf{r}\|} = -4\pi \delta(\mathbf{r})$$

45.2.3 quantum state

<https://www.bilibili.com/video/BV1qu411578c/?t=7m45s>

<https://www.bilibili.com/video/BV1xu4y1S7W2>

45.2.4 rigorous definition of delta function

$$\int_{-\infty}^{\infty} \delta(x - 0) \cdot f(x) dx = f(0)$$

$$\int \delta(x) f(x) dx = f(0)$$

$$\int \delta(x)(\cdot) dx : f(x) \rightarrow f(0)$$

$$f(x) \xrightarrow{\int \delta(x)(\cdot) dx} f(0)$$

f must be with good properties

$$\text{continuous linear functional} \begin{cases} \text{function-like} & \eta_\varphi(f) = \int \varphi(x) f(x) dx \\ \text{non-function-like, e.g.} & \delta_{x'}(f) = f(x') \end{cases}$$

general function with “formally” integral notation

$$\delta_{x'}(f) = \int \delta(x - x') f(x) dx = f(x')$$

<https://www.bilibili.com/video/BV1qu411578c/?t=11m36s>

Quantum mechanics is founded on Hilbert space; However, 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.

量子力学本不需 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/?t=3m2s>

45.3.1 Taylor vs. Fourier

$$f(x) = x_0 x^0 + x_1 x^1 + x_2 x^2 + \dots = \sum_{k=0}^{\infty} x_k x^k, x_k = \frac{f^{(k)}(0)}{k!}$$

$$f(r) = r_0 r^0 + r_1 r^1 + r_2 r^2 + \dots = \sum_{k=0}^{\infty} r_k r^k, r_k = \frac{f^{(k)}(0)}{k!}$$

$$f(\theta) = a_0 \cos(0\theta) + a_1 \cos(1\theta) + a_2 \cos(2\theta) + \dots = \sum_{k=0}^{\infty} a_k \cos(k\theta)$$

$$a_k = \int_{-\infty}^{\infty} \cos(k\theta) f(\theta) d\theta$$

$$\begin{aligned} f(\theta) &= a_0 \cos(0\theta) + a_1 \cos(1\theta) + a_2 \cos(2\theta) + \dots = \sum_{k=0}^{\infty} a_k \cos(k\theta) \\ &= a_0 \Re[e^{i0\theta}] + a_1 \Re[e^{i1\theta}] + a_2 \Re[e^{i2\theta}] + \dots = \sum_{k=0}^{\infty} a_k \Re[e^{ik\theta}] = \sum_{k=0}^{\infty} a_k \Re[(e^{i\theta})^k] \\ &= \Re[a_0 e^{i0\theta} + a_1 e^{i1\theta} + a_2 e^{i2\theta} + \dots] = \Re \left[\sum_{k=0}^{\infty} a_k e^{ik\theta} \right] = \Re \left[\sum_{k=0}^{\infty} a_k (e^{i\theta})^k \right] \end{aligned}$$

$$f(x) = x_0 x^0 + x_1 x^1 + x_2 x^2 + \dots = \sum_{k=0}^{\infty} x_k x^k$$

$$f(z) = z_0 z^0 + z_1 z^1 + z_2 z^2 + \dots = \sum_{k=0}^{\infty} z_k z^k$$

$$z = r e^{i\theta}$$

$$\begin{aligned}
f(z) &= z_0 z^0 + z_1 z^1 + z_2 z^2 + \cdots = \sum_{k=0}^{\infty} z_k z^k \\
f(re^{i\theta}) &= z_0 (re^{i\theta})^0 + z_1 (re^{i\theta})^1 + z_2 (re^{i\theta})^2 + \cdots = \sum_{k=0}^{\infty} z_k (re^{i\theta})^k \\
&= f(r) = z_0 (e^{i\theta})^0 r^0 + z_1 (e^{i\theta})^1 r^1 + z_2 (e^{i\theta})^2 r^2 + \cdots = \sum_{k=0}^{\infty} z_k (e^{i\theta})^k r^k \\
&= r_0 r^0 + r_1 r^1 + r_2 r^2 + \cdots = \sum_{k=0}^{\infty} r_k r^k, r_k = z_k (e^{i\theta})^k \\
&= f(e^{i\theta}) = z_0 r^0 (e^{i\theta})^0 + z_1 r^1 (e^{i\theta})^1 + z_2 r^2 (e^{i\theta})^2 + \cdots = \sum_{k=0}^{\infty} z_k r^k (e^{i\theta})^k \\
&= (e^{i\theta})_0 (e^{i\theta})^0 + (e^{i\theta})_1 (e^{i\theta})^1 + (e^{i\theta})_2 (e^{i\theta})^2 + \cdots = \sum_{k=0}^{\infty} (e^{i\theta})_k (e^{i\theta})^k \\
&, (e^{i\theta})_k = z_k r^k
\end{aligned}$$

45.4 Hilbert space

Definition 45.1. sequence limit

$$\lim_{n \rightarrow \infty} a_n = a \Leftrightarrow \forall \epsilon > 0, \exists N \in \mathbb{N}, \forall n \in \mathbb{N} (n > N \Rightarrow |a_n - a| < \epsilon)$$

$$x_1, x_2, \dots, x_n, x_{n+1}, \dots$$

- 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

$$\begin{gathered}
\{x_n\}_{n \in \mathbb{N}} \text{ is convergent} \\
\Leftrightarrow \\
\forall \epsilon > 0, \exists N \in \mathbb{N}, \forall m, n \in \mathbb{N} (m, n > N \Rightarrow |x_m - x_n| < \epsilon) \Leftrightarrow \{x_n\}_{n \in \mathbb{N}} \text{ is a Cauchy sequence}
\end{gathered}$$

Definition 45.2. Cauchy sequence

$$\begin{gathered}
\{x_n\}_{n \in \mathbb{N}} \text{ is a Cauchy sequence} \\
\Leftrightarrow \\
\forall \epsilon > 0, \exists N \in \mathbb{N}, \forall m, n \in \mathbb{N} (m, n > N \Rightarrow |x_m - x_n| < \epsilon)
\end{gathered}$$

$$\text{convergence } \begin{cases} |x_n - x| < \epsilon & \text{convergent sequence} \\ |x_m - x_n| < \epsilon & \text{Cauchy sequence} \end{cases}$$

$$\text{convergence } \begin{cases} |x_n - x| < \epsilon & \text{convergent sequence} \\ \downarrow \text{easy proof} & \uparrow \text{hard proof, weak condition needs all real number as supporting base} \\ |x_m - x_n| < \epsilon & \text{Cauchy sequence} \end{cases}$$

$$\text{convergence } \begin{cases} |x_n - x| < \epsilon & \text{convergent sequence} \\ \downarrow & \uparrow x_n, x_m, x \in \mathbb{R} \text{ with real completeness} \\ |x_m - x_n| < \epsilon & \text{Cauchy sequence} \end{cases}$$

45.4.1 inner product space

vector space definition^[40.2]

Fig: 40.1

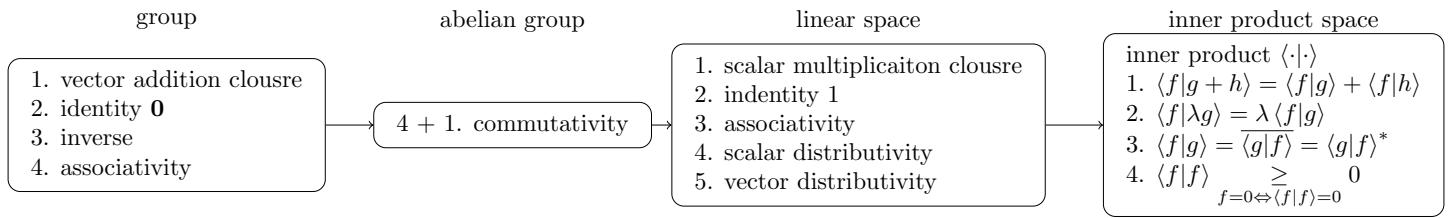


Figure 45.6: inner product space construction

$$\langle \cdot | \cdot \rangle : \begin{cases} \langle f|g+h \rangle = \langle f|g \rangle + \langle f|h \rangle \\ \langle f|cg \rangle = c \langle f|g \rangle \\ \langle f|g \rangle = \overline{\langle g|f \rangle} = \langle g|f \rangle^* \\ \langle f|f \rangle \geq 0 \\ f=0 \Leftrightarrow \langle f|f \rangle = 0 \end{cases}$$

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,

$$\begin{cases} d(f, g) = d\langle f|g \rangle = \sqrt{\langle f-g|f-g \rangle} \\ f, g \in \mathbb{R} \\ \langle f|g \rangle = fg = f \cdot g \\ \Downarrow \\ d\langle f|g \rangle = \sqrt{(f-g)^2} = |f-g| \end{cases}$$

For n -dimensional vector

$$\begin{cases} d(f, g) = d\langle f|g \rangle = \sqrt{\langle f-g|f-g \rangle} \\ f, g \in F^n \\ \langle f|g \rangle = \sum_{k=1}^n f_k g_k \\ \Downarrow \\ d\langle f|g \rangle = \sqrt{\sum_{k=1}^n (f_k - g_k)^2} \end{cases}$$

For continuous complex function on $[a, b]$,

$$\begin{cases} d(f, g) = d\langle f|g \rangle = \sqrt{\langle f-g|f-g \rangle} \\ f, g \in C_c([a, b]) \\ \langle f|g \rangle = \int_a^b \overline{f(x)} g(x) dx \\ \Downarrow \\ d\langle f|g \rangle = \sqrt{\int_a^b \overline{[f-g](x)} [f-g](x) dx} = \sqrt{\int_a^b |[f-g](x)|^2 dx} \end{cases}$$

For real number sequence,

$$\begin{aligned}
 & \left\{ \begin{array}{l} d(f, g) = d \langle f | g \rangle = \sqrt{\langle f - g | f - g \rangle} \\ f, g \in \mathbb{R} \\ \langle f | g \rangle = fg = f \cdot g \end{array} \right. \downarrow \\
 & d \langle f | g \rangle = \sqrt{(f - g)^2} = |f - g| \\
 & \text{convergence } \left\{ \begin{array}{ll} |x_n - x| < \epsilon & \text{convergent sequence} \\ |x_m - x_n| < \epsilon & \text{Cauchy sequence} \end{array} \right. \\
 & \text{convergence } \left\{ \begin{array}{ll} |x_n - x| < \epsilon & \text{convergent sequence} \\ \downarrow & \uparrow x_n, x_m, x \in \mathbb{R} \text{ with real completeness} \\ |x_m - x_n| < \epsilon & \text{Cauchy sequence} \end{array} \right.
 \end{aligned}$$

For complex function sequence,

$$\begin{aligned}
 & \left\{ \begin{array}{l} d(f, g) = d \langle f | g \rangle = \sqrt{\langle f - g | f - g \rangle} \\ f, g \in C_c([a, b]) \\ \langle f | g \rangle = \int_a^b \overline{f(x)} g(x) dx \end{array} \right. \text{continuous complex function} \\
 & \quad \downarrow \\
 & d \langle f | g \rangle = \sqrt{\int_a^b |\overline{f(x)}(x)|^2 dx} = \sqrt{\int_a^b |[f - g](x)|^2 dx} \\
 & \text{function convergence } \left\{ \begin{array}{ll} d \langle f_n | f \rangle < \epsilon & \text{convergent function sequence} \\ d \langle f_m | f_n \rangle < \epsilon & \text{Cauchy function sequence} \end{array} \right. \\
 & \text{function convergence } \left\{ \begin{array}{ll} d \langle f_n | f \rangle < \epsilon & \text{convergent function sequence} \\ \downarrow & \nparallel \text{not always continuous} \\ d \langle f_m | f_n \rangle < \epsilon & \text{Cauchy function sequence} \end{array} \right.
 \end{aligned}$$

