Appendix A

Optional Topics

A.1 Dealing with Infinity*

A.1.1 The Axiom of Choice

The axiom of choice, formulated by Zermelo in 1904, is innocent-looking. However, one can prove theorems with its aid that some mathematicians were originally reluctant to accept in the past.

**Definition A.1.1 (The Axiom of Choice).** Given a collection $\mathcal{X}$ of disjoint nonempty sets, there exists a set $C$ having exactly one element in common with each element of $\mathcal{X}$. That is, for each $X \in \mathcal{X}$ the set $C \cap X$ contains a single element.

Most mathematicians today accept the axiom of choice as part of the set theory on which they base their mathematics. A straightforward consequence of the axiom of choice is the existence of a choice function.

**Lemma A.1.2 (Existence of a Choice Function).** Given a collection $\mathcal{Y}$ of non-empty sets, there exists a function $c : \mathcal{Y} \to \bigcup_{Y \in \mathcal{Y}} Y$ satisfying $c(Y) \in Y$ for every $Y \in \mathcal{Y}$.

**Proof.** The difference between the axiom of choice and the lemma is that in the latter statement the sets of the collection $\mathcal{Y}$ need not be disjoint. Given an element $Y \in \mathcal{Y}$, define the set $Y'$ by

$$Y' = \{(Y, y) \mid y \in Y\}.$$
That is, $Y'$ is the collection of all ordered pairs where the first coordinate of the ordered pair is the set $Y$, and the second coordinate is an element of $Y$. Because $Y$ contains at least one element, the set $Y'$ is nonempty. Furthermore, $Y'$ is a subset of the cartesian product

$$Y \times \bigcup_{Y \in \mathcal{Y}} Y.$$ 

If $Y_1$ and $Y_2$ are two different sets in $\mathcal{Y}$, then the sets $Y_1'$ and $Y_2'$ are disjoint; specifically, the elements of $Y_1'$ and $Y_2'$ differ at least in their first coordinates.

Consider the collection

$$\mathcal{Z} = \{Y'|Y \in \mathcal{Y}\}.$$ 

This is a collection of disjoint nonempty subsets of

$$\mathcal{Y} \times \bigcup_{Y \in \mathcal{Y}} Y.$$ 

By the axiom of choice, there exists a set $Z$ having exactly one element in common with each element of $\mathcal{Z}$. Define the function

$$c : \mathcal{Z} \to \mathcal{Y} \times \bigcup_{Y \in \mathcal{Y}} Y$$

by $c(Y') = Y' \cap Z$. This function $c$ implicitly provides the rule for a function from $\mathcal{Y}$ to the set $\bigcup_{Y \in \mathcal{Y}} Y$ such that $y$ belongs to $Y$ whenever $(Y, y) \in Z$. This rule is the desired choice function.

\[ \square \]

### A.1.2 Well-Ordered Sets

A simple order $<$ on a set $X$ is a relation such that, for all $x, y, z \in X$,

1. if $x \neq y$ then either $x < y$ or $y < x$
2. if $x < y$ then $x \neq y$
3. if $x < y$ and $y < z$ then $x < z$.

**Definition A.1.3.** A set $X$ with an order relation $<$ is said to be well-ordered if every nonempty subset of $X$ has a smallest element.

The set of natural numbers, for example, is well-ordered. On the other hand, the set of integers is not well-ordered.
Fact A.1.4 (Well-ordering theorem). If \( X \) is a set, there exists an order relation on \( X \) that is a well-ordering.

This theorem was proved by Zermelo using the axiom of choice. It startled the mathematical community in 1904 and spurred much controversy about the axiom of choice. It is given here without a proof.

Corollary A.1.5. There exists an uncountable well-ordered set.

Definition A.1.6. Let \( X \) be an ordered set. Given \( x \in X \), the set

\[ Y_x = \{ y \in Y | y < x \} \]

is called the section of \( X \) by \( x \).

Corollary A.1.7. There exists an uncountable well-ordered set, every section of which is countable.

The well-ordering principle is a necessary tool in proofs by induction when the set over which the induction process is applied is not a segment of the natural numbers; this is the so-called transfinite induction.

A.1.3 The Maximum Principle

A strict partial order \( \prec \) on a set \( X \) is a relation such that for all \( x, y, z \in X \)

1. if \( x \prec y \) then \( x \neq y \)
2. if \( x \prec y \) and \( y \prec z \) then \( x \prec z \).

A strict partial order is similar to a simple order, except that it need not be true that for every distinct \( x, y \in X \), either \( x \prec y \) or \( y \prec x \).

Fact A.1.8 (The maximum principle). Let \( X \) be a set and suppose that \( \prec \) is a strict partial order on \( X \). If \( Y \) is a subset of \( X \) that is simply ordered by \( \prec \), then there exists a maximal simply ordered subset \( Z \) of \( X \) containing \( Y \).

The maximum principle is given here without a proof. It is interesting to note that the well-ordering theorem and the maximum principle are equivalent; either of
them implies the other. Furthermore, each of them is equivalent to the axiom of choice.

