MINICOURSE NOTES based on

CONVEX ANALYSIS AND NONLINEAR OPTIMIZATION

(Springer Verlag, in press)

J.M. Borwein[1]
Centre for Experimental and Constructive Mathematics
Department of Mathematics and Statistics
Simon Fraser University, Burnaby, B.C., Canada V5A 1S6
jborwein@cs.sfu.ca

and

A.S. Lewis[1]
Department of Combinatorics and Optimization
University of Waterloo, Waterloo, Ont., Canada N2L 3G1
aslewis@orion.uwaterloo.ca

April 5, 1999

[1]Research partially supported by the Natural Sciences and Engineering Research Council of Canada

Contents

0.1 Preface 3

1 Background **7**
1.1 Euclidean spaces 7
1.2 Symmetric matrices 15

2 Inequality constraints **19**
2.1 Optimality conditions 19
2.2 Theorems of the alternative 26
2.3 Max-functions and first order conditions 31

3 Fenchel duality **37**
3.1 Subgradients and convex functions 37
3.2 The value function 46
3.3 The Fenchel conjugate 51

4 Convex analysis **61**
4.1 Continuity of convex functions 61
4.2 Fenchel biconjugation 68
4.3 Lagrangian duality 75

5 Special cases **81**
5.1 Polyhedral convex sets and functions 81
5.2 Functions of eigenvalues 90
5.3 Duality for linear and semidefinite programming 95
5.4 Convex process duality 102

6 The Variational Principle **112**
6.1 An introduction to metric regularity 112

7 Fixed points **119**
7.1 Brouwer's fixed point theorem 119
7.2 Selection results and the Kakutani-Fan fixed point theorem . . 126
7.3 Variational inequalities 133

0.1 Preface

Optimization is a rich and thriving mathematical discipline. Properties of minimizers and maximizers of functions rely intimately on a wealth of techniques from mathematical analysis, including tools from calculus and its generalizations, topological notions, and more geometric ideas. The theory underlying current computational optimization techniques grows ever more sophisticated – duality-based algorithms and interior point methods are typical examples. The powerful and elegant language of convex analysis unifies much of this theory. Hence our aim of writing a concise, accessible account of convex analysis and its applications and extensions, for a broad audience.

For students of optimization and analysis, there is great benefit to blurring the distinction between the two disciplines. Many important analytic problems have illuminating optimization formulations and hence can be approached through our main variational tools: subgradients and optimality conditions, the many guises of duality, metric regularity and so forth. More generally, the idea of convexity is central to the transition from classical analysis to various branches of modern analysis: from linear to nonlinear analysis, from smooth to nonsmooth, and from the study of functions to multifunctions. Thus although we use certain optimization models repeatedly to illustrate the main results (models such as linear and semidefinite programming duality and cone polarity), we constantly emphasize the power of abstract models and notation.

Good reference works on finite-dimensional convex analysis already exist. Rockafellar's classic *Convex Analysis*has been indispensable and ubiquitous since the 1970's, and a more general sequel with Wets, *Variational Analysis*appeared recently. Hiriart-Urruty and Lemaréchal's *Convex Analysis and Minimization Algorithms* is a comprehensive but gentler introduction. Our goal is not to supplant these works, but on the contrary to promote them, and thereby to motivate future researchers. This book aims to make converts.

We try to be succinct rather than systematic, avoiding becoming bogged down in technical details. Our style is relatively informal: for example, the text of each section sets the context for many of the result statements. We value the variety of independent, self-contained approaches over a single, unified, sequential development. We hope to showcase a few memorable principles rather than to develop the theory to its limits. We discuss no algorithms. We point out a few important references as we go, but we make no attempt at comprehensive historical surveys.

Infinite-dimensional optimization lies beyond our immediate scope. This is for reasons of space and accessibility rather than history or application: convex analysis developed historically from the calculus of variations, and has important applications in optimal control and other areas of infinite-dimensional optimization. However, rather like Halmos's *Finite Dimensional Vector Spaces* ease of extension beyond finite dimensions substantially motivates our choice of results and techniques. We would, in part, like this book to be an entrée for mathematicians to a valuable and intrinsic part of modern analysis. The final chapter illustrates some of the challenges arising in infinite dimensions.

This book can serve as a teaching text, at roughly the level of first year graduate students. In principle we assume no knowledge of real analysis, although in practice we expect a certain mathematical maturity. While the main body of the text is self-contained, each section concludes with an often extensive set of optional exercises. These exercises fall into three categories, marked with zero, one or two asterisks respectively: examples which illustrate the ideas in the text or easy expansions of sketched proofs; important pieces of additional theory or more testing examples; longer, harder examples or peripheral theory.

Many people have helped improve the presentation of this material. We would like to thank all of them, but in particular Heinz Bauschke, Guillaume Haberer, Claude Lemaréchal, Olivier Ley, Hristo Sendov, Mike Todd and Xianfu Wang.

Chapter 1

Background

1.1 Euclidean spaces

We begin by reviewing some of the fundamental algebraic, geometric and analytic ideas we use throughout the book. **Our setting, for most of the book, is an arbitrary Euclidean space $\mathbf{E}$**, by which we mean a finite-dimensional vector space over the reals $\mathbf{R}$ equipped with an inner product $\langle \cdot, \cdot \rangle$. We would lose no generality if we considered only the space $\mathbf{R}^n$ of real (column) n-vectors (with its standard inner product), but a more abstract, coordinate-free notation is often more flexible and elegant.

We define the *norm* of any point x in $\mathbf{E}$ by $\|x\| = \sqrt{\langle x, x \rangle}$, and the *unit ball* is the set

$$B = \{x \in \mathbf{E} \mid \|x\| \leq 1\}.$$

Any two points x and y in $\mathbf{E}$ satisfy the *Cauchy-Schwarz* *inequality*

$$|\langle x, y\rangle| \le \|x\|\|y\|.$$

We define the sum of two sets C and D in $\mathbf{E}$ by

$$C + D = \{x + y \mid x \in C,\ y \in D\}.$$

The definition of $C - D$ is analogous, and for a subset Λ of $\mathbf{R}$ we define

$$\Lambda C = \{\lambda x \mid \lambda \in \Lambda,\ x \in C\}.$$

Given another Euclidean space $\mathbf{Y}$, we can consider the Cartesian product Euclidean space $\mathbf{E} \times \mathbf{Y}$, with inner product defined by $\langle (e, x), (f, y)\rangle = \langle e, f\rangle + \langle x, y\rangle$.

We denote the nonnegative reals by $\mathbf{R}_+$. If C is nonempty and satisfies $\mathbf{R}_+ C = C$ we call it a *cone*. (Notice we require that cones contain 0.) Examples are the positive orthant

$$\mathbf{R}^n_+ = \{x \in \mathbf{R}^n \mid \text{each } x_i \ge 0\},$$

and the cone of vectors with nonincreasing components

$$\mathbf{R}^n_\ge = \{x \in \mathbf{R}^n \mid x_1 \ge x_2 \ge \ldots \ge x_n\}.$$

The smallest cone containing a given set $D \subset \mathbf{E}$ is clearly $\mathbf{R}_+ D$.

The fundamental geometric idea of this book is *convexity*. A set C in $\mathbf{E}$ is *convex* if the line segment joining any two points x and y in C is contained in C: algebraically, $\lambda x + (1-\lambda)y \in C$ whenever $0 \le \lambda \le 1$. An easy exercise shows that intersections of convex sets are convex.

Given any set $D \subset \mathbf{E}$, the *linear span* of D, denoted $\mathrm{span}\,(D)$, is the smallest linear space containing D. It consists exactly of all linear combinations of elements of D. Analogously, the *convex hull* of D, denoted $\mathrm{conv}\,(D)$, is the smallest convex set containing D. It consists exactly of all *convex combinations* of elements of D, that is to say points of the form $\Sigma_{i=1}^{m} \lambda_i x^i$, where $\lambda_i \in \mathbf{R}_+$ and $x^i \in D$ for each i, and $\Sigma\, \lambda_i = 1$ (see Exercise 2.

The language of elementary point-set topology is fundamental in optimization. A point x lies in the *interior* of the set $D \subset \mathbf{E}$ (denoted $\mathrm{int}\, D$) if there is a real $\delta > 0$ satisfying $x + \delta B \subset D$. In this case we say D is a *neigbourhood* of x. For example, the interior of $\mathbf{R}^n_+$ is

$$\mathbf{R}^n_{++} = \{x \in \mathbf{R}^n \mid \text{each } x_i > 0\}.$$

We say the point x in $\mathbf{E}$ is the *limit* of the sequence of points $x^1, x^2, \ldots$ in $\mathbf{E}$, written $x^i \to x$ as $i \to \infty$ (or $\lim_{i\to\infty} x^i = x$), if $\|x^i - x\| \to 0$. The *closure* of D is the set of limits of sequences of points in D, written $\mathrm{cl}\, D$, and the *boundary* of D is $\mathrm{cl}\, D \setminus \mathrm{int}\, D$, written $\mathrm{bd}\, D$. The set D is *open* if $D = \mathrm{int}\, D$, and is *closed* if $D = \mathrm{cl}\, D$.

Linear subspaces of $\mathbf{E}$ are important examples of closed sets. Easy exercises show that D is open exactly when its complement D^c is closed, and that arbitrary unions and finite intersections of open sets are open. The interior of D is just the largest open set contained in D, while $\operatorname{cl} D$ is the smallest closed set containing D. Finally, a subset G of D is *open in* D if there is an open set $U \subset \mathbf{E}$ with $G = D \cap U$.

Much of the beauty of convexity comes from *duality* ideas, interweaving geometry and topology. The following result, which we prove a little later, is both typical and fundamental.

Theorem 1.1.1 (Basic separation) *Suppose that the set $C \subset \mathbf{E}$ is closed and convex, and that the point y does not lie in C. Then there exist real b and a nonzero element a of $\mathbf{E}$ satisfying $\langle a, y \rangle > b \geq \langle a, x \rangle$ for all points x in C.*

Sets in $\mathbf{E}$ of the form $\{x \mid \langle a, x \rangle = b\}$ and $\{x \mid \langle a, x \rangle \leq b\}$ (for a nonzero element a of $\mathbf{E}$ and real b) are called *hyperplanes* and *closed halfspaces* respectively. In this language the above result states that the point y is *separated* from the set C by a hyperplane: in other words, C is contained in a certain closed halfspace whereas y is not. Thus there is a 'dual' representation of C as the intersection of all closed halfspaces containing it.

The set D is *bounded* if there is a real k satisfying $kB \supset D$, and is *compact* if it is closed and bounded. The following result is a central tool in real analysis.

Theorem 1.1.2 (Bolzano-Weierstrass) *Any bounded sequence in* $\mathbf{E}$ *has a convergent subsequence.*

Just as for sets, geometric and topological ideas also intermingle for the functions we study. Given a set D in $\mathbf{E}$, we call a function $f : D \to \mathbf{R}$ *continuous* (on D) if $f(x^i) \to f(x)$ for any sequence $x^i \to x$ in D. In this case it easy to check, for example, that for any real α the *level set* $\{x \in D \mid f(x) \leq \alpha\}$ is closed providing D is closed.

Given another Euclidean space $\mathbf{Y}$, we call a map $A : \mathbf{E} \to \mathbf{Y}$ *linear* if any points x and z in $\mathbf{E}$ and any reals λ and μ satisfy $A(\lambda x + \mu z) = \lambda Ax + \mu Az$. In fact any linear function from $\mathbf{E}$ to $\mathbf{R}$ has the form $\langle a, \cdot \rangle$ for some element a of $\mathbf{E}$. Linear maps and *affine* functions (linear functions plus constants) are continuous. Thus, for example, closed halfspaces are indeed closed. A *polyhedron* is a finite intersection of closed halfspaces, and is therefore both closed and convex. The *adjoint* of the map A above is the linear map $A^* : \mathbf{Y} \to \mathbf{E}$ defined by the property

$$\langle A^* y, x \rangle = \langle y, Ax \rangle, \quad \text{for all points } x \text{ in } \mathbf{E} \text{ and } y \text{ in } \mathbf{Y}$$

(whence $A^{**} = A$). The *null space* of A is $N(A) = \{x \in \mathbf{E} \mid Ax = 0\}$. The *inverse image* of a set $H \subset \mathbf{Y}$ is the set $A^{-1}H = \{x \in \mathbf{E} \mid Ax \in H\}$ (so for example $N(A) = A^{-1}\{0\}$). Given a subspace G of $\mathbf{E}$, the *orthogonal complement* of G is the subspace

$$G^{\perp} = \{y \in \mathbf{E} \mid \langle x, y \rangle = 0 \text{ for all } x \in G\},$$

so called because we can write $\mathbf{E}$ as a direct sum $G \oplus G^{\perp}$. (In other words, any element of $\mathbf{E}$ can be written uniquely as the sum of an element of G and an element of $G^{\perp}$.) Any subspace satisfies $G^{\perp\perp} = G$. The range of any linear map A coincides with $N(A^*)^{\perp}$.

Optimization studies properties of minimizers and maximizers of functions. Given a set $\Lambda \subset \mathbf{R}$, the *infimum* of Λ (written $\inf \Lambda$) is the greatest lower bound on Λ, and the *supremum* (written $\sup \Lambda$) is the least upper bound. To ensure these are always defined, it is natural to append $-\infty$ and $+\infty$ to the real numbers, and allow their use in the usual notation for open and closed intervals. Hence $\inf \emptyset = +\infty$ and $\sup \emptyset = -\infty$, and for example $(-\infty, +\infty]$ denotes the interval $\mathbf{R} \cup \{+\infty\}$. We try to avoid the appearance of $+\infty - \infty$, but when necessary we use the convention $+\infty - \infty = +\infty$, so that any two sets C and D in $\mathbf{R}$ satisfy $\inf C + \inf D = \inf(C + D)$. We also adopt the conventions $0 \cdot (\pm\infty) = (\pm\infty) \cdot 0 = 0$.

A *(global) minimizer* of a function $f : D \to \mathbf{R}$ is a point $\bar{x}$ in D at which f attains its infimum

$$\inf_D f = \inf f(D) = \inf\{f(x) \mid x \in D\}.$$

In this case we refer to $\bar{x}$ as an *optimal solution* of the *optimization problem* $\inf_D f$.

For a positive real δ and a function $g : (0, \delta) \to \mathbf{R}$, we define

$$\begin{aligned} \liminf_{t \downarrow 0} g(t) &= \lim_{t \downarrow 0} \inf_{(0,t)} g, \quad \text{and} \\ \limsup_{t \downarrow 0} g(t) &= \lim_{t \downarrow 0} \sup_{(0,t)} g. \end{aligned}$$

The limit $\lim_{t \downarrow 0} g(t)$ exists if and only if the above expressions are equal.

The question of the *existence* of an optimal solution for an optimization problem is typically topological. The following result is a prototype. The proof is a standard application of the Bolzano-Weierstrass Theorem above.

Proposition 1.1.3 (Weierstrass) *Suppose that the set $D \subset \mathbf{E}$ is nonempty and closed, and that all the level sets of the continuous function $f : D \to \mathbf{R}$ are bounded. Then f has a global minimizer.*

Just as for sets, convexity of functions will be crucial for us. Given a convex set $C \subset \mathbf{E}$, we say that the function $f : C \to \mathbf{R}$ is *convex* if

$$f(\lambda x + (1-\lambda)y) \leq \lambda f(x) + (1-\lambda)f(y)$$

for all points x and y in C and $0 \leq \lambda \leq 1$. The function f is *strictly convex* if the inequality holds strictly whenever x and y are distinct in C and $0 < \lambda < 1$. It is easy to see that a strictly convex function can have at most one minimizer.

Requiring the function f to have bounded level sets is a 'growth condition'. Another example is the stronger condition

$$\liminf_{\|x\|\to\infty} \frac{f(x)}{\|x\|} \left(= \lim_{r\to+\infty} \inf \left\{ \frac{f(x)}{\|x\|} \;\middle|\; 0 \neq x \in C \cap rB \right\} \right) > 0.$$

Surprisingly, for *convex* functions these two growth conditions are equivalent.

Proposition 1.1.4 *For a convex set $C \subset \mathbf{E}$, a convex function $f : C \to \mathbf{R}$ has bounded level sets if and only if it satisfies the growth condition (1.1.4).*

1.2 Symmetric matrices

Throughout most of this book our setting is an abstract Euclidean space $\mathbf{E}$. This has a number of advantages over always working in $\mathbf{R}^n$: the basis-independent notation is more elegant and often clearer, and it encourages techniques which extend beyond finite dimensions. But more concretely, identifying $\mathbf{E}$ with $\mathbf{R}^n$ may obscure properties of a space beyond its simple Euclidean structure. As an example, in this short section we describe a Euclidean space which 'feels' very different from $\mathbf{R}^n$: the space $\mathbf{S}^n$ of $n \times n$ real symmetric matrices.

The nonnegative orthant $\mathbf{R}^n_+$ is a cone in $\mathbf{R}^n$ which plays a central role in our development. In a variety of contexts the analogous role in $\mathbf{S}^n$ is played by the cone of positive semidefinite matrices, $\mathbf{S}^n_+$. These two cones have some important differences: in particular, $\mathbf{R}^n_+$ is a polyhedron whereas the cone of positive semidefinite matrices $\mathbf{S}^n_+$ is not, even for $n = 2$. The cones $\mathbf{R}^n_+$ and $\mathbf{S}^n_+$ are important largely because of the orderings they induce. For points x and y in $\mathbf{R}^n$ we write $x \leq y$ if $y - x \in \mathbf{R}^n_+$, and $x < y$ if $y - x \in \mathbf{R}^n_{++}$ (with analogous definitions for $\geq$ and $>$). The cone $\mathbf{R}^n_+$ is a *lattice cone*: for any points x and y in $\mathbf{R}^n$ there is a point z satisfying

$$w \geq x \text{ and } w \geq y \;\Leftrightarrow\; w \geq z.$$

(The point z is just the componentwise maximum of x and y.) Analogously, for matrices X and Y in $\mathbf{S}^n$ we write $X \preceq Y$ if $Y - X \in \mathbf{S}^n_+$, and $X \prec Y$ if $Y - X$ lies in $\mathbf{S}^n_{++}$, the set of positive definite matrices (with analogous definitions for $\succeq$ and $\succ$). By contrast, $\mathbf{S}^n_+$ is *not* a lattice cone (see Exercise 4).

We denote the identity matrix by I. The *trace* of a square matrix Z is the sum of the diagonal entries, written $\operatorname{tr} Z$. It has the important property $\operatorname{tr}(VW) = \operatorname{tr}(WV)$ for any matrices V and W for which VW is well-defined and square. We make the vector space $\mathbf{S}^n$ into a Euclidean space by defining the inner product

$$\langle X, Y \rangle = \operatorname{tr}(XY), \quad \text{for } X, Y \in \mathbf{S}^n.$$

Any matrix X in $\mathbf{S}^n$ has n real eigenvalues (counted by multiplicity), which we write in nonincreasing order $\lambda_1(X) \geq \lambda_2(X) \geq \ldots \geq \lambda_n(X)$. In this way we define a function $\lambda : \mathbf{S}^n \to \mathbf{R}^n$. We also define a linear map $\operatorname{Diag} : \mathbf{R}^n \to \mathbf{S}^n$, where for a vector x in $\mathbf{R}^n$, $\operatorname{Diag} x$ is an $n \times n$ diagonal matrix with diagonal entries x_i. This map embeds $\mathbf{R}^n$ as a subspace of $\mathbf{S}^n$ and the cone $\mathbf{R}^n_+$ as a subcone of $\mathbf{S}^n_+$. The determinant of a square matrix Z is written $\det Z$.

We write $\mathbf{O}^n$ for the group of $n \times n$ *orthogonal* matrices (those matrices U satisfying $U^T U = I$). Then any matrix X in $\mathbf{S}^n$ has an *ordered spectral decomposition* $X = U^T(\mathrm{Diag}\,\lambda(X))U$, for some matrix U in $\mathbf{O}^n$. This shows, for example, that the function λ is *norm-preserving*: $\|X\| = \|\lambda(X)\|$ for all X in $\mathbf{S}^n$. For any X in $\mathbf{S}^n_+$, the spectral decomposition also shows there is a unique matrix $X^{1/2}$ in $\mathbf{S}^n_+$ whose square is X.

The Cauchy-Schwarz inequality has an interesting refinement in $\mathbf{S}^n$ which is crucial for variational properties of eigenvalues, as we shall see.

Theorem 1.2.1 (Fan) *Any matrices X and Y in $\mathbf{S}^n$ satisfy the inequality*

$$\mathrm{tr}\,(XY) \leq \lambda(X)^T\lambda(Y). \tag{1.2.2}$$

Equality holds if and only if X and Y have a **simultaneous ordered spectral decomposition***: there is a matrix U in $\mathbf{O}^n$ with*

$$X = U^T(\mathrm{Diag}\,\lambda(X))U \quad \textit{and} \quad Y = U^T(\mathrm{Diag}\,\lambda(Y))U.i \tag{1.2.3}$$

A standard result in linear algebra states that matrices X and Y have a simultaneous (*unordered*) spectral decomposition if and only if they commute. Notice condition (1.2.3) is a stronger property.

The special case of Fan's inequality where both matrices are diagonal gives the following classical inequality.

For a vector x in $\mathbf{R}^n$, we denote by $[x]$ the vector with the same components permuted into nondecreasing order. We leave the proof of this result as an exercise.

Proposition 1.2.3 (Hardy-Littlewood-Polya) *Any vectors x and y in $\mathbf{R}^n$ satisfy the inequality*

$$x^T y \leq [x]^T [y].$$

We describe a proof of Fan's Theorem in the exercises, using the above proposition and the following classical relationship between the set $\mathbf{\Gamma}^n$ of *doubly stochastic* matrices (square matrices with all nonnegative entries, and each row and column summing to 1) and the set $\mathbf{P}^n$ of *permutation* matrices (square matrices with all entries 0 or 1, and with exactly one entry 1 in each row and in each column).

Theorem 1.2.4 (Birkhoff) *Any doubly stochastic matrix is a convex combination of permutation matrices.*

We defer the proof to a later section (§4.1, Exercise 21).

Chapter 2

Inequality constraints

2.1 Optimality conditions

Early in multivariate calculus we learn the significance of differentiability in finding minimizers. In this section we begin our study of the interplay between convexity and differentiability in optimality conditions.