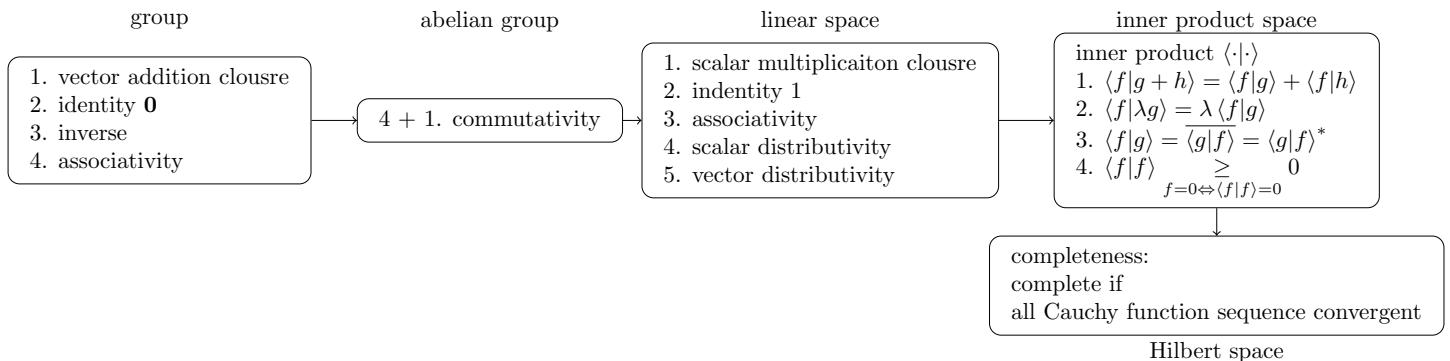


Figure 45.7: Hilbert space construction

45.5 Dirac bracket

reciprocal space^[40.4.1]

double dual concept^[40.4.1.2]

$$\begin{aligned}
 V &= \{ \mathbf{e}_1 \quad \mathbf{e}_2 \quad \mathbf{e}_3 \quad \mathbf{v} \quad \dots \} \\
 \mathbf{e}^1 : \downarrow &\quad \downarrow \quad \downarrow \quad \downarrow \quad \downarrow \\
 V^* = \{ \mathbf{e}^2 \} F &\supseteq \{ 1 \quad 0 \quad 0 \quad v_1 \quad \dots \} \\
 \mathbf{e}^3 : \downarrow &\quad \downarrow \quad \downarrow \quad \downarrow \quad \downarrow \\
 V &= \{ \mathbf{e}_1 \quad \mathbf{e}_2 \quad \mathbf{e}_3 \quad \mathbf{v} \quad \dots \} \\
 \mathbf{v}^* : \downarrow &\quad \downarrow \quad \downarrow \quad \downarrow \quad \downarrow \\
 \vdots &\quad F \supseteq \{ v^{*1} \quad v^{*2} \quad v^{*3} \quad v^{*i} v_i \quad \dots \}
 \end{aligned}$$

$$\begin{array}{ccccccc}
& V^* & = \{ & e^1 & e^2 & e^3 & \omega & \dots \} \\
V^{**} = \{ & \begin{matrix} e^{1*} \\ e^{2*} \\ e^{3*} \\ \omega^* \end{matrix} : & \begin{matrix} \downarrow \\ F \\ V^* \\ \vdots \end{matrix} & \supseteq \{ & \begin{matrix} 1 & 0 & 0 & \omega^1 & \dots \end{matrix} \\
& \begin{matrix} e^1 & e^2 & e^3 & \omega & \dots \end{matrix} \\
& \begin{matrix} \downarrow \\ \vdots \end{matrix} \\
& \begin{matrix} F \\ V^* \\ \vdots \end{matrix} & \supseteq \{ & \begin{matrix} \omega^{1*} & \omega^{2*} & \omega^{3*} & \omega^{\mu*} \omega^{\mu} & \dots \end{matrix} \\
& \begin{matrix} e^1 & e^2 & e^3 & v^* & \dots \end{matrix} \\
& \begin{matrix} \downarrow \\ \vdots \end{matrix} \\
& \begin{matrix} e_1 \\ e_2 \\ e_3 \\ v \end{matrix} : & \begin{matrix} \downarrow \\ \vdots \end{matrix} & \supseteq \{ & \begin{matrix} 1 & 0 & 0 & v^{*1} & \dots \end{matrix} \\
& \begin{matrix} F \\ V^* \\ \vdots \end{matrix} & \supseteq \{ & \begin{matrix} e^1 & e^2 & e^3 & v^* & \dots \end{matrix} \\
& \begin{matrix} \downarrow \\ \vdots \end{matrix} \\
& \begin{matrix} F \\ \vdots \end{matrix} & \supseteq \{ & \begin{matrix} v_1 & v_2 & v_3 & v_{\mu} v^{*\mu} & \dots \end{matrix}
\end{array}$$

dual space of span of partials^[41.2]

coefficient of linear combination for vector space and dual space^[41.4]

$$\begin{array}{ccccccc}
& V & = \{ & e_1 & e_2 & e_3 & v & \dots \} \\
V^* = \{ & \begin{matrix} e^1 \\ e^2 \\ e^3 \\ v^* \end{matrix} : & \begin{matrix} \downarrow \\ F \\ V \\ \vdots \end{matrix} & \supseteq \{ & \begin{matrix} 1 & 0 & 0 & v_1 & \dots \end{matrix} \\
& \begin{matrix} e_1 & e_2 & e_3 & v & \dots \end{matrix} \\
& \begin{matrix} \downarrow \\ \vdots \end{matrix} \\
& \begin{matrix} F \\ \vdots \end{matrix} & \supseteq \{ & \begin{matrix} v^{*1} & v^{*2} & v^{*3} & v^{*i} v_i & \dots \end{matrix}
\end{array}$$

inner product space^[45.4.1]

$$\begin{cases} d(f, g) = d \langle f | g \rangle = \sqrt{\langle f - g | f - g \rangle} \\ f, g \in C_c([a, b]) \\ (f, g) = \langle f | g \rangle = \int_a^b \overline{f(x)} g(x) dx = \int_a^b f^*(x) g(x) dx \end{cases} \quad \text{continuous complex function}$$

$$\downarrow$$

$$d \langle f | g \rangle = \sqrt{\int_a^b [f - g](x) [f - g](x) dx} = \sqrt{\int_a^b |[f - g](x)|^2 dx}$$

$$v^*(f) = \eta_f(\cdot) = (f, \cdot) = \langle f | \cdot \rangle = \int_a^b \overline{f(x)}(\cdot) dx = \int_a^b f^*(x)(\cdot) dx$$

$$\eta_f(g) = (f, g) = \langle f | g \rangle = \int_a^b \overline{f(x)} g(x) dx = \int_a^b f^*(x) g(x) dx$$

https://en.wikipedia.org/wiki/Antilinear_map

$$\begin{cases} v^*(f + g) = v^*(f) + v^*(g) & \text{additivity} = \text{superposition} \\ v^*(\lambda f) = \lambda^* v^*(f) = \bar{\lambda} v^*(f) & \text{conjugate homogeneity} = \text{complex conjugate} \end{cases}$$

$$\begin{cases} v^*(f + g) = v^*(f) + v^*(g) & \text{additivity} = \text{superposition} \\ v^*(\lambda f) = \lambda^* v^*(f) = \bar{\lambda} v^*(f) & \text{conjugate homogeneity} = \text{complex conjugate} \end{cases}$$

$$\Leftrightarrow \begin{cases} v^*(f + g) = \lambda^* v^*(f) + v^*(g) = \bar{\lambda} v^*(f) + v^*(g) & \text{antilinearity} \end{cases}$$

$$\begin{array}{ccccccc}
& V & = \{ & \hat{f}_1 & \hat{f}_2 & \hat{f}_3 & f & \dots \} \\
\hat{f}^1 = \left(\begin{matrix} \hat{f}_1, \cdot \end{matrix} \right) & : & \begin{matrix} \downarrow \\ F \\ V \\ \vdots \end{matrix} & \supseteq \{ & \begin{matrix} (\hat{f}_1, \hat{f}_1) = 1 & (\hat{f}_1, \hat{f}_2) = 0 & (\hat{f}_1, \hat{f}_3) = 0 & (\hat{f}_1, f) = \dots \end{matrix} \\
& \begin{matrix} \hat{f}_1 & \hat{f}_2 & \hat{f}_3 & f & \dots \end{matrix} \\
& \begin{matrix} \downarrow \\ \vdots \end{matrix} \\
& \begin{matrix} \hat{f}^2 = \left(\begin{matrix} \hat{f}_2, \cdot \end{matrix} \right) \\ \hat{f}^3 = \left(\begin{matrix} \hat{f}_3, \cdot \end{matrix} \right) \\ f^* = (f, \cdot) \end{matrix} & \begin{matrix} \vdots \\ F \\ V \\ \vdots \end{matrix} & \supseteq \{ & \begin{matrix} (\hat{f}_2, \hat{f}_1) \\ (\hat{f}_3, \hat{f}_1) \\ (f, \hat{f}_1) \end{matrix} & \begin{matrix} (\hat{f}_2, \hat{f}_2) \\ (\hat{f}_3, \hat{f}_2) \\ (f, \hat{f}_2) \end{matrix} & \begin{matrix} (\hat{f}_3, \hat{f}_3) \\ (\hat{f}_3, f) \\ (f, \hat{f}_3) \end{matrix} & \begin{matrix} (f, f) \\ \dots \end{matrix}
\end{array}$$

$$\begin{array}{ccccccc}
V & = \{ & \hat{f}_1 & & \hat{f}_2 & & \hat{f}_3 & & f & \cdots \} \\
\hat{f}^1 = \left\langle \hat{f}_1 \middle| \cdot \right\rangle & : & \downarrow & & \downarrow & & \downarrow & & \downarrow \\
\hat{f}^2 = \left\langle \hat{f}_2 \middle| \cdot \right\rangle & } & F & \supseteq \{ & \left\langle \hat{f}_1 \middle| \hat{f}_1 \right\rangle = 1 & \left\langle \hat{f}_1 \middle| \hat{f}_2 \right\rangle = 0 & \left\langle \hat{f}_1 \middle| \hat{f}_3 \right\rangle = 0 & \left\langle \hat{f}_1 \middle| f \right\rangle & \cdots \} \\
\hat{f}^3 = \left\langle \hat{f}_3 \middle| \cdot \right\rangle & V & = \{ & \hat{f}_1 & \hat{f}_2 & \hat{f}_3 & f & \cdots \} \\
f^* = \langle f | \cdot \rangle & : & \downarrow & & \downarrow & & \downarrow & & \downarrow \\
& \vdots & F & \supseteq \{ & \langle f | \hat{f}_1 \rangle & \langle f | \hat{f}_2 \rangle & \langle f | \hat{f}_3 \rangle & \langle f | f \rangle & \cdots \}
\end{array}$$

row vector is dual vector of column vector

$$f = \begin{pmatrix} a \\ b \end{pmatrix} \rightarrow v^*(f) = \eta_f = (a^* \quad b^*) = (\bar{a} \quad \bar{b}) = \overline{\begin{pmatrix} a \\ b \end{pmatrix}}^\top = \overline{\begin{pmatrix} a \\ b \end{pmatrix}}^\top = \begin{pmatrix} a \\ b \end{pmatrix}^\dagger$$

$$\eta_f \cdot f = (a^* \quad b^*) \begin{pmatrix} a \\ b \end{pmatrix} = a^*a + b^*b$$

<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>

$$\text{Hilbert space } V \left\{ \begin{array}{l} B = \left\{ \hat{f}_i \right\}_{i \in I} = \left\{ \dots, \hat{f}_i, \dots \right\} \\ \sum_i c_i \hat{f}_i = 0 \Rightarrow c_i = 0 \quad \text{bases are linear independent} \\ f = \sum_i c_i \hat{f}_i, c_i \in \mathbb{C} \quad \text{expansion over bases} \\ \left\{ \left(\hat{f}_i, \hat{f}_j \right) = \left\langle \hat{f}_i \middle| \hat{f}_j \right\rangle = \delta_{ij} \right. \\ \left. \text{completeness: } \forall \hat{f}_i \in B, \exists f \in V - B \left[\left(\hat{f}_i, f \right) = \left\langle \hat{f}_i \middle| f \right\rangle = 0 \right] \right. \end{array} \right.$$

$$f = \sum_i \left(\hat{f}_i, f \right) \hat{f}_i = \sum_i \left(\hat{f}_i, \hat{f}_j \right) \hat{f}_i = \sum_i \left\langle \hat{f}_i \middle| \hat{f}_j \right\rangle \hat{f}_i$$

<https://www.bilibili.com/video/BV1GN411m7A9/?t=7m55s>

45.5.1 Dirac bracket symmetry

$$f = \sum_i \left(\hat{f}_i, f \right) \hat{f}_i = \sum_i \hat{f}_i \left(\hat{f}_i, f \right)$$

$$f^* = (f, \cdot) = \left(\sum_i \left(\hat{f}_i, f \right) \hat{f}_i, \cdot \right) = \left(\sum_i \hat{f}_i \left(\hat{f}_i, f \right), \cdot \right)$$

$$\xrightarrow{\text{antilinearity}} \overline{\sum_i \left(\hat{f}_i, f \right)} \left(\hat{f}_i, \cdot \right) = \sum_i \overline{\left(\hat{f}_i, f \right)} \left(\hat{f}_i, \cdot \right)$$

$$= \sum_i \left(f, \hat{f}_i \right) \left(\hat{f}_i, \cdot \right)$$

$$\xrightarrow{\text{Dirac bracket}} \sum_i \left\langle f \middle| \hat{f}_i \right\rangle \left\langle \hat{f}_i \middle| \cdot \right\rangle = \langle f |$$