Let $\prec$ be a strict partial order on $X$. For $x, y \in X$, the relation $x \preceq y$ holds if $x \prec y$ or $x = y$. The relation $\preceq$ so defined is called a partial order on $X$. For example, the inclusion relation $\subset$ on a collection of sets is a partial order, whereas proper inclusion is a strict partial order.
Bibliography

Index

applications, 91
   linear regression, 91
   Wiener-Hopf, 93

Banach space
   strictly convex, 111

eigenvalues, 151
   algebraic multiplicity, 152
   characteristic polynomial, 152
   defective, 157
   diagonalizable, 153
   eigenspace, 152
   eigenvalue, 152
   eigenvector, 151
   generalized eigenspace, 158
   generalized eigenvector, 159
   geometric multiplicity, 152
   Jordan chain, 159
   Jordan normal form, 158
   similar, 154
   spectrum, 151

domain, 19

global minimum value, 111
image, 20
injective, 20
inverse function, 20
inverse image, 20
one-to-one, 20
one-to-one correspondence, 20
onto, 20
preimage, 20
surjective, 20

inner-product space, 64
   adjoint, 135
   best approximation, 79
   Cauchy-Schwarz inequality, 68
   dual approximation, 96
   Euclidean space, 67
   Gram-Schmidt orthogonalization, 71
   half space, 101
   Hilbert space, 73
   induced norm, 67
   inner product, 64
   least-squares, 87
   normal equations, 85
   orthogonal, 67, 69, 73
   orthogonal complement, 72
   orthogonal set, 69

field, 43
functions, 19
   bijective, 20
   codomain, 19
   concave, 111
   convex, 111
INDEX

orthonormal basis, 73
Parseval identity, 74
projection, 67
Riesz representation theorem, 134
standard inner product, 64
unitary, 73
integral, 61
almost everywhere, 60
Lebesgue integral, 60, 61
Lebesgue measure, 60
Riemann integral, 61
linear transform, 54
algebra, 125
Banach algebra, 125
bounded, 130
coordinate matrix, 56
idempotent, 82
invertible, 125
linear operator, 124
non-singular, 55
nullity, 57
nullspace, 57
operator norm, 129
orthogonal projection, 83
projection, 82
pseudoinverse, 140
range, 56
rank, 57
singular, 55
transpose, 128
vector product, 125
logic, 1
biconditional, 4

complete, 9
conditional connective, 3, 4
conjecture, 1
conjunction, 2
consistent, 9
contradiction, 5
contrapositive, 9
converse, 7
corollary, 12
decidable, 9
disjunction, 3
decisional quantifier, 10
fallacy, 9
free variable, 10
implication relation, 4
lemma, 12
logical equivalence, 8
logical implication, 6
mathematical induction, 13
negation, 3
predicate, 10
proof, 1
proposition, 12
semidecidable, 11
tautology, 5
theorem, 12
universal quantifier, 10

matrix
compact SVD, 167
convergent, 160
elementary column operation, 46
elementary row operation, 46
Frobenius norm, 134
Gramian, 86  
Hermitian transpose, 45  
inverse, 46  
invertible, 46  
matrix product, 45  
orthogonal, 150  
positive-definite, 86  
positive-semidefinite, 86  
projection matrix, 89  
pseudoinverse, 89, 140  
reduced row echelon form, 46  
row echelon form, 45  
spectral radius, 133  
trace, 75  
transpose, 45  
unitary, 150  
matrix factorization, 141  
backward substitution, 142  
Cholesky factorization, 148  
forward substitution, 141  
LDLT decomposition, 148  
lower triangular, 141  
LU decomposition, 143  
orthogonal, 149  
unit triangular, 141  
unitary, 149  
upper triangular, 141  
metric space, 23  
d-open ball, 24  
boundary, 27  
Cauchy sequence, 25  
closed, 26  
closure, 27  
compact, 32  
complete, 29  
completion, 30  
continuous, 27, 28  
contraction, 30  
converges, 25  
dense, 30  
distance, 24  
Euclidean metric, 24  
interior, 26  
isometry, 30  
limit, 27  
limit point, 26  
Lipschitz continuous, 28  
metric, 23  
open, 25  
points, 24  
pointwise convergence, 35  
sequence, 24  
totally bounded, 32  
uniform convergence, 35  
uniformly continuous, 28  
optimization  
active, 115  
convex, 120  
feasible, 114  
Fréchet derivative, 107  
Fréchet differentiable, 107  
Gâteaux differentiable, 106  
Gâteaux differential, 106  
gradient, 107  
Jacobian matrix, 107  
Lagrange multiplier, 115  
Lagrangian, 115
INDEX

Lagrangian dual, 118
linear program, 114
local minimum value, 111
locally optimal, 115
objective function, 114
optimal value, 114
Slater’s condition, 120
standard form, 114
strong duality, 119
weak duality, 118

set theory
axiom of choice, 169
cardinality, 16
Cartesian Product, 18
complement, 17
complex numbers, 15
countably infinite, 16
disjoint, 17
elements, 15
empty set, 15
equivalence classes, 18
equivalence relation, 18
infimum, 33
integers, 15
intersection, 17
maximum, 33
minimum, 33
naive set theory, 14
natural numbers, 15
partial order, 172
quotient set, 18
rational numbers, 15
real numbers, 15
Russell’s Paradox, 16
set, 15
set difference, 17
set-builder notation, 15
simple order <, 170
singleton, 15
strict partial order, 171
subset, 17
supremum, 33
uncountably infinite, 16
union, 17
well-ordered, 170
topology, 36
basis, 36
closed, 38
closure, 38
continuous, 39
converge, 41
dense, 39
extended real numbers, 33
interior, 38
limit point, 38
metric topology, 37
metrizable, 37
neighborhood, 38
open set, 36
separable, 39

vector space, 47
affine hyperplane, 100
Banach space, 61
best approximation, 79
closed subspace, 62
convex set, 97
coordinate vector, 53
dimension, 52
direct sum, 49
dual basis, 127
dual space, 126
finite-dimensional, 50
functional, 110
Hamel basis, 50
homomorphism, 126
hyperplane, 100
isomorphism, 126
linear combination, 48
linear functional, 75
linear transform, 54
linearly dependent, 49
linearly independent, 49
norm, 58
normalized, 61
ordered basis, 53
Schauder basis, 62
span, 49
standard basis, 50
standard Schauder basis, 62
subspace, 48
unit vector, 61