Convex Optimization over Probability Measures

Christopher Jordan-Squire

A dissertation submitted in partial fulfillment of the
requirements for the degree of

Doctor of Philosophy

University of Washington

2015

Reading Committee:

James V. Burke, Chair

Rekha Thomas

Dmitriy Drusvyatskiy

Peter Hoff

Program Authorized to Offer Degree:
Mathematics

University of Washington

Abstract

Convex Optimization over Probability Measures

Christopher Jordan-Squire

Chair of the Supervisory Committee:
Professor James V. Burke
Mathematics

The thesis studies convex optimization over the Banach space of regular Borel measures on a compact set. The focus is on problems where the variables are constrained to be probability measures. Applications include non-parametric maximum likelihood estimation of mixture densities, optimal experimental design, and distributionally robust stochastic programming. The theoretical study begins by developing the duality theory for optimization problems having non-finite-valued convex objectives over the set of probability measures. It is then shown that the infinite-dimensional problems can be posed as non-convex optimization problems in finite dimensions. The duality theory and constraint qualifications for these finite dimensional problems are derived and applied in each of the applications studied. It is then shown that the non-convex finite-dimensional problems can be decomposed by first optimizing over a subset of the variables, called the "weights", for which the associated optimization problem is convex. Optimization over the remaining variables, called the "support points", is then considered. The objective function in the reduced problem is neither convex nor finite-valued. For these reasons, a smoothing of this function is introduced. The epi-continuity of this smoothing and the convergence of critical points is established. It is shown that all known constraint qualifications fail for this problem, requiring a detailed analysis of the behavior of optimal solutions as the smoothing converges to the original problem. Finally, proof of concept numerical results are given.

TABLE OF CONTENTS

Page

NOTATION

This section gives an overview of notation used throughout the thesis. Some of the notation introduced in Chapter 2 only to define further notation does not appear in the list below.

$\mathbb{Z}$	13	The integers
$\mathbb{Z}_+$	13	The non-negative integers
$\mathbb{R}$	13	The real numbers
$\overline{\mathbb{R}}$	13	The extended real numbers $\overline{\mathbb{R}} = \mathbb{R} \cup \{+\infty\} \cup \{-\infty\}$.
$\mathbb{R}_+$	13	The non-negative reals
$\mathbb{R}_{++}$	13	The strictly positive reals
$\mathbb{R}_-$	13	The non-positive reals
$\mathbb{R}_{--}$	13	The strictly negative reals
$\mathbb{S}^m$	14	The $m \times m$ real symmetric matrices
$\mathbb{S}^m_+$	14	The $m \times m$ positive semi-definite real symmetric matrices
$\mathbb{S}^m_{++}$	14	The $m \times m$ positive definite real symmetric matrices
$\mathbb{S}^m_-$	14	The $m \times m$ negative semi-definite real symmetric matrices
$\mathbb{S}^m_{--}$	14	The $m \times m$ negative definite real symmetric matrices
$A \preceq B$	14	If $A, B \in \mathbb{S}^m_+$, then $A \preceq B$ if and only if $B - A \in \mathbb{S}^m_+$
$A \succeq B$	14	If $A, B \in \mathbb{S}^m_+$, then $A \succeq B$ if and only if $A - B \in \mathbb{S}^m_+$
$\mathrm{Lin}(V_1, V_2)$	14	The linear operators from a vector space V_1 to a vector space V_2
$\mathrm{Bilin}(V_1, V_2)$	14	The bilinear forms from $V_1 \times V_2$ to $\mathbb{R}$
$\mathrm{Bilin}(V_1, V_2, U)$	14	The bilinear forms from $V_1 \times V_2$ to U