For an initial example, consider the problem of minimizing a function $f : C \to \mathbf{R}$ on a set C in $\mathbf{E}$. We say a point $\bar{x}$ in C is a *local minimizer of f on C* if $f(x) \geq f(\bar{x})$ for all points x in C close to $\bar{x}$. The *directional derivative* of a function f at $\bar{x}$ in a direction $d \in \mathbf{E}$ is

$$f'(\bar{x}; d) = \lim_{t \downarrow 0} \frac{f(\bar{x} + td) - f(\bar{x})}{t},$$

when this limit exists. When the directional derivative $f'(\bar{x}; d)$ is actually linear in d (that is, $f'(\bar{x}; d) = \langle a, d \rangle$

for some element a of $\mathbf{E}$) then we say f is *(Gâteaux) differentiable* at $\bar{x}$, with *(Gâteaux) derivative* $\nabla f(\bar{x}) = a$. If f is differentiable at every point in C then we simply say f is differentiable (on C). An example we use quite extensively is the function $X \in \mathbf{S}^n_{++} \mapsto \log\det X$: an exercise shows this function is differentiable on $\mathbf{S}^n_{++}$ with derivative X^{-1}.

A convex cone which arises frequently in optimization is the *normal cone* to a convex set C at a point $\bar{x} \in C$, written $N_C(\bar{x})$. This is the convex cone of *normal vectors*: vectors d in $\mathbf{E}$ such that $\langle d, x - \bar{x}\rangle \leq 0$ for all points x in C.

Proposition 2.1.1 (First order necessary condition) *Suppose that C is a convex set in $\mathbf{E}$, and that the point $\bar{x}$ is a local minimizer of the function $f : C \to \mathbf{R}$. Then for any point x in C, the directional derivative, if it exists, satisfies $f'(\bar{x}; x - \bar{x}) \geq 0$. In particular, if f is differentiable at $\bar{x}$ then the condition $-\nabla f(\bar{x}) \in N_C(\bar{x})$ holds.*

Proof. If some point x in C satisfies $f'(\bar{x}; x - \bar{x}) < 0$ then all small real $t > 0$ satisfy $f(\bar{x} + t(x - \bar{x})) < f(\bar{x})$, contradicting the local minimality of $\bar{x}$. ♠

The case of this result where C is an open set is the canonical introduction to the use of calculus in optimization: local minimizers $\bar{x}$ must be *critical points* (that is, $\nabla f(\bar{x}) = 0$). This book is largely devoted to the study of first order necessary conditions for a local minimizer of a function subject to constraints. In that case local minimizers $\bar{x}$ may not lie in the interior of the set C of interest, so the normal cone $N_C(\bar{x})$ is not simply $\{0\}$.

The next result shows that when f is convex the first order condition above is *sufficient* for $\bar{x}$ to be a global minimizer of f on C. The proof is outlined in Exercise 4).

Proposition 2.1.2 (First order sufficient condition) *Suppose that the set $C \subset \mathbf{E}$ is convex and that the function $f : C \to \mathbf{R}$ is convex. Then for any points $\bar{x}$ and x in C, the directional derivative $f'(\bar{x}; x - \bar{x})$ exists in $[-\infty, +\infty)$. If the condition $f'(\bar{x}; x - \bar{x}) \geq 0$ holds for all x in C, or in particular if the condition $-\nabla f(\bar{x}) \in N_C(\bar{x})$ holds, then $\bar{x}$ is a global minimizer of f on C.*

Proof. A straightforward exercise using the convexity of f shows the function

$$t \in (0, 1] \mapsto \frac{f(\bar{x} + t(x - \bar{x})) - f(\bar{x})}{t}$$

is nondecreasing. The result then follows easily (Exercise 7). ♠

In particular, any critical point of a convex function is a global minimizer.

The following useful result illustrates what the first order conditions become for a more concrete optimization problem.

Corollary 2.1.3 (First order conditions for linear constraints) *Given a convex set $C \subset \mathbf{E}$, a function $f : C \to \mathbf{R}$, a linear map $A : \mathbf{E} \to \mathbf{Y}$ (where $\mathbf{Y}$ is a Euclidean space) and a point b in $\mathbf{Y}$, consider the optimization problem*

$$\inf\{f(x) \mid x \in C,\ Ax = b\}. \tag{2.1.4}$$

Suppose the point $\bar{x} \in \operatorname{int} C$ satisfies $A\bar{x} = b$.

(a) If $\bar{x}$ is a local minimizer for the problem (2.1.4) and f is differentiable at $\bar{x}$ then $\nabla f(\bar{x}) \in A^*\mathbf{Y}$.

(b) Conversely, if $\nabla f(\bar{x}) \in A^*\mathbf{Y}$ and f is convex then $\bar{x}$ is a global minimizer for (2.1.4).

The element $y \in \mathbf{Y}$ satisfying $\nabla f(\bar{x}) = A^*y$ in the above result is called a *Lagrange multiplier*. This kind of construction recurs in many different forms in our development.

In the absence of convexity, we need second order information to tell us more about minimizers. The following elementary result from multivariate calculus is typical.

Theorem 2.1.5 (Second order conditions) *Suppose the twice continuously differentiable function* $f : \mathbf{R}^n \to \mathbf{R}$ *has a critical point* $\bar{x}$*. If* $\bar{x}$ *is a local minimizer then the Hessian* $\nabla^2 f(\bar{x})$ *is positive semidefinite. Conversely, if the Hessian is positive definite then* $\bar{x}$ *is a local minimizer.*

(In fact for $\bar{x}$ to be a local minimizer it is sufficient for the Hessian to be positive semidefinite locally: the function $x \in \mathbf{R} \mapsto x^4$ highlights the distinction.)

To illustrate the effect of constraints on second order conditions, consider the framework of Corollary 2.1.3 (First order conditions for linear constraints) in the case $\mathbf{E} = \mathbf{R}^n$, and suppose $\nabla f(\bar{x}) \in A^*\mathbf{Y}$ and f is twice continuously differentiable near $\bar{x}$. If $\bar{x}$ is a local minimizer then $y^T \nabla^2 f(\bar{x}) y \geq 0$ for all vectors y in $N(A)$. Conversely, if $y^T \nabla^2 f(\bar{x}) y > 0$ for all nonzero y in $N(A)$ then $\bar{x}$ is a local minimizer.

We are already beginning to see the broad interplay between analytic, geometric and topological ideas in optimization theory. A good illustration is the separation result of §1.1, which we now prove.

Theorem 2.1.6 (Basic separation) *Suppose that the set $C \subset \mathbf{E}$ is closed and convex, and that the point y does not lie in C. Then there exist a real b and a nonzero element a of $\mathbf{E}$ such that $\langle a, y\rangle > b \geq \langle a, x\rangle$ for all points x in C.*

Proof. We may assume C is nonempty, and define a function $f : \mathbf{E} \to \mathbf{R}$ by $f(x) = \|x - y\|^2/2$. Now by the Weierstrass Proposition (1.1.3) there exists a minimizer $\bar{x}$ for f on C, which by the First order necessary condition (2.1.1) satisfies $-\nabla f(\bar{x}) = y - \bar{x} \in N_C(\bar{x})$. Thus $\langle y - \bar{x}, x - \bar{x}\rangle \leq 0$ holds for all points x in C. Now setting $a = y - \bar{x}$ and $b = \langle y - \bar{x}, \bar{x}\rangle$ gives the result. ♠

We end this section with a rather less standard result, illustrating another idea which is important later: the use of 'variational principles' to treat problems where minimizers may not exist, but which nonetheless have 'approximate' critical points. This result is a precursor of a principle due to Ekeland, which we develop in §6.1.

Proposition 2.1.7 *If the function $f : \mathbf{E} \to \mathbf{R}$ is differentiable and bounded below then there are points where f has small derivative.*

Proof. Fix any real $\epsilon > 0$. The function $f + \epsilon\|\cdot\|$ has bounded level sets, so has a global minimizer x^ϵ by the

Weierstrass Proposition (1.1.3). If the vector $d = \nabla f(x^\epsilon)$ satisfies $\|d\| > \epsilon$ then from the inequality

$$\begin{aligned}\lim_{t \downarrow 0} \frac{f(x^\epsilon - td) - f(x^\epsilon)}{t} &= -\langle \nabla f(x^\epsilon), d \rangle \\ &= -\|d\|^2 < -\epsilon \|d\|,\end{aligned}$$

we would have, for small $t > 0$, the contradiction

$$\begin{aligned}-t\epsilon\|d\| &> f(x^\epsilon - td) - f(x^\epsilon) \\ &= \left(f(x^\epsilon - td) + \epsilon\|x^\epsilon - td\|\right) \\ &\quad - \left(f(x^\epsilon) + \epsilon\|x^\epsilon\|\right) + \epsilon(\|x^\epsilon\| - \|x^\epsilon - td\|) \\ &\geq -\epsilon t \|d\|,\end{aligned}$$

by definition of x^ϵ, and the triangle inequality. Hence $\|\nabla f(x^\epsilon)\| \leq \epsilon$. ♠

Notice that the proof relies on consideration of a *non-differentiable* function, even though the result concerns derivatives.

2.2 Theorems of the alternative

One well-trodden route to the study of first order conditions uses a class of results called 'theorems of the alternative', and in particular the Farkas Lemma (which we derive at the end of this section). Our first approach, however, relies on a different theorem of the alternative.

Theorem 2.2.1 (Gordan) *For any elements* $a^0, a^1, \ldots, a^m$ *of* $\mathbf{E}$*, exactly one of the following systems has a solution:*

$$\sum_{i=0}^{m} \lambda_i a^i \;=\; 0, \quad \sum_{i=0}^{m} \lambda_i = 1, \quad 0 \le \lambda_0, \lambda_1, \ldots, \lambda_m \in \mathbf{R}; \tag{2.2.2}$$

$$\langle a^i, x\rangle \;<\; 0 \;\; \textit{for } i = 0, 1, \ldots, m, \quad x \in \mathbf{E}. \tag{2.2.3}$$

Geometrically, Gordan's Theorem says that 0 does not lie in the convex hull of the set $\{a^0, a^1, \ldots, a^m\}$ if and only if there is an open halfspace $\{y \mid \langle y, x\rangle < 0\}$ containing $\{a^0, a^1, \ldots, a^m\}$ (and hence its convex hull). This is another illustration of the idea of separation (in this case we separate 0 and the convex hull).

Theorems of the alternative like Gordan's Theorem may be proved in a variety of ways, including separation and algorithmic approaches. We employ a less standard technique, using our earlier analytic ideas, and leading to a rather unified treatment. It relies on the relationship between the optimization problem

$$\inf\{f(x) \mid x \in \mathbf{E}\}, \tag{2.2.4}$$

where the function f is defined by

$$f(x) = \log\left(\sum_{i=0}^{m} \exp\langle a^i, x\rangle\right), \tag{2.2.5}$$

and the two systems (2.2.2) and (2.2.3). We return to the surprising function (2.2.5) when we discuss conjugacy in §3.3.

Theorem 2.2.6 *The following statements are equivalent:*

(i) The function defined by (2.2.5) is bounded below.

(ii) System (2.2.2) is solvable.

(iii) System (2.2.3) is unsolvable.

Proof. The implications (ii) $\Rightarrow$ (iii) $\Rightarrow$ (i) are easy exercises, so it remains to show (i) $\Rightarrow$ (ii). To see this we apply Proposition 2.1.7. We deduce that for each $k = 1, 2, \ldots,$ there is a point x^k in $\mathbf{E}$ satisfying

$$\|\nabla f(x^k)\| = \left\|\sum_{i=0}^{m} \lambda_i^k a^i\right\| < 1/k,$$

where the scalars

$$\lambda_i^k = \frac{\exp\langle a^i, x^k\rangle}{\Sigma_{r=0}^m \exp\langle a^r, x^k\rangle} > 0$$

satisfy $\Sigma_{i=0}^m \lambda_i^k = 1$. Now the limit λ of any convergent subsequence of the the bounded sequence (λ^k) solves system (2.2.2). ♠

The equivalence of (ii) and (iii) now gives Gordan's Theorem.

We now proceed by using Gordan's Theorem to derive the Farkas Lemma, one of the cornerstones of many approaches to optimality conditions. The proof uses the idea of the *projection* onto a vector subspace $\mathbf{Y}$ of $\mathbf{E}$. Notice first that $\mathbf{Y}$ becomes a Euclidean space by equipping it with the same inner product. The projection of a point x in $\mathbf{E}$ onto $\mathbf{Y}$, written $P_{\mathbf{Y}}x$, is simply the nearest point to x in $\mathbf{Y}$. This is well-defined (see Exercise 8 in §2.1), and is characterized by the fact that $x - P_{\mathbf{Y}}x$ is orthogonal to $\mathbf{Y}$. A standard exercise shows $P_{\mathbf{Y}}$ is a linear map.

Lemma 2.2.7 (Farkas) *For any points $a^1, a^2, \ldots, a^m$ and c in $\mathbf{E}$, exactly one of the following systems has a solution:*

$$\sum_{i=1}^{m} \mu_i a^i = c, \quad 0 \le \mu_1, \mu_2, \ldots, \mu_m \in \mathbf{R}; \tag{2.2.8}$$

$$\langle a^i, x \rangle \le 0 \ \ \textit{for} \ \ i = 1, 2, \ldots, m, \quad \langle c, x \rangle > 0, \quad x \in \mathbf{E}. \tag{2.2.9}$$

Proof. Again, it is immediate that if system (2.2.8) has a solution then system (2.2.9) has no solution. Conversely, we assume (2.2.9) has no solution, and deduce that (2.2.8) has a solution by using induction on the number of elements m. The result is clear for $m = 0$.

Suppose then that the result holds in any Euclidean space and for any set of $m - 1$ elements and any element c. Define $a^0 = -c$. Applying Gordan's Theorem (2.2.1) to the unsolvability of (2.2.9) shows there are scalars $\lambda_0, \lambda_1, \ldots, \lambda_m \ge 0$ in $\mathbf{R}$, not all zero, satisfying $\lambda_0 c = \sum_1^m \lambda_i a^i$. If $\lambda_0 > 0$ the proof is complete, so suppose $\lambda_0 = 0$ and without loss of generality $\lambda_m > 0$.

Define a subspace of $\mathbf{E}$ by $\mathbf{Y} = \{y \mid \langle a^m, y \rangle = 0\}$, so by assumption the system

$$\langle a^i, y \rangle \le 0 \ \text{ for } i = 1, 2, \ldots, m-1, \quad \langle c, y \rangle > 0, \quad y \in \mathbf{Y},$$

or equivalently

$$\langle P_{\mathbf{Y}} a^i, y \rangle \le 0 \qquad \text{for } i = 1, 2, \ldots, m-1, \langle P_{\mathbf{Y}} c, y \rangle > 0, \quad y \in \mathbf{Y}, \tag{2.2.10}$$

has no solution.

By the induction hypothesis applied to the subspace $\mathbf{Y}$, there are nonnegative reals $\mu_1, \mu_2, \ldots, \mu_{m-1}$ satisfying $\Sigma_{i=1}^{m-1} \mu_i P_{\mathbf{Y}} a^i = P_{\mathbf{Y}} c$, so the vector $c - \Sigma_1^{m-1} \mu_i a^i$ is orthogonal to the subspace $\mathbf{Y} = (\mathrm{span}\,(a^m))^{\perp}$. Thus some real μ_m satisfies

$$(2.2.10) \qquad \mu_m a^m = c - \sum_1^{m-1} \mu_i a^i.$$

If μ_m is nonnegative we immediately obtain a solution of (2.2.8), and if not then we can substitute $a^m = -\lambda_m^{-1} \Sigma_1^{m-1} \lambda_i a^i$ in equation (2.2.10) to obtain a solution. ♠

Just like Gordan's Theorem, the Farkas Lemma has an important geometric interpretation which gives an alternative approach to its proof (Exercise 6): any point c not lying in the *finitely generated cone*

$$(2.2.11) \quad C = \left\{ \sum_1^m \mu_i a^i \mid 0 \le \mu_1, \mu_2, \ldots, \mu_m \in \mathbf{R} \right\}$$

can be separated from C by a hyperplane. If x solves system (2.2.9) then C is contained in the closed halfspace $\{a \mid \langle a, x\rangle \le 0\}$, whereas c is contained in the complementary open halfspace. In particular, it follows that any finitely generated cone is closed.

2.3 Max-functions and first order conditions

This section is an elementary exposition of the first order necessary conditions for a local minimizer of a differentiable function subject to differentiable inequality constraints. Throughout this section we use the term 'differentiable' in the Gâteaux sense, defined in §2.1. Our approach, which relies on considering the local minimizers of a 'max-function'

$$g(x) = \max_{i=0,1,\ldots,m} \{g_i(x)\}, \tag{2.3.1}$$

illustrates a pervasive analytic idea in optimization: *non-smoothness*. Even if the functions $g_0, g_1, \ldots, g_m$ are smooth, g may not be, and hence the gradient may no longer be a useful notion.

Proposition 2.3.2 (Directional derivatives of max-functions) *Let $\bar{x}$ be a point in the interior of a set $C \subset \mathbf{E}$. Suppose that continuous functions $g_0, g_1, \ldots, g_m : C \to \mathbf{R}$ are differentiable at $\bar{x}$, that g is the max-function (2.3.1), and define the index set $K = \{i \mid g_i(\bar{x}) = g(\bar{x})\}$. Then for all directions d in $\mathbf{E}$, the directional derivative of g is given by*

$$g'(\bar{x}; d) = \max_{i \in K} \{\langle \nabla g_i(\bar{x}), d \rangle\}. \tag{2.3.3}$$

Proof. By continuity we can assume, without loss of generality, $K = \{0, 1, \ldots, m\}$: those g_i not attaining the maximum in (2.3.1) will not affect $g'(\bar{x}; d)$. Now for each i, we have the inequality

$$\begin{aligned}\liminf_{t \downarrow 0} \frac{g(\bar{x} + td) - g(\bar{x})}{t} &\geq \lim_{t \downarrow 0} \frac{g_i(\bar{x} + td) - g_i(\bar{x})}{t} \\ &= \langle \nabla g_i(\bar{x}), d \rangle.\end{aligned}$$

Suppose

$$\limsup_{t \downarrow 0} \frac{g(\bar{x} + td) - g(\bar{x})}{t} > \max_i \{\langle \nabla g_i(\bar{x}), d \rangle\}.$$

Then some real sequence $t_k \downarrow 0$ and real $\epsilon > 0$ satisfy

$$\begin{aligned}\frac{g(\bar{x} + t_k d) - g(\bar{x})}{t_k} &\geq \max_i \{\langle \nabla g_i(\bar{x}), d \rangle\} + \epsilon, \\ &\quad \text{for all } k \in \mathbf{N}\end{aligned}$$

(where $\mathbf{N}$ denotes the sequence of natural numbers). We can now choose a subsequence R of $\mathbf{N}$ and a fixed index j so that all integers k in R satisfy $g(\bar{x}+t_k d) = g_j(\bar{x}+t_k d)$. In the limit we obtain the contradiction

$$\langle \nabla g_j(\bar{x}), d \rangle \geq \max_i \{\langle \nabla g_i(\bar{x}), d \rangle\} + \epsilon.$$

Hence

$$\limsup_{t \downarrow 0} \frac{g(\bar{x} + td) - g(\bar{x})}{t} \leq \max_i \{\langle \nabla g_i(\bar{x}), d \rangle\},$$

and the result follows. ♠

For most of this book we consider optimization problems of the form

$$(2.3.4)\quad \begin{cases} \inf & f(x) \\ \text{subject to} & g_i(x) \;\le\; 0, \quad \text{for } i \in I, \\ & h_j(x) \;=\; 0, \quad \text{for } j \in J, \\ & \quad\; x \;\in\; C, \end{cases}$$

where C is a subset of $\mathbf{E}$, I and J are finite index sets, and the *objective function* f and *inequality* and *equality constraint functions* g_i $(i \in I)$ and h_j $(j \in J)$ respectively are continuous from C to $\mathbf{R}$. A point x in C is *feasible* if it satisfies the constraints, and the set of all feasible x is called the *feasible region*. If the problem has no feasible points, we call it *inconsistent*. We say a feasible point $\bar{x}$ is a *local minimizer* if $f(x) \ge f(\bar{x})$ for all feasible x close to $\bar{x}$. We aim to derive first order necessary conditions for local minimizers.

We begin in this section with the differentiable, inequality constrained problem

$$(2.3.5)\begin{cases} \inf & f(x) \\ \text{subject to} & g_i(x) \;\le\; 0 \;\text{ for } i = 1, 2, \ldots, m, \\ & \quad\; x \;\in\; C. \end{cases}$$

For a feasible point $\bar{x}$ we define the *active set* $I(\bar{x}) = \{i \mid g_i(\bar{x}) = 0\}$. For this problem, assuming $\bar{x} \in \operatorname{int} C$, we call a vector $\lambda \in \mathbf{R}^m_+$ a *Lagrange multiplier vector* for $\bar{x}$ if $\bar{x}$ is a critical point of the *Lagrangian*

$$L(x;\lambda) = f(x) + \sum_{i=1}^{m} \lambda_i g_i(x)$$

(in other words, $\nabla f(\bar{x}) + \sum \lambda_i \nabla g_i(\bar{x}) = 0$) and *complementary slackness* holds: $\lambda_i = 0$ for indices i not in $I(\bar{x})$.

Theorem 2.3.6 (Fritz John conditions) *Suppose problem (2.3.5) has a local minimizer $\bar{x} \in \operatorname{int} C$. If the functions f, g_i $(i \in I(\bar{x}))$ are differentiable at $\bar{x}$ then there exist $\lambda_0, \lambda_i \in \mathbf{R}_+$, $(i \in I(\bar{x}))$, not all zero, satisfying*

$$\lambda_0 \nabla f(\bar{x}) + \sum_{i \in I(\bar{x})} \lambda_i \nabla g_i(\bar{x}) = 0.$$

Proof. Consider the function

$$g(x) = \max\{f(x) - f(\bar{x}),\ g_i(x)\ (i \in I(\bar{x}))\}.$$

Since $\bar{x}$ is a local minimizer for the problem (2.3.5), it is a local minimizer of the function g, so all directions $d \in \mathbf{E}$ satisfy the inequality

$$g'(\bar{x}; d) = \max\{\langle \nabla f(\bar{x}), d\rangle, \langle \nabla g_i(\bar{x}), d\rangle\ \ (i \in I(\bar{x}))\} \geq 0,$$

by the First order necessary conditions (2.1.1) and Proposition 2.3.2 (Directional derivatives of max-functions). Thus the system

$$\langle \nabla f(\bar{x}), d \rangle < 0, \quad \langle \nabla g_i(\bar{x}), d \rangle < 0 \quad (i \in I(\bar{x}))$$

has no solution, and the result follows by Gordan's Theorem (2.2.1). ♠

One obvious disadvantage remains with the Fritz John first order conditions above: if $\lambda_0 = 0$ then the conditions are independent of the objective function f. To rule out this possibility we need to impose a regularity condition or 'constraint qualification', an approach which is another recurring theme. The easiest such condition in this context is simply the linear independence of the gradients of the active constraints $\{\nabla g_i(\bar{x}) \mid i \in I(\bar{x})\}$. The culminating result of this section uses the following weaker condition.