⇓

$$f = \sum_i \hat{f}_i \left(\hat{f}_i, f \right)$$

$$\xrightarrow{\text{Dirac bracket}} \sum_i \left| \hat{f}_i \right\rangle \left\langle \hat{f}_i \middle| f \right\rangle = |f\rangle$$

$$\begin{cases} |f\rangle = \sum_i |\hat{f}_i\rangle \langle \hat{f}_i| f \rangle = \sum_i \langle \hat{f}_i| f \rangle |\hat{f}_i\rangle \\ \langle f| = \sum_i \langle f| \hat{f}_i \rangle \langle \hat{f}_i| \\ \langle \hat{f}_i| \hat{f}_j \rangle = \delta_{ij} \end{cases}$$

$$V^* = \{ \begin{array}{c} V = \{ |\hat{f}_1\rangle, |\hat{f}_2\rangle, |\hat{f}_3\rangle, |f\rangle, \dots \} \\ \begin{array}{c} \left\langle \begin{array}{c} \hat{f}_1 \\ \hat{f}_2 \\ \hat{f}_3 \\ f \end{array} \right| \end{array} \end{array} \} \quad F \supseteq \{ \langle \hat{f}_1| \hat{f}_1 \rangle = 1, \langle \hat{f}_1| \hat{f}_2 \rangle = 0, \langle \hat{f}_1| \hat{f}_3 \rangle = 0, \langle \hat{f}_1| f \rangle = 0, \dots \} \\ \begin{array}{c} V = \{ |\hat{f}_1\rangle, |\hat{f}_2\rangle, |\hat{f}_3\rangle, |f\rangle, \dots \} \\ \begin{array}{c} \left\langle \begin{array}{c} \hat{f}_1 \\ \hat{f}_2 \\ \hat{f}_3 \\ f \end{array} \right| \end{array} \end{array} \} \\ \begin{array}{c} \vdots \quad F \supseteq \{ \langle f| \hat{f}_1 \rangle, \langle f| \hat{f}_2 \rangle, \langle f| \hat{f}_3 \rangle, \langle f| f \rangle, \dots \} \\ \vdots \end{array}$$

$$\begin{cases} v^*(f+g) = v^*(f) + v^*(g) & \text{additivity = superposition} \\ v^*(\lambda f) = \lambda^* v^*(f) = \bar{\lambda} v^*(f) & \text{conjugate homogeneity = complex conjugate} \end{cases}$$

$\Leftrightarrow \{ v^*(f+g) = \lambda^* v^*(f) + v^*(g) = \bar{\lambda} v^*(f) + v^*(g) \text{ antilinearity}$

$$\begin{cases} (c\psi, \cdot) \rightarrow \langle c\psi| = c^* \langle \psi| & c\psi \rightarrow |c\psi\rangle = c|\psi\rangle \\ \langle \psi| \leftrightarrow |\psi\rangle \\ \langle \psi| \phi \rangle = \overline{\langle \phi| \psi \rangle} & \langle \phi| \psi \rangle = \overline{\langle \psi| \phi \rangle} \\ \langle c\psi| \phi \rangle = c^* \langle \psi| \phi \rangle & \langle \phi| c\psi \rangle = c \langle \phi| \psi \rangle \end{cases}$$

$|f\rangle \langle g| = ?$

$|f\rangle \langle g| h\rangle = |f\rangle \langle g|h\rangle = \langle g|h\rangle |f\rangle$

\Downarrow

 $|h\rangle \xrightarrow{|f\rangle \langle g|} |f\rangle \langle g| h\rangle = \langle g|h\rangle |f\rangle$

$\langle h| |f\rangle \langle g| = \langle h| f\rangle \langle g|$

\Downarrow

 $\langle h| \xrightarrow{|f\rangle \langle g|} \langle h| |f\rangle \langle g| = \langle h| f\rangle \langle g|$

$|f\rangle \langle g|$ is a linear transform or rank-2 tensor

$$\begin{aligned} |f\rangle &= \sum_i |\hat{f}_i\rangle \langle \hat{f}_i| f \rangle \\ &= \sum_i |\hat{f}_i\rangle \langle \hat{f}_i| |f\rangle \\ &= \left(\sum_i |\hat{f}_i\rangle \langle \hat{f}_i| \right) |f\rangle = 1 |f\rangle = |f\rangle \\ 1 &= \sum_i |\hat{f}_i\rangle \langle \hat{f}_i| \\ \sum_i |\hat{f}_i\rangle \langle \hat{f}_i| &= 1 \end{aligned}$$

$\sum_i |\hat{f}_i\rangle \langle \hat{f}_i| = 1$

the above is completeness relationship.

$$\begin{cases} |f\rangle = 1|f\rangle = \left(\sum_i |\hat{f}_i\rangle \langle \hat{f}_i| \right) |f\rangle = \sum_i |\hat{f}_i\rangle \langle \hat{f}_i| f \rangle \\ \langle f| = \langle f| 1 = \langle f| \left(\sum_i |\hat{f}_i\rangle \langle \hat{f}_i| \right) = \sum_i \langle f| \hat{f}_i \rangle \langle \hat{f}_i| \end{cases}$$

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

46.1 Newton

<https://www.bilibili.com/video/BV1ge411e7K7>

46.2 Lagrange

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

$$\int_{-\frac{\tau}{2}+\psi}^{\frac{\tau}{2}+\psi} \cos\left(n\frac{2\pi}{\tau}t + \phi\right) dt = \left[\frac{\sin\left(n\frac{2\pi}{\tau}t + \phi\right)}{n\frac{2\pi}{\tau}} \right]_{-\frac{\tau}{2}+\psi}^{\frac{\tau}{2}+\psi} = 0$$

$$\cos \alpha \cos \beta = \frac{\cos(\alpha + \beta) + \cos(\alpha - \beta)}{2}$$

$$\sin \alpha \sin \beta = \frac{-\cos(\alpha + \beta) + \cos(\alpha - \beta)}{2}$$

$$\sin \alpha \cos \beta = \frac{\sin(\alpha + \beta) + \sin(\alpha - \beta)}{2}$$

$$\cos \alpha \sin \beta = \frac{\sin(\alpha + \beta) - \sin(\alpha - \beta)}{2}$$

$$\begin{aligned} & \int \cos\left(m\frac{2\pi}{\tau}t\right) \cos\left(n\frac{2\pi}{\tau}t + \phi_n\right) dt \\ &= \begin{cases} \int \cos \phi_0 dt & m = n = 0 \\ \int \frac{\cos((m+n)\frac{2\pi}{\tau}t + \phi_n) + \cos((m-n)\frac{2\pi}{\tau}t - \phi_n)}{2} dt & mn \neq 0 \end{cases} \\ &= \begin{cases} t \cos \phi_0 & m = n = 0 \\ \begin{cases} \int \frac{\cos(2n\frac{2\pi}{\tau}t + \phi_n) + \cos(-\phi_n)}{2} dt & m = n \\ \int \frac{\cos((m+n)\frac{2\pi}{\tau}t + \phi_n) + \cos((m-n)\frac{2\pi}{\tau}t - \phi_n)}{2} dt & m \neq n \end{cases} & mn \neq 0 \end{cases} \\ &= \begin{cases} t \cos \phi_0 & m = n = 0 \\ \begin{cases} \int \frac{\cos(2n\frac{2\pi}{\tau}t + \phi_n) + \cos \phi_n}{2} dt & m = n \\ \int \frac{\cos((m+n)\frac{2\pi}{\tau}t + \phi_n) + \cos((m-n)\frac{2\pi}{\tau}t - \phi_n)}{2} dt & m \neq n \end{cases} & mn \neq 0 \end{cases} \\ &= \begin{cases} t \cos \phi_0 & m = n = 0 \\ \begin{cases} \int \frac{\cos(2n\frac{2\pi}{\tau}t + \phi_n)}{2} dt + \frac{t \cos \phi_n}{2} & m = n \\ \int \frac{\cos((m+n)\frac{2\pi}{\tau}t + \phi_n) + \cos((m-n)\frac{2\pi}{\tau}t - \phi_n)}{2} dt & m \neq n \end{cases} & mn \neq 0 \end{cases} \end{aligned}$$

$$\cos \alpha \cos \beta = \frac{\cos(\alpha + \beta) + \cos(\alpha - \beta)}{2}$$

$$\sin \alpha \sin \beta = \frac{-\cos(\alpha + \beta) + \cos(\alpha - \beta)}{2}$$

$$\sin \alpha \cos \beta = \frac{\sin(\alpha + \beta) + \sin(\alpha - \beta)}{2}$$

$$\cos \alpha \sin \beta = \frac{\sin(\alpha + \beta) - \sin(\alpha - \beta)}{2}$$

$$\begin{aligned}
& \int_{-\frac{\tau}{2}+\psi}^{\frac{\tau}{2}+\psi} \cos\left(m \frac{2\pi}{\tau} t\right) \cos\left(n \frac{2\pi}{\tau} t + \phi_n\right) dt \\
&= \begin{cases} \tau \cos \phi_0 & m = n = 0 \\ \begin{cases} \frac{\tau \cos \phi_n}{2} & m = n \\ 0 & m \neq n \end{cases} & mn \neq 0 \end{cases} \\
&= \begin{cases} \begin{cases} \tau \cos \phi_0 & n = 0 \\ \frac{\tau \cos \phi_n}{2} & n \neq 0 \end{cases} & m = n \\ 0 & m \neq n \end{cases} \\
& \frac{2}{\tau \cos \phi_n} \int_{-\frac{\tau}{2}+\psi}^{\frac{\tau}{2}+\psi} \cos\left(m \frac{2\pi}{\tau} t\right) \cos\left(n \frac{2\pi}{\tau} t + \phi_n\right) dt \\
&= \begin{cases} 2 & n = 0 \\ 1 & n \neq 0 \\ 0 & m \neq n \end{cases} \\
& \int_{-\frac{\tau}{2}+\psi}^{\frac{\tau}{2}+\psi} \sin\left(m \frac{2\pi}{\tau} t\right) \cos\left(n \frac{2\pi}{\tau} t + \phi_n\right) dt \\
&= \begin{cases} 0 & m = n \\ 0 & m \neq n \end{cases} \\
& \int_{-\frac{\tau}{2}+\psi}^{\frac{\tau}{2}+\psi} \cos\left(m \frac{2\pi}{\tau} t\right) \sin\left(n \frac{2\pi}{\tau} t + \phi_n\right) dt \\
&= \begin{cases} 0 & m = n \\ 0 & m \neq n \end{cases}
\end{aligned}$$

$$\begin{aligned}
& \frac{2}{\tau \cos \phi_n} \int_{-\frac{\tau}{2}+\psi}^{\frac{\tau}{2}+\psi} \cos\left(m \frac{2\pi}{\tau} t\right) \cos\left(n \frac{2\pi}{\tau} t + \phi_n\right) dt = \begin{cases} \begin{cases} 2 & n = 0 \\ 1 & n \neq 0 \\ 0 & m \neq n \end{cases} & m = n \\ 0 & m \neq n \end{cases} \\
& \frac{2}{\tau \cos \phi_n} \int_{-\frac{\tau}{2}+\psi}^{\frac{\tau}{2}+\psi} \sin\left(m \frac{2\pi}{\tau} t\right) \sin\left(n \frac{2\pi}{\tau} t + \phi_n\right) dt = \begin{cases} 1 & m = n \neq 0 \\ 0 & \neg(m = n \neq 0) \end{cases} \\
& \int_{-\frac{\tau}{2}+\psi}^{\frac{\tau}{2}+\psi} \sin\left(m \frac{2\pi}{\tau} t\right) \cos\left(n \frac{2\pi}{\tau} t + \phi_n\right) dt = \begin{cases} 0 & m = n \\ 0 & m \neq n \end{cases} \\
& \int_{-\frac{\tau}{2}+\psi}^{\frac{\tau}{2}+\psi} \cos\left(m \frac{2\pi}{\tau} t\right) \sin\left(n \frac{2\pi}{\tau} t + \phi_n\right) dt = \begin{cases} 0 & m = n \\ 0 & m \neq n \end{cases}
\end{aligned}$$