$\operatorname{Aff}(C)$	14	The smallest affine space containing the set C
$\operatorname{Par}(C)$	14	The subspace parallel to $\operatorname{Aff}(C)$
$\operatorname{ri} C$	14	The relative interior of a set C
$v_1 \circ v_2$	15	The Hadamard product between two vectors $v_1, v_2 \in \mathbb{R}^n$ for some n
Ran	14	The image of a linear operator under its domain
δ_C	15	The convex indicator function of a set C, $\delta_C(c) = \begin{cases} 0 & c \in C \\ +\infty & c \notin C \end{cases}$
$\mathbf{1}_C$	15	The (set) indicator function of a set C, $\mathbf{1}_C(c) = \begin{cases} 1 & c \in C \\ 0 & c \notin C \end{cases}$
$\operatorname{lev}_f^{\leq}(\alpha)$	15	The α lower level set of f, i.e. $\operatorname{lev}_f^{\leq}(\alpha) := \{x \in X \mid f(x) \leq \alpha\}$
$\operatorname{lev}_f^{=}(\alpha)$	15	The α equality level set of a function f, i.e. $\operatorname{lev}_f^{=}(\alpha) := \{x \in X \mid f(x) = \alpha\}$
C^∞	17	The horizon cone of a set C
K°	18	The polar cone of a cone K
$\operatorname{cone} C$	18	The cone containing C, i.e. $\operatorname{cone}(C) = \{\alpha c \mid \alpha \in \mathbb{R}_+,\ c \in C\}$
$\operatorname{co} C$	18	The convex hull of a set C
$\operatorname{cl} M$	19	The closure of a set M in a given topology
$\operatorname{cl}^* M$	19	The weak* closure of a set M
$T(c\|C)$	19	The tangent cone to C at c
$\operatorname{epi} f$	19	The epigraph of a function f
$\operatorname{dom} f$	19	The domain of a function f
f^*	20	The convex conjugate $f^* : U \to \overline{\mathbb{R}}$ of a function $f : V \to \overline{\mathbb{R}}$, where V and U are paired in duality.

$\operatorname{ext} C$	22	The extreme points of a convex set C
a_β	23	The atomic measure, or point mass, at β
Δ^{d-1}	23	The $(d-1)$-dimensional unit simplex in $\mathbb{R}^d$
e		The vector of all ones
$\widehat{\partial} f$	26	The regular subdifferential of a function $f : \mathbb{R}^d \to \overline{\mathbb{R}}$
∂f	20, 26	The subdifferential of a function $f : \mathbb{R}^d \to \overline{\mathbb{R}}$
∂f	26	The subdifferential of a function $f : \mathbb{R}^d \to \overline{\mathbb{R}}$
$\overline{\partial} f$	27	The Clarke subdifferential of a function $f : \mathbb{R}^d \to \overline{\mathbb{R}}$
Ω	36	A convex, compact space $\Omega \subset \mathbb{R}^p$
β		An element of Ω, i.e. $\beta \in \Omega$
x		A point $x = (\beta_1, \ldots, \beta_d) \in \Omega^d$
$\mathcal{B}(\Omega)$	36	The set of all (regular) Borel measures on Ω
$\mathcal{P}(\Omega)$	36	The set of all (regular Borel) probability measures on Ω
$C(\Omega)$		The set of all continuous real-valued functions on Ω
$\mathcal{E}$	36	A finite dimensional real inner product space
ψ	36	A proper lsc convex function $\psi : \mathcal{E} \to \overline{\mathbb{R}}$
F	36	A differentiable function $F : \Omega \to \mathcal{E}$, the integral kernel of S
S	36	A linear function $S : \mathcal{B}(\Omega) \to \mathcal{E}$ given by $S\mu = \int_\Omega F(\beta)\mu(d\beta)$
$\widehat{n}$		Some integer $\widehat{n}$ with $\widehat{n} > n$
$\operatorname{expit}(\alpha)$	39	The function $\operatorname{expit} : \mathbb{R} \to \mathbb{R}$ given by $\operatorname{expit}(\alpha) = \operatorname{logit}^{-1}(\alpha) = (1 + \exp(-\alpha))^{-1}$
$\operatorname{logit}(p)$	45	The function $\operatorname{logit} : (0,1) \to \mathbb{R}$ given by $\operatorname{logit}(p) = \log\left(\frac{p}{1-p}\right)$
`Primal`	36	The problem $\inf_{\mu \in \mathcal{P}(\Omega)} \psi(S\mu)$
`Dual`	56	The convex dual to `Primal`
`F-Primal`	59	A non-convex finite-dimensional embedding of `Primal`