Assumption 2.3.7 (The Mangasarian-Fromowitz constraint qualification) *There is a direction d in $\mathbf{E}$ satisfying $\langle \nabla g_i(\bar{x}), d \rangle < 0$ for all indices i in the active set $I(\bar{x})$.*

Theorem 2.3.8 (Karush-Kuhn-Tucker conditions) *Suppose the problem (2.3.5) has a local minimizer $\bar{x}$*

in $\operatorname{int} C$. *If the functions* f, g_i *(for* $i \in I(\bar{x})$*) are differentiable at* $\bar{x}$*, and if the Mangasarian-Fromowitz constraint qualification (2.3.7) holds, then there is a Lagrange multiplier vector for* $\bar{x}$.

Proof. By the trivial implication in Gordan's Theorem (2.2.1), the constraint qualification ensures $\lambda_0 \neq 0$ in the Fritz John conditions (2.3.6). ♠

Chapter 3

Fenchel duality

3.1 Subgradients and convex functions

We have already seen, in the First order sufficient conditions (2.1.2), one benefit of convexity in optimization: critical points of convex functions are global minimizers. In this section we extend the types of functions we consider in two important ways:

(i) We do not require f to be differentiable;

(ii) We allow f to take the value $+\infty$.

Our derivation of first order conditions in §2.3 illustrates the utility of considering nonsmooth functions even in the context of smooth problems. Allowing the value $+\infty$ lets us rephrase a problem like $\inf\{g(x) \mid x \in C\}$ as $\inf g + \delta_C$, where the *indicator function* $\delta_C(x)$ is 0 for x in C and $+\infty$ otherwise.

The *domain* of a function $f : \mathbf{E} \to (-\infty, +\infty]$ is the set

$$\operatorname{dom} f = \{x \in \mathbf{E} \mid f(x) < +\infty\}.$$

We say f is *convex* if it is convex on its domain, and *proper* if its domain is nonempty. We call a function $g : \mathbf{E} \to [-\infty, +\infty)$ *concave* if $-g$ is convex, although for reasons of simplicity we will consider primarily convex functions. If a convex function f satisfies the stronger condition

$$f(\lambda x + \mu y) \leq \lambda f(x) + \mu f(y) \text{ for all } x, y \in \mathbf{E},\ \lambda, \mu \in \mathbf{R}_+$$

we say f is *sublinear*. If $f(\lambda x) = \lambda f(x)$ for all x in $\mathbf{E}$ and λ in $\mathbf{R}_+$ then f is *positively homogeneous*: in particular this implies $f(0) = 0$. (Recall the convention $0 \cdot (+\infty)) = 0$.) If $f(x+y) \leq f(x) + f(y)$ for all x and y in $\mathbf{E}$ then we say f is *subadditive*. It is immediate that if the function f is sublinear then $-f(x) \leq f(-x)$ for all x in $\mathbf{E}$. The *lineality space* of a sublinear function f is the set

$$\operatorname{lin} f = \{x \in \mathbf{E} \mid -f(x) = f(-x)\}.$$

The following result (left as an exercise) shows this set is a subspace.

Proposition 3.1.1 (Sublinearity) *A function* $f : \mathbf{E} \to (-\infty, +\infty]$ *is sublinear if and only if it is positively homogeneous and subadditive. For a sublinear function* f*, the lineality space* $\operatorname{lin} f$ *is the largest subspace of* $\mathbf{E}$ *on which* f *is linear.*

As in the First order sufficient condition (2.1.2), it is easy to check that if the point $\bar{x}$ lies in the domain of the convex function f then the directional derivative $f'(\bar{x}; \cdot)$ is well-defined and positively homogeneous, taking values in $[-\infty, +\infty]$. The *core* of a set C (written $\operatorname{core}(C)$) is the set of points x in C such that for any direction d in $\mathbf{E}$, $x + td$ lies in C for all small real t. This set clearly contains the interior of C, although it may be larger (Exercise 2).

Proposition 3.1.2 (Sublinearity of the directional derivative) *If the function* $f : \mathbf{E} \to (-\infty, +\infty]$ *is convex then for any point* $\bar{x}$ *in* $\operatorname{core}(\operatorname{dom} f)$ *the directional derivative* $f'(\bar{x}; \cdot)$ *is everywhere finite and sublinear.*

Proof. For d in $\mathbf{E}$ and nonzero t in $\mathbf{R}$, define

$$g(d; t) = \frac{f(\bar{x} + td) - f(\bar{x})}{t}.$$

By convexity we deduce, for $0 < t \leq s \in \mathbf{R}$, the inequality

$$g(d; -s) \leq g(d; -t) \leq g(d; t) \leq g(d; s).$$

Since $\bar{x}$ lies in core (dom f), for small $s > 0$ both $g(d; -s)$ and $g(d; s)$ are finite, so as $t \downarrow 0$ we have

$$+\infty > g(d; s) \geq g(d; t) \downarrow f'(\bar{x}; d) \geq g(d; -s) > -\infty. \tag{3.1.3}$$

Again by convexity we have, for any directions d and e in $\mathbf{E}$ and real $t > 0$,

$$g(d + e; t) \leq g(d; 2t) + g(e; 2t).$$

Now letting $t \downarrow 0$ gives subadditivity of $f'(\bar{x}; \cdot)$. The positive homogeneity is easy to check. ♠

The idea of the derivative is fundamental in analysis because it allows us to approximate a wide class of functions using *linear functions*. In optimization we are concerned specifically with the minimization of functions, and hence often a *one-sided approximation* is sufficient. In place of the gradient we therefore consider *subgradients*: those elements ϕ of $\mathbf{E}$ satisfying

$$\langle \phi, x - \bar{x} \rangle \ \leq \ f(x) - f(\bar{x}), \quad \text{for all points } x \text{ in } \mathbf{E}. \tag{3.1.4}$$

We denote the set of subgradients (called the *subdifferential*) by $\partial f(\bar{x})$, defining $\partial f(\bar{x}) = \emptyset$ for $\bar{x}$ not in dom f. The subdifferential is always a closed convex set. We can think of $\partial f(\bar{x})$ as the value at $\bar{x}$ of the 'multifunction' or 'set-valued mapping' $\partial f : \mathbf{E} \to \mathbf{E}$. The importance of such mappings is another of our themes: we define its *domain*

$$\operatorname{dom} \partial f = \{x \in \mathbf{E} \mid \partial f(x) \neq \emptyset\}$$

(See Exercise 19.) We say f is *essentially strictly convex* if it is strictly convex on any convex subset of $\operatorname{dom} \partial f$.

The following very easy observation suggests the fundamental significance of subgradients in optimization.

Proposition 3.1.4 (Subgradients at optimality) *For any proper function $f : \mathbf{E} \to (-\infty, +\infty]$, the point $\bar{x}$ is a (global) minimizer of f if and only if the condition $0 \in \partial f(\bar{x})$ holds.*

Alternatively put, minimizers of f correspond exactly to 'zeroes' of ∂f.

The derivative is a local property whereas the subgradient definition (3.1.4) describes a global property. The main result of this section shows that the set of subgradients of a convex function is usually *nonempty*, and that we can describe it locally in terms of the directional derivative. We begin with another simple exercise.

Proposition 3.1.5 (Subgradients and directional derivatives) *If the function $f : \mathbf{E} \to (-\infty, +\infty]$ is convex and the point $\bar{x}$ lies in* $\operatorname{dom} f$, *then an element*

ϕ *of* $\mathbf{E}$ *is a subgradient of* f *at* $\bar{x}$ *if and only if it satisfies* $\langle \phi, \cdot \rangle \leq f'(\bar{x}; \cdot)$.

The idea behind the construction of a subgradient for a function f that we present here is rather simple. We recursively construct a decreasing sequence of sublinear functions which, after translation, minorize f. At each step we guarantee one extra direction of linearity. The basic step is summarized in the following exercise.

Lemma 3.1.6 *Suppose that the function* $p : \mathbf{E} \to (-\infty, +\infty]$ *is sublinear, and that the point* $\bar{x}$ *lies in* $\mathrm{core}\,(\mathrm{dom}\, p)$. *Then the function* $q(\cdot) = p'(\bar{x}; \cdot)$ *satisfies the conditions*

(i) $q(\lambda \bar{x}) = \lambda p(\bar{x})$ *for all real* λ,

(ii) $q \leq p$, *and*

(iii) $\mathrm{lin}\, q \supset \mathrm{lin}\, p + \mathrm{span}\, \{\bar{x}\}$.

With this tool we are now ready for the main result, giving conditions guaranteeing the existence of a subgradient. Proposition 3.1.5 showed how to identify subgradients from directional derivatives: this next result shows how to move in the reverse direction.

Theorem 3.1.7 (Max Formula) *If the function* $f : \mathbf{E} \to (-\infty, +\infty]$ *is convex then any point* $\bar{x}$ *in* core (dom f) *and any direction* d *in* $\mathbf{E}$ *satisfy*

$$f'(\bar{x}; d) = \max\{\langle \phi, d\rangle \mid \phi \in \partial f(\bar{x})\}. \tag{3.1.8}$$

In particular, the subdifferential $\partial f(\bar{x})$ *is nonempty.*

Proof. In view of Proposition 3.1.5, we simply have to show that for any fixed d in $\mathbf{E}$ there is a subgradient ϕ satisfying $\langle \phi, d\rangle = f'(\bar{x}; d)$. Choose a basis $\{e_1, e_2, \ldots, e_n\}$ for $\mathbf{E}$ with $e_1 = d$ if d is nonzero. Now define a sequence of functions $p_0, p_1, \ldots, p_n$ recursively by $p_0(\cdot) = f'(\bar{x}; \cdot)$, and $p_k(\cdot) = p'_{k-1}(e_k; \cdot)$, for $k = 1, 2, \ldots, n$. We essentially show that $p_n(\cdot)$ is the required subgradient.

First note that, by Proposition 3.1.2, each p_k is everywhere finite and sublinear. By part (iii) of Lemma 3.1.6 we know

$$\operatorname{lin} p_k \supset \operatorname{lin} p_{k-1} + \operatorname{span}\{e_k\}, \quad \text{for } k = 1, 2, \ldots, n,$$

so p_n is linear. Thus there is an element ϕ of $\mathbf{E}$ satisfying $\langle \phi, \cdot\rangle = p_n(\cdot)$.

Part (ii) of Lemma 3.1.6 implies $p_n \leq p_{n-1} \leq \ldots \leq p_0$, so certainly, by Proposition 3.1.5, any point x in $\mathbf{E}$ satisfies

$$p_n(x - \bar{x}) \leq p_0(x - \bar{x}) = f'(\bar{x}; x - \bar{x}) \leq f(x) - f(\bar{x}).$$

Thus ϕ is a subgradient. If d is 0 then we have $p_n(0) = 0 = f'(\bar{x}; 0)$. Finally, if d is nonzero then by part (i) of

Lemma 3.1.6 we see

$$p_n(d) \le p_0(d) = p_0(e_1) = -p_0'(e_1; -e_1) = \\ -p_1(-e_1) = -p_1(-d) \le -p_n(-d) = p_n(d),$$

whence $p_n(d) = p_0(d) = f'(\bar{x}; d)$. ♠

Corollary 3.1.9 (Differentiability of convex functions) *Suppose that the function* $f : \mathbf{E} \to (-\infty, +\infty]$ *is convex, and that the point* $\bar{x}$ *lies in* $\operatorname{core}(\operatorname{dom} f)$*. Then* f *is Gâteaux differentiable at* $\bar{x}$ *exactly when* f *has a unique subgradient at* $\bar{x}$ *(in which case this subgradient is the derivative).*

We say the convex function f is *essentially smooth* if it is Gâteaux differentiable on $\operatorname{dom} \partial f$. In other words, essentially smooth functions have subdifferentials which are either singleton or empty.

The Max Formula (Theorem 3.1.7) shows that convex functions typically have subgradients. In fact this property characterizes convexity (see Exercise 12). This leads to a number of important ways of recognizing convex functions, of which the following is an example. Notice how a locally defined analytic condition results in a global geometric conclusion. The proof is outlined in the exercises.

Theorem 3.1.10 (Hessian characterization of convexity) *Given an open convex set $S \subset \mathbf{R}^n$, suppose the continuous function $f : \operatorname{cl} S \to \mathbf{R}$ is twice continuously differentiable on S. Then f is convex if and only if its Hessian matrix is positive semidefinite everywhere on S.*

3.2 The value function

In this section we describe another approach to the Karush-Kuhn-Tucker conditions (2.3.8) in the convex case, using the existence of subgradients we established in the previous section. We consider the *(inequality-constrained) convex program*

$$(3.2.1)\quad \begin{cases} \inf & f(x) \\ \text{subject to} & g_i(x) \ \leq \ 0, \quad \text{for } i = 1, 2, \ldots, m, \\ & \quad\ x \ \in \ \mathbf{E}, \end{cases}$$

where the functions $f, g_1, g_2, \ldots, g_m : \mathbf{E} \to (-\infty, +\infty]$ are convex and satisfy $\emptyset \neq \operatorname{dom} f \subset \cap_i \operatorname{dom} g_i$. Denoting the vector with components $g_i(x)$ by $g(x)$, the function $L : \mathbf{E} \times \mathbf{R}^m_+ \to (-\infty, +\infty]$ defined by

$$(3.2.2) \qquad L(x; \lambda) = f(x) + \lambda^T g(x),$$

is called the *Lagrangian*. A *feasible solution* is a point x in $\operatorname{dom} f$ satisfying the constraints.

We should emphasize that the term 'Lagrange multiplier' has different meanings in different contexts. In the present context we say a vector $\bar{\lambda} \in \mathbf{R}^m_+$ is a *Lagrange multiplier vector* for a feasible solution $\bar{x}$ if $\bar{x}$ minimizes the function $L(\cdot; \bar{\lambda})$ over $\mathbf{E}$ and $\bar{\lambda}$ satisfies the complementary slackness conditions: $\bar{\lambda}_i = 0$ whenever $g_i(\bar{x}) < 0$.

We can often use the following principle to solve simple optimization problems.

Proposition 3.2.3 (Lagrangian sufficient conditions) *If the point $\bar{x}$ is feasible for the convex program (3.2.1) and there is a Lagrange multiplier vector, then $\bar{x}$ is optimal.*

The proof is immediate, and in fact does not rely on convexity.

The Karush-Kuhn-Tucker conditions (2.3.8) are a converse to the above result when the functions $f, g_1, g_2, \ldots, g_m$ are convex and differentiable. We next follow a very different, and surprising route to this result, circumventing differentiability. We *perturb* the problem (3.2.1), and analyze the resulting *value function* $v : \mathbf{R}^m \to [-\infty, +\infty]$, defined by the equation

$$v(b) = \inf\{f(x) \mid g(x) \le b\}. \tag{3.2.4}$$

We show that Lagrange multiplier vectors $\bar{\lambda}$ correspond to subgradients of v (see Exercise 9).

Our old definition of convexity for functions does not naturally extend to functions $h : \mathbf{E} \to [-\infty, +\infty]$ (due to the possible occurrence of $\infty - \infty$). To generalize it we introduce the idea of the *epigraph* of h

$$\operatorname{epi}(h) = \{(y, r) \in \mathbf{E} \times \mathbf{R} \mid h(y) \le r\}, \tag{3.2.5}$$

and we say h is a *convex function* if epi (h) is a convex set. An exercise shows in this case that the domain

$$\mathrm{dom}\,(h) = \{y \mid h(y) < +\infty\}$$

is convex, and further that the value function v defined by equation (3.2.4) is convex. We say h is *proper* if dom h is nonempty and h never takes the value $-\infty$: if we wish to demonstrate the existence of subgradients for v using the results in the previous section then we need to exclude values $-\infty$.

Lemma 3.2.6 *If the function* $h : \mathbf{E} \to [-\infty, +\infty]$ *is convex and some point* $\hat{y}$ *in* core (dom h) *satisfies* $h(\hat{y}) > -\infty$*, then* h *never takes the value* $-\infty$*.*

Proof. Suppose some point y in $\mathbf{E}$ satisfies $h(y) = -\infty$. Since $\hat{y}$ lies in core (dom h) there is a real $t > 0$ with $\hat{y} + t(\hat{y} - y)$ in dom (h), and hence a real r with $(\hat{y} + t(\hat{y} - y), r)$ in epi (h). Now for any real s, (y, s) lies in epi (h), so we know

$$\left(\hat{y}, \frac{r + ts}{1 + t}\right) = \frac{1}{1 + t}(\hat{y} + t(\hat{y} - y), r) + \frac{t}{1 + t}(y, s) \in \mathrm{epi}\,(h),$$

Letting $s \to -\infty$ gives a contradiction. ♠

In §2.3 we saw that the Karush-Kuhn-Tucker conditions needed a regularity condition. In this approach we will apply a different condition, known as the *Slater Constraint Qualification* for the problem (3.2.1):

(3.2.7) There exists $\hat{x}$ in $\mathrm{dom}\,(f)$ with $g_i(\hat{x}) < 0$

$$\text{for } i = 1, 2, \ldots, m.$$

Theorem 3.2.8 (Lagrangian necessary conditions) *Suppose that the point $\bar{x}$ in* $\mathrm{dom}\,(f)$ *is optimal for the convex program (3.2.1), and that the Slater condition (3.2.7) holds. Then there is a Lagrange multiplier vector for $\bar{x}$.*

Proof. Defining the value function v by equation (3.2.4), certainly $v(0) > -\infty$, and the Slater condition shows $0 \in \mathrm{core}\,(\mathrm{dom}\, v)$, so in particular Lemma 3.2.6 shows that v never takes the value $-\infty$. (An incidental consequence, from §4.1, is the continuity of v at 0.) We now deduce the existence of a subgradient $-\bar{\lambda}$ of v at 0, by the Max Formula (3.1.7).

Any vector b in $\mathbf{R}^m_+$ obviously satisfies $g(\bar{x}) \le b$, whence the inequality

$$f(\bar{x}) = v(0) \le v(b) + \bar{\lambda}^T b \le f(\bar{x}) + \bar{\lambda}^T b.$$

Hence $\bar{\lambda}$ lies in $\mathbf{R}^m_+$. Furthermore, any point x in $\mathrm{dom}\, f$ clearly satisfies

$$f(x) \ge v(g(x)) \ge v(0) - \bar{\lambda}^T g(x) = f(\bar{x}) - \bar{\lambda}^T g(x).$$

The case $x = \bar{x}$, using the inequalities $\bar{\lambda} \geq 0$ and $g(\bar{x}) \leq 0$, shows $\bar{\lambda}^T g(\bar{x}) = 0$, which yields the complementary slackness conditions. Finally, all points x in dom f must satisfy $f(x) + \bar{\lambda}^T g(x) \geq f(\bar{x}) = f(\bar{x}) + \bar{\lambda}^T g(\bar{x})$. ♠

In particular, if in the above result $\bar{x}$ lies in core $(\cap_i \mathrm{dom}\, g_i \cap \mathrm{dom}\, f)$ and the functions $f, g_1, g_2, \ldots, g_m$ are differentiable at $\bar{x}$ then

$$\nabla f(\bar{x}) + \sum_{i=1}^{m} \bar{\lambda}_i \nabla g_i(\bar{x}) = 0,$$

so we recapture the Karush-Kuhn-Tucker conditions (2.3.8). In fact in this case it is easy to see that the Slater condition is equivalent to the Mangasarian-Fromowitz constraint qualification (Assumption 2.3.7).

3.3 The Fenchel conjugate

In the next few sections we sketch a little of the elegant and concise theory of Fenchel conjugation, and we use it to gain a deeper understanding of the Lagrangian necessary conditions for convex programs (3.2.8). The *Fenchel conjugate* of a function $h : \mathbf{E} \to [-\infty, +\infty]$ is the function $h^* : \mathbf{E} \to [-\infty, +\infty]$ defined by

$$h^*(\phi) = \sup_{x \in \mathbf{E}} \{\langle \phi, x \rangle - h(x)\}.$$

The function h^* is convex and if the domain of h is nonempty then h^* never takes the value $-\infty$. Clearly the conjugacy operation is *order-reversing*: for functions $f, g : \mathbf{E} \to [-\infty, +\infty]$, the inequality $f \geq g$ implies $f^* \leq g^*$.

Conjugate functions are ubiquitous in optimization. For example, we have already seen the conjugate of the exponential, defined by

$$\exp^*(t) = \begin{cases} t \log t - t & (t > 0) \\ 0 & (t = 0) \\ +\infty & (t < 0) \end{cases}$$

(see §3.1, Exercise 27). A rather more subtle example is the function $g : \mathbf{E} \to (-\infty, +\infty]$ defined, for points $a^0, a^1, \ldots, a^m$ in $\mathbf{E}$, by

(3.3.8)

$$g(z) \;=\; \inf_{x\in\mathbf{R}^{m+1}} \left\{ \sum_i \exp^*(x_i) \;\middle|\; \sum_i x_i = 1,\; \sum_i x_i a^i = z \right\}.$$

The conjugate is the function we used in §2.2 to prove various theorems of the alternative:

$$(3.3.1) \qquad g^*(y) = \log\left(\sum_i \exp\langle a^i, y\rangle\right)$$

(see Exercise 7).

As we shall see later (§4.2), many important convex functions h equal their *biconjugates* h^{**}. Such functions thus occur as natural pairs, h and h^*. The table in this section shows some elegant examples on $\mathbf{R}$.

The following result summarizes the properties of two particularly important convex functions.

Proposition 3.3.2 (Log barriers) *The functions* $\mathrm{lb} : \mathbf{R}^n \to (-\infty, +\infty]$ *and* $\mathrm{ld} : \mathbf{S}^n \to (-\infty, +\infty]$ *defined by*

$$\mathrm{lb}\,(x) \;=\; \begin{cases} -\Sigma_{i=1}^n \log x_i, & \text{if } x \in \mathbf{R}^n_{++}, \\ +\infty, & \text{otherwise, and} \end{cases}$$

$$\mathrm{ld}\,(X) \;=\; \begin{cases} -\log\det X, & \text{if } X \in \mathbf{S}^n_{++}, \\ +\infty, & \text{otherwise} \end{cases}$$

are essentially smooth, and strictly convex on their domains. They satisfy the conjugacy relations

$$\mathrm{lb}^*(x) \;=\; \mathrm{lb}\,(-x) - n, \quad \text{for all } x \in \mathbf{R}^n, \text{ and}$$

$$\text{ld}^*(X) = \text{ld}(-X) - n, \quad \textit{for all } X \in \mathbf{S}^n.$$

The perturbed functions $\text{lb} + \langle c, \cdot \rangle$ *and* $\text{ld} + \langle C, \cdot \rangle$ *have compact level sets for any vector* $c \in \mathbf{R}^n_{++}$ *and matrix* $C \in \mathbf{S}^n_{++}$ *respectively.*

(See §3.1, Exercise 21 (The log barrier), and §1.2, Exercise 14 (Level sets of perturbed log barriers): the conjugacy formulas are simple calculations.) Notice the simple relationships $\text{lb} = \text{ld} \circ \text{Diag}$ and $\text{ld} = \text{lb} \circ \lambda$ between these two functions.