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

14

$$\mathbf{x} = \langle x_1, x_2, \dots, x_n \rangle = [x_1 \ x_2 \ \dots \ x_n]^\top = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}$$

$$f(x_1, x_2, \dots, x_n) = f(\langle x_1, x_2, \dots, x_n \rangle) = f(\mathbf{x})$$

$$\mathbf{y} = \langle y_1, y_2, \dots, y_m \rangle = [y_1 \ y_2 \ \dots \ y_m]^\top = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_m \end{bmatrix}$$

54.1 vector-by-scalar

$$\frac{\partial \mathbf{y}}{\partial x} = \begin{bmatrix} \frac{\partial y_1}{\partial x} \\ \frac{\partial y_2}{\partial x} \\ \vdots \\ \frac{\partial y_m}{\partial x} \end{bmatrix}$$

54.2 scalar-by-vector

$$\nabla f = \frac{\partial}{\partial \mathbf{x}} f = \frac{\partial f}{\partial \mathbf{x}} = \left\langle \frac{\partial f}{\partial x_1}, \frac{\partial f}{\partial x_2}, \dots, \frac{\partial f}{\partial x_n} \right\rangle = \begin{bmatrix} \frac{\partial f}{\partial x_1} & \frac{\partial f}{\partial x_2} & \dots & \frac{\partial f}{\partial x_n} \end{bmatrix}^\top = \begin{bmatrix} \frac{\partial f}{\partial x_1} \\ \frac{\partial f}{\partial x_2} \\ \vdots \\ \frac{\partial f}{\partial x_n} \end{bmatrix}$$

$$f_i = f_i(x_1, x_2, \dots, x_n) = f_i(\mathbf{x})$$

$$\mathbf{f} = \langle f_1, f_2, \dots, f_m \rangle = [f_1 \ f_2 \ \dots \ f_m]^\top = \begin{bmatrix} f_1 \\ f_2 \\ \vdots \\ f_m \end{bmatrix}$$

54.3 vector-by-vector

54.3.1 numerator-layout notation

分子布局

$$\frac{\partial \mathbf{y}}{\partial \mathbf{x}} = \begin{bmatrix} \frac{\partial y_1}{\partial \mathbf{x}} \\ \frac{\partial y_2}{\partial \mathbf{x}} \\ \vdots \\ \frac{\partial y_p}{\partial \mathbf{x}} \end{bmatrix} = \begin{bmatrix} \frac{\partial \mathbf{y}}{\partial x_1} & \frac{\partial \mathbf{y}}{\partial x_2} & \cdots & \frac{\partial \mathbf{y}}{\partial x_n} \end{bmatrix} = \begin{bmatrix} \frac{\partial y_1}{\partial x_1} & \frac{\partial y_1}{\partial x_2} & \cdots & \frac{\partial y_1}{\partial x_n} \\ \frac{\partial y_2}{\partial x_1} & \frac{\partial y_2}{\partial x_2} & \cdots & \frac{\partial y_2}{\partial x_n} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial y_m}{\partial x_1} & \frac{\partial y_m}{\partial x_2} & \cdots & \frac{\partial y_m}{\partial x_n} \end{bmatrix} = \begin{bmatrix} \frac{\partial y_i}{\partial x_j} \end{bmatrix}_{m \times n}$$

54.3.2 denominator-layout notation

分母布局

$$\frac{\partial \mathbf{y}}{\partial \mathbf{x}} = \begin{bmatrix} \frac{\partial \mathbf{y}}{\partial x_1} \\ \frac{\partial \mathbf{y}}{\partial x_2} \\ \vdots \\ \frac{\partial \mathbf{y}}{\partial x_n} \end{bmatrix} = \begin{bmatrix} \frac{\partial y_1}{\partial \mathbf{x}} & \frac{\partial y_2}{\partial \mathbf{x}} & \cdots & \frac{\partial y_m}{\partial \mathbf{x}} \end{bmatrix} = \begin{bmatrix} \frac{\partial y_1}{\partial x_1} & \frac{\partial y_1}{\partial x_2} & \cdots & \frac{\partial y_1}{\partial x_n} \\ \frac{\partial y_2}{\partial x_1} & \frac{\partial y_2}{\partial x_2} & \cdots & \frac{\partial y_2}{\partial x_n} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial y_m}{\partial x_1} & \frac{\partial y_m}{\partial x_2} & \cdots & \frac{\partial y_m}{\partial x_n} \end{bmatrix} = \begin{bmatrix} \frac{\partial y_j}{\partial x_i} \end{bmatrix}_{n \times m}$$

Chapter 55

genetics

[楊欣洲: genetics](<https://www.youtube.com/playlist?list=PLTp0eSi9MdkPp0swo8-VVplaG8bateq7q>)

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 1st & 2nd 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 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: VCLLA = 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
 - * $p = P(A)$
 - * $q = P(a) = 1 - p$
- penetrance
 - * $f_{AA} = P(\text{affected} \mid AA)$
 - * $f_{Aa} = P(\text{affected} \mid Aa) = f_{aA}$
 - * $f_{aa} = P(\text{affected} \mid aa)$
- 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)

$$\mathcal{N}(x | \mu, \sigma^2) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left\{-\frac{(x-\mu)^2}{2\sigma^2}\right\}$$

If a continuous random variable X follows a normal distribution with mean μ and variance σ^2

$$\begin{aligned} X &\sim n(\mu, \sigma^2) = \mathcal{N}(\mu, \sigma^2) \\ \Leftrightarrow f_X(x) &= \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2} = \frac{e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2}}{\sigma\sqrt{2\pi}} = n(x | \mu, \sigma^2) \\ &= \frac{1}{\sqrt{2\pi}\sigma} \exp\left\{-\frac{(x-\mu)^2}{2\sigma^2}\right\} = \mathcal{N}(x | \mu, \sigma^2) \end{aligned}$$

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

$$\begin{aligned} Z &\sim n(0, 1^2) = \mathcal{N}(0, 1^2) \\ \Leftrightarrow f_Z(z) &= \frac{e^{-\frac{1}{2}\left(\frac{z-0}{1}\right)^2}}{1 \cdot \sqrt{2\pi}} = \frac{1}{\sqrt{2\pi}} \exp\left\{-\frac{z^2}{2}\right\} \\ &= n(z | 0, 1^2) = \mathcal{N}(z | 0, 1^2) \end{aligned}$$

To generalize from univariate random variables to multivariate random vectors, a random vector¹⁵

$$\mathbf{Z} = \langle Z_1, Z_2, \dots, Z_p \rangle = [Z_1 \ Z_2 \ \dots \ Z_p]^\top = \begin{bmatrix} Z_1 \\ Z_2 \\ \vdots \\ Z_p \end{bmatrix}$$

with p random variable components is said to follow the standard multivariate normal distribution if and only if its joint PDF is given by

$$f_{\mathbf{Z}}(\mathbf{z}) = \frac{1}{(2\pi)^{p/2}} \exp\left\{-\frac{\mathbf{z}^\top \mathbf{z}}{2}\right\} = \frac{1}{(2\pi)^{p/2}} \exp\left\{-\frac{\mathbf{z} \cdot \mathbf{z}}{2}\right\} \quad (56.1)$$

(56.1) can be rewritten as the following

$$\begin{aligned}
f_{\mathbf{Z}}(\mathbf{z}) &= \underbrace{\frac{1}{\sqrt{2\pi}} \frac{1}{\sqrt{2\pi}} \cdots \frac{1}{\sqrt{2\pi}}}_{p \text{ times}} \exp \left\{ -\frac{z_1^2}{2} - \frac{z_2^2}{2} - \cdots - \frac{z_p^2}{2} \right\} \\
&= \frac{1}{\sqrt{2\pi}} \exp \left\{ -\frac{z_1^2}{2} \right\} \frac{1}{\sqrt{2\pi}} \exp \left\{ -\frac{z_2^2}{2} \right\} \cdots \frac{1}{\sqrt{2\pi}} \exp \left\{ -\frac{z_p^2}{2} \right\} \\
&= f(z_1) f(z_2) \cdots f(z_p)
\end{aligned}$$

where

$$f_{Z_i}(z_i) = \frac{1}{\sqrt{2\pi}} \exp \left\{ -\frac{z_i^2}{2} \right\} = f(z_i) \Rightarrow Z_i \sim \mathcal{N}(0, 1^2) = \text{n}(0, 1^2) \quad (56.2)$$

$$\begin{aligned}
f_{Z_i}(z_i) &= \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} f_{\mathbf{Z}}(z_1, \dots, z_{i-1}, z_i, z_{i+1}, \dots, z_p) dz_1 \cdots dz_{i-1} dz_{i+1} \cdots dz_p \\
&= \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} f(z_1) \cdots f(z_{i-1}) f(z_i) f(z_{i+1}) \cdots f(z_p) dz_1 \cdots dz_{i-1} dz_{i+1} \cdots dz_p \\
&= f(z_i) \int_{-\infty}^{\infty} f(z_1) dz_1 \cdots \int_{-\infty}^{\infty} f(z_{i-1}) dz_{i-1} \int_{-\infty}^{\infty} f(z_{i+1}) dz_{i+1} \cdots \int_{-\infty}^{\infty} f(z_p) dz_p \\
&= f(z_i) \stackrel{(56.2)}{=} \frac{1}{\sqrt{2\pi}} \exp \left\{ -\frac{z_i^2}{2} \right\}
\end{aligned}$$

covariance matrix^[16]

Definition 56.2. covariance matrix of a random vector⁸

$$\text{C}[\mathbf{X}] = \text{Cov}[\mathbf{X}] = \text{V}[\mathbf{X}] = \mathbb{E}[(\mathbf{X} - \mathbb{E}(\mathbf{X}))(\mathbf{X} - \mathbb{E}(\mathbf{X}))^\top]$$

$$\mathbf{X} = \langle X_1, X_2, \dots, X_p \rangle = [X_1 \ X_2 \ \cdots \ X_p]^\top = \begin{bmatrix} X_1 \\ X_2 \\ \vdots \\ X_p \end{bmatrix}$$

$$\mathbb{E}[\mathbf{X}] = \langle \mathbb{E}[X_1], \mathbb{E}[X_2], \dots, \mathbb{E}[X_p] \rangle = [\mathbb{E}[X_1] \ \mathbb{E}[X_2] \ \cdots \ \mathbb{E}[X_p]]^\top = \begin{bmatrix} \mathbb{E}[X_1] \\ \mathbb{E}[X_2] \\ \vdots \\ \mathbb{E}[X_p] \end{bmatrix}$$

$$\mathbf{X} - \mathbb{E}[\mathbf{X}] = \begin{bmatrix} X_1 - \mathbb{E}[X_1] \\ X_2 - \mathbb{E}[X_2] \\ \vdots \\ X_p - \mathbb{E}[X_p] \end{bmatrix}$$

$$\begin{aligned}
&[\mathbf{X} - \mathbb{E}(\mathbf{X})][\mathbf{X} - \mathbb{E}(\mathbf{X})]^\top \\
&= \begin{bmatrix} X_1 - \mathbb{E}[X_1] \\ X_2 - \mathbb{E}[X_2] \\ \vdots \\ X_p - \mathbb{E}[X_p] \end{bmatrix} [X_1 - \mathbb{E}[X_1] \ X_2 - \mathbb{E}[X_2] \ \cdots \ X_p - \mathbb{E}[X_p]] \\
&= \begin{bmatrix} (X_1 - \mathbb{E}[X_1])(X_1 - \mathbb{E}[X_1]) & (X_1 - \mathbb{E}[X_1])(X_2 - \mathbb{E}[X_2]) & \cdots & (X_1 - \mathbb{E}[X_1])(X_p - \mathbb{E}[X_p]) \\ (X_2 - \mathbb{E}[X_2])(X_1 - \mathbb{E}[X_1]) & (X_2 - \mathbb{E}[X_2])(X_2 - \mathbb{E}[X_2]) & \cdots & (X_2 - \mathbb{E}[X_2])(X_p - \mathbb{E}[X_p]) \\ \vdots & \vdots & \ddots & \vdots \\ (X_p - \mathbb{E}[X_p])(X_1 - \mathbb{E}[X_1]) & (X_p - \mathbb{E}[X_p])(X_2 - \mathbb{E}[X_2]) & \cdots & (X_p - \mathbb{E}[X_p])(X_p - \mathbb{E}[X_p]) \end{bmatrix} \\
&= \begin{bmatrix} (X_1 - \mathbb{E}[X_1])^2 & \cdots & (X_1 - \mathbb{E}[X_1])(X_p - \mathbb{E}[X_p]) \\ \vdots & \ddots & \vdots \\ (X_p - \mathbb{E}[X_p])(X_1 - \mathbb{E}[X_1]) & \cdots & (X_p - \mathbb{E}[X_p])^2 \end{bmatrix}
\end{aligned}$$