`F-Dual`	61	The convex composite dual to `F-Primal`
λ	59, 85,87	Weight vector $\lambda \in \mathbb{R}^{\widehat{n}}$ for columns of $A_F(x)$ in `F-Primal`. Also the primal variable $\lambda \in \mathbb{R}^d$ of weights for columns of $A_\phi(x)$ in $\texttt{BendersPrimal}(x)$ and $\texttt{BendersPrimal}(x, t)$
$\mathscr{S}$	36	The set $\mathscr{S} \subset \mathcal{B}(\Omega)$ of optimal solutions to `Primal`
`NPMLE`	42	The non-parametric maximum likelihood problem
`NPMLEDual`	72	The dual to `NPMLE`
`OptD`	44	The optimal experimental design problem
`OptDDual`	76	The dual to `OptD`
`DRSP`	46	The distributionally robust stochastic programming problem
`DRPSDual`	78	The dual to `DRSP`
`MER`	49	A relaxation of the maximum entropy problem
$A_F(x)$	58	For $x \in \Omega^{\widehat{n}}$ where $x = (\beta_1, \ldots, \beta_{\widehat{n}})$, then $A_F(x)$ is the $n \times \widehat{n}$ matrix $A_F(x) = \begin{bmatrix} F(\beta_1) & F(\beta_2) & \cdots & F(\beta_{\widehat{n}}) \end{bmatrix}$.
ϕ	83	A proper convex lsc function $\phi : \mathbb{R}^n \to \overline{\mathbb{R}}$
$\mathbb{E}$	83	The Euclidian space $\mathbb{R}^d \times \mathbb{R}^{n_I} \times \mathbb{S}^{n_M}$
$\mathbb{F}$	83	The Euclidian space $\mathbb{R}^p \times \{0\}^{n_E} \times \mathbb{R}_-^{n_I} \times \mathbb{S}_-^{n_M}$
$\mathbb{K}$	83	The cone $\{0\}^{n_E} \times \mathbb{R}_-^{n_I} \times \mathbb{S}_-^{n_M}$
$f_\phi(\beta)$	82	A smooth function $f_\phi : \Omega \to \mathbb{R}^n$
$f_E(\beta)$	82	A smooth function $f_E : \Omega \to \mathbb{R}^{n_E}$
$f_I(\beta)$	82	A smooth function $f_I : \Omega \to \mathbb{R}^{n_I}$
$f_M(\beta)$	82	A smooth function $f_M : \Omega \to \mathbb{S}^{n_M}$
$A_\phi(x)$	82	For $x \in \Omega^d$, the linear operator $A_\phi(x) \in \mathrm{Lin}(\mathbb{R}^d, \mathbb{R}^n)$ is given by $A_\phi(x)\lambda = \sum_{i=1}^d \lambda_i f_\phi(\beta_i)$
$A_E(x)$	82	For $x \in \Omega^d$, the linear operator $A_\phi(x) \in \mathrm{Lin}(\mathbb{R}^d, \mathbb{R}^{n_E})$ is given by $A_\phi(x)\lambda = \sum_{i=1}^d \lambda_i f_E(\beta_i)$

$A_I(x)$	82	For $x \in \Omega^d$, the linear operator $A_\phi(x) \in \text{Lin}(\mathbb{R}^d, \mathbb{R}^{n_I})$is given by $A_\phi(x)\lambda = \sum_{i=1}^d \lambda_i f_I(\beta_i)$
$A_M(x)$	82	For $x \in \Omega^d$, the linear operator $A_\phi(x) \in \text{Lin}(\mathbb{R}^d, \mathbb{S}^{n_M})$ is given by $A_\phi(x)\lambda = \sum_{i=1}^d \lambda_i f_M(\beta_i)$
$\text{lb}(x)$	87	The function $\text{lb} : \mathbb{R}^M \to \overline{\mathbb{R}}$ given by $\text{lb}(x) = \begin{cases} -\sum_{i=1}^M \log \beta_i & \beta_i > 0 \text{ for each } i, \\ \infty & \text{else} \end{cases}$
t	87	The relaxation/homotopy parameter in $\texttt{BendersPrimal}(x,t)$
s	85,87	Slack variable $s \in \mathbb{R}^{n_I}$ for inequality constraints in $\texttt{BendersPrimal}(x)$ and $\texttt{BendersPrimal}(x,t)$
$s(x,t)$	98	For $t > 0$, the unique optimal solution s to $\texttt{BendersPrimal}(x,t)$
S	85,87	Slack variable $S \in \mathbb{S}^{n_M}$ for semidefinite constraints in $\texttt{BendersPrimal}(x)$ and $\texttt{BendersPrimal}(x,t)$
$S(x,t)$	98	For $t > 0$, the unique optimal solution S to $\texttt{BendersPrimal}(x,t)$
z	85,87	Defined by the constraint $z = A_\phi(x)\lambda$
z_E	85,87	Defined by the constraint $z_E = A_E(x)\lambda$
z_I	85,87	Defined by the constraint $z_I = A_I(x)\lambda$
z_M	85,87	Defined by the constraint $z_M = A_M(x)\lambda$
w	88,88	Dual variable to the constraint $z = A_\phi(x)\lambda$
$w(x,t)$	98	For $t > 0$, the unique optimal solution w to $\texttt{BendersDual}(x,t)$
w_E	88,88	Dual variable to the constraint $z_E = A_E(x)\lambda$
$w_E(x,t)$	98	For $t > 0$, the unique optimal solution w_E to $\texttt{BendersDual}(x,t)$
w_I	88,88	Dual variable to the constraint $z_I = A_I(x)\lambda$
$w_I(x,t)$	98	For $t > 0$, the unique optimal solution w_I to $\texttt{BendersDual}(x,t)$