The next elementary but important result relates conjugation with the subgradient. The proof is an exercise.

Proposition 3.3.3 (Fenchel-Young inequality) *Any points* ϕ *in* $\mathbf{E}$ *and* x *in the domain of a function* $h : \mathbf{E} \to (-\infty, +\infty]$ *satisfy the inequality*

$$h(x) + h^*(\phi) \geq \langle \phi, x \rangle.$$

Equality holds if and only if $\phi \in \partial h(x)$.

In §3.2 we analyzed the standard inequality constrained convex program by studying its optimal value under perturbations. A similar approach works for another model for convex programming, particularly suited to problems with linear constraints. An interesting byproduct is a convex analogue of the chain rule for differentiable functions, $\nabla(f + g \circ A)(x) = \nabla f(x) + A^*\nabla g(Ax)$ (for a linear map A).

In this section we fix a Euclidean space $\mathbf{Y}$. We denote the set of points where a function $g : \mathbf{Y} \to [-\infty, +\infty]$ is finite and continuous by $\operatorname{cont} g$.

Theorem 3.3.4 (Fenchel duality and convex calculus) *For given functions $f : \mathbf{E} \to (-\infty, +\infty]$ and $g : \mathbf{Y} \to (-\infty, +\infty]$ and a linear map $A : \mathbf{E} \to \mathbf{Y}$, let $p, d \in [-\infty, +\infty]$ be primal and dual values defined respectively by the optimization problems*

$$(3.3.5) \qquad p = \inf_{x \in \mathbf{E}} \{f(x) + g(Ax)\}$$

$$(3.3.6) \qquad d = \sup_{\phi \in \mathbf{Y}} \{-f^*(A^*\phi) - g^*(-\phi)\}.$$

These values satisfy the **weak duality** *inequality $p \geq d$. If furthermore f and g are convex and satisfy the condition*

$$(3.3.7) \qquad 0 \in \operatorname{core}(\operatorname{dom} g - A\operatorname{dom} f),$$

or the stronger condition

$$(3.3.8) \qquad A\operatorname{dom} f \cap \operatorname{cont} g \neq \emptyset,$$

then the values are equal ($p = d$), and the supremum in the dual problem (3.3.6) is attained if finite.

At any point x in $\mathbf{E}$, the calculus rule

$$\partial(f + g \circ A)(x) \supset \partial f(x) + A^* \partial g(Ax) \tag{3.3.9}$$

holds, with equality if f and g are convex and condition (3.3.7) or (3.3.8) holds.

Proof. The weak duality inequality follows immediately from the Fenchel-Young inequality (3.3.3). To prove equality we define an optimal value function $h : \mathbf{Y} \to [-\infty, +\infty]$ by

$$h(u) = \inf_{x \in \mathbf{E}} \{f(x) + g(Ax + u)\}.$$

It is easy to check h is convex, and $\operatorname{dom} h = \operatorname{dom} g - A \operatorname{dom} f$. If p is $-\infty$ there is nothing to prove, while if condition (3.3.7) holds and p is finite then Lemma 3.2.6 and the Max Formula (3.1.7) show there is a subgradient $-\phi \in \partial h(0)$. Hence we deduce

$$\begin{aligned} h(0) &\leq h(u) + \langle \phi, u \rangle, \quad \text{for all } u \in \mathbf{Y}, \\ &\leq f(x) + g(Ax + u) + \langle \phi, u \rangle, \quad \text{for all } u \in \mathbf{Y},\ x \in \mathbf{E}, \\ &= \{f(x) - \langle A^*\phi, x \rangle\} + \{g(Ax + u) - \langle -\phi, Ax + u \rangle\}. \end{aligned}$$

Taking the infimum over all points u, and then over all points x gives the inequalities

$$h(0) \leq -f^*(A^*\phi) - g^*(-\phi) \leq d \leq p = h(0).$$

Thus ϕ attains the supremum in problem (3.3.6), and $p = d$. An easy exercise shows that condition (3.3.8) implies condition (3.3.7). The proof of the calculus rule in the second part of the theorem is a simple consequence of the first part: see Exercise 9. ♠

The case of the Fenchel theorem above when the function g is simply the indicator function of a point gives the following particularly elegant and useful corollary.

Corollary 3.3.10 (Fenchel duality for linear constraints) *Given any function* $f : \mathbf{E} \to (-\infty, +\infty]$, *any linear map* $A : \mathbf{E} \to \mathbf{Y}$, *and any element* b *of* $\mathbf{Y}$, *the weak duality inequality*

$$\inf_{x \in \mathbf{E}} \{f(x) \mid Ax = b\} \geq \sup_{\phi \in \mathbf{Y}} \{\langle b, \phi \rangle - f^*(A^*\phi)\}$$

holds. If f *is convex and* b *belongs to* $\mathrm{core}\,(A\mathrm{dom}\, f)$ *then equality holds, and the supremum is attained when finite.*

A pretty application of the Fenchel duality circle of ideas is the calculation of polar cones. The *(negative) polar cone* of the set $K \subset \mathbf{E}$ is the convex cone

$$K^- = \{\phi \in \mathbf{E} \mid \langle \phi, x \rangle \leq 0, \ \ \text{for all } x \in K\},$$

and the cone K^{--} is called the *bipolar*. A particularly important example of the polar cone is the normal cone to a convex set $C \subset \mathbf{E}$ at a point x in C, since $N_C(x) = (C - x)^-$.

We use the following two examples extensively: the proofs are simple exercises.

Proposition 3.3.11 (Self-dual cones)

$$\begin{aligned} (\mathbf{R}^n_+)^- &= -\mathbf{R}^n_+, \quad \textit{and} \\ (\mathbf{S}^n_+)^- &= -\mathbf{S}^n_+. \end{aligned}$$

The next result shows how the calculus rules above can be used to derive geometric consequences.

Corollary 3.3.12 (Krein-Rutman polar cone calculus) *For any cones $H \subset \mathbf{Y}$ and $K \subset \mathbf{E}$ and any linear map $A : \mathbf{E} \to \mathbf{Y}$, the relation*

$$(K \cap A^{-1}H)^- \supset A^*H^- + K^-$$

holds. Equality holds if H and K are convex and satisfy $H - AK = Y$ (or in particular $AK \cap \operatorname{int} H \neq \emptyset$).

Proof. Rephrasing the definition of the polar cone shows that for any cone $K \subset \mathbf{E}$, the polar cone K^- is just $\partial\delta_K(0)$. The result now follows by the Fenchel theorem above. ♠

The polarity operation arises naturally from Fenchel conjugation, since for any cone $K \subset \mathbf{E}$, we have $\delta_{K^-} =$

δ_K^*, whence $\delta_{K^{--}} = \delta_K^{**}$. The next result, which is an elementary application of the Basic separation theorem (2.1.6), leads naturally into the development of the next chapter by identifying K^{--} as the closed convex cone generated by K.

Theorem 3.3.13 (Bipolar cone) *The bipolar cone of any nonempty set $K \subset \mathbf{E}$ is given by $K^{--} = \mathrm{cl}\,(\mathrm{conv}\,(\mathbf{R}_+ K))$.*

For example, we deduce immediately that the normal cone $N_C(x)$ to a convex set C at a point x in C, and the *(convex) tangent cone* to C at x defined by $T_C(x) = \mathrm{cl}\,\mathbf{R}_+(C - x)$, are polars of each other.

Exercise 20 outlines how to use these two results about cones to characterize *pointed* cones (those closed convex cones K satisfying $K \cap -K = \{0\}$).

Theorem 3.3.14 (Pointed cones) *A closed convex cone $K \subset \mathbf{E}$ is pointed if and only if there is an element y of $\mathbf{E}$ for which the set*

$$C = \{x \in K \mid \langle x, y\rangle = 1\}$$

is compact and generates K (that is, $K = \mathbf{R}_+ C$).

$f(x) = g^*(x)$	$\mathrm{dom}\, f$	$g(y) = f^*(y)$	$\mathrm{dom}\, g$
0	$\mathbf{R}$	0	$\{0\}$
0	$\mathbf{R}_+$	0	$-\mathbf{R}_+$
0	$[-1, 1]$	$\lvert y \rvert$	$\mathbf{R}$
0	$[0, 1]$	y^+	$\mathbf{R}$
$\lvert x \rvert^p/p \quad (1 < p \in \mathbf{R})$	$\mathbf{R}$	$\lvert y \rvert^q/q \quad (\frac{1}{p} + \frac{1}{q} = 1)$	$\mathbf{R}$
$\lvert x \rvert^p/p \quad (1 < p \in \mathbf{R})$	$\mathbf{R}_+$	$\lvert y^+ \rvert^q/q \quad (\frac{1}{p} + \frac{1}{q} = 1)$	$\mathbf{R}$
$-x^p/p \quad (p \in (0, 1))$	$\mathbf{R}_+$	$-(-y)^q/q \quad (\frac{1}{p} + \frac{1}{q} = 1)$	$-\mathbf{R}_{++}$
$\sqrt{1 + x^2}$	$\mathbf{R}$	$-\sqrt{1 - y^2}$	$[-1, 1]$
$-\log x$	$\mathbf{R}_{++}$	$-1 - \log(-y)$	$-\mathbf{R}_{++}$
$\cosh x$	$\mathbf{R}$	$y \sinh^{-1}(y) - \sqrt{1 + y^2}$	$\mathbf{R}$
$-\log(\cos x)$	$(-\frac{\pi}{2}, \frac{\pi}{2})$	$y \tan^{-1}(y) - \frac{1}{2} \log(1 + y^2)$	$\mathbf{R}$
$\log(\cosh x)$	$\mathbf{R}$	$y \tanh^{-1}(y) + \frac{1}{2} \log(1 - y^2)$	$(-1, 1)$
e^x	$\mathbf{R}$	$\begin{cases} y \log y - y & (y > 0) \\ 0 & (y = 0) \end{cases}$	$\mathbf{R}_+$
$\log(1 + e^x)$	$\mathbf{R}$	$\begin{cases} y \log y + (1 - y) \log(1 - y) & (y \in (0, 1)) \\ 0 & (y = 0, 1) \end{cases}$	$[0, 1]$
$-\log(1 - e^x)$	$\mathbf{R}$	$\begin{cases} y \log y - (1 + y) \log(1 + y) & (y > 0) \\ 0 & (y = 0) \end{cases}$	$\mathbf{R}_+$

Table 3.1: Conjugate pairs of convex functions on $\mathbf{R}$

$f = g^*$	$g = f^*$
$f(x)$	$g(y)$
$h(ax) \quad (a \neq 0)$	$h^*(y/a)$
$h(x+b)$	$h^*(y) - by$
$ah(x) \quad (a > 0)$	$ah^*(y/a)$

Table 3.2: Transformed conjugates

Chapter 4

Convex analysis

4.1 Continuity of convex functions

We have already seen that linear functions are always continuous. More generally, a remarkable feature of convex functions on $\mathbf{E}$ is that they must be continuous on the interior of their domains. Part of the surprise is that an *algebraic/geometric* assumption (convexity) leads to a *topological* conclusion (continuity). It is this powerful fact that guarantees the usefulness of regularity conditions like $A \operatorname{dom} f \cap \operatorname{cont} g \neq \emptyset$ (3.3.8) that we studied in the previous section.

Clearly an arbitrary function f is bounded above on some neighbourhood of any point in $\operatorname{cont} f$. In fact the converse is also true, and in a rather strong sense, needing the following definition. For a real $L \geq 0$, we say that a function $f : \mathbf{E} \to (-\infty, +\infty]$ is *Lipschitz (with* *constant* L*)* on a subset C of dom f if $|f(x) - f(y)| \leq L\|x - y\|$ for any points x and y in C. If f is Lipschitz on a neighbourhood of a point z then we say that f is *locally Lipschitz around* z. If $\mathbf{Y}$ is another Euclidean space we make analogous definitions for functions $F : \mathbf{E} \to \mathbf{Y}$, with $\|F(x) - F(y)\|$ replacing $|f(x) - f(y)|$.

Theorem 4.1.1 (Local boundedness) *Let* $f : \mathbf{E} \to (-\infty, +\infty]$ *be a convex function. Then* f *is locally Lipschitz around a point* z *in its domain if and only if it is bounded above on a neighbourhood of* z*.*

Proof. One direction is clear, so let us without loss of generality take $z = 0$, $f(0) = 0$, and suppose $f \leq 1$ on $2B$: we shall deduce f is Lipschitz on B.

Notice first the bound $f \geq -1$ on $2B$, since convexity implies $f(-x) \geq -f(x)$ on $2B$. Now for any distinct points x and y in B, define $\alpha = \|y - x\|$ and fix a point $w = y + \alpha^{-1}(y - x)$, which lies in $2B$. By convexity we obtain

$$\begin{aligned} f(y) - f(x) &\leq \frac{1}{1+\alpha} f(x) + \frac{\alpha}{1+\alpha} f(w) - f(x) \\ &\leq \frac{2\alpha}{1+\alpha} \leq 2\|y - x\|, \end{aligned}$$

and the result now follows, since x and y may be interchanged. ♠

This result makes it easy to identify the set of points at which a convex function on $\mathbf{E}$ is continuous. First we prove a key lemma.

Lemma 4.1.2 *Let* Δ *be the* **simplex** $\{x \in \mathbf{R}^n_+ \,|\, \Sigma x_i \leq 1\}$. *If the function* $g : \Delta \to \mathbf{R}$ *is convex then it is continuous on* $\operatorname{int} \Delta$.

Proof. By the above result, we just need to show g is bounded above on Δ. But any point x in Δ satisfies

$$\begin{aligned} g(x) &= g\Big(\sum_1^n x_i e^i + (1 - \Sigma x_i) 0\Big) \\ &\leq \sum_1^n x_i g(e^i) + (1 - \Sigma x_i) g(0) \\ &\leq \max\{g(e^1), g(e^2), \ldots, g(e^n), g(0)\} \end{aligned}$$

(where $\{e^1, e^2, \ldots, e^n\}$ is the standard basis in $\mathbf{R}^n$). ♠

Theorem 4.1.3 (Convexity and continuity) *Let* $f : \mathbf{E} \to (-\infty, +\infty]$ *be a convex function. Then* f *is continuous (in fact locally Lipschitz) on the interior of its domain.*

Proof. We lose no generality if we restrict ourselves to the case $\mathbf{E} = \mathbf{R}^n$. For any point x in $\operatorname{int}(\operatorname{dom} f)$ we can choose a neighbourhood of x in $\operatorname{dom} f$ which is a scaled-down, translated copy of the simplex (since the simplex is bounded, with nonempty interior). The proof of the preceding lemma now shows f is bounded above on a neighbourhood of x, and the result follows by Theorem 4.1.1 (Local boundedness). ♠

Since it is easy to see that if the convex function f is locally Lipschitz around a point $\bar{x}$ in $\mathrm{int}\,(\mathrm{dom}\, f)$ with constant L then $\partial f(\bar{x}) \subset LB$, we can also conclude that $\partial f(\bar{x})$ is a nonempty compact convex set. Furthermore, this result allows us to conclude quickly that 'all norms on $\mathbf{E}$ are equivalent' (see Exercise 2).

We have seen that for a function f which is convex, the two sets $\mathrm{cont}\, f$ and $\mathrm{int}\,(\mathrm{dom}\, f)$ are identical. By contrast, our algebraic approach to the existence of subgradients involved $\mathrm{core}\,(\mathrm{dom}\, f)$. It transpires that this is the same set. To see this we introduce the idea of the *gauge function* $\gamma_C : \mathbf{E} \to (-\infty, +\infty]$ associated with a nonempty set C in $\mathbf{E}$:

$$\gamma_C(x) = \inf\{\lambda \in \mathbf{R}_+ \mid x \in \lambda C\}.$$

It is easy to check γ_C is sublinear (and in particular convex) when C is convex. Notice $\gamma_B = \|\cdot\|$.

Theorem 4.1.4 (Core and interior) *The core and the interior of any convex set in* $\mathbf{E}$ *are identical and convex.*

Proof. Any convex set $C \subset \mathbf{E}$ clearly satisfies $\operatorname{int} C \subset \operatorname{core} C$. If we suppose, without loss of generality, $0 \in \operatorname{core} C$, then γ_C is everywhere finite, and hence continuous by the previous result. We claim

$$\operatorname{int} C = \{x \mid \gamma_C(x) < 1\}.$$

To see this, observe that the right hand side is contained in C, and is open by continuity, and hence is contained in $\operatorname{int} C$. The reverse inclusion is easy, and we deduce $\operatorname{int} C$ is convex. Finally, since $\gamma_C(0) = 0$, we see $0 \in \operatorname{int} C$, which completes the proof. ♠

The conjugate of the gauge function γ_C is the indicator function of a set $C^\circ \subset \mathbf{E}$ defined by

$$C^\circ = \{\phi \in \mathbf{E} \mid \langle \phi, x \rangle \leq 1 \text{ for all } x \in C\}.$$

We call C° the *polar set* for C. Clearly it is a closed convex set containing 0, and when C is a cone it coincides with the polar cone C^-. The following result therefore generalizes the Bipolar cone theorem (3.3.13).

Theorem 4.1.5 (Bipolar set) *The bipolar set of any subset C of $\mathbf{E}$ is given by*

$$C^{\circ\circ} = \operatorname{cl}(\operatorname{conv}(C \cup \{0\})).$$

The ideas of polarity and separating hyperplanes are intimately related. The separation-based proof of the above result (which we leave as an exercise) is a good example, as is the next theorem, whose proof is outlined in Exercise 6.

Theorem 4.1.6 (Supporting hyperplane) *Suppose that the convex set $C \subset \mathbf{E}$ has nonempty interior, and that the point $\bar{x}$ lies on the boundary of C. Then there is a* **supporting hyperplane** *to C at $\bar{x}$: there is a nonzero element a of $\mathbf{E}$ satisfying $\langle a, x\rangle \geq \langle a, \bar{x}\rangle$ for all points x in C.*

(The set $\{x \in \mathbf{E} \mid \langle a, x - \bar{x}\rangle = 0\}$ is the supporting hyperplane.)

To end this section we use this result to prove a remarkable theorem of Minkowski describing an extremal representation of finite-dimensional compact convex sets. An *extreme point* of a convex set $C \subset \mathbf{E}$ is a point x in C whose complement $C \setminus \{x\}$ is convex. We denote the set of extreme points by $\operatorname{ext} C$.

Lemma 4.1.7 *Given a supporting hyperplane H of a convex set $C \subset \mathbf{E}$, any extreme point of $C \cap H$ is also an extreme point of C.*

Our proof of Minkowski's theorem depends on two facts: first, any convex set which spans $\mathbf{E}$ and contains

0 has nonempty interior (see §1.1, Exercise 13(b))); secondly, we can define the *dimension* of a set $C \subset \mathbf{E}$ (written $\dim C$) as the dimension of $\operatorname{span}(C - x)$ for *any* point x in C (see §1.1, Exercise 12 (Affine sets)).

Theorem 4.1.8 (Minkowski) *Any compact convex set $C \subset \mathbf{E}$ is the convex hull of its extreme points.*

Proof. Our proof is by induction on $\dim C$: clearly the result holds when $\dim C = 0$. Assume the result holds for all sets of dimension less than $\dim C$. We will deduce it for the set C.

By translating C, and redefining $\mathbf{E}$, we can assume $0 \in C$ and $\operatorname{span} C = \mathbf{E}$. Thus C has nonempty interior.

Given any point x in $\operatorname{bd} C$, the Supporting hyperplane theorem (4.1.6) shows C has a supporting hyperplane H at x. By the induction hypothesis applied to the set $C \cap H$ we deduce, using Lemma 4.1.7,

$$x \in \operatorname{conv}(\operatorname{ext}(C \cap H)) \subset \operatorname{conv}(\operatorname{ext} C).$$

So we have proved $\operatorname{bd} C \subset \operatorname{conv}(\operatorname{ext} C)$, whence $\operatorname{conv}(\operatorname{bd} C) \subset \operatorname{conv}(\operatorname{ext} C)$. But since C is compact it is easy to see $\operatorname{conv}(\operatorname{bd} C) = C$, and the result now follows. ♠

4.2 Fenchel biconjugation

We have seen that for many important convex functions $h : \mathbf{E} \to (-\infty, +\infty]$, the biconjugate h^{**} agrees identically with h. The table in §3.3 lists many one-dimensional examples, and the Bipolar cone theorem (3.3.13) shows $\delta_K = \delta_K^{**}$ for any closed convex cone K. In this section we isolate exactly the circumstances when $h = h^{**}$.

We can easily check that h^{**} is a minorant of h (that is, $h^{**} \leq h$ pointwise). Our specific aim in this section is to find conditions on a point x in $\mathbf{E}$ guaranteeing $h^{**}(x) = h(x)$. This becomes the key relationship for the study of duality in optimization. As we see in this section, the conditions we need are both geometric and topological. This is neither particularly surprising or stringent. Since any conjugate function must have a closed convex epigraph, we cannot expect a function to agree with its biconjugate unless it itself has a closed convex epigraph. On the other hand, this restriction is not particularly strong since, as the previous section showed, convex functions automatically have strong continuity properties.

We say the function $h : \mathbf{E} \to [-\infty, +\infty]$ is *closed* if its epigraph is a closed set. We say h is *lower semicontinuous* at a point x in $\mathbf{E}$ if

$$\liminf h(x^r) \; \Big(= \lim_{s\to\infty} \inf_{r \geq s} h(x^r)\Big) \; \geq h(x)$$

for any sequence $x^r \to x$. A function $h : \mathbf{E} \to [-\infty, +\infty]$ is *lower semicontinuous* if it is lower semicontinuous at every point in $\mathbf{E}$: this is in fact equivalent to h being closed, which in turn holds if and only if h has closed level sets. Any two functions h and g satisfying $h \le g$ (in which case we call h a *minorant* of g) must satisfy $h^* \ge g^*$, and hence $h^{**} \le g^{**}$.

Theorem 4.2.1 (Fenchel biconjugation) *The properties below are equivalent, for any function* $h : \mathbf{E} \to (-\infty, +\infty]$*:*

(a) h *is closed and convex;*

(b) $h = h^{**}$*;*

(c) for all points x *in* $\mathbf{E}$,

$$h(x) = \sup\{\alpha(x) \mid \alpha \text{ an affine minorant of } h\}.$$

Hence the conjugacy operation induces a bijection between proper closed convex functions.

