

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*

$$c : \mathcal{Y} \rightarrow \bigcup_{Y \in \mathcal{Y}} Y$$

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

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

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

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

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

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

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*

$$Y_x = \{y \in Y \mid y < x\}$$

*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

- [BV04] S. P. Boyd, L. Vandenberghe. *Convex Optimization*. Cambridge University Press, 2004.

Index

- applications, 76
 - linear regression, 76
 - Wiener-Hopf, 78
- Banach space
 - best approximation, 65
 - closed subspace, 52
 - Schauder basis, 52
 - standard Schauder basis, 52
 - strictly convex, 89
- eigenvalues, 127
 - algebraic multiplicity, 128
 - characteristic polynomial, 128
 - defective, 133
 - diagonalizable, 129
 - eigenspace, 128
 - eigenvalue, 127
 - eigenvector, 127
 - generalized eigenspace, 134
 - generalized eigenvector, 135
 - geometric multiplicity, 128
 - Jordan chain, 135
 - Jordan normal form, 134
 - similar, 130
 - spectrum, 127
- field, 35
- functions, 16
- bijjective, 17
- codomain, 16
- convex, 89
- domain, 16
- global minimum value, 89
- image, 17
- injective, 17
- inverse function, 17
- inverse image, 17
- one-to-one, 17
- one-to-one correspondence, 17
- onto, 17
- preimage, 17
- surjective, 17
- inner-product space, 53
 - adjoint, 111
 - best approximation, 65
 - Cauchy-Schwarz inequality, 56
 - dual approximation, 81
 - Euclidean space, 56
 - Gram-Schmidt orthogonalization, 59
 - Hilbert space, 61
 - induced norm, 56
 - inner product, 53
 - least-squares, 73
 - normal equations, 71
 - orthogonal, 57

- orthogonal complement, 60
- orthogonal set, 57
- orthonormal basis, 61
- Parseval identity, 61
- Riesz representation theorem, 109
- standard inner product, 54
- integral, 50
 - almost everywhere, 51
 - Lebesgue integral, 50, 51
 - Lebesgue measure, 51
 - Riemann integral, 50
- linear transform, 45
 - bounded, 105
 - idempotent, 67
 - invertible, 101
 - linear operator, 100
 - non-singular, 45
 - nullity, 47
 - nullspace, 47
 - operator norm, 104
 - orthogonal projection, 68
 - pseudoinverse, 115
 - range, 46
 - rank, 47
 - singular, 45
 - transpose, 104
- logic, 1
 - biconditional, 4
 - conditional, 3
 - conditional connective, 4
 - conjecture, 1
 - conjunction, 2
 - contradiction, 5
 - contrapositive, 9
 - converse, 7
 - corollary, 10
 - disjunction, 3
 - existential quantifier, 10
 - fallacy, 9
 - implication relation, 4
 - lemma, 10
 - logical equivalence, 8
 - logical implication, 6
 - negation, 3
 - proof, 1
 - proof by contradiction, 9
 - proposition, 10
 - tautology, 5
 - theorem, 10
 - universal quantifier, 9
- matrix
 - compact SVD, 143
 - convergent, 136
 - Frobenius norm, 109
 - Grammian, 71
 - Hermitian transpose, 38
 - inverse, 37
 - invertible, 37
 - matrix product, 37
 - orthogonal, 126
 - positive-definite, 72
 - positive-semidefinite, 72
 - projection matrix, 75
 - pseudoinverse, 75, 115
 - spectral radius, 108
 - trace, 63

- unitary, 126
- matrix factorization, 117
 - backward substitution, 118
 - Cholesky factorization, 124
 - forward substitution, 117
 - LDLT decomposition, 124
 - lower triangular, 117
 - LU decomposition, 119
 - orthogonal, 125
 - unit triangular, 117
 - unitary, 125
 - upper triangular, 117
- metric space, 19
 - d -open ball, 20
 - Cauchy sequence, 21
 - closure, 22
 - compact, 26
 - complete, 24
 - completion, 25
 - continuous, 22, 23
 - contraction, 25
 - converges, 21
 - dense, 24
 - distance, 20
 - Euclidean metric, 20
 - interior, 22
 - isometry, 24
 - limit point, 22
 - Lipschitz continuous, 23
 - metric, 19
 - open, 21
 - points, 20
 - pointwise convergence, 29
 - sequence, 20
 - totally bounded, 26
 - uniform convergence, 29
 - uniformly continuous, 23
- optimization
 - active, 92
 - convex, 97
 - feasible, 91
 - Fréchet differentiable, 88
 - Gâteaux differentiable, 88
 - Gâteaux differential, 87
 - gradient, 88
 - Jacobian matrix, 88
 - Lagrange multiplier, 92
 - Lagrangian, 92
 - Lagrangian dual, 96
 - linear program, 92
 - local minimum value, 89
 - locally optimal, 92
 - objective function, 91
 - optimal value, 91
 - Slater's condition, 97
 - standard form, 91
 - strong duality, 96
 - weak duality, 96
- set theory
 - axiom of choice, 145
 - Cartesian Product, 15
 - complement, 14
 - complex numbers, 13
 - disjoint, 14
 - elements, 12
 - empty set, 13
 - equivalence classes, 15

- equivalence relation, 15
 - infimum, 27
 - integers, 13
 - intersection, 14
 - maximum, 27
 - minimum, 27
 - naive set theory, 12
 - natural numbers, 13
 - partial order, 148
 - quotient set, 15
 - rational numbers, 13
 - real numbers, 13
 - Russell's Paradox, 13
 - set, 12
 - set difference, 14
 - set-builder notation, 13
 - simple order $<$, 146
 - singleton, 13
 - strict partial order, 147
 - subset, 14
 - supremum, 27
 - union, 14
 - well-ordered, 146
- topology, 29
- basis, 30
 - closed, 31
 - closure, 31
 - continuous, 32
 - converge, 34
 - dense, 32
 - extended real numbers, 27
 - interior, 31
 - limit point, 32
 - metric topology, 30
 - metrizable, 31
 - neighborhood, 32
 - open set, 30
 - separable, 32
- vector space, 37
- Banach space, 52
 - convex set, 83
 - dimension, 42
 - direct sum, 39
 - dual basis, 103
 - dual space, 102
 - finite-dimensional, 40
 - functional, 88
 - Hamel basis, 40
 - homomorphism, 102
 - isomorphism, 102
 - linear combination, 38
 - linear functional, 62
 - linear transform, 45
 - linearly dependent, 40
 - linearly independent, 40
 - norm, 48
 - normalized, 52
 - ordered basis, 43
 - span, 39
 - standard basis, 41
 - subspace, 39
 - unit vector, 52