$$\mathbb{E}[(\mathbf{X} - \mathbb{E}(\mathbf{X}))(\mathbf{X} - \mathbb{E}(\mathbf{X}))^\top] \quad (56.3)$$

$$= \mathbb{E} \begin{bmatrix} (X_1 - \mathbb{E}[X_1])^2 & \cdots & (X_1 - \mathbb{E}[X_1])(X_p - \mathbb{E}[X_p]) \\ \vdots & \ddots & \vdots \\ (X_p - \mathbb{E}[X_p])(X_1 - \mathbb{E}[X_1]) & \cdots & (X_p - \mathbb{E}[X_p])^2 \end{bmatrix} \quad (56.4)$$

$$= \begin{bmatrix} \mathbb{E}[(X_1 - \mathbb{E}[X_1])^2] & \cdots & \mathbb{E}[(X_1 - \mathbb{E}[X_1])(X_p - \mathbb{E}[X_p])] \\ \vdots & \ddots & \vdots \\ \mathbb{E}[(X_p - \mathbb{E}[X_p])(X_1 - \mathbb{E}[X_1])] & \cdots & \mathbb{E}[(X_p - \mathbb{E}[X_p])^2] \end{bmatrix} \quad (56.5)$$

$$= \begin{bmatrix} V(X_1, X_1) & \cdots & V(X_1, X_p) \\ \vdots & \ddots & \vdots \\ V(X_p, X_1) & \cdots & V(X_p, X_p) \end{bmatrix} = \begin{bmatrix} V(X_1, X_1) & V(X_1, X_2) & \cdots & V(X_1, X_p) \\ V(X_2, X_1) & V(X_2, X_2) & \cdots & V(X_2, X_p) \\ \vdots & \vdots & \ddots & \vdots \\ V(X_p, X_1) & V(X_p, X_2) & \cdots & V(X_p, X_p) \end{bmatrix} \quad (56.6)$$

$$= \begin{bmatrix} V(X_1) & \cdots & V(X_1, X_p) \\ \vdots & \ddots & \vdots \\ V(X_p, X_1) & \cdots & V(X_p) \end{bmatrix} = \begin{bmatrix} V(X_1) & V(X_1, X_2) & \cdots & V(X_1, X_p) \\ V(X_2, X_1) & V(X_2) & \cdots & V(X_2, X_p) \\ \vdots & \vdots & \ddots & \vdots \\ V(X_p, X_1) & V(X_p, X_2) & \cdots & V(X_p) \end{bmatrix} \quad (56.7)$$

$$= \begin{bmatrix} V(X_1) & \cdots & C(X_1, X_p) \\ \vdots & \ddots & \vdots \\ C(X_p, X_1) & \cdots & V(X_p) \end{bmatrix} = \begin{bmatrix} V(X_1) & C(X_1, X_2) & \cdots & C(X_1, X_p) \\ C(X_2, X_1) & V(X_2) & \cdots & C(X_2, X_p) \\ \vdots & \vdots & \ddots & \vdots \\ C(X_p, X_1) & C(X_p, X_2) & \cdots & V(X_p) \end{bmatrix} \quad (56.8)$$

$$= \begin{bmatrix} \sigma_1^2 & \cdots & \sigma_{1p} \\ \vdots & \ddots & \vdots \\ \sigma_{p1} & \cdots & \sigma_p^2 \end{bmatrix} = \begin{bmatrix} \sigma_1^2 & \sigma_{12} & \cdots & \sigma_{1p} \\ \sigma_{21} & \sigma_2^2 & \cdots & \sigma_{2p} \\ \vdots & \vdots & \ddots & \vdots \\ \sigma_{p1} & \sigma_{p2} & \cdots & \sigma_p^2 \end{bmatrix} = \begin{bmatrix} \sigma_{11} & \sigma_{12} & \cdots & \sigma_{1p} \\ \sigma_{21} & \sigma_{22} & \cdots & \sigma_{2p} \\ \vdots & \vdots & \ddots & \vdots \\ \sigma_{p1} & \sigma_{p2} & \cdots & \sigma_{pp} \end{bmatrix} = [\sigma_{ij}]_{p \times p} = \Sigma \quad (56.9)$$

$$\mathbf{X} \sim \mathcal{D}(\boldsymbol{\mu}, \Sigma) = d(\boldsymbol{\mu}_{\mathbf{X}}, \Sigma_{\mathbf{X}}) = d(\mathbb{E}[\mathbf{X}], \mathbb{C}[\mathbf{X}]) = d(\mathbb{E}[\mathbf{X}], \mathbb{V}[\mathbf{X}])$$

$$\mathbf{Z} \sim \mathcal{N}(\boldsymbol{\mu}_{\mathbf{Z}}, \Sigma_{\mathbf{Z}}) = n(\mathbb{E}[\mathbf{Z}], \mathbb{V}[\mathbf{Z}])$$

$$\mathbb{E}[\mathbf{Z}] = \begin{bmatrix} \mathbb{E}[Z_1] \\ \mathbb{E}[Z_2] \\ \vdots \\ \mathbb{E}[Z_p] \end{bmatrix} = [\mathbb{E}[Z_i]]_{p \times 1}$$

$$\Rightarrow \mathbb{E}[Z_i] = \int_{-\infty}^{\infty} z_i f_{Z_i}(z_i) dz_i \stackrel{(56.2)}{=} \int_{-\infty}^{\infty} z_i \frac{e^{-\frac{1}{2}z_i^2}}{\sqrt{2\pi}} dz_i = 0$$

$$\Rightarrow \mathbb{E}[\mathbf{Z}] = \mathbf{0}$$

$$\Rightarrow \mathbf{Z} \sim \mathcal{N}(\boldsymbol{\mu}_{\mathbf{Z}} = \mathbf{0}, \Sigma_{\mathbf{Z}}) = n(\mathbf{0}, \mathbb{V}[\mathbf{Z}])$$

$$\mathbb{V}(Z_i) = \int_{-\infty}^{\infty} (z_i - \mu_{Z_i})^2 f_{Z_i}(z_i) dz_i \stackrel{(56.2)}{=} \int_{-\infty}^{\infty} (z_i - 0)^2 \frac{e^{-\frac{1}{2}z_i^2}}{\sqrt{2\pi}} dz_i = 1$$

$$\mathbb{V}(Z_i, Z_j) \stackrel{i \neq j \Rightarrow Z_i, Z_j \text{ are independent}}{=} 0 \quad (56.10)$$

$$\begin{aligned} V[\mathbf{Z}] &= \begin{bmatrix} V(Z_1) & V(Z_1, Z_2) & \cdots & V(Z_1, Z_p) \\ V(Z_2, Z_1) & V(Z_2) & \cdots & V(Z_2, Z_p) \\ \vdots & \vdots & \ddots & \vdots \\ V(Z_p, Z_1) & V(Z_p, Z_2) & \cdots & V(Z_p) \end{bmatrix} = \begin{bmatrix} \sigma_{11} & \sigma_{12} & \cdots & \sigma_{1p} \\ \sigma_{21} & \sigma_{22} & \cdots & \sigma_{2p} \\ \vdots & \vdots & \ddots & \vdots \\ \sigma_{p1} & \sigma_{p2} & \cdots & \sigma_{pp} \end{bmatrix} \\ &= \begin{bmatrix} \sigma_1^2 & \sigma_{12} & \cdots & \sigma_{1p} \\ \sigma_{21} & \sigma_2^2 & \cdots & \sigma_{2p} \\ \vdots & \vdots & \ddots & \vdots \\ \sigma_{p1} & \sigma_{p2} & \cdots & \sigma_p^2 \end{bmatrix} \stackrel{56.10}{=} \begin{bmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 \end{bmatrix} = I_{p \times p} = I_p = I \end{aligned}$$

$$\mathbf{Z} \sim \mathcal{N}(\boldsymbol{\mu}_\mathbf{Z}, \Sigma_\mathbf{Z}) = n(E[\mathbf{Z}], V[\mathbf{Z}]) = \mathcal{N}(\mathbf{0}, I) \Leftrightarrow \begin{cases} \boldsymbol{\mu}_\mathbf{Z} = E[\mathbf{Z}] = \mathbf{0} = [0]_p = [0]_{p \times 1} \\ \Sigma_\mathbf{Z} = V[\mathbf{Z}] = I = I_p = I_{p \times p} \end{cases}$$

$$\begin{aligned} \mathbf{Z} &= \begin{bmatrix} Z_1 \\ Z_2 \\ \vdots \\ Z_p \end{bmatrix} = \begin{bmatrix} \frac{X_1 - \mu_1}{\sigma_1} \\ \frac{X_2 - \mu_2}{\sigma_2} \\ \vdots \\ \frac{X_p - \mu_p}{\sigma_p} \end{bmatrix} = \begin{bmatrix} \frac{1}{\sigma_1} & 0 & \cdots & 0 \\ 0 & \frac{1}{\sigma_2} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \frac{1}{\sigma_p} \end{bmatrix} \begin{bmatrix} X_1 - \mu_1 \\ X_2 - \mu_2 \\ \vdots \\ X_p - \mu_p \end{bmatrix} \\ &= \begin{bmatrix} \frac{1}{\sigma_1} & 0 & \cdots & 0 \\ 0 & \frac{1}{\sigma_2} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \frac{1}{\sigma_p} \end{bmatrix} \left(\begin{bmatrix} X_1 \\ X_2 \\ \vdots \\ X_p \end{bmatrix} - \begin{bmatrix} \mu_1 \\ \mu_2 \\ \vdots \\ \mu_p \end{bmatrix} \right) = B^{-1}(\mathbf{X} - \boldsymbol{\mu}) \end{aligned}$$

$$\Rightarrow \mathbf{X} = B\mathbf{Z} + \boldsymbol{\mu}$$

$$\mathbf{X} = B\mathbf{Z} + \boldsymbol{\mu} = T(\mathbf{Z})$$

by $V[A\mathbf{X} + \mathbf{b}] = A V[\mathbf{X}] A^\top$

$$\Sigma = \Sigma_\mathbf{X} = V[\mathbf{X}] = V[B\mathbf{Z} + \boldsymbol{\mu}] = BV[\mathbf{Z}]B^\top = BIB^\top = BB^\top \quad (56.11)$$

Consider two infinitesimal volumes of p -dimensional parallelepipeds in the different \mathbb{R}^p spaces¹⁶

$$V_x = [x_1, x_1 + dx_1] \times [x_2, x_2 + dx_2] \times \cdots \times [x_p, x_p + dx_p]$$

and

$$V_z = [z_1, z_1 + dz_1] \times [z_2, z_2 + dz_2] \times \cdots \times [z_p, z_p + dz_p]$$

Their relationship under linear transformation is

$$\begin{aligned} V_x &= T(V_z) = [T(z_1), T(z_1) + T(dz_1)] \\ &\quad \times [T(z_2), T(z_2) + T(dz_2)] \\ &\quad \times \cdots \\ &\quad \times [T(z_p), T(z_p) + T(dz_p)] \end{aligned}$$

and

$$dx_i = \sum_j \frac{\partial x_i}{\partial z_j} dz_j$$

For examples in 2 dimension,

$$\begin{bmatrix} dx_1 \\ dx_2 \end{bmatrix} = \begin{bmatrix} \frac{\partial x_1}{\partial z_1} & \frac{\partial x_1}{\partial z_2} \\ \frac{\partial x_2}{\partial z_1} & \frac{\partial x_2}{\partial z_2} \end{bmatrix} \begin{bmatrix} dz_1 \\ dz_2 \end{bmatrix}$$

Two element infinitesimal one-directional vectors of \mathbf{Z} transformed into another space of \mathbf{X} are

$$T(dz_1) = \begin{bmatrix} \frac{\partial x_1}{\partial z_1} & \frac{\partial x_1}{\partial z_2} \\ \frac{\partial x_2}{\partial z_1} & \frac{\partial x_2}{\partial z_2} \end{bmatrix} \begin{bmatrix} dz_1 \\ 0 \end{bmatrix} = \begin{bmatrix} \frac{\partial x_1}{\partial z_1} dz_1 \\ \frac{\partial x_2}{\partial z_1} dz_1 \end{bmatrix}$$

and

$$T(dz_2) = \begin{bmatrix} \frac{\partial x_1}{\partial z_1} & \frac{\partial x_1}{\partial z_2} \\ \frac{\partial x_2}{\partial z_2} & \frac{\partial x_2}{\partial z_2} \end{bmatrix} \begin{bmatrix} 0 \\ dz_2 \end{bmatrix} = \begin{bmatrix} \frac{\partial x_1}{\partial z_2} dz_2 \\ \frac{\partial x_2}{\partial z_2} dz_2 \end{bmatrix}$$