Proof. We can assume h is proper. Since conjugate functions are always closed and convex we know property (b) implies property (a). Also, any affine minorant α of h satisfies $\alpha = \alpha^{**} \le h^{**} \le h$, and hence property (c) implies (b). It remains to show (a) implies (c).

Fix a point x^0 in $\mathbf{E}$. Assume first $x^0 \in \mathrm{cl}\,(\mathrm{dom}\, h)$, and fix any real $r < h(x^0)$. Since h is closed, the set $\{x \mid h(x) > r\}$ is open, so there is an open convex neighbourhood U of x^0 with $h(x) > r$ on U. Now note that the set $\mathrm{dom}\, h \cap \mathrm{cont}\, \delta_U$ is nonempty, so we can apply the Fenchel theorem (3.3.4) to deduce that some element ϕ of $\mathbf{E}$ satisfies

$$(4.2.2) \quad r \leq \inf_x \{h(x) + \delta_U(x)\} = \{-h^*(\phi) - \delta_U^*(-\phi)\}.$$

Now define an affine function $\alpha(\cdot) = \langle \phi, \cdot \rangle + \delta_U^*(-\phi) + r$. Inequality (4.2.2) shows that α minorizes h, and by definition we know $\alpha(x^0) \geq r$. Since r was arbitrary, (c) follows at the point $x = x^0$.

Suppose on the other hand x^0 does not lie in $\mathrm{cl}\,(\mathrm{dom}\, h)$. By the Basic separation theorem (2.1.6) there is a real b and a nonzero element a of $\mathbf{E}$ satisfying

$$\langle a, x^0 \rangle > b \geq \langle a, x \rangle, \quad \text{for all points } x \text{ in } \mathrm{dom}\, h.$$

The argument in the preceding paragraph shows there is an affine minorant α of h. But now the affine function $\alpha(\cdot) + k(\langle a, \cdot \rangle - b)$ is a minorant of h for all $k = 1, 2, \ldots$. Evaluating these functions at $x = x^0$ proves property (c) at x^0. The final remark follows easily. ♠

We can immediately deduce that a closed convex function

$h : \mathbf{E} \to [-\infty, +\infty]$ equals its biconjugate if and only if it is proper or identically $+\infty$ or $-\infty$.

Restricting the conjugacy bijection to finite sublinear functions gives the following result.

Corollary 4.2.3 (Support functions) *Fenchel conjugacy induces a bijection between everywhere-finite sublinear functions and nonempty compact convex sets in* $\mathbf{E}$*:*

(a) If the set $C \subset \mathbf{E}$ *is compact, convex and nonempty then the* **support function** δ_C^* *is everywhere finite and sublinear.*

(b) If the function $h : \mathbf{E} \to \mathbf{R}$ *is sublinear then* $h^* = \delta_C$*, where the set*

$$C = \{\phi \in \mathbf{E} \mid \langle \phi, d \rangle \leq h(d) \text{ for all } d \in \mathbf{E}\}$$

is nonempty, compact and convex.

Proof. See Exercise 9. ♠

Conjugacy offers a convenient way to recognize when a convex function has bounded level sets.

Theorem 4.2.4 (Moreau-Rockafellar) *A closed convex proper function on* $\mathbf{E}$ *has bounded level sets if and only if its conjugate is continuous at 0.*

Proof. By Proposition 1.1.4, a convex function $f : \mathbf{E} \to (-\infty, +\infty]$ has bounded level sets if and only if it satisfies the growth condition

$$\liminf_{\|x\| \to \infty} \frac{f(x)}{\|x\|} > 0.$$

Since f is closed we can check that this is equivalent to the existence of a minorant of the form $\epsilon\|\cdot\| + k \le f(\cdot)$, for some constants $\epsilon > 0$ and k. Taking conjugates, this is in turn equivalent to f^* being bounded above near 0, and the result then follows by Theorem 4.1.1 (Local boundedness). ♠

Strict convexity is also easy to recognize via conjugacy, using the following result — see Exercise 19 for the proof.

Theorem 4.2.5 (Strict-smooth duality) *A proper closed convex function on* $\mathbf{E}$ *is essentially strictly convex if and only if its conjugate is essentially smooth.*

What can we say about h^{**} when the function $h : \mathbf{E} \to [-\infty, +\infty]$ is not necessarily closed? To answer this question we introduce the idea of the *closure* of h, denoted $\mathrm{cl}\, h$, defined by

$$\text{(4.2.6)} \qquad \mathrm{epi}\,(\mathrm{cl}\, h) = \mathrm{cl}\,(\mathrm{epi}\, h).$$

It is easy to verify that $\mathrm{cl}\, h$ is then well-defined. The definition immediately implies $\mathrm{cl}\, h$ is the largest closed function minorizing h. Clearly if h is convex, so is $\mathrm{cl}\, h$. We leave the proof of the next simple result as an exercise.

Proposition 4.2.7 (Lower semicontinuity and closure) *A convex function $f : \mathbf{E} \to [-\infty, +\infty]$ is lower semicontinuous at a point x where it is finite if and only if $f(x) = (\mathrm{cl}\, f)(x)$. In this case f is proper.*

We can now answer the question we posed at the beginning of the section.

Theorem 4.2.8 *Suppose the function $h : \mathbf{E} \to [-\infty, +\infty]$ is convex.*

(a) If h^{**} is somewhere finite then $h^{**} = \mathrm{cl}\, h$.

(b) For any point x where h is finite, $h(x) = h^{**}(x)$ if and only if h is lower semicontinuous at x.

Proof. Observe first that since h^{**} is closed and minorizes h, we know $h^{**} \le \mathrm{cl}\, h \le h$. If h^{**} is somewhere finite then h^{**} (and hence $\mathrm{cl}\, h$) is never $-\infty$, by applying Proposition 4.2.7 (Lower semicontinuity and closure) to h^{**}. On the other hand, if h is finite and lower semicontinuous at x then Proposition 4.2.7 shows $\mathrm{cl}\, h(x)$ is finite, and applying the proposition again to $\mathrm{cl}\, h$ shows once more that $\mathrm{cl}\, h$ is never $-\infty$. In either case, the Fenchel biconjugation theorem implies $\mathrm{cl}\, h = (\mathrm{cl}\, h)^{**} \le h^{**} \le \mathrm{cl}\, h$, so $\mathrm{cl}\, h = h^{**}$. Part (a) is now immediate, while part (b) follows by using Proposition 4.2.7 once more. ♠

Any proper convex function h with an affine minorant has its biconjugate h^{**} somewhere finite. (In fact, because $\mathbf{E}$ is finite-dimensional, h^{**} is somewhere finite if and only if h is proper — see exercise 25.)

4.3 Lagrangian duality

The duality between a convex function h and its Fenchel conjugate h^* that we outlined earlier is an elegant piece of theory. The real significance, however, lies in its power to describe duality theory for convex programs, one of the most far-reaching ideas in the study of optimization.

We return to the convex program that we studied in §3.2:

$$(4.3.1) \qquad \begin{cases} \inf & f(x) \\ \text{subject to} & g(x) \leq 0, \\ & x \in \mathbf{E}. \end{cases}$$

Here, the function f and the components $g_1, g_2, \ldots, g_m : \mathbf{E} \to (-\infty, +\infty]$ are convex, and satisfy $\emptyset \neq \operatorname{dom} f \subset \cap_1^m \operatorname{dom} g_i$. As before, the Lagrangian function $L : \mathbf{E} \times \mathbf{R}_+^m \to (-\infty, +\infty]$ is defined by $L(x; \lambda) = f(x) + \lambda^T g(x)$.

Notice that the Lagrangian encapsulates all the information of the *primal problem* (4.3.1): clearly

$$\sup_{\lambda \in \mathbf{R}_+^m} L(x; \lambda) = \begin{cases} f(x), & \text{if } x \text{ is feasible,} \\ +\infty, & \text{otherwise,} \end{cases}$$

so if we denote the optimal value of (4.3.1) by $p \in [-\infty, +\infty]$, we could rewrite the problem in the following form:

$$(4.3.2) \qquad p = \inf_{x \in \mathbf{E}} \sup_{\lambda \in \mathbf{R}_+^m} L(x; \lambda).$$

This makes it rather natural to consider an associated problem:

$$d = \sup_{\lambda \in \mathbf{R}^m_+} \inf_{x \in \mathbf{E}} L(x;\lambda), \tag{4.3.3}$$

where $d \in [-\infty, +\infty]$ is called the *dual value*. Thus the *dual problem* consists of maximizing over vectors λ in $\mathbf{R}^m_+$ the *dual function* $\Phi(\lambda) = \inf_x L(x;\lambda)$. This dual problem is perfectly well-defined without any assumptions on the functions f and g. It is an easy exercise to show the 'weak duality inequality' $p \geq d$. Notice Φ is concave.

It can happen that the primal value p is strictly larger than the dual value d (see Exercise 5). In this case we say there is a *duality gap*. In this section we investigate conditions ensuring there is no duality gap. As in §3.2, the chief tool in our analysis is the primal value function $v : \mathbf{R}^m \to [-\infty, +\infty]$, defined by

$$v(b) = \inf\{f(x) \mid g(x) \leq b\}. \tag{4.3.4}$$

Below we summarize the relationships between these various ideas and pieces of notation.

Proposition 4.3.5 (Dual optimal value)

(a) The primal optimal value p is $v(0)$.

(b) The conjugate of the value function satisfies

$$v^*(-\lambda) = \begin{cases} -\Phi(\lambda), & \textit{if } \lambda \geq 0, \\ +\infty, & \textit{otherwise.} \end{cases}$$

(c) The dual optimal value d is $v^{**}(0)$.

Proof. Part (a) is just the definition of p. Part (b) follows from the identities

$$\begin{aligned} v^*(-\lambda) &= \sup\{-\lambda^T b - v(b) \mid b \in \mathbf{R}^m\} \\ &= \sup\{-\lambda^T b - f(x) \mid g(x) + z = b, \\ &\qquad x \in \operatorname{dom} f,\ b \in \mathbf{R}^m,\ z \in \mathbf{R}^m_+\} \\ &= \sup\{-\lambda^T (g(x) + z) - f(x) \mid x \in \operatorname{dom} f, \\ &\qquad z \in \mathbf{R}^m_+\} \\ &= -\inf\{f(x) + \lambda^T g(x) \mid x \in \operatorname{dom} f\} \\ &\qquad + \sup\{-\lambda^T z \mid z \in \mathbf{R}^m_+\} \\ &= \begin{cases} -\Phi(\lambda), & \text{if } \lambda \geq 0, \\ +\infty, & \text{otherwise.} \end{cases} \end{aligned}$$

Finally, we observe

$$\begin{aligned} d &= \sup_{\lambda \in \mathbf{R}^m_+} \Phi(\lambda) = -\inf_{\lambda \in \mathbf{R}^m_+} -\Phi(\lambda) \\ &= -\inf_{\lambda \in \mathbf{R}^m_+} v^*(-\lambda) = v^{**}(0), \end{aligned}$$

so part (c) follows. ♠

Notice the above result does not use convexity.

The reason for our interest in the relationship between a convex function and its biconjugate should now be clear, in light of parts (a) and (c) above.

Corollary 4.3.6 (Zero duality gap) *Suppose the value of the primal problem (4.3.1) is finite. Then the primal and dual values are equal if and only if the value function v is lower semicontinuous at 0. In this case the set of optimal dual solutions is $-\partial v(0)$.*

Proof. By the previous result, there is no duality gap exactly when the values function satisfies $v(0) = v^{**}(0)$, so Theorem 4.2.8 proves the first assertion. By part (b) of the previous result, dual optimal solutions λ are characterized by the property $0 \in \partial v^*(-\lambda)$, or equivalently, $v^*(-\lambda) + v^{**}(0) = 0$. But we know $v(0) = v^{**}(0)$, so this property is equivalent to $-\lambda \in \partial v(0)$. ♠

This result sheds new light on our proof of the Lagrangian necessary conditions (3.2.8): the proof in fact demonstrates the existence of a dual optimal solution. We consider below two distinct approaches to proving the absence of a duality gap. The first uses the Slater condition, as in Theorem 3.2.8, to force attainment in the dual problem. The second (dual) approach uses compactness to force attainment in the primal problem.

Theorem 4.3.7 (Dual attainment) *If the Slater condition holds for the primal problem (4.3.1) then the primal and dual values are equal, and the dual* *value is attained if finite.*

Proof. If p is $-\infty$ there is nothing to prove, since we know $p \geq d$. If on the other hand p is finite then, as in the proof of the Lagrangian necessary conditions (3.2.8), the Slater condition forces $\partial v(0) \neq \emptyset$. Hence v is finite and lower semicontinuous at 0 (§4.2, exercise 15), and the result follows by Corollary 4.3.6 (Zero duality gap). ♠

An indirect way of stating the Slater condition is that there is a point $\hat{x}$ in $\mathbf{E}$ for which the set $\{\lambda \in \mathbf{R}^m_+ | L(\hat{x};\lambda) \geq \alpha\}$ is compact for all real α. The second approach uses a 'dual' condition to ensure the value function is closed.

Theorem 4.3.8 (Primal attainment) *Suppose that the functions*

$$f, g_1, g_2, \ldots, g_m : \mathbf{E} \to (-\infty, +\infty]$$

are closed, and that for some real $\hat{\lambda}_0 \geq 0$ *and some vector* $\hat{\lambda}$ *in* $\mathbf{R}^m_+$, *the function* $\hat{\lambda}_0 f + \hat{\lambda}^T g$ *has compact level sets. Then the value function* v *defined by equation (4.3.4) is closed, and the infimum in this equation is attained when finite. Consequently, if the functions* $f, g_1, g_2, \ldots, g_m$ *are in addition convex and the dual value for the problem (4.3.1) is not* $-\infty$, *then the primal and dual values,* p *and* d, *are equal, and the primal value is attained when finite.*

Proof. If the points (b^r, s_r) lie in $\operatorname{epi} v$ for $r = 1, 2, \ldots$, and approach the point (b, s), then for each integer r there is a point x^r in $\mathbf{E}$ satisfying $f(x^r) \leq s_r + r^{-1}$ and $g(x^r) \leq b^r$. Hence we deduce

$$(\hat{\lambda}_0 f + \hat{\lambda}^T g)(x^r) \leq \hat{\lambda}_0(s_r + r^{-1}) + \hat{\lambda}^T b^r \to \hat{\lambda}_0 s + \hat{\lambda}^T b.$$

By the compact level set assumption, the sequence (x^r) has a subsequence converging to some point $\bar{x}$, and since all the functions are closed, we know $f(\bar{x}) \leq s$ and $g(\bar{x}) \leq b$. We deduce $v(b) \leq s$, so (b, s) lies in $\operatorname{epi} v$ as we required. When $v(b)$ is finite, the same argument with (b^r, s_r) replaced by $(b, v(b))$ for each r shows the infimum is attained.

If the functions $f, g_1, g_2, \ldots, g_m$ are convex then we know (from §3.2) v is convex. If d is $+\infty$ then, then again from the inequality $p \geq d$, there is nothing to prove. If d $(= v^{**}(0))$ is finite then Theorem 4.2.8 shows $v^{**} = \operatorname{cl} v$, and the above argument shows $\operatorname{cl} v = v$. Hence $p = v(0) = v^{**}(0) = d$, and the result follows.

Notice that if either the objective function f or any one of the constraint functions $g_1, g_2, \ldots, g_m$ has compact level sets then the compact level set condition in the above result holds.

Chapter 5

Special cases

5.1 Polyhedral convex sets and functions

In our earlier section on theorems of the alternative (§2.2), we observed that finitely generated cones are closed. Remarkably, a finite linear-algebraic assumption leads to a topological conclusion. In this section we pursue the consequences of this type of assumption in convex analysis.

There are two natural ways to impose a finite linear structure on the sets and functions we consider. The first we have already seen: a 'polyhedron' (or *polyhedral* set) is a finite intersection of closed halfspaces in $\mathbf{E}$, and we say a function $f : \mathbf{E} \to [-\infty, +\infty]$ is *polyhedral* if its epigraph is polyhedral. On the other hand, a *polytope* is the convex hull of a finite subset of $\mathbf{E}$, and we call a subset of $\mathbf{E}$ *finitely generated* if it is the sum of a polytope and a finitely generated cone (in the sense of formula (2.2.11)). Notice we do not yet know if a cone which is a finitely generated set in this sense is finitely generated in the sense of (2.2.11): we return to this point later in the section. The function f is *finitely generated* if its epigraph is finitely generated. A central result of this section is that polyhedra and finitely generated sets in fact coincide.

We begin with some easy observations collected together in the following two results.

Proposition 5.1.1 (Polyhedral functions) *Suppose the function* $f : \mathbf{E} \to [-\infty, +\infty]$ *is polyhedral. Then* f *is closed and convex, and can be decomposed in the form*

$$f = \max_{i \in I} g_i + \delta_P, \tag{5.1.2}$$

where the index set I *is finite (and possibly empty), the functions* g_i *are affine, and the set* $P \subset \mathbf{E}$ *is polyhedral (and possibly empty). Thus the domain of* f *is polyhedral, and coincides with* $\operatorname{dom} \partial f$ *if* f *is proper.*

Proof. Since any polyhedron is closed and convex, so is f, and the decomposition (5.1.2) follows directly from the definition. If f is proper then both the sets I and P are nonempty in this decomposition. At any point x in P ($= \operatorname{dom} f$) we know $0 \in \partial \delta_P(x)$, and the function $\max_i g_i$

certainly has a subgradient at x since it is everywhere finite. Hence we deduce $\partial f(x) \neq \emptyset$. ♠

Proposition 5.1.3 (Finitely generated functions) *Suppose the function $f : \mathbf{E} \to [-\infty, +\infty]$ is finitely generated. Then f is closed and convex, and* $\operatorname{dom} f$ *is finitely generated. Furthermore, f^* is polyhedral.*

Proof. Polytopes are compact and convex (by Caratheodory's theorem (§2.2, Exercise 5)), and finitely generated cones are closed and convex, so finitely generated sets (and therefore functions) are closed and convex, by §1.1, Exercise 5(a). We leave the remainder of the proof as an exercise. ♠

An easy exercise shows that a set $P \subset \mathbf{E}$ is polyhedral (respectively, finitely generated) if and only if δ_P is likewise.

To prove that polyhedra and finitely generated sets in fact coincide, we consider the two extreme special cases: first, compact sets, and secondly, cones. Observe first that compact, finitely generated sets are just polytopes, directly from the definition.

Lemma 5.1.4 *Any polyhedron has at most finitely many extreme points.*

Proof. Fix a finite set of affine functions $\{g_i \mid i \in I\}$ on $\mathbf{E}$, and consider the polyhedron

$$P = \{x \in \mathbf{E} \mid g_i(x) \leq 0 \text{ for } i \in I\}.$$

For any point x in P, the 'active set' is $\{i \in I | g_i(x) = 0\}$. Suppose two distinct extreme points x and y of P have the same active set. Then, for any small real ϵ, the points $x \pm \epsilon(y - x)$ both lie in P. But this contradicts the assumption that x is extreme. Hence different extreme points have different active sets, and the result follows. ♠

This lemma, together with Minkowski's theorem (4.1.8) reveals the nature of compact polyhedra.

Theorem 5.1.5 *Any compact polyhedron is a polytope.*

We next turn to cones.

Lemma 5.1.6 *Any polyhedral cone is a finitely generated cone (in the sense of (2.2.11)).*

Proof. Given a polyhedral cone $P \subset \mathbf{E}$, define a subspace $L = P \cap -P$, and a pointed polyhedral cone $K = P \cap L^{\perp}$. Observe the decomposition $P = K \oplus L$.

By the Pointed cone theorem (3.3.14), there is an element y of $\mathbf{E}$ for which the set

$$C = \{x \in K \mid \langle x, y \rangle = 1\}$$

is compact and satisfies $K = \mathbf{R}_+ C$. Since C is polyhedral, the previous result shows it is a polytope. Thus K is finitely generated, whence so is P. ♠

Theorem 5.1.7 (Polyhedrality) *A set or function is polyhedral if and only if it is finitely generated.*

Proof. For finite sets $\{a_i \mid i \in I\} \subset \mathbf{E}$ and $\{b_i \mid i \in I\} \subset \mathbf{R}$, consider the polyhedron in $\mathbf{E}$ defined by

$$P = \{x \in \mathbf{E} \mid \langle a_i, x \rangle \le b_i \text{ for } i \in I\}.$$

The polyhedral cone in $\mathbf{E} \times \mathbf{R}$ defined by

$$Q = \{(x, r) \in \mathbf{E} \times \mathbf{R} \mid \langle a_i, x \rangle - b_i r \le 0 \text{ for } i \in I\}$$

is finitely generated, by the previous lemma, so there are finite subsets $\{x_j \mid j \in J\}$ and $\{y_t \mid t \in T\}$ of $\mathbf{E}$ with

$$\begin{aligned} Q = \Big\{ & \sum_{j \in J} \lambda_j (x_j, 1) + \sum_{t \in T} \mu_t (y_t, 0) \;\Big|\; \\ & \lambda_j \in \mathbf{R}_+ \text{ for } j \in J,\ \mu_t \in \mathbf{R}_+ \text{ for } t \in T \Big\}. \end{aligned}$$

We deduce

$$\begin{aligned} P &= \{(x,1) \in Q\} \\ &= \operatorname{conv}\{x_j \mid j \in J\} + \left\{ \sum_{t \in T} \mu_t y_y \;\middle|\; \mu_t \in \mathbf{R}_+ \text{ for } t \in T \right\}, \end{aligned}$$

so P is finitely generated. We have thus shown that any polyhedral set (and hence function) is finitely generated.

Conversely, suppose the function $f : \mathbf{E} \to [-\infty, +\infty]$ is finitely generated. Consider first the case when f is proper. By Proposition 5.1.3, f^* is polyhedral, and hence (by the above argument) finitely generated. But f is closed and convex, by Proposition 5.1.3, so the Fenchel biconjugation theorem (4.2.1) implies $f = f^{**}$. By applying Proposition 5.1.3 once again we see f^{**} (and hence f) is polyhedral. We leave the improper case as an exercise. ♠

Notice these two results show our two notions of a finitely generated cone do indeed coincide.

The following collection of exercises shows that many linear-algebraic operations preserve polyhedrality.

Proposition 5.1.8 (Polyhedral algebra) *Consider a Euclidean space* $\mathbf{Y}$ *and a linear map* $A : \mathbf{E} \to \mathbf{Y}$.