$w_{\scriptscriptstyle M}$	88,88	Dual variable to the constraint $z_{\scriptscriptstyle M} = A_{\scriptscriptstyle M}(x)\lambda$
$w_{\scriptscriptstyle M}(x,t)$	98	For $t > 0$, the unique optimal solution $w_{\scriptscriptstyle M}$ to $\texttt{BendersDual}(x,t)$
γ	88,88	Dual variable to the constraint $e^T\lambda = 1$
$G(x)$	85	Optimal value function for $\texttt{BendersPrimal}(x)$
$G(x,t)$	87	Optimal value function for $\texttt{BendersPrimal}(x,t)$
$\texttt{BendersPrimal}(x)$	85	An instance of $\texttt{Primal}$ with $\psi(z_1,z_2) = \phi(z_1) + \delta_{\mathbb{K}}(z_2)$ and fixed support points, so that $\mu = \sum_{i=1}^{d} \lambda_i \mathsf{a}_{\beta_i}$
$\texttt{BendersPrimal}(x,t)$	87	A convex relaxation of $\texttt{BendersPrimal}(x)$ with homotopy parameter t
$D(x)$	88	Optimal value function for $\texttt{BendersDual}(x)$
$D(x,t)$	88	Optimal value function for $\texttt{BendersDual}(x,t)$
$\texttt{BendersDual}(x)$	88	The (convex) dual problem to $\texttt{BendersPrimal}(x)$
$\texttt{BendersDual}(x,t)$	88	The (convex) dual problem to $\texttt{BendersPrimal}(x,t)$
$p(\lambda,s,S;x,t)$	90	Objective function for $\texttt{BendersPrimal}(x,t)$
$\text{Distinct}(x)$	95	The set $\{\beta \in \Omega \mid \beta = x_i \text{ for some } i \in \{1,\ldots,d\}\}$
$\text{Sol}_\lambda(x)$	95	The set $\{\lambda \mid \lambda$ optimal solution to $\texttt{BendersPrimal}(x)\}$
$\mathcal{I}(x,\beta)$	95	The indices $\{i \mid x_i = \beta,\ 1 \le i \le d\}$
$\lambda_{\mathcal{I}(x,\beta)}$	95	Defined by $\lambda_{\mathcal{I}(x,\beta)} = \sum_{i\in\mathcal{I}(x,\beta)} \lambda_i$
$\mathcal{A}(x)$	126	Superset of the subgradient $\partial G(x)$

ACKNOWLEDGMENTS

Thanks go to the many people that taught me about statistics, optimization and computation, including Ting Kei Pong, Sasha Aravkin, Phillip Keung, George Dahl, Andy Terrel, Anthony Scopatz, and Mark Wiebe. Also to my advisor's patience while I kept getting distracted by computational issues and related subjects.

DEDICATION

to my family, who've had to deal with me gallivanting across the country and frequently forgetting to call them or do anything interesting I can tell them about.