Their corresponding area(volume) in the space of \mathbf{X} is

$$\begin{aligned} \int_{A_x} dA_x &= \int_{A_x} dx_1 dx_2 = \int_{T(A_z)} dx_1 dx_2 \\ &= \int_{A_z} |[T(dz_1) \ T(dz_2)]| = \int_{A_z} \left| \begin{bmatrix} \frac{\partial x_1}{\partial z_1} dz_1 & \frac{\partial x_1}{\partial z_2} dz_2 \\ \frac{\partial x_2}{\partial z_1} dz_1 & \frac{\partial x_2}{\partial z_2} dz_2 \end{bmatrix} \right| \\ &= \int_{A_z} \left| \begin{bmatrix} \frac{\partial x_1}{\partial z_1} & \frac{\partial x_1}{\partial z_2} \\ \frac{\partial x_2}{\partial z_1} & \frac{\partial x_2}{\partial z_2} \end{bmatrix} \right| dz_1 dz_2 = \int_{A_z} |J| dA_z \end{aligned}$$

To generalize for volumes in p dimension,

$$\begin{aligned} \int_{V_x} dV_x &= \int_{V_x} dx_1 dx_2 \cdots dx_p = \int_{T(V_z)} dx_1 dx_2 \cdots dx_p \\ &= \int_{A_z} |[T(dz_1) \ T(dz_2) \ \cdots \ T(dz_p)]| = \int_{V_z} \left| \left[\frac{\partial x_i}{\partial z_j} dz_j \right]_{p \times p} \right| \\ &= \int_{V_z} \left| \left[\frac{\partial x_i}{\partial z_j} \right]_{p \times p} \right| dz_1 dz_2 \cdots dz_p = \int_{V_z} |J| dV_z \end{aligned}$$

i.e.

$$\int_{V_x} dV_x = \int_{V_z} |J| dV_z \tag{56.12}$$

where J is a Jacobian matrix

$$J = \left[\frac{\partial x_i}{\partial z_j} \right]_{p \times p} = \frac{\partial \mathbf{x}}{\partial \mathbf{z}}$$

or $|J|$ is a Jacobian determinant or simply Jacobian

$$|J| = \left| \frac{\partial x_i}{\partial z_j} \right|_{p \times p} = \left| \frac{\partial \mathbf{x}}{\partial \mathbf{z}} \right|$$

The probability of the same event should be invariant under transformation.

$$\begin{aligned}
\int_{V_x} f_{\mathbf{X}}(\mathbf{x}) dV_{\mathbf{x}} &= \int_{V_x} f_{\mathbf{X}}(\mathbf{x}) dx_1 dx_2 \cdots dx_p \\
&= \int_{T(V_z)} f_{\mathbf{X}}(\mathbf{x}) dx_1 dx_2 \cdots dx_p \\
&= \int_{V_z} f_{\mathbf{Z}}(\mathbf{z}) dV_{\mathbf{z}} = \int_{V_z} f_{\mathbf{Z}}(\mathbf{z}) dz_1 dz_2 \cdots dz_p
\end{aligned}$$

i.e.

$$\int_{V_x} f_{\mathbf{X}}(\mathbf{x}) dV_{\mathbf{x}} = \int_{V_z} f_{\mathbf{Z}}(\mathbf{z}) dV_{\mathbf{z}} \quad (56.13)$$

$$\begin{cases} \int_{V_x} f_{\mathbf{X}}(\mathbf{x}) dV_{\mathbf{x}} = \int_{V_z} f_{\mathbf{Z}}(\mathbf{z}) dV_{\mathbf{z}} & 56.13 \\ \int_{V_x} dV_{\mathbf{x}} = \int_{V_z} |J| dV_{\mathbf{z}} & 56.12 \end{cases}$$

$$\begin{aligned}
\mathbf{Z} &= B^{-1}(\mathbf{X} - \boldsymbol{\mu}) \\
\mathbf{z} &= B^{-1}(\mathbf{x} - \boldsymbol{\mu}) \\
\mathbf{X} &= B\mathbf{Z} + \boldsymbol{\mu} \\
\mathbf{x} &= B\mathbf{z} + \boldsymbol{\mu}
\end{aligned} \quad (56.14)$$

$$\begin{aligned}
J &= \left[\frac{\partial x_i}{\partial z_j} \right]_{p \times p} = \frac{\partial \mathbf{x}}{\partial \mathbf{z}} = B \\
|J| &= \left| \frac{\partial x_i}{\partial z_j} \right|_{p \times p} = \left| \frac{\partial \mathbf{x}}{\partial \mathbf{z}} \right| = |B|
\end{aligned} \quad (56.15)$$

$$\begin{aligned}
\int_{V_z} f_{\mathbf{Z}}(\mathbf{z}) dV_{\mathbf{z}} &\stackrel{56.13}{=} \int_{V_x} f_{\mathbf{X}}(\mathbf{x}) dV_{\mathbf{x}} = \int_{V_x} dV_{\mathbf{x}} f_{\mathbf{X}}(\mathbf{x}) \\
&\stackrel{56.12}{=} \int_{V_z} |J| dV_{\mathbf{z}} f_{\mathbf{X}}(\mathbf{x}(z)) = \int_{V_z} f_{\mathbf{X}}(\mathbf{x}(z)) |J| dV_{\mathbf{z}} \\
f_{\mathbf{Z}}(\mathbf{z}) &\stackrel{\ddagger}{=} f_{\mathbf{X}}(\mathbf{x}(z)) |J| \\
f_{\mathbf{X}}(\mathbf{x}(z)) &\stackrel{\ddagger}{=} |J|^{-1} f_{\mathbf{Z}}(\mathbf{z}) \stackrel{56.1}{=} |J|^{-1} \frac{1}{(2\pi)^{p/2}} \exp \left\{ \frac{-\mathbf{z}^T \mathbf{z}}{2} \right\} \\
f_{\mathbf{X}}(\mathbf{x}) &\stackrel{\ddagger}{=} |J|^{-1} f_{\mathbf{Z}}(\mathbf{z}(x)) \stackrel{56.15, 56.14}{=} |B|^{-1} f_{\mathbf{Z}}(B^{-1}(\mathbf{x} - \boldsymbol{\mu})) \\
&= |B|^{-1} (2\pi)^{-p/2} \exp \left\{ \frac{-1}{2} [B^{-1}(\mathbf{x} - \boldsymbol{\mu})]^T [B^{-1}(\mathbf{x} - \boldsymbol{\mu})] \right\} \\
&= |B|^{-1/2} |B|^{-1/2} (2\pi)^{-p/2} \exp \left\{ \frac{-1}{2} (\mathbf{x} - \boldsymbol{\mu})^T (B^{-1})^T B^{-1} (\mathbf{x} - \boldsymbol{\mu}) \right\} \\
&= |B|^{-1/2} |B^T|^{-1/2} (2\pi)^{-p/2} \exp \left\{ \frac{-1}{2} (\mathbf{x} - \boldsymbol{\mu})^T (B^T)^{-1} B^{-1} (\mathbf{x} - \boldsymbol{\mu}) \right\} \\
&= |BB^T|^{-1/2} (2\pi)^{-p/2} \exp \left\{ \frac{-1}{2} (\mathbf{x} - \boldsymbol{\mu})^T (BB^T)^{-1} (\mathbf{x} - \boldsymbol{\mu}) \right\} \\
&\stackrel{56.11}{=} |\Sigma|^{-1/2} (2\pi)^{-p/2} \exp \left\{ \frac{-1}{2} (\mathbf{x} - \boldsymbol{\mu})^T \Sigma^{-1} (\mathbf{x} - \boldsymbol{\mu}) \right\} \\
&= (|\Sigma| (2\pi)^p)^{-1/2} \exp \left\{ \frac{-1}{2} (\mathbf{x} - \boldsymbol{\mu})^T \Sigma^{-1} (\mathbf{x} - \boldsymbol{\mu}) \right\}
\end{aligned}$$

Definition 56.3. probability density function (PDF) of multivariate normal distribution (= multivariate Gaussian distribution)

$$\mathcal{N}(\mathbf{x} | \boldsymbol{\mu}, \Sigma) = f_{\mathbf{X}}(\mathbf{x}) = (|\Sigma| (2\pi)^p)^{-1/2} \exp \left\{ \frac{-1}{2} (\mathbf{x} - \boldsymbol{\mu})^T \Sigma^{-1} (\mathbf{x} - \boldsymbol{\mu}) \right\}$$

Definition 56.4. correlation coefficient

$$\rho_{ij} = \frac{\sigma_{ij}}{\sqrt{\sigma_{ii}}\sqrt{\sigma_{jj}}} = \frac{\sigma_{ij}}{\sqrt{\sigma_i^2}\sqrt{\sigma_j^2}} = \frac{\sigma_{ij}}{\sigma_i\sigma_j} = \frac{V(X_i, X_j)}{\sqrt{V(X_i)}\sqrt{V(X_j)}} = R(X_i, X_j)$$

56.1 bivariate normal distribution

$p = 2$ is the case of bivariate normal distribution

$$\Sigma = [\sigma_{ij}]_{2 \times 2} = \begin{bmatrix} \sigma_{11} & \sigma_{12} \\ \sigma_{21} & \sigma_{22} \end{bmatrix} = \begin{bmatrix} \sigma_1^2 & \sigma_{12} \\ \sigma_{21} & \sigma_2^2 \end{bmatrix} = \begin{bmatrix} \sigma_1^2 & \sigma_1\sigma_2\rho_{12} \\ \sigma_2\sigma_1\rho_{21} & \sigma_2^2 \end{bmatrix} = \begin{bmatrix} \sigma_1^2 & \sigma_1\sigma_2\rho \\ \sigma_2\sigma_1\rho & \sigma_2^2 \end{bmatrix}$$

$$\rho_{12} = \rho = \rho_{21}$$

$$|\Sigma| = \begin{vmatrix} \sigma_1^2 & \sigma_1\sigma_2\rho_{12} \\ \sigma_2\sigma_1\rho_{21} & \sigma_2^2 \end{vmatrix} = \sigma_1^2\sigma_2^2(1 - \rho_{12}\rho_{21}) = \sigma_1^2\sigma_2^2(1 - \rho^2)$$

$$\begin{bmatrix} a & b \\ c & d \end{bmatrix}^{-1} = \frac{1}{\begin{vmatrix} a & b \\ c & d \end{vmatrix}} \begin{bmatrix} d & -b \\ -c & a \end{bmatrix}$$

$$\begin{aligned} \Sigma^{-1} &= \frac{1}{|\Sigma|} \begin{bmatrix} \sigma_2^2 & -\sigma_1\sigma_2\rho \\ -\sigma_2\sigma_1\rho & \sigma_1^2 \end{bmatrix} \\ &= \frac{1}{\sigma_1^2\sigma_2^2(1 - \rho^2)} \begin{bmatrix} \sigma_2^2 & -\sigma_1\sigma_2\rho \\ -\sigma_2\sigma_1\rho & \sigma_1^2 \end{bmatrix} \\ &= \frac{1}{(1 - \rho^2)} \begin{bmatrix} \frac{1}{\sigma_1^2} & \frac{-\rho}{\sigma_1\sigma_2} \\ \frac{-\rho}{\sigma_2\sigma_1} & \frac{1}{\sigma_2^2} \end{bmatrix} \end{aligned}$$

$$\begin{aligned} &\mathcal{N}\left(\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \middle| \boldsymbol{\mu} = \begin{bmatrix} \mu_1 \\ \mu_2 \end{bmatrix}, \Sigma = \begin{bmatrix} \sigma_1^2 & \sigma_1\sigma_2\rho \\ \sigma_2\sigma_1\rho & \sigma_2^2 \end{bmatrix}\right) \\ &= \left(|\Sigma|(2\pi)^{p=2}\right)^{-1/2} \exp\left\{\frac{-1}{2}(\mathbf{x} - \boldsymbol{\mu})^\top \Sigma^{-1} (\mathbf{x} - \boldsymbol{\mu})\right\} \\ &= \left(\sigma_1^2\sigma_2^2(1 - \rho^2)(2\pi)^2\right)^{-1/2} \exp\left\{\frac{-1}{2} \begin{bmatrix} x_1 - \mu_1 \\ x_2 - \mu_2 \end{bmatrix}^\top \Sigma^{-1} \begin{bmatrix} x_1 - \mu_1 \\ x_2 - \mu_2 \end{bmatrix}\right\} \\ &= \frac{1}{2\pi\sigma_1\sigma_2\sqrt{1 - \rho^2}} \exp\left\{\frac{-1}{2} \begin{bmatrix} x_1 - \mu_1 \\ x_2 - \mu_2 \end{bmatrix}^\top \frac{1}{(1 - \rho^2)} \begin{bmatrix} \frac{1}{\sigma_1^2} & \frac{-\rho}{\sigma_1\sigma_2} \\ \frac{-\rho}{\sigma_2\sigma_1} & \frac{1}{\sigma_2^2} \end{bmatrix} \begin{bmatrix} x_1 - \mu_1 \\ x_2 - \mu_2 \end{bmatrix}\right\} \\ &= \frac{1}{2\pi\sigma_1\sigma_2\sqrt{1 - \rho^2}} \exp\left\{\frac{-1}{2(1 - \rho^2)} \begin{bmatrix} x_1 - \mu_1 \\ x_2 - \mu_2 \end{bmatrix}^\top \begin{bmatrix} \frac{1}{\sigma_1^2} & \frac{-\rho}{\sigma_1\sigma_2} \\ \frac{-\rho}{\sigma_2\sigma_1} & \frac{1}{\sigma_2^2} \end{bmatrix} \begin{bmatrix} x_1 - \mu_1 \\ x_2 - \mu_2 \end{bmatrix}\right\} \\ &= \frac{1}{2\pi\sigma_1\sigma_2\sqrt{1 - \rho^2}} \exp\left\{\frac{-1}{2(1 - \rho^2)} \left[\left(\frac{x_1 - \mu_1}{\sigma_1}\right)^2 - 2\rho \left(\frac{x_1 - \mu_1}{\sigma_1}\right) \left(\frac{x_2 - \mu_2}{\sigma_2}\right) + \left(\frac{x_2 - \mu_2}{\sigma_2}\right)^2 \right]\right\} \end{aligned}$$