(a) *If the set* $P \subset \mathbf{E}$ *is polyhedral then so is its image* AP.

(b) If the set $K \subset \mathbf{Y}$ is polyhedral then so is its preimage $A^{-1}K$.

(c) The sum and pointwise maximum of finitely many polyhedral functions are polyhedral.

(d) If the function $g : \mathbf{Y} \to [-\infty, +\infty]$ is polyhedral then so is the composite function $g \circ A$.

(e) If the function $q : \mathbf{E} \times \mathbf{Y} \to [-\infty, +\infty]$ is polyhedral then so is the function $h : \mathbf{Y} \to [-\infty, +\infty]$ defined by $h(u) = \inf_{x \in \mathbf{E}} q(x, u)$.

Corollary 5.1.9 (Polyhedral Fenchel duality) *All the conclusions of the Fenchel duality theorem (3.3.4) remain valid if the regularity condition (3.3.7) is replaced by the assumption that the functions f and g are polyhedral with $\mathrm{dom}\, g \cap A \mathrm{dom}\, f$ nonempty.*

Proof. We follow the original proof, simply observing that the value function h defined in the proof is polyhedral, by the Polyhedral algebra proposition above. Thus when the optimal value is finite, h has a subgradient at 0. ♠

We conclude this section with a result emphasizing the power of Fenchel duality for convex problems with linear constraints.

Corollary 5.1.10 (Mixed Fenchel duality) *All the conclusions of the Fenchel duality theorem (3.3.4) remain valid if the regularity condition (3.3.7) is replaced by the assumption that* $\mathrm{dom}\, g \cap A \mathrm{cont}\, f$ *is nonempty and the function* g *is polyhedral.*

Proof. Assume without loss of generality the primal optimal value

$$p = \inf_{x \in \mathbf{E}} \{f(x) + g(Ax)\} = \inf_{x \in \mathbf{E},\ r \in \mathbf{R}} \{f(x) + r \,|\, g(Ax) \leq r\}$$

is finite. By assumption there is a feasible point for the problem on the right at which the objective function is continuous, so there is an affine function $\alpha : \mathbf{E} \times \mathbf{R} \to \mathbf{R}$ minorizing the function $(x, r) \mapsto f(x) + r$ such that

$$p = \inf_{x \in \mathbf{E},\ r \in \mathbf{R}} \{\alpha(x, r) \mid g(Ax) \leq r\}$$

(see §3.3, Exercise 13(c)). Clearly α has the form $\alpha(x, r) = \beta(x) + r$ for some affine minorant β of f, so

$$p = \inf_{x \in \mathbf{E}} \{\beta(x) + g(Ax)\}.$$

Now we apply the Polyhedral Fenchel duality theorem to deduce the existence of an element ϕ of $\mathbf{Y}$ such that

$$p = -\beta^*(A^*\phi) - g^*(-\phi) \leq -f^*(A^*\phi) - g^*(-\phi) \leq p$$

(using the weak duality inequality), and the duality result follows. The calculus rules follow as before. ♠

It is interesting to compare this result with the version of Fenchel duality using the Open mapping theorem (§4.1, Exercise 9), where the assumption that g is polyhedral is replaced by surjectivity of A.

5.2 Functions of eigenvalues

Fenchel conjugacy gives a concise and beautiful avenue to many eigenvalue inequalities in classical matrix analysis. In this section we outline this approach.

The two cones $\mathbf{R}^n_+$ and $\mathbf{S}^n_+$ appear repeatedly in applications, as do their corresponding logarithmic barriers lb and ld, which we defined in §3.3. We can relate the vector and matrix examples, using the notation of §1.2, through the identities

$$(5.2.1) \qquad \delta_{\mathbf{S}^n_+} = \delta_{\mathbf{R}^n_+} \circ \lambda, \text{ and } \mathrm{ld} = \mathrm{lb} \circ \lambda.$$

We see in this section that these identities fall into a broader pattern.

Recall the function $[\cdot] : \mathbf{R}^n \to \mathbf{R}^n$ rearranges components into nonincreasing order. We say a function f on $\mathbf{R}^n$ is *symmetric* if $f(x) = f([x])$ for all vectors x in $\mathbf{R}^n$: in other words, permuting components does not change the function value. The following formula is crucial.

Theorem 5.2.2 (Spectral conjugacy) *Any function $f : \mathbf{R}^n \to [-\infty, +\infty]$ which is symmetric satisfies the formula*

$$(f \circ \lambda)^* = f^* \circ \lambda.$$

Proof. By Fan's inequality (1.2.2), any matrix Y in $\mathbf{S}^n$

satisfies the inequalities

$$\begin{aligned}
(f \circ \lambda)^*(Y) &= \sup_{X \in \mathbf{S}^n} \{\operatorname{tr}(XY) - f(\lambda(X))\} \\
&\leq \sup_X \{\lambda(X)^T \lambda(Y) - f(\lambda(X))\} \\
&\leq \sup_{x \in \mathbf{R}^n} \{x^T \lambda(Y) - f(x)\} \\
&= f^*(\lambda(Y)).
\end{aligned}$$

On the other hand, fixing a spectral decomposition $Y = U^T(\operatorname{Diag}\lambda(Y))U$ for some matrix U in $\mathbf{O}^n$ leads to the reverse inequality:

$$\begin{aligned}
& f^*(\lambda(Y)) \\
=\ & \sup_{x \in \mathbf{R}^n} \{x^T \lambda(Y) - f(x)\} \\
=\ & \sup_x \{\operatorname{tr}((\operatorname{Diag} x)UYU^T) - f(x)\} \\
=\ & \sup_x \{\operatorname{tr}(U^T(\operatorname{Diag} x)UY) - f(\lambda(U^T \operatorname{Diag} xU))\} \\
\leq\ & \sup_{X \in \mathbf{S}^n} \{\operatorname{tr}(XY) - f(\lambda(X))\} \\
=\ & (f \circ \lambda)^*(Y). \qquad \spadesuit
\end{aligned}$$

This formula, for example, makes it very easy to calculate ld^* (see the Log barriers proposition (3.3.2)), and to check the self-duality of the cone $\mathbf{S}^n_+$.

Once we can compute conjugates easily, we can also recognize closed convex functions easily, using the Fenchel biconjugation theorem (4.2.1).

Corollary 5.2.3 (Davis) *Suppose the function $f : \mathbf{R}^n \to (-\infty, +\infty]$ is symmetric. Then the 'spectral function' $f \circ \lambda$ is closed and convex if and only if f is closed and convex.*

We deduce immediately that the logarithmic barrier ld is closed and convex, as well as the function $X \mapsto \mathrm{tr}\,(X^{-1})$ on $\mathbf{S}^n_{++}$, for example.

Identifying subgradients is also easy using the conjugacy formula and the Fenchel-Young inequality (3.3.3).

Corollary 5.2.4 (Spectral subgradients) *Suppose $f : \mathbf{R}^n \to (-\infty, +\infty]$ is a symmetric function. Then for any two matrices X and Y in $\mathbf{S}^n$, the following properties are equivalent:*

(i) $Y \in \partial(f \circ \lambda)(X)$;

(ii) X *and* Y *have a simultaneous ordered spectral decomposition and satisfy* $\lambda(Y) \in \partial f(\lambda(X))$;

(iii) $X = U^T(\mathrm{Diag}\,x)U$ *and* $Y = U^T(\mathrm{Diag}\,y)U$ *for some matrix* U *in* $\mathbf{O}^n$ *and vectors* x *and* y *in* $\mathbf{R}^n$ *satisfying* $y \in \partial f(x)$.

Proof. Notice the inequalities

$$\begin{aligned}(f \circ \lambda)(X) + (f \circ \lambda)^*(Y) &= f(\lambda(X)) + f^*(\lambda(Y)) \\ &\geq \lambda(X)^T \lambda(Y) \geq \mathrm{tr}\,(XY).\end{aligned}$$

The condition $Y \in \partial(f \circ \lambda)(X)$ is equivalent to equality between the left- and right-hand-sides (and hence throughout), and the equivalence of properties (i) and (ii) follows, using Fan's inequality (1.2.1). For the remainder of the proof, see Exercise 9. ♠

Corollary 5.2.5 (Spectral differentiability) *Suppose that the function $f : \mathbf{R}^n \to (-\infty, +\infty]$ is symmetric, closed and convex. Then $f \circ \lambda$ is differentiable at a matrix X in $\mathbf{S}^n$ if and only if f is differentiable at $\lambda(X)$.*

Proof. If $\partial(f \circ \lambda)(X)$ is a singleton, so is $\partial f(\lambda(X))$, by the Spectral subgradients corollary above. Conversely, suppose $\partial f(\lambda(X))$ consists only of the vector $y \in \mathbf{R}^n$. Using Exercise 9(b), we see the components of y are nonincreasing, so by the same corollary, $\partial(f \circ \lambda)(X)$ is the nonempty convex set

$$\{U^T(\mathrm{Diag}\, y)U \mid U \in \mathbf{O}^n,\ U^T\mathrm{Diag}\,(\lambda(X))U = X\}.$$

But every element of this set has the same norm (namely $\|y\|$), so the set must be a singleton. ♠

Notice that the proof in fact shows that when f is differentiable at $\lambda(X)$ we have the formula

$$\nabla(f \circ \lambda)(X) = U^T(\mathrm{Diag}\, \nabla f(\lambda(X)))U, \quad (5.2.6)$$

for any matrix U in $\mathbf{O}^n$ satisfying $U^T(\mathrm{Diag}\,\lambda(X))U = X$.

The pattern of these results is clear: many analytic and geometric properties of the matrix function $f \circ \lambda$ parallel the corresponding properties of the underlying function f. The following exercise is another example.

Corollary 5.2.7 *Suppose the function* $f : \mathbf{R}^n \to (-\infty, +\infty]$ *is symmetric, closed and convex. Then* $f \circ \lambda$ *is essentially strictly convex (respectively, essentially smooth) if and only if* f *is likewise.*

For example, the logarithmic barrier ld is both essentially smooth and essentially strictly convex.

5.3 Duality for linear and semidefinite programming

Linear programming is the study of optimization problems involving a linear objective function subject to linear constraints. This simple optimization model has proved enormously powerful in both theory and practice, so we devote this section to deriving linear programming duality theory from our convex-analytic perspective. We contrast this theory with the corresponding results for 'semidefinite programming', a class of matrix optimization problems analogous to linear programs but involving the positive semidefinite cone.

Linear programs are inherently polyhedral, so our main development follows directly from the polyhedrality section (§5.1). But to begin, we sketch an alternative development directly from the Farkas lemma (2.2.7). Given vectors $a^1, a^2, \ldots, a^m$ and c in $\mathbf{R}^n$ and a vector b in $\mathbf{R}^m$, consider the *primal linear program*

$$\left\{\begin{array}{lrcl} \inf & \langle c, x\rangle & & \\ \text{subject to} & \langle a^i, x\rangle - b_i & \leq & 0, \quad \text{for } i = 1, 2, \ldots, m, \\ & x & \in & \mathbf{R}^n. \end{array}\right. \tag{5.3.1}$$

Denote the primal optimal value by $p \in [-\infty, +\infty]$. In the Lagrangian duality framework (§4.3), the dual problem is

$$\text{(5.3.2)} \qquad \begin{cases} \sup & -b^T\mu \\ \text{subject to} & \Sigma_{i=1}^m \mu_i a^i \;=\; -c \\ & \mu \;\in\; \mathbf{R}^m_+, \end{cases}$$

with dual optimal value $d \in [-\infty, +\infty]$. From §4.3 we know the weak duality inequality $p \geq d$. If the primal problem (5.3.1) satisfies the Slater condition then the Dual attainment theorem (4.3.7) shows $p = d$ with dual attainment when the values are finite. However, as we shall see, the Slater condition is superfluous here.

Suppose the primal value p is finite. Then it is easy to see that the 'homogenized' system of inequalities in $\mathbf{R}^{n+1}$,

$$\text{(5.3.3)} \begin{cases} \langle a^i, x\rangle - b_i z \;\leq\; 0, & \text{for } i = 1, 2, \ldots, m, \\ -z \;\leq\; 0, & \text{and} \\ \langle -c, x\rangle + pz \;>\; 0, & x \in \mathbf{R}^n, \quad z \in \mathbf{R}, \end{cases}$$

has no solution. Applying the Farkas lemma (2.2.7) to this system, we deduce there is a vector $\bar{\mu}$ in $\mathbf{R}^n_+$ and a scalar β in $\mathbf{R}_+$ satisfying

$$\sum_{i=1}^m \bar{\mu}_i(a^i, -b_i) + \beta(0, -1) = (-c, p).$$

Thus $\bar{\mu}$ is a feasible solution for the dual problem (5.3.2), with objective value at least p. The weak duality inequality now implies $\bar{\mu}$ is optimal and $p = d$. We needed no Slater condition: the assumption of a finite primal optimal value alone implies zero duality gap and dual attainment.

We can be more systematic using our polyhedral theory. Suppose that $\mathbf{Y}$ is a Euclidean space, that the map $A : \mathbf{E} \to \mathbf{Y}$ is linear, and consider cones $H \subset \mathbf{Y}$ and $K \subset \mathbf{E}$. For given elements c of $\mathbf{E}$ and b of $\mathbf{Y}$, consider the primal 'abstract linear program'

$$(5.3.4) \qquad \left\{ \begin{array}{lrcl} \inf & \langle c, x \rangle & & \\ \text{subject to} & Ax - b & \in & H, \\ & x & \in & K. \end{array} \right.$$

As usual, denote the optimal value by p. We can write this problem in Fenchel form (3.3.5) if we define functions f on $\mathbf{E}$ and g on $\mathbf{Y}$ by $f(x) = \langle c, x \rangle + \delta_K(x)$ and $g(y) = \delta_H(y - b)$. Then the Fenchel dual problem (3.3.6) is

$$(5.3.5) \qquad \left\{ \begin{array}{lrcl} \sup & \langle b, \phi \rangle & & \\ \text{subject to} & A^*\phi - c & \in & K^-, \\ & \phi & \in & -H^-, \end{array} \right.$$

with dual optimal value d. If we now apply the Fenchel duality theorem (3.3.4) in turn to problem (5.3.4), and then to problem (5.3.5) (using the Bipolar cone theorem (3.3.13)), we obtain the following general result.

Corollary 5.3.6 (Cone programming duality) *Suppose the cones H and K in problem (5.3.4) are convex.*

(a) If any of the conditions

$$\text{(i) } b \in \operatorname{int}(AK - H),$$

$$\text{(ii) } b \in AK - \operatorname{int} H, \text{ or}$$

$$\text{(iii) } b \in A(\operatorname{int} K) - H, \text{ and } \begin{array}{l} H \text{ is polyhedral or} \\ A \text{ is surjective} \end{array}$$

hold then there is no duality gap ($p = d$) and the dual optimal value d is attained if finite.

(b) Suppose H and K are also closed. If any of the conditions

$$\text{(i) } -c \in \operatorname{int}(A^*H^- + K^-),$$

$$\text{(ii) } -c \in A^*H^- + \operatorname{int} K^-, \text{ or}$$

$$\text{(iii) } -c \in A^*(\operatorname{int} H^-) + K^-, \text{ and } \begin{array}{l} K \text{ is polyhedral or} \\ A^* \text{ is surjective} \end{array}$$

hold then there is no duality gap and the primal optimal value p is attained if finite.

In both parts (a) and (b), the sufficiency of condition (iii) follows by applying the Mixed Fenchel duality corollary (5.1.10), or the Open mapping theorem (§4.1, Exercise 9). In the fully polyhedral case we obtain the following result.

Corollary 5.3.7 (Linear programming duality) *Suppose the cones H and K in the the dual pair of problems (5.3.4) and (5.3.5) are polyhedral. If either problem has finite optimal value then there is no duality gap and both problems have optimal solutions.*

Proof. We apply the Polyhedral Fenchel duality corollary (5.1.9) to each problem in turn. ♠

Our earlier result, for the linear program (5.3.1), is clearly just a special case of this corollary.

Linear programming has an interesting matrix analogue. Given matrices $A_1, A_2, \ldots, A_m$ and C in $\mathbf{S}^n_+$ and a vector b in $\mathbf{R}^m$, consider the primal *semidefinite program*

$$\left\{ \begin{array}{lrcl} \inf & \operatorname{tr}(CX) & & \\ \text{subject to} & \operatorname{tr}(A_i X) & = & b_i, \quad \text{for } i = 1, 2, \ldots, m, \\ & X & \in & \mathbf{S}^n_+. \end{array} \right. \tag{5.3.8}$$

This is a special case of the abstract linear program (5.3.4), so the dual problem is

$$(5.3.9) \quad \begin{cases} \sup & b^T\phi \\ \text{subject to} & C - \Sigma_{i=1}^m \phi_i A_i \in \mathbf{S}_+^n, \\ & \phi \in \mathbf{R}^m, \end{cases}$$

since $(\mathbf{S}_+^n)^- = -\mathbf{S}_+^n$, by the Self-dual cones proposition (3.3.11), and we obtain the following duality theorem from the general result above.

Corollary 5.3.10 (Semidefinite programming duality)
When the primal problem (5.3.8) has a positive definite feasible solution, there is no duality gap and the dual optimal value is attained when finite. On the other hand, if there is a vector ϕ in $\mathbf{R}^m$ with $C - \Sigma_i \phi_i A_i$ positive definite then once again there is no duality gap and the primal optimal value is attained when finite.

Unlike linear programming, we need a condition stronger than mere consistency to guarantee no duality gap. For example, if we consider the primal semidefinite program (5.3.8) with

$$n = 2,\ m = 1,\ C = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix},\ A_1 = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix},\ \text{and } b = 0,$$

the primal optimal value is 0 (and is attained), whereas the dual problem (5.3.9) is inconsistent.

5.4 Convex process duality

In this section we introduce the idea of a 'closed convex process'. These are set-valued maps whose graphs are closed convex cones. As such, they provide a powerful unifying formulation for the study of linear maps, convex cones, and linear programming. The exercises show the elegance of this approach in a range of applications.

Throughout this section we fix a Euclidean space $\mathbf{Y}$. For clarity, we denote the closed unit balls in $\mathbf{E}$ and $\mathbf{Y}$ by $B_{\mathbf{E}}$ and $B_{\mathbf{Y}}$ respectively. A *multifunction* $\Phi : \mathbf{E} \to \mathbf{Y}$ is a map from $\mathbf{E}$ to the set of subsets of $\mathbf{Y}$. The *domain* of Φ is the set

$$D(\Phi) = \{x \in \mathbf{E} \mid \Phi(x) \neq \emptyset\}.$$

We say Φ *has nonempty images* if its domain is $\mathbf{E}$. For any subset C of $\mathbf{E}$ we write $\Phi(C)$ for the image $\cup_{x \in C} \Phi(x)$, and the *range* of Φ is the set $R(\Phi) = \Phi(\mathbf{E})$. We say Φ is *surjective* if its range is $\mathbf{Y}$. The *graph* of Φ is the set

$$G(\Phi) = \{(x, y) \in \mathbf{E} \times \mathbf{Y} \mid y \in \Phi(x)\},$$

and we define the *inverse* multifunction $\Phi^{-1} : \mathbf{Y} \to \mathbf{E}$ by the relationship

$$x \in \Phi^{-1}(y) \Leftrightarrow y \in \Phi(x), \quad \text{for } x \text{ in } \mathbf{E} \text{ and } y \text{ in } \mathbf{Y}.$$

A multifunction is *convex*, or *closed*, or *polyhedral*, if its graph is likewise. A *process* is a multifunction whose graph is a cone. For example, we can interpret linear maps as closed convex processes in the obvious way.

Closure is one example of a variety of continuity properties of multifunctions we study in this section. We say the multifunction Φ is *LSC* at a point (x_0, y) in its graph if, for all neighbourhoods V of y, the image $\Phi(x)$ intersects V for all points x close to x_0. (In particular, x_0 must lie in $\mathrm{int}\,(D(\Phi))$.) Equivalently, for any sequence of points (x_n) approaching x_0 there is a sequence of points $y_n \in \Phi(x_n)$ approaching y. If, for x_0 in the domain, this property holds for all points y in $\Phi(x_0)$, we say Φ is *LSC at* x_0. (The notation comes from 'lower semicontinuous', a name we avoid in this context because of incompatibility with the single-valued case — see Exercise 5.)

On the other hand, we say Φ is *open* at a point (x, y_0) in its graph if, for all neighbourhoods U of x, the point y_0 lies in $\mathrm{int}\,(\Phi(U))$. (In particular, y_0 must lie in $\mathrm{int}\,(R(\Phi))$.) Equivalently, for any sequence of points (y_n) approaching y_0 there is a sequence of points (x_n) approaching x such that $y_n \in \Phi(x_n)$ for all n. If, for y_0 in the range, this property holds for all points x in $\Phi^{-1}(y_0)$, we say Φ is *open* at y_0. These properties are inverse to each other, in the following sense.

Proposition 5.4.1 (Openness and lower semicontinuity) *Any multifunction* $\Phi : \mathbf{E} \to \mathbf{Y}$ *is LSC* *at a point (x, y) in its graph if and only if Φ^{-1} is open at (y, x).*

We leave the proof as an exercise.

For convex multifunctions, openness at a point in the graph has strong global implications: the following result is another exercise.

Proposition 5.4.2 *If a convex multifunction is open at some point in its graph then it is open throughout the interior of its range.*

In particular, a convex process $\Phi : \mathbf{E} \to \mathbf{Y}$ is open at $(0, 0) \in \mathbf{E} \times \mathbf{Y}$ if and only if it is open at $0 \in \mathbf{Y}$: we just say Φ is *open at zero* (or, dually, Φ^{-1} is *LSC at zero*).

There is a natural duality for convex processes which generalizes the adjoint operation for linear maps. Specifically, for a convex process $\Phi : \mathbf{E} \to \mathbf{Y}$, we define the *adjoint* process $\Phi^* : \mathbf{Y} \to \mathbf{E}$ by

$$G(\Phi^*) = \{(\mu, \nu) \mid (\nu, -\mu) \in G(\Phi)^-\}.$$

Then an easy consequence of the Bipolar cone theorem (3.3.13) is

$$G(\Phi^{**}) = -G(\Phi),$$

providing Φ is closed. (We could define a 'lower' adjoint by the relationship $\Phi_*(\mu) = -\Phi^*(-\mu)$, in which case $(\Phi^*)_* = \Phi$.)