Chapter 1

INTRODUCTION

The general problem studied is the constrained optimization problem

$$\min_{\mu \in \mathcal{B}(\Omega)} \psi\left(S\mu\right) + \delta_{\mathcal{P}(\Omega)}(\mu) \tag{1.1}$$

where $\Omega \subset \mathbb{R}^p$ is compact, $\mathcal{B}(\Omega)$ is the linear space of (regular) Borel measures over Ω, $\mathcal{P}(\Omega) \subset \mathcal{B}(\Omega)$ are the probability measures, $\psi : \mathcal{E} \to \overline{\mathbb{R}}$ is convex with $\mathcal{E}$ a finite dimensional real inner product space, $S : \mathcal{B}(\Omega) \to \mathcal{E}$ is the linear transformation $S\mu = \int_\Omega F(\beta)\, \mu(d\beta)$, with $F : \Omega \to \mathcal{E}$ continuous. Specific applications are captured by specifying the functions ψ and F. In many of these applications, ψ takes the form

$$\psi(z_1, z_2) = \phi(z_1) + \delta_K(z_2),$$

where $\mathcal{E} = \mathcal{E}_1 \times \mathcal{E}_2$, the function $\phi : \mathcal{E}_1 \to \overline{\mathbb{R}}$ is convex and $K \subset \mathcal{E}_2$ is a non-empty closed convex cone. Note that any convex set constraint of the form

$$\int_\Omega F(\beta)\, \mu(d\beta) \in C$$

can be replaced by the convex cone constraint

$$\int_\Omega \begin{bmatrix} F(\beta) \\ 1 \end{bmatrix} \mu(d\beta) \in \operatorname{cl}\left(\operatorname{cone}(C \times \{1\})\right),$$

for $\mu \in \mathcal{P}(\Omega)$. Therefore we need only consider cone constraints.

The problem structure (1.1) includes many applications, for example non-parametric maximum likelihood for mixture models, optimal experimental design, and distributionally robust stochastic programming. In non-parametric maximum likelihood estimation for mixture densities, ψ is the negative log-likelihood of the marginal density $\int_\Omega f(y|\beta)\,\mu(d\beta)$ of the observed values y. In optimal design the objective function is a convex permutation-invariant function of the eigenvalues of the Fisher information matrix, e.g. the log-determinant or the trace of the inverse matrix. In distributionally robust stochastic programming, ψ is the identity and F gives the loss for a given realization of β distributed as μ. In addition, the function ψ is not necessarily finite-valued, and can encode constraints on μ. Potential constraints on μ include mean constraints, mean-variance constraints, quantile constraints, and chance constraints. These applications are discussed in Section 3.2.

Problem (1.1) is a convex optimization problem on $\mathcal{P}(\Omega) \subset \mathcal{B}(\Omega)$. In convex optimization, duality theory plays a central role. In this regard we are interested in computing and analyzing an appropriately specified dual convex program to (1.1). Importantly, any duality theory should answer when the primal and dual optimal values are finite and when they are attained. When the optimal values coincide, the primal-dual pair of problems is said to exhibit *strong duality*, that is, there is *zero duality gap*. Furthermore, testable certificates of optimality should be given. The duality theory developed here requires some care, as $\mathcal{B}(\Omega)$ is a non-reflexive Banach space. In convex analysis, duality theory is based on spaces that are *paired in duality* [53]. The natural duality pairing for $\mathcal{B}(\Omega)$ is with $C(\Omega)$, the space of continuous functions on Ω, with the sup-norm topology on $C(\Omega)$ and the weak* topology on $\mathcal{B}(\Omega)$. The duality pairing occurs through integration of functions in $C(\Omega)$ over measures in $\mathcal{B}(\Omega)$. This duality pairing ensures that every continuous linear functional on $\mathcal{B}(\Omega)$ can be represented as integration against a (fixed) continuous function. This limits, for instance, the possible separating hyperplanes on $\mathcal{B}(\Omega)$.

The dual problem to (1.1) is developed using similar techniques to those used in finite dimensions. A comparison with the derivation of the duality theory for finite-dimensional linear programming is instructive. This duality theory can be derived through a convex perturbation function $f : \mathbb{R}^p \times \mathbb{R}^q \times \mathbb{R}^p \to \overline{\mathbb{R}}$ given by

$$f(x, y^*, s^*) := \langle c,\, x\rangle + \delta(Ax - b + y|\{0\}) + \delta(x|\mathbb{R}^p_+)$$

with $(x, y^*, s^*) \in \mathbb{R}^p \times \mathbb{R}^q \times \mathbb{R}^p$ and $A \in \mathbb{R}^{q\times p}$. The inf-projection of f in x is defined to be the convex function $p(y^*, s^*) := \inf_x f(x, y^*, s^*)