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=PLplnkTzzqsZS3R5DjmCQsqpu43oS9CFN>

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]

$$\dots, Y_{t-2}, Y_{t-1}, Y_t, Y_{t+1}, Y_{t+2}, \dots$$

- lag
– 1st lag = lag 1

$$Y_{t-1}$$

- k^{th} lag = lag k

$$Y_{t-k}$$

1st difference

$$\Delta Y_t = Y_t - Y_{t-1}$$

approximation for RoR = rate of return

$$\begin{aligned}\Delta \ln Y_t &= \ln Y_t - \ln Y_{t-1} = \ln \frac{Y_t}{Y_{t-1}} = \ln \left(1 + \frac{Y_t - Y_{t-1}}{Y_{t-1}} \right) \\ &= \frac{Y_t - Y_{t-1}}{Y_{t-1}} + O\left(\left(\frac{Y_t - Y_{t-1}}{Y_{t-1}}\right)^2\right) \\ &\approx \frac{Y_t - Y_{t-1}}{Y_{t-1}} = \text{RoR}\end{aligned}$$

$$\begin{cases} \ln(1+x) = \sum_{n=1}^{\infty} (-1)^{n+1} \frac{x^n}{n} = x - \frac{x^2}{2} + \frac{x^3}{3} - \dots \\ \ln(1-x) = - \sum_{n=1}^{\infty} \frac{x^n}{n} = -x - \frac{x^2}{2} - \frac{x^3}{3} - \dots \end{cases}$$

$$\begin{aligned}\frac{1}{1+t} &= 1 - t + t^2 - t^3 + \dots \\ \ln(1+x) &= \int_0^x \frac{1}{1+t} dt \\ &= x - \frac{x^2}{2} + \frac{x^3}{3} - \frac{x^4}{4} + \dots\end{aligned}$$

Definition 60.1. autocorrelation = serial correlation

$$\exists t_1 \neq t_2 [V(Y_{t_1}, Y_{t_2}) \neq 0]$$

$$\begin{cases} \mathbb{E}(Y_t) \triangleq \mu_t \\ \text{V}(Y_t) \triangleq \sigma_t^2 \end{cases}$$

Definition 60.2. k^{th} order autocovariance

$$\text{V}(Y_t, Y_{t-k}) \triangleq \gamma(t, k)$$

$$\sigma_t^2 \triangleq \text{V}(Y_t) = \text{V}(Y_t, Y_t) = \text{V}(Y_t, Y_{t-0}) = \gamma(t, k=0) = \gamma(t, 0)$$

Definition 60.3. k^{th} order autocorrelation

$$\rho_{t,t-k} = \frac{\sigma_{t,t-k}}{\sqrt{\sigma_{tt}}\sqrt{\sigma_{t-k,t-k}}} = \frac{\sigma_{t,t-k}}{\sqrt{\sigma_t^2}\sqrt{\sigma_{t-k}^2}} = \frac{\sigma_{t,t-k}}{\sigma_t\sigma_{t-k}} = \frac{\text{V}(Y_t, Y_{t-k})}{\sqrt{\text{V}(Y_t)}\sqrt{\text{V}(Y_{t-k})}} \triangleq \text{R}(Y_t, Y_{t-k}) \triangleq \rho(t, k)$$

Definition 60.4. stationary time series

$$\begin{cases} \mathbb{E}(Y_t) \triangleq \mu_t = \mu & (1) \text{ independent of } t \\ \text{V}(Y_t) \triangleq \sigma_t^2 = \sigma^2 < \infty & (2) \text{ independent of } t \\ \text{V}(Y_t, Y_{t-k}) \triangleq \gamma(t, k) = \gamma(k) & (3) \text{ independent of } t \end{cases}$$

properties

$$\begin{cases} \text{V}(Y_t) \triangleq \sigma_t^2 = \gamma(t, k=0) = \gamma(k=0) \stackrel{(3)}{=} \gamma(0) \stackrel{(2)}{=} \gamma_0 \stackrel{(2)}{=} \sigma^2 & (4) \Rightarrow \text{V}(Y_{t-k}) = \gamma(0) \\ \text{R}(Y_t, Y_{t-k}) \triangleq \rho(t, k) \triangleq \frac{\text{V}(Y_t, Y_{t-k})}{\sqrt{\text{V}(Y_t)}\sqrt{\text{V}(Y_{t-k})}} \\ = \frac{\gamma(t, k)}{\sqrt{\gamma(0)}\sqrt{\gamma(0)}} \stackrel{(3)}{=} \frac{\gamma(k)}{\gamma(0)} \stackrel{(4)}{=} \frac{\gamma(k)}{\gamma_0} \stackrel{(2)}{=} \frac{\gamma_k}{\gamma_0} \triangleq \rho(k) \triangleq \rho_k & (5) \\ \gamma(k) = \gamma(-k) & (6) \Rightarrow \rho(k) = \rho(-k) \end{cases}$$

$$\gamma(k) \stackrel{(3)}{=} \text{V}(Y_t, Y_{t-k}) = \text{V}(Y_{t-k}, Y_t) = \text{V}(Y_{t'}, Y_{t'+k}) = \text{V}(Y_{t'}, Y_{t'-(k)}) \stackrel{(3)}{=} \gamma(-k) \Rightarrow (6)$$

point estimation

$$\begin{cases} \widehat{\mathbb{E}}(Y_t) \triangleq \bar{Y} \triangleq \widehat{\mu} = \frac{1}{T} \sum_{t=1}^T Y_t & \rightarrow \mathbb{E}(Y_t) = \mu \\ \widehat{\text{V}}(Y_t) \triangleq \widehat{\gamma}_0 = \frac{1}{T} \sum_{t=1}^T (Y_t - \bar{Y})^2 & \rightarrow \text{V}(Y_t) = \sigma^2 = \gamma_0 \\ \widehat{\text{V}}(Y_t, Y_{t-k}) \triangleq \widehat{\gamma}_k \\ = \frac{1}{T} \sum_{t=k+1}^T (Y_t - \bar{Y})(Y_{t-k} - \bar{Y}) = \frac{1}{T} \sum_{t=1}^{T-k} (Y_t - \bar{Y})(Y_{t+k} - \bar{Y}) & \rightarrow \text{V}(Y_t, Y_{t-k}) = \gamma_k \\ \frac{\widehat{\text{V}}(Y_t, Y_{t-k})}{\sqrt{\widehat{\text{V}}(Y_t)}\sqrt{\widehat{\text{V}}(Y_{t-k})}} \triangleq \widehat{\rho}_k = \frac{\widehat{\gamma}_k}{\widehat{\gamma}_0} = \frac{\frac{1}{T} \sum_{t=k+1}^T (Y_t - \bar{Y})(Y_{t-k} - \bar{Y})}{\frac{1}{T} \sum_{t=1}^T (Y_t - \bar{Y})^2} & \rightarrow \text{R}(Y_t, Y_{t-k}) = \rho_k \end{cases}$$

$$Y_1 Y_{1+k} + Y_2 Y_{2+k} + \dots + Y_t Y_{t+k} + \dots + Y_{T-k} Y_T$$

1st-order autocorrelation estimation

$$\widehat{\rho}_1 = \frac{\widehat{\gamma}_1}{\widehat{\gamma}_0} = \frac{\frac{1}{T} \sum_{t=1+1=2}^T (Y_t - \bar{Y})(Y_{t-1} - \bar{Y})}{\frac{1}{T} \sum_{t=1}^T (Y_t - \bar{Y})^2}$$

60.1 AR(1) = 1st-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) = 1st-order autoregressive model\$

$$Y_t = \beta_0 + \beta_1 Y_{t-1} + \varepsilon_t$$

$$\varepsilon_t \sim \mathcal{N}(0, \sigma_\varepsilon^2)$$

$$\begin{aligned} Y_t &= \beta_0 + \beta_1 Y_{t-1} + \varepsilon_t \\ &= \beta_0 + \beta_1 (\beta_0 + \beta_1 Y_{t-2} + \varepsilon_{t-1}) + \varepsilon_t \\ &= \beta_0 (1 + \beta_1) + \beta_1^2 Y_{t-2} + \varepsilon_t + \beta_1 \varepsilon_{t-1} \\ &= \beta_0 (1 + \beta_1) + \beta_1^2 (\beta_0 + \beta_1 Y_{t-3} + \varepsilon_{t-2}) + \varepsilon_t + \beta_1 \varepsilon_{t-1} \\ &= \beta_0 (1 + \beta_1 + \beta_1^2) + \beta_1^3 Y_{t-3} + \varepsilon_t + \beta_1 \varepsilon_{t-1} + \beta_1^2 \varepsilon_{t-2} \\ &\quad \vdots \\ &= \beta_0 (1 + \beta_1 + \beta_1^2 + \cdots) + \varepsilon_t + \beta_1 \varepsilon_{t-1} + \beta_1^2 \varepsilon_{t-2} + \cdots \\ &\stackrel{|\beta_1| < 1}{=} \frac{\beta_0}{1 - \beta_1} + \sum_{k=0}^{\infty} \beta_1^k \varepsilon_{t-k} \end{aligned}$$

$$\left\{ \begin{array}{l} \mu = E(Y_t) = E\left(\frac{\beta_0}{1 - \beta_1} + \sum_{i=0}^{\infty} \beta_1^i \varepsilon_{t-i}\right) \\ = \frac{\beta_0}{1 - \beta_1} + \sum_{i=0}^{\infty} \beta_1^i E(\varepsilon_{t-i}) \stackrel{\varepsilon_t \sim \mathcal{N}(0, \sigma_\varepsilon^2)}{=} \frac{\beta_0}{1 - \beta_1} \quad \mu = \frac{\beta_0}{1 - \beta_1} \\ \gamma_0 = \sigma^2 = V(Y_t) = V\left(\mu + \sum_{i=0}^{\infty} \beta_1^i \varepsilon_{t-i}\right) = V\left(\sum_{i=0}^{\infty} \beta_1^i \varepsilon_{t-i}\right) \\ = \sum_{i=0}^{\infty} (\beta_1^i)^2 V(\varepsilon_{t-i}) \stackrel{\varepsilon_t \sim \mathcal{N}(0, \sigma_\varepsilon^2)}{=} \sum_{i=0}^{\infty} (\beta_1^2)^i \sigma_\varepsilon^2 = \frac{\sigma_\varepsilon^2}{1 - \beta_1^2} \quad \gamma_0 = \sigma^2 = \frac{\sigma_\varepsilon^2}{1 - \beta_1^2} \\ \gamma_k = V(Y_t, Y_{t-k}) = V\left(\mu + \sum_{i=0}^{\infty} \beta_1^i \varepsilon_{t-i}, \mu + \sum_{i=0}^{\infty} \beta_1^i \varepsilon_{t-k-i}\right) \\ = V\left(\sum_{i=0}^{\infty} \beta_1^i \varepsilon_{t-i}, \sum_{i=0}^{\infty} \beta_1^i \varepsilon_{t-k-i}\right) = \cdots \quad \gamma_k = \frac{\beta_1^k \sigma_\varepsilon^2}{1 - \beta_1^2} \\ \rho_k = R(Y_t, Y_{t-k}) = \frac{\gamma_k}{\gamma_0} = \frac{1 - \beta_1^2}{\sigma_\varepsilon^2} = \beta_1^k \quad \rho_k = \frac{\gamma_k}{\gamma_0} = \beta_1^k \end{array} \right.$$

$$\rho_1 = \frac{\gamma_1}{\gamma_0} = \beta_1^1 = \beta_1$$

$$\begin{aligned} \gamma_k &= V(Y_t, Y_{t-k}) = V\left(\mu + \sum_{i=0}^{\infty} \beta_1^i \varepsilon_{t-i}, \mu + \sum_{i=0}^{\infty} \beta_1^i \varepsilon_{t-k-i}\right) \\ &= V\left(\sum_{i=0}^{\infty} \beta_1^i \varepsilon_{t-i}, \sum_{i=0}^{\infty} \beta_1^i \varepsilon_{t-k-i}\right) = \cdots \end{aligned}$$

$$\mathbf{Y} = \langle Y_1, Y_2, \dots, Y_T \rangle = \begin{bmatrix} Y_1 & Y_2 & \dots & Y_T \end{bmatrix}^T = \begin{bmatrix} Y_1 \\ Y_2 \\ \vdots \\ Y_T \end{bmatrix}$$

AR(1) covariance matrix

by (56.3)

$$\begin{aligned}
V(\mathbf{Y}) &= E \left[[\mathbf{Y} - E(\mathbf{Y})] [\mathbf{Y} - E(\mathbf{Y})]^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} = \Sigma \\
&= \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_{\varepsilon}^2}{1-\beta_1^2} & \frac{\beta_1^1 \sigma_{\varepsilon}^2}{1-\beta_1^2} & \cdots & \frac{\beta_1^{T-1} \sigma_{\varepsilon}^2}{1-\beta_1^2} \\ \frac{\beta_1^1 \sigma_{\varepsilon}^2}{1-\beta_1^2} & \frac{\sigma_{\varepsilon}^2}{1-\beta_1^2} & \cdots & \frac{\beta_1^{T-2} \sigma_{\varepsilon}^2}{1-\beta_1^2} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\beta_1^{T-1} \sigma_{\varepsilon}^2}{1-\beta_1^2} & \frac{\beta_1^{T-2} \sigma_{\varepsilon}^2}{1-\beta_1^2} & \cdots & \frac{\sigma_{\varepsilon}^2}{1-\beta_1^2} \end{bmatrix} \\
&= \frac{\sigma_{\varepsilon}^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}$$

$$\mathbf{Y} \sim \mathcal{D}(\boldsymbol{\mu}, \Sigma) = d(\boldsymbol{\mu}_{\mathbf{Y}}, \Sigma_{\mathbf{Y}}) = d(E[\mathbf{Y}], V[\mathbf{Y}])$$

where

$$\Sigma = \begin{bmatrix} \gamma_0 & \gamma_1 & \cdots & \gamma_{T-1} \\ \gamma_1 & \gamma_0 & \cdots & \gamma_{T-2} \\ \vdots & \vdots & \ddots & \vdots \\ \gamma_{T-1} & \gamma_{T-2} & \cdots & \gamma_0 \end{bmatrix} = \sigma^2 \begin{bmatrix} 1 & \rho & \rho^2 & \cdots & \rho^{T-1} \\ \rho & 1 & \rho & \cdots & \rho^{T-2} \\ \rho^2 & \rho & 1 & \cdots & \rho^{T-3} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \rho^{T-1} & \rho^{T-2} & \rho^{T-3} & \cdots & 1 \end{bmatrix}$$

60.2 AR(2) = 2nd-order autoregressive model = second-order autoregressive model

Definition 60.6. AR(2) = 2nd-order autoregressive model

$$Y_t = \beta_0 + \beta_1 Y_{t-1} + \beta_2 Y_{t-2} + \varepsilon_t$$

$$\varepsilon_t \sim \mathcal{N}(0, \sigma_{\varepsilon}^2)$$

60.3 AR(p) = p^{th} -order autoregressive model

$$Y_t = \beta_0 + \sum_{i=1}^p \beta_i Y_{t-i} + \varepsilon_t$$

60.4 MA(q) = q^{th} -order moving-average model

$$Y_t = \mu + \varepsilon_t + \sum_{i=1}^q \alpha_i \varepsilon_{t-i}$$

60.5 ARMA(p, q) = $p^{\text{th}}, q^{\text{th}}$ -order autoregressive-moving-average model

$$Y_t = \varepsilon_t + \sum_{i=1}^q \alpha_i \varepsilon_{t-i} + \sum_{i=1}^p \beta_i Y_{t-i}$$

Chapter 61

conductivity

61.1 resistance

https://www.bilibili.com/video/BV1tN41147SK/?spm_id_from=333.999.section.playall&vd_source=2b462dd516a0bbedcb7d97d0fc60b38

61.2 semiconductor

<https://www.youtube.com/watch?v=-lHXZk5M6cI>

Chapter 62

LaTeX annotation by TikZ

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

$$\begin{aligned}f(x) &= x^x \\ \ln f(x) &= x \ln x \\ \frac{d}{dx} \ln f(x) &= \frac{d}{dx} [x \ln x] \\ \frac{f'(x)}{f(x)} &= [x \ln x]' \\ f'(x) &= f(x) [x \ln x]' = x^x [(x)' \ln x + x (\ln x)'] \\ &= x^x \left[1 \ln x + x \frac{1}{x} \right] = x^x [\ln x + 1] = x^x [1 + \ln x] \\ &= x^x + x^x \ln x\end{aligned}$$

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= \emptyset,
    xticklabels= \emptyset,
    %extra x ticks={0},
    ytick = \emptyset,
    yticklabels= \emptyset,
    x = 5cm, y = 5cm, % x y scaling
    %grid = major,
    domain = -1:1,
    %samples = 1000
]
\node at (axis cs: 0,0) {
    $\begin{aligned}
    f(x) &= x^x \\
    \ln f(x) &= x \ln x \\
    \frac{d}{dx} \ln f(x) &= \frac{d}{dx} [x \ln x] \\
    \frac{f'(x)}{f(x)} &= [x \ln x]' \\
    f'(x) &= f(x) [x \ln x]' = x^x [(x)' \ln x + x (\ln x)'] \\
    &= x^x \left[ 1 \ln x + x \frac{1}{x} \right] = x^x [\ln x + 1] = x^x [1 + \ln x] \\
    &= x^x + x^x \ln x
    \end{aligned}$}
    \begin{aligned}
    f(x) &= x^x \\
    &\ln f(x) = x \ln x \\
    \frac{d}{dx} \ln f(x) &= \frac{d}{dx} [x \ln x] \\
    \frac{f'(x)}{f(x)} &= [x \ln x]' \\
    f'(x) &= f(x) [x \ln x]' = x^x [(x)' \ln x + x (\ln x)'] \\
    &= 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}$
    \begin{aligned}
    f(x) &= x^x \\
    &\ln f(x) = x \ln x \\
    \frac{d}{dx} \ln f(x) &= \frac{d}{dx} [x \ln x] \\
    \frac{f'(x)}{f(x)} &= [x \ln x]' \\
    f'(x) &= f(x) [x \ln x]' = x^x [(x)' \ln x + x (\ln x)'] \\
    &= 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}$
};
\end{axis}
\end{tikzpicture}
```

Figure 62.2: test

$$\begin{aligned}
 f(x) &= x^x \\
 \ln f(x) &= x \ln x \\
 \frac{d}{dx} \ln f(x) &= \frac{d}{dx} [x \ln x] \\
 \frac{f'(x)}{f(x)} &= [x \ln x]' \\
 f'(x) &= f(x) [x \ln x]' = x^x [(x)' \ln x + x (\ln x)'] \\
 &= x^x \left[1 \ln x + x \frac{1}{x} \right] = x^x [\ln x + 1] = x^x [1 + \ln x] \\
 &= x^x + x^x \ln x
 \end{aligned}$$

<https://tex.stackexchange.com/questions/670153/how-to-annotate-calculations>

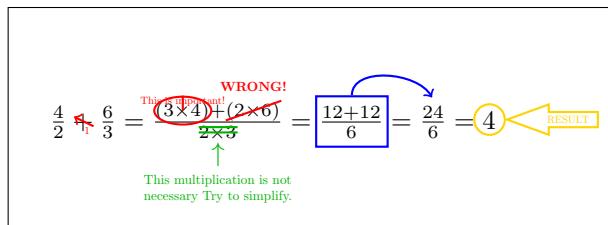


Figure 62.3: test

```

\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= \emptyset,
    xticklabels= \emptyset,
    %extra x ticks={0},
    ytick = \emptyset,
    yticklabels= \emptyset,
    x = 5cm, y = 5cm, % x y scaling
    grid = major,
    domain = -1:1,
    %samples = 1000
]
\node at (axis cs:0,0) {
    → $\tikzmarknode{a1}{\frac{4}{2}}+\tikzmarknode{a2}{\frac{6}{3}}=$
    → $\frac{\tikzmarknode{b1}{(3\times 4)}+\tikzmarknode{b2}{(2\times 6)}}{6}=$
    → $\tikzmarknode{c1}{\frac{12+12}{6}}=\tikzmarknode{c2}{\frac{24}{6}}=\tikzmarknode{d1}{4}$;
    \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}

```

$$\frac{4}{2} + \frac{6}{3} = \frac{(3 \times 4) + (2 \times 6)}{6} = \frac{12+12}{6} = \frac{24}{6} = 4$$

WRONG!

This multiplication is not necessary Try to simplify.

Figure 62.4: test

<https://tex.stackexchange.com/questions/494884/anchor-alignment-in-tikzmarknode>

$$\begin{aligned} \sum_{i=1}^3 \sum_{j=1}^4 a_{ij} &= \sum_{j=1}^4 a_{1j} + \sum_{j=1}^4 a_{2j} + \sum_{j=1}^4 a_{3j} \\ &= [a_{11} + a_{12} + a_{13} + a_{14}] + [a_{21} + a_{22} + a_{23} + a_{24}] + [a_{31} + a_{32} + a_{33} + a_{34}] \end{aligned}$$

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= \emptyset,
    xticklabels= \emptyset,
    %extra x ticks={0},
    ytick = \emptyset,
    yticklabels= \emptyset,
    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} &= \sum\limits_{j=1}^4 a_{1j} + \sum\limits_{j=1}^4 a_{2j} + \sum\limits_{j=1}^4 a_{3j} \\
    &= a_{11} + a_{12} + a_{13} + a_{14} + a_{21} + a_{22} + a_{23} + a_{24} + a_{31} + a_{32} + a_{33} + a_{34}
    \end{aligned}
}
\end{tikzpicture}
```

$$\sum_{i=1}^3 \sum_{j=1}^4 a_{ij} = \sum_{j=1}^4 a_{1j} + \sum_{j=1}^4 a_{2j} + \sum_{j=1}^4 a_{3j}$$

$$= a_{11} + a_{12} + a_{13} + a_{14} + a_{21} + a_{22} + a_{23} + a_{24} + a_{31} + a_{32} + a_{33} + a_{34}$$

Figure 62.6: test

```

    +
    \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]{red1}{\sum\limits_{j=1}^4 a_{1j}} +
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}
```

https://github.com/synergys/annotated_latex_equations

https://github.com/synergys/annotated_latex_equations/blob/main/example_prob.tex

$$\begin{aligned}f(x) &= x^x \\ \ln f(x) &= x \ln x \\ \frac{d}{dx} \ln f(x) &= \frac{d}{dx} [x \ln x] \\ \frac{f'(x)}{f(x)} &= [x \ln x]' \\ f'(x) &= f(x) [x \ln x]' = x^x [(x)' \ln x + x (\ln x)'] \\ &= x^x \left[1 \ln x + x \frac{1}{x} \right] = x^x [\ln x + 1] = x^x [1 + \ln x] \\ &= x^x + x^x \ln x\end{aligned}$$

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. Bernstein, M. A., King, K. F. & Zhou, X. J. *Handbook of MRI Pulse Sequences*. (Academic Press, Amsterdam ; Boston, 2004).
14. ccjou. 矩陣導數. (2013).
15. ccjou. 多變量常態分布. (2014).
16. ccjou. Jacobian 矩陣與行列式. (2012).