The language of adjoint processes is elegant and concise for many variational problems involving cones. A good example is the cone program (5.3.4). We can write this problem as

$$\inf_{x \in \mathbf{E}} \{\langle c, x\rangle \mid b \in \Psi(x)\},$$

where Ψ is the closed convex process defined by

$$\Psi(x) = \begin{cases} Ax - H, & \text{if } x \in K, \\ \emptyset, & \text{otherwise,} \end{cases} \tag{5.4.3}$$

for points c in $\mathbf{E}$, b in $\mathbf{Y}$, and closed convex cones $H \subset \mathbf{Y}$ and $K \subset \mathbf{E}$. An easy calculation shows the adjoint process is

$$\Psi^*(\mu) = \begin{cases} A^*\mu + K^-, & \text{if } \mu \in H^-, \\ \emptyset, & \text{otherwise,} \end{cases} \tag{5.4.4}$$

so we can write the dual problem (5.3.5) as

$$\sup_{\mu \in \mathbf{Y}} \{\langle b, \mu\rangle \mid -c \in \Psi^*(-\mu)\}. \tag{5.4.5}$$

Furthermore the constraint qualifications in the Cone programming duality corollary (5.3.6) become simply $b \in \operatorname{int} R(\Psi)$ and $-c \in \operatorname{int} R(\Psi^*)$.

In §1.1 we mentioned the fundamental linear-algebraic fact that the nullspace of any linear map A and the range of its adjoint satisfy the relationship

$$(A^{-1}(0))^- = R(A^*). \tag{5.4.6}$$

Our next step is to generalize this to processes. We begin with an easy lemma.

Lemma 5.4.7 *Any convex process* $\Phi : \mathbf{E} \to \mathbf{Y}$ *and subset* C *of* $\mathbf{Y}$ *satisfy* $\Phi^*(C^\circ) \subset (\Phi^{-1}(C))^\circ$.

Equality in this relationship requires more structure.

Theorem 5.4.8 (Adjoint process duality) *Let* $\Phi : \mathbf{E} \to \mathbf{Y}$ *be a convex process, and suppose the set* $C \subset \mathbf{Y}$ *is convex, with* $R(\Phi) \cap C$ *nonempty.*

(a) Either of the assumptions

(i) the multifunction $x \in \mathbf{E} \mapsto \Phi(x) - C$ *is open at zero (or, in particular,* $\operatorname{int} C$ *contains zero), or*

(ii) Φ *and* C *are polyhedral,*

imply

$$(\Phi^{-1}(C))^\circ = \Phi^*(C^\circ).$$

(b) On the other hand, if C *is compact and* Φ *is closed then*

$$(\Phi^{-1}(C))^\circ = \operatorname{cl}(\Phi^*(C^\circ)).$$

Proof. Suppose assumption (i) holds in part (a). For a fixed element ϕ of $(\Phi^{-1}(C))^\circ$, we can check that the 'value function' $v : \mathbf{Y} \to [-\infty, +\infty]$ defined, for elements y of $\mathbf{Y}$, by

$$(5.4.9) \qquad v(y) = \inf_{x \in \mathbf{E}} \{-\langle \phi, x \rangle \mid y \in \Phi(x) - C\}$$

is convex. The assumption $\phi \in (\Phi^{-1}(C))^\circ$ is equivalent to $v(0) \geq -1$, while the openness assumption implies $0 \in \mathrm{core}\,(\mathrm{dom}\, v)$. Thus v is proper, by Lemma 3.2.6, and so the Max formula (3.1.7) shows v has a subgradient $-\lambda \in \mathbf{Y}$ at 0. A simple calculation now shows $\lambda \in C^\circ$ and $\phi \in \Phi^*(\lambda)$, which, together with Lemma 5.4.7, proves the result.

If Φ and C are polyhedral, the Polyhedral algebra proposition (5.1.8) shows v is also polyhedral, so again has a subgradient, and our argument proceeds as before.

Turning to part (b), we can rewrite $\phi \in (\Phi^{-1}(C))^\circ$ as

$$(\phi, 0) \in (G(\Phi) \cap (\mathbf{E} \times C))^\circ,$$

and apply the polarity formula in §4.1, Exercise 8 to deduce

$$(\phi, 0) \in \mathrm{cl}\,(G(\Phi)^- + (0 \times C^\circ)).$$

Hence there are sequences $(\phi_n, -\rho_n)$ in $G(\Phi)^-$ and μ_n in C° with ϕ_n approaching ϕ and $\mu_n - \rho_n$ approaching 0. We deduce

$$\phi_n \in \Phi^*(\rho_n) \subset \Phi^*(C^\circ + \epsilon_n B_{\mathbf{Y}}),$$

where the real sequence $\epsilon_n = \|\mu_n - \rho_n\|$ approaches 0. Since C is bounded we know int (C°) contains 0 (by §4.1, Exercise 5), and the result follows using the the positive homogeneity of Φ^*. ♠

The nullspace/range formula (5.4.6) thus generalizes to a closed convex process Φ:

$$(\Phi^{-1}(0))^\circ = \mathrm{cl}\,(R(\Phi^*)),$$

and the closure is not required if Φ is open at zero.

We are mainly interested in using these polarity formulae to relate two 'norms' for a convex process $\Phi : \mathbf{E} \to \mathbf{Y}$. The 'lower norm'

$$\|\Phi\|_l = \inf\{r \in \mathbf{R}_{++} \mid \Phi(x) \cap rB_{\mathbf{Y}} \neq \emptyset,\ \forall x \in B_{\mathbf{E}}\}$$

quantifies Φ being LSC at zero: it is easy to check that Φ is LSC at zero if and only if its lower norm is finite. The 'upper norm'

$$\|\Phi\|_u = \inf\{r \in \mathbf{R}_{++} \mid \Phi(B_{\mathbf{E}}) \subset rB_{\mathbf{Y}}\}$$

quantifies a form of 'upper semicontinuity' (see §7.2). Clearly Φ is *bounded* (that is, bounded sets have bounded images), if and only if its upper norm is finite. Both norms generalize the *norm* of a linear map $A : \mathbf{E} \to \mathbf{Y}$, defined by

$$\|A\| = \sup\{\|Ax\| \mid \|x\| \leq 1\}.$$

Theorem 5.4.10 (Norm duality) *Any closed convex process Φ satisfies*

$$\|\Phi\|_l = \|\Phi^*\|_u.$$

Proof. For any real $r > \|\Phi\|_l$ we know $B_{\mathbf{E}} \subset \Phi^{-1}(rB_{\mathbf{Y}})$, by definition. Taking polars implies $B_{\mathbf{E}} \supset r^{-1}\Phi^*(B_{\mathbf{Y}})$, by the Adjoint process duality theorem (5.4.8), whence $\|\Phi^*\|_u < r$.

Conversely, $\|\Phi^*\|_u < r$ implies $\Phi^*(B_{\mathbf{Y}}) \subset rB_{\mathbf{E}}$. Taking polars and applying the Adjoint process duality theorem again followed by the Bipolar set theorem (4.1.5) shows $B_{\mathbf{E}} \subset r(\mathrm{cl}\,(\Phi^{-1}(B_{\mathbf{Y}})))$. But since $B_{\mathbf{Y}}$ is compact we can check $\Phi^{-1}(B_{\mathbf{Y}})$ is closed, and the result follows. ♠

The values of the upper and lower norms of course depend on the spaces $\mathbf{E}$ and $\mathbf{Y}$. Our proof of the Norm duality theorem above shows that it remains valid when $B_{\mathbf{E}}$ and $B_{\mathbf{Y}}$ denote unit balls for arbitrary norms (see §4.1, exercise 2), providing we replace them by their polars $B_{\mathbf{E}}^\circ$ and $B_{\mathbf{Y}}^\circ$ in the definition of $\|\Phi^*\|_u$.

The next result is an immediate consequence of the Norm duality theorem.

Corollary 5.4.11 *A closed convex process is LSC at zero if and only if its adjoint is bounded.*

We are now ready to prove the main result of this section.

Theorem 5.4.12 (Open mapping) *The following properties of a closed convex process* Φ *are equivalent:*

(a) Φ *is open at zero;*

(b) $(\Phi^*)^{-1}$ *is bounded.*

(c) Φ *is surjective.*

Proof. The equivalence of parts (a) and (b) is just Corollary 5.4.11 (after taking inverses and observing the identity $G((\Phi^*)^{-1}) = -G((\Phi^{-1})^*)$. Part (a) clearly implies part (c), so it remains to prove the converse. But if Φ is surjective then we know

$$Y = \bigcup_{n=1}^{\infty} \Phi(nB_{\mathbf{E}}) = \bigcup_{n=1}^{\infty} n\Phi(B_{\mathbf{E}}),$$

so 0 lies in the core, and hence the interior, of the convex set $\Phi(B_{\mathbf{E}})$. Thus Φ is open at zero. ♠

Taking inverses gives the following equivalent result.

Theorem 5.4.13 (Closed graph) *The following properties of a closed convex process* Φ *are equivalent:*

(a) Φ *is LSC at zero;*

(b) Φ^* *is bounded.*

(c) Φ *has nonempty images.*

Chapter 6

The Variational Principle

6.1 An introduction to metric regularity

Our main optimization models so far are inequality-constrained. A little thought shows our techniques are not useful for equality-constrained problems like

$$\inf\{f(x) \mid h(x) = 0\}.$$

In this section we study such problems by linearizing the feasible region $h^{-1}(0)$, using the contingent cone.

Throughout this section we consider an open set $U \subset \mathbf{E}$, a closed set $S \subset U$, a Euclidean space $\mathbf{Y}$, and a continuous map $h : U \to \mathbf{Y}$. The restriction of h to S we denote $h|_S$. The following easy result (see Exercise 1) suggests our direction.

Proposition 6.1.1 *If h is Fréchet differentiable at* *the point* $x \in U$ *then*

$$K_{h^{-1}(h(x))}(x) \subset N(\nabla h(x)).$$

Our aim in this section is to find conditions guaranteeing equality in this result.

Our key tool is the next result. It states that if a closed function attains a value close to its infimum at some point, then a nearby point minimizes a slightly perturbed function.

Theorem 6.1.2 (Ekeland variational principle) *Suppose the function* $f : \mathbf{E} \to (-\infty, +\infty]$ *is closed and the point* $x \in \mathbf{E}$ *satisfies* $f(x) \le \inf f + \epsilon$, *for some real* $\epsilon > 0$. *Then for any real* $\lambda > 0$ *there is a point* $v \in \mathbf{E}$ *satsifying the conditions*

(a) $\|x - v\| \le \lambda$,

(b) $f(v) \le f(x)$, *and*

(c) v *is the unique minimizer of the function* $f(\cdot) + (\epsilon/\lambda)\|\cdot - v\|$.

Proof. We can assume f is proper, and by assumption it is bounded below. Since the function

$$f(\cdot) + \frac{\epsilon}{\lambda}\|\cdot - x\|$$

therefore has compact level sets, its set of minimizers $M \subset \mathbf{E}$ is nonempty and compact. Choose a minimizer v for f on M. Then for points $z \neq v$ in M we know

$$f(v) \leq f(z) < f(z) + \frac{\epsilon}{\lambda}\|z - v\|,$$

while for z not in M we have

$$f(v) + \frac{\epsilon}{\lambda}\|v - x\| < f(z) + \frac{\epsilon}{\lambda}\|z - x\|.$$

Part (c) follows by the triangle inequality. Since v lies in M we have

$$f(z) + \frac{\epsilon}{\lambda}\|z - x\| \geq f(v) + \frac{\epsilon}{\lambda}\|v - x\| \text{ for all } z \text{ in } \mathbf{E}.$$

Setting $z = x$ shows the inequalities

$$f(v) + \epsilon \geq \inf f + \epsilon \geq f(x) \geq f(v) + \frac{\epsilon}{\lambda}\|v - x\|.$$

Properties (a) and (b) follow. ♠

As we shall see, a precise calculation of the contingent cone $K_{h^{-1}(h(x))}(x)$ requires us first to bound the distance of a point z to the set $h^{-1}(h(x))$ in terms of the function value $h(z)$. This leads us to the notion of 'metric regularity'. In this section we present a somewhat simplified version of this idea, which suffices for most of our purposes: we defer a more comprehensive treatment to a later section. We say h is *weakly metrically regular* on S at the point x in S if there is a real constant k such that

$$d_{S\cap h^{-1}(h(x))}(z) \le k\|h(z)-h(x)\| \text{ for all } z \text{ in } S \text{ close to } x.$$

Lemma 6.1.3 *Suppose* $0 \in S$ *and* $h(0) = 0$*. If* h *is not weakly metrically regular on* S *at* 0*, there is a sequence* $v_r \to 0$ *in* S *such that* $h(v_r) \neq 0$ *for all* r*, and a strictly positive sequence* $\delta_r \downarrow 0$ *such that the function*

$$\|h(\cdot)\| + \delta_r\| \cdot - v_r\|$$

is minimized on S *at* v_r*.*

Proof. By definition there is a sequence $x_r \to 0$ in S such that

$$(6.1.4) \qquad d_{S\cap h^{-1}(0)}(x_r) > r\|h(x_r)\| \text{ for all } r.$$

For each index r we apply the Ekeland principle with

$$f = \|h\| + \delta_S, \quad \epsilon = \|h(x_r)\|, \quad \lambda = \min\{r\epsilon, \sqrt{\epsilon}\},$$

and $x = x_r$ to deduce the existence of a point v_r in S such that

(a) $\|x_r - v_r\| \le \min\left\{r\|h(x_r)\|, \sqrt{\|h(x_r)\|}\right\}$, and

(c) v_r minimizes the function

$$\|h(\cdot)\| + \max\left\{r^{-1}, \sqrt{\|h(x_r)\|}\right\} \| \cdot - v_r\|$$

on S.

Property (a) shows $v_r \to 0$, while (c) reveals the minimizing property of v_r. Finally, inequality (6.1.4) and property (a) prove $h(v_r) \neq 0$. ♠

We can now present a convenient condition for weak metric regularity.

Theorem 6.1.5 (Surjectivity and metric regularity) *If h is strictly differentiable at the point x in S and*

$$\nabla h(x)(T_S(x)) = \mathbf{Y}$$

then h is weakly metrically regular on S at x.

Proof. Notice first h is locally Lipschitz around x. Without loss of generality, suppose $x = 0$ and $h(0) = 0$. If h is not weakly metrically regular on S at 0 then by Lemma 6.1.3 there is a sequence $v_r \to 0$ in S such that $h(v_r) \neq 0$ for all r, and a real sequence $\delta_r \downarrow 0$ such that the function

$$\|h(\cdot)\| + \delta_r \| \cdot - v_r\|$$

is minimized on S at v_r. Denoting the local Lipschitz constant by L, we deduce, from the sum rule and the Exact penalization proposition for Clarke subgradients, the condition

$$0 \in \partial_\circ(\|h\|)(v_r) + \delta_r B + L\partial_\circ d_S(v_r).$$

Hence there are elements u_r of $\partial_\circ(\|h\|)(v_r)$ and w_r of $L\partial_\circ d_S(v_r)$ such that $u_r + w_r$ approaches 0.

By choosing a subsequence we can assume

$$\|h(v_r)\|^{-1}h(v_r) \to y \neq 0$$

and an exercise then shows $u_r \to (\nabla h(0))^* y$. Since the Clarke subdifferential is closed at 0 we deduce

$$-(\nabla h(0))^* y \in L\partial_\circ d_S(0) \subset N_S(0).$$

But by assumption there is a nonzero element p of $T_S(0)$ such that $\nabla h(0)p = -y$, so we arrive at the contradiction

$$0 \geq \langle p, -(\nabla h(0))^* y\rangle = \langle \nabla h(0)p, -y\rangle = \|y\|^2 > 0.$$

♠

We can now prove the main result of this section.

Theorem 6.1.6 (Liusternik) *If h is strictly differentiable at the point x and $\nabla h(x)$ is surjective, then the set $h^{-1}(h(x))$ is tangentially regular at x and*

$$K_{h^{-1}(h(x))}(x) = N(\nabla h(x)).$$

Proof. Assume without loss of generality $x = 0$ and $h(0) = 0$. In light of Proposition 6.1.1, it suffices to prove

$$N(\nabla h(0)) \subset T_{h^{-1}(0)}(0).$$

Fix any element p of $N(\nabla h(0))$ and consider a sequence $x^r \to 0$ in $h^{-1}(0)$ and $t_r \downarrow 0$ in $\mathbf{R}_{++}$. The previous result shows h is weakly metrically regular at 0, so there is a constant k such that

$$d_{h^{-1}(0)}(x^r + t_r p) \le k\|h(x^r + t_r p)\|$$

holds for all large r, and hence there are points z^r in $h^{-1}(0)$ satisfying

$$\|x^r + t_r p - z^r\| \le k\|h(x^r + t_r p)\|.$$

If we define directions $p^r = t_r^{-1}(z^r - x^r)$ then clearly the points $x^r + t_r p^r$ lie in $h^{-1}(0)$ for large r, and since

$$\begin{aligned}
\|p - p^r\| &= \|x^r + t_r p - z^r\|/t_r \\
&\le k\|h(x^r + t_r p) - h(x^r)\|/t_r \\
&\to k\|(\nabla h(0))p\| \\
&= 0,
\end{aligned}$$

we deduce $p \in T_{h^{-1}(0)}$. ♠

Chapter 7

Fixed points

7.1 Brouwer's fixed point theorem

Many questions in optimization and analysis reduce to solving a nonlinear equation $h(x) = 0$, for some function $h : \mathbf{E} \to \mathbf{E}$. Equivalently, if we define another map $f = I - h$ (where I is the identity map), we seek a point x in $\mathbf{E}$ satisfying $f(x) = x$: we call x a *fixed point* of f.

The most potent fixed point existence theorems fall into three categories: 'geometric' results, devolving from the Banach contraction principle (which we state below), order-theoretic results (to which we briefly return in §7.3), and 'topological' results, for which the prototype is the theorem of Brouwer forming the main body of this section. We begin with Banach's result.

Given a set $C \subset \mathbf{E}$ and a continuous *self map* $f : C \to C$, we ask whether f has a fixed point. We call f a *contraction* if there is a real constant $\gamma_f < 1$ such that

(7.1.1) $$\|f(x) - f(y)\| \leq \gamma_f \|x - y\| \text{ for all } x, y \in C.$$

Theorem 7.1.2 (Banach contraction) *Any contraction on a closed subset of* $\mathbf{E}$ *has a unique fixed point.*

Proof. Suppose the set $C \subset \mathbf{E}$ is closed and the function $f : C \to C$ satisfies the contraction condition (7.1.1). We apply the Ekeland variational principle (6.1.2) to the function

$$z \in \mathbf{E} \mapsto \begin{cases} \|z - f(z)\|, & \text{if } z \in C, \\ +\infty, & \text{otherwise,} \end{cases}$$

at an arbitrary point x in C, with the choice of constants

$$\epsilon = \|x - f(x)\| \text{ and } \lambda = \frac{\epsilon}{1 - \gamma_f}.$$

This shows there is a point v in C satisfying

$$\|v - f(v)\| < \|z - f(z)\| + (1 - \gamma_f)\|z - v\|$$

for all points $z \neq v$ in C. Hence v is a fixed point, since otherwise choosing $z = f(v)$ gives a contradiction. The uniqueness is easy. ♠

What if the map f is not a contraction? A very useful weakening of the notion is the idea of a *nonexpansive* map, which is to say a self map f satisfying

$$\|f(x) - f(y)\| \leq \|x - y\| \quad \text{for all } x, y$$

(see Exercise 2). A nonexpansive map on a nonempty compact set or a nonempty closed convex set may not have a fixed point, as simple examples like translations on $\mathbf{R}$ or rotations of the unit circle show. On the other hand, a straightforward argument using the Banach contraction theorem shows this cannot happen if the set is nonempty, compact *and* convex. However, in this case we have the following more fundamental result.

Theorem 7.1.3 (Brouwer) *Any continuous self map of a nonempty compact convex subset of* $\mathbf{E}$ *has a fixed point.*

In this section we present an 'analyst's approach' to Brouwer's theorem. We use the two following important analytic tools, concerning $C^{(1)}$ (continuously differentiable) functions on the closed unit ball $B \subset \mathbf{R}^n$.

Theorem 7.1.4 (Stone-Weierstrass) *For any continuous map* $f : B \to \mathbf{R}^n$*, there is a sequence of* $C^{(1)}$ *maps* $f_r : B \to \mathbf{R}^n$ *converging uniformly to* f*.*

An easy exercise shows that, in this result, if f is a self map then we can assume each f_r is also a self map.

Theorem 7.1.5 (Change of variable) *Suppose that the set $W \subset \mathbf{R}^n$ is open and that the $C^{(1)}$ map $g : W \to \mathbf{R}^n$ is one-to-one with ∇g invertible throughout W. Then the image $g(W)$ is open, with measure*

$$\int_W |\det \nabla g|.$$

We also use the elementary topological fact that the open unit ball int B is *connected*: that is, it cannot be written as the disjoint union of two nonempty open sets.

The key step in our argument is the following topological result.

Theorem 7.1.6 (Retraction) *The unit sphere S is not a $C^{(1)}$* **retract** *of the unit ball B: that is, there is no $C^{(1)}$ map from B to S whose restriction to S is the identity.*

Proof. Suppose there is such a retraction map $p : B \to S$. For real t in $[0,1]$, define a self map of B by $p_t = tp + (1-t)I$. As a function of the variables $x \in B$ and t, the function $\det \nabla p_t(x)$ is continuous, and hence strictly positive for small t. Furthermore, p_t is one-to-one for small t (see Exercise 7).

If we denote the open unit ball $B \setminus S$ by U, then the change of variables theorem above shows, for small t, that $p_t(U)$ is open, with measure

$$(7.1.7) \qquad \nu(t) = \int_U \det \nabla p_t.$$

On the other hand, by compactness, $p_t(B)$ is a closed subset of B, and we also know $p_t(S) = S$. A little manipulation now shows we can write U as a disjoint union of two open sets:

$$(7.1.8) \qquad U = (p_t(U) \cap U) \cup (p_t(B)^c \cap U).$$

The first set is nonempty, since $p_t(0) = tp(0) \in U$. But as we observed, U is connected, so the second set must be empty, which shows $p_t(B) = B$. Thus the function $\nu(t)$ defined by equation (7.1.7) equals the volume of the unit ball B for all small t.

However, as a function of $t \in [0, 1]$, $\nu(t)$ is a polynomial, so it must be constant. Since p is a retraction we know that all points x in U satisfy $\|p(x)\|^2 = 1$. Differentiating implies $(\nabla p(x))p(x) = 0$, from which we deduce $\det \nabla p(x) = 0$, since $p(x)$ is nonzero. Thus $\nu(1)$ is zero, which is a contradiction. ♠

Proof of Brouwer's theorem Consider first a $C^{(1)}$ self map f on the unit ball B. Suppose f has no fixed point. A straightforward exercise shows there are unique functions $\alpha : B \to \mathbf{R}_+$ and $p : B \to S$ satisfying the relationship

(7.1.9) $p(x) = x + \alpha(x)(x - f(x))$, for all x in B.

Geometrically, $p(x)$ is the point where the line extending from the point $f(x)$ through the point x meets the unit sphere S. In fact p must then be a $C^{(1)}$ retraction, contradicting the retraction theorem above. Thus we have proved that any $C^{(1)}$ self map of B has a fixed point.

Now suppose the function f is just continuous. By the Stone-Weierstrass theorem (7.1.4), there is a sequence of $C^{(1)}$ maps $f_r : B \to \mathbf{R}^n$ converging uniformly to f, and by Exercise 4 we can assume each f_r is a self map. Our argument above shows each f_r has a fixed point x^r. Since B is compact, the sequence (x^r) has a subsequence converging to some point x in B, which it is easy to see must be a fixed point of f. So any continuous self map of B has a fixed point.

Finally, consider a nonempty compact convex set $C \subset \mathbf{E}$ and a continuous self map g on C. Just as in our proof of Minkowski's theorem (4.1.8), we may as well assume C has nonempty interior. Thus there is a *homeomorphism* (a continuous onto map with continuous inverse) $h : C \to B$ — see Exercise 11 . Since $h \circ g \circ h^{-1}$ is a continuous self map of B, our argument above shows it has a fixed point x in B, and therefore $h^{-1}(x)$ is a fixed point of g. ♠

7.2 Selection results and the Kakutani-Fan fixed point theorem

The Brouwer fixed point theorem in the previous section concerns functions from a nonempty compact convex set to itself. In optimization, as we have already seen in §5.4, it may be convenient to broaden our language to consider *multifunctions* Ω from the set to itself and seek a *fixed point* — a point x satisfying $x \in \Omega(x)$.

To begin this section we summarize some definitions for future reference. We consider a subset $K \subset \mathbf{E}$, a Euclidean space $\mathbf{Y}$, and a multifunction $\Omega : K \to \mathbf{Y}$. We say Ω is *USC* at a point x in K if every open set U containing $\Omega(x)$ also contains $\Omega(z)$ for all points z in K close to x. Equivalently, for any sequence of points (x_n) in K approaching x, any sequence of elements $y_n \in \Omega(x_n)$, is eventually close to $\Omega(x)$. If Ω is USC at every point in K we simply call it *USC*. On the other hand, as in §5.4, we say Ω is *LSC* if, for every x in K, every neighbourhood V of any point in $\Omega(x)$ intersects $\Omega(z)$ for all points z in K close to x.

We refer to the sets $\Omega(x)$ $(x \in K)$ as the *images* of Ω. The multifunction Ω is a *cusco* if it is USC with nonempty compact convex images. Clearly such multifunctions are *locally bounded*: any point in K has a neighbourhood whose image is bounded. Cuscos appear in several important optimization contexts. For example, the Clarke subdifferential of a locally Lipschitz function is a cusco (see Exercise 5).

To see another important class of examples we need a further definition. We say a multifunction $\Phi : \mathbf{E} \to \mathbf{E}$ is *monotone* if it satisfies the condition

$$\langle u - v, x - y \rangle \geq 0 \quad \text{whenever } u \in \Phi(x) \text{ and } v \in \Phi(y).$$

In particular, any (not necessarily self-adjoint) positive semidefinite linear operator is monotone, as is the subdifferential of any convex function. One multifunction *contains* another if the graph of the first contains the graph of the second. We say a monotone multifunction is *maximal* if the only monotone multifunction containing it is itself. The subdifferentials of closed proper convex functions are examples (see Exercise 16). Zorn's lemma (which lies outside our immediate scope) shows any monotone multifunction is contained in a maximal monotone multifunction.

Theorem 7.2.1 (Maximal monotonicity) *Any maximal monotone multifunction is a cusco on the interior of its domain.*

Proof. See Exercise 16. ♠

Maximal monotone multifunctions in fact have to be single-valued *generically*, that is on sets which are 'large' in a topological sense, specifically on a dense set which is a 'G_δ' (a countable intersection of open sets) — see Exercise 17.

Returning to our main theme, the central result of this section extends Brouwer's theorem to the multifunction case.

Theorem 7.2.2 (Kakutani-Fan) *If the set $C \subset \mathbf{E}$ is nonempty, compact and convex, then any cusco $\Omega : C \to C$ has a fixed point.*

Before we prove this result, we outline a little more topology. A *cover* of a set $K \subset \mathbf{E}$ is a collection of sets in $\mathbf{E}$ whose union contains K. The cover is *open* if each set in the collection is open. A *subcover* is just a subcollection of the sets which is also a cover. The following result, which we state as a theorem, is in truth the definition of compactness in spaces more general than $\mathbf{E}$.

Theorem 7.2.3 (General definition of compactness)

Any open cover of a compact set in $\mathbf{E}$ has a finite subcover.

Given a finite open cover $\{O_1, O_2, \ldots, O_m\}$ of a set $K \subset \mathbf{E}$, a *partition of unity subordinate to* this cover is a set of continuous functions $p_1, p_2, \ldots, p_m : K \to \mathbf{R}_+$ whose sum is identically 1 and satisfying $p_i(x) = 0$ for all points x outside O_i (for each index i). We outline the proof of the next result, a central topological tool, in the exercises.

Theorem 7.2.4 (Partition of unity) *There is a partition of unity subordinate to any finite open cover of a compact subset of* $\mathbf{E}$.

Besides fixed points, the other main theme of this section is the idea of a *continuous selection* of a multifunction Ω on a set $K \subset \mathbf{E}$, by which we mean a continuous map f on K satisfying $f(x) \in \Omega(x)$ for all points x in K. The central step in our proof of the Kakutani-Fan theorem is the following 'approximate selection' theorem.

Theorem 7.2.5 (Cellina) *Given any compact set $K \subset \mathbf{E}$, suppose the multifunction $\Omega : K \to \mathbf{Y}$ is USC with nonempty convex images. Then for any real $\epsilon > 0$ there is a continuous map $f : K \to \mathbf{Y}$ which is an 'approximate selection' of Ω :*

$$d_{G(\Omega)}(x, f(x)) < \epsilon \quad \text{for all points } x \text{ in } K. \tag{7.2.6}$$

Furthermore the range of f is contained in the convex hull of the range of Ω.

Proof. We can assume the norm on $\mathbf{E} \times \mathbf{Y}$ is given by

$$\|(x, y)\|_{\mathbf{E}\times\mathbf{Y}} = \|x\|_{\mathbf{E}} + \|y\|_{\mathbf{Y}} \text{ for all } x \in \mathbf{E} \text{ and } y \in \mathbf{Y}$$

(since all norms are equivalent — see §4.1, Exercise 2). Now, since Ω is USC, for each point x in K there is a real δ_x in the interval $(0, \epsilon/2)$ satisfying

$$\Omega(x + \delta_x B_{\mathbf{E}}) \subset \Omega(x) + \frac{\epsilon}{2} B_{\mathbf{Y}}.$$

Since the sets $x + (\delta_x/2)\mathrm{int}\, B_{\mathbf{E}}$ (as the point x ranges over K) comprise an open cover of the compact set K, there is a finite subset $\{x_1, x_2, \ldots, x_m\}$ of K with the sets $x_i + (\delta_i/2)\mathrm{int}\, B_{\mathbf{E}}$ comprising a finite subcover (where δ_i is shorthand for δ_{x_i} for each index i).

Theorem 7.2.4 shows there is a partition of unity $p_1, p_2, \ldots, p_m : K \to \mathbf{R}_+$ subordinate to this subcover. We now construct our desired approximate selection f by choosing a point y_i from $\Omega(x_i)$ for each i and defining

$$(7.2.7) \quad f(x) = \sum_{i=1}^{m} p_i(x) y_i, \quad \text{for all points } x \text{ in } K.$$

Fix any point x in K and define the set $I = \{i \,|\, p_i(x) \neq 0\}$. By definition, x satisfies $\|x - x_i\| < \delta_i/2$ for each i in I. If we choose an index j in I maximizing δ_j, the triangle inequality shows $\|x_j - x_i\| < \delta_j$, whence we deduce the inclusions

$$y_i \in \Omega(x_i) \subset \Omega(x_j + \delta_j B_{\mathbf{E}}) \subset \Omega(x_j) + \frac{\epsilon}{2} B_{\mathbf{Y}}$$

for all i in I. In other words, for each i in I we know $d_{\Omega(x_j)}(y_i) \leq \epsilon/2$. Since the distance function is convex, equation (7.2.7) shows $d_{\Omega(x_j)}(f(x)) \leq \epsilon/2$. Since we also know $\|x - x_j\| < \epsilon/2$, this proves inequality (7.2.6). The final claim follows immediately from equation (7.2.7). ♠

Proof of the Kakutani-Fan theorem With the assumption of the theorem, Cellina's result above shows, for each positive integer r, there is a continuous self map f_r of C satisfying

$$d_{G(\Omega)}(x, f_r(x)) < \frac{1}{r} \text{ for all points } x \text{ in } C.$$

By Brouwer's theorem (7.1.3), each f_r has a fixed point x^r in C, which therefore satisfies

$$d_{G(\Omega)}(x^r, x^r) < \frac{1}{r} \text{ for each } r.$$

Since C is compact, the sequence (x^r) has a convergent subsequence, and its limit must be a fixed point of Ω because Ω is closed, by Exercise 3(c) (Closed versus USC). ♠

In the next section we describe some variational applications of the Kakutani-Fan theorem. But we end this section with an *exact* selection theorem parallel to Cellina's result but assuming a LSC rather than an USC multifunction.

Theorem 7.2.8 (Michael) *Given any closed set* $K \subset \mathbf{E}$*, suppose the multifunction* $\Omega : K \to \mathbf{Y}$ *is LSC with nonempty closed convex images. Then, given any point* $(\bar{x}, \bar{y})$ *in* $G(\Omega)$*, there is a continuous selection* f *of* Ω *satisfying* $f(\bar{x}) = \bar{y}$*.*

We outline the proof in the exercises.

7.3 Variational inequalities

At the very beginning of this book we considered the problem of minimizing a differentiable function $f : \mathbf{E} \to \mathbf{R}$ over a convex set $C \subset \mathbf{E}$. A necessary optimality condition for a point x_0 in C to be a local minimizer is

(7.3.1) $\langle \nabla f(x_0), x - x_0 \rangle \geq 0$ for all points x in C,

or equivalently

$$0 \in \nabla f(x_0) + N_C(x_0).$$

If the function f is convex instead of differentiable, the necessary and sufficient condition for optimality (assuming a constraint qualification) is

$$0 \in \partial f(x_0) + N_C(x_0),$$

and there are analogous nonsmooth necessary conditions.

We call problems like (7.3.1) 'variational inequalities'. Let us fix a multifunction $\Omega : C \to \mathbf{E}$. In this section we use the fixed point theory we have developed to study the *multivalued variational inequality*

$VI(\Omega, C)$: Find points x_0 in C and y_0 in $\Omega(x_0)$ satisfying $\langle y_0, x - x_0 \rangle \geq 0$ for all points x in C.

A more concise way to write the problem is:

Find a point x_0 in C satisfying $0 \in \Omega(x_0) + N_C(x_0)$.
(7.3.2)

Suppose the set C is closed, convex and nonempty. Recall that the projection $P_C : \mathbf{E} \to C$ is the (continuous) map which sends points in $\mathbf{E}$ to their unique nearest points in C (see §2.1, Exercise 8). Using this notation we can also write the variational inequality as a fixed point problem:

(7.3.3) Find a fixed point of $P_C \circ (I - \Omega) : C \to C$.

This reformulation is useful if the multifunction Ω is single-valued, but less so in general because the composition will often not have convex images.

A more versatile approach is to define the (multivalued) *normal mapping* $\Omega_C = (\Omega \circ P_C) + I - P_C$, and repose the problem as:

(7.3.4) Find a point $\bar{x}$ in $\mathbf{E}$ satisfying $0 \in \Omega_C(\bar{x})$;

then setting $x_0 = P_C(\bar{x})$ gives a solution to the original problem. Equivalently, we could phrase this as:

(7.3.5) Find a fixed point of $(I - \Omega) \circ P_C : \mathbf{E} \to \mathbf{E}$.

As we shall see, this last formulation lets us immediately use the fixed point theory of the previous section.

The basic result guaranteeing the existence of solutions to variational inequalities is the following.

Theorem 7.3.6 (Solvability of variational inequalities) *If the subset C of $\mathbf{E}$ is compact, convex and nonempty, then for any cusco $\Omega : C \to \mathbf{E}$ the variational inequality $VI(\Omega, C)$ has a solution.*

Proof. We in fact prove **Theorem 7.3.6 is equivalent to the Kakutani-Fan fixed point theorem (7.2.2)**.

When Ω is a cusco its range $\Omega(C)$ is compact — we outline the proof in §7.2, Exercise 6. We can easily check that the multifunction $(I-\Omega)\circ P_C$ is also a cusco, because the projection P_C is continuous. Since this multifunction maps the compact convex set $\operatorname{conv}(C - \Omega(C))$ into itself, the Kakutani-Fan theorem shows it has a fixed point, which, as we have already observed, implies the solvability of $VI(\Omega, C)$.

Conversely, suppose the set $C \subset \mathbf{E}$ is nonempty, compact and convex. For any cusco $\Omega : C \to C$, the Solvability theorem (7.3.6) implies we can solve the variational inequality $VI(I - \Omega, C)$, so there are points x_0 in C and z_0 in $\Omega(x_0)$ satisfying

$$\langle x_0 - z_0, x - x_0 \rangle \geq 0 \ \text{ for all points } x \text{ in } C.$$

Setting $x = z_0$ shows $x_0 = z_0$, so x_0 is a fixed point. ♠

An elegant application is von Neumann's minimax theorem, which we proved by a Fenchel duality argument in §4.2, Exercise 16. Consider Euclidean spaces $\mathbf{Y}$ and $\mathbf{Z}$, nonempty compact convex subsets $F \subset \mathbf{Y}$ and $G \subset \mathbf{Z}$, and a linear map $A : \mathbf{Y} \to \mathbf{Z}$. If we define a function $\Omega : F \times G \to \mathbf{Y} \times \mathbf{Z}$ by $\Omega(y, z) = (-A^* z, Ay)$, then it is easy to see that a point (y_0, z_0) in $F \times G$ solves the variational inequality $VI(\Omega, F \times G)$ if and only if it is a *saddlepoint*:

$$\langle z_0, Ay \rangle \le \langle z_0, Ay_0 \rangle \le \langle z, Ay_0 \rangle \quad \text{for all } y \in F,\ z \in G.$$

In particular, by the Solvability of variational inequalities theorem, there exists a saddlepoint, so

$$\min_{z \in G} \max_{y \in F} \langle z, Ay \rangle = \max_{y \in F} \min_{z \in G} \langle z, Ay \rangle.$$

Many interesting variational inequalities involve a noncompact set C. In such cases we need to impose a growth condition on the multifunction to guarantee solvability. The following result is an example.

Theorem 7.3.7 (Noncompact variational inequalities) *If the subset C of $\mathbf{E}$ is nonempty, closed and convex, and the cusco $\Omega : C \to \mathbf{E}$ is* **coercive**, *that is, it satisfies the condition*

$$\liminf_{\|x\| \to \infty,\ x \in C} \inf \langle x, \Omega(x) + N_C(x) \rangle > 0, \tag{7.3.8}$$

then the variational inequality $VI(\Omega, C)$ has a solution.

Proof. For any large integer r, we can apply the solvability theorem (7.3.6) to the variational inequality $VI(\Omega, C \cap rB)$ to find a point x_r in $C \cap rB$ satisfying

$$\begin{aligned} 0 &\in \Omega(x_r) + N_{C \cap rB}(x_r) \\ &= \Omega(x_r) + N_C(x_r) + N_{rB}(x_r) \\ &\subset \Omega(x_r) + N_C(x_r) + \mathbf{R}_+ x_r \end{aligned}$$

(using §3.3, Exercise 10). Hence for all large r, the point x_r satisfies

$$\inf \langle x_r, \Omega(x_r) + N_C(x_r) \rangle \leq 0.$$

This sequence of points (x_r) must therefore remain bounded, by the coercivity condition (7.3.8), and so x_r lies in $\operatorname{int} rB$ for large r and hence satisfies $0 \in \Omega(x_r) + N_C(x_r)$, as required. ♠

A straightforward exercise shows in particular that the growth condition (7.3.8) holds whenever the cusco Ω is defined by $x \in \mathbf{R}^n \mapsto x^T A x$ for a matrix A in $\mathbf{S}^n_{++}$.

The most important example of a noncompact variational inequality is the case when the set C is a closed convex cone $S \subset \mathbf{E}$. In this case $VI(\Omega, S)$ becomes the *multivalued complementarity problem*:

$$\text{Find points } x_0 \text{ in } S \text{ and } y_0 \text{ in } \Omega(x_0) \cap (-S^-) \text{ satisfying } \langle x_0, y_0 \rangle = 0. \tag{7.3.9}$$

As a particular example, we consider the dual pair of abstract linear programs (5.3.4) and (5.3.5):

$$\begin{cases} \inf & \langle c, z \rangle \\ \text{subject to} & Az - b \in H, \\ & z \in K, \end{cases} \tag{7.3.10}$$

(where $\mathbf{Y}$ is a Euclidean space, the map $A : \mathbf{E} \to \mathbf{Y}$ is linear, the cones $H \subset \mathbf{Y}$ and $K \subset \mathbf{E}$ are closed and convex, and b and c are given elements of $\mathbf{Y}$ and $\mathbf{E}$ respectively), and

$$\begin{cases} \sup & \langle b, \phi \rangle \\ \text{subject to} & A^*\phi - c \in K^-, \\ & \phi \in -H^-. \end{cases} \tag{7.3.11}$$

As usual, we denote the corresponding primal and dual optimal values by p and d. We consider the corresponding variational inequality on the space $\mathbf{E} \times \mathbf{Y}$,

$$\begin{array}{c} VI(\Omega, K \times (-H^-)), \quad \text{where} \\ \Omega(z, \phi) = (c - A^*\phi, Ax - b). \end{array} \tag{7.3.12}$$

Theorem 7.3.13 (Linear programming and variational inequalities) *Any solution of the above variational inequality (7.3.12) consists of a pair of optimal solutions for the linear programming dual pair*

(7.3.10) and (7.3.11). The converse is also true, providing there is no duality gap ($p = d$).

We leave the proof as an exercise.

Notice that the linear map appearing in the above example, $M : \mathbf{E} \times \mathbf{Y} \to \mathbf{E} \times \mathbf{Y}$ defined by $M(z, \phi) = (-A^*\phi, Az)$, is monotone. We study monotone complementarity problems further in Exercise 7.

To end this section we return to the complementarity problem (7.3.9) in the special case where $\mathbf{E}$ is $\mathbf{R}^n$, the cone S is $\mathbf{R}^n_+$, and the multifunction Ω is single-valued: $\Omega(x) = \{F(x)\}$ for all points x in $\mathbf{R}^n_+$. In other words, we consider the following problem:

$$\text{Find a point } x_0 \text{ in } \mathbf{R}^n_+ \text{ satisfying } F(x_0) \in \mathbf{R}^n_+$$
$$\text{and } \langle x_0, F(x_0) \rangle = 0.$$

The lattice operation $\wedge$ is defined on $\mathbf{R}^n$ by $(x \wedge y)_i = \min\{x_i, y_i\}$ for points x and y in $\mathbf{R}^n$ and each index i. With this notation we can rewrite the above problem as an *order complementarity problem*:

$$OCP(F): \qquad \text{Find a point } x_0 \text{ in } \mathbf{R}^n_+ \text{ satisfying}$$
$$x_0 \wedge F(x_0) = 0.$$

The map $x \in \mathbf{R}^n \mapsto x \wedge F(x) \in \mathbf{R}^n$ is sometimes amenable to fixed point methods.

As an example, let us fix a real $\alpha > 0$, a vector $q \in \mathbf{R}^n$, and an $n \times n$ matrix P with nonnegative entries, and define the map $F : \mathbf{R}^n \to \mathbf{R}^n$ by $F(x) = \alpha x - Px + q$. Then the complementarity problem $OCP(F)$ is equivalent to finding a fixed point of the map $\Phi : \mathbf{R}^n \to \mathbf{R}^n$ defined by

$$\Phi(x) = \frac{1}{\alpha}(0 \vee (Px - q)), \tag{7.3.14}$$

a problem which can be solved iteratively — see Exercise 8.

Bibliography

[1] M.S. Bazaraa and C.M. Shetty. *Nonlinear Programming*. Wiley, New York, 1979.

[2] G. Birkhoff. Tres observaciones sobre el algebra lineal. *Universidad Nacionale Tucamán Revista*, 5:147–151, 1946.

[3] V. Chvatal. *Linear programming*. Freeman, New York, 1983.

[4] K. Fan. On a theorem of Weyl concerning eigenvalues of linear transformations. *Proceedings of the National Academy of Sciences of U.S.A.*, 35:652–655, 1949.

[5] J. Farkas. Theorie der einfachen Ungleichungen. *Journal für die reine und angewandte Mathematik*, 124:1–27, 1902.

[6] P. Gordan. Uber die Auflösung linearer Gleichungen mit reelen Coefficienten. *Mathematische Annalen*, 6:23–28, 1873.

[7] P.R. Halmos. *Finite-Dimensional Vector Spaces.* Van Nostrand, Princeton, N.J., 1958.

[8] J.-B. Hiriart-Urruty. What conditions are satisfied at points minimizing the maximum of a finite number of differentiable functions. In *Nonsmooth optimization: methods and applications.* Gordan and Breach, New York, 1992.

[9] J.-B. Hiriart-Urruty and C. Lemaréchal. *Convex analysis and minimization algorithms.* Springer-Verlag, Berlin, 1993.

[10] D. König. *Theorie der Endlichen und Unendlichen Graphen.* Akademische Verlagsgesellschaft, Leipzig, 1936.

[11] H. Minkowski. Theorie der konvexen Körper, insbesondere Begründung ihres Oberflächenbegriffs. In *Gesammelte Abhandlungen II.* , Leipzig, 1911.

[12] R.T. Rockafellar. *Convex Analysis.* Princeton University Press, Princeton, N.J., 1970.

[13] R.T. Rockafellar and R.J.-B. Wets. *Variational analysis.* Springer, Berlin, 1998.

[14] C.M. Theobald. An inequality for the trace of the product of two symmetric matrices. *Mathematical*

Proceedings of the Cambridge Philosophical Society, 77:265–266, 1975.

[15] J. von Neumann. Some matrix inequalities and metrization of matric-space. *Tomsk University Review*, 1:286–300, 1937. In: *Collected Works*, Pergamon, Oxford, 1962, Volume IV, 205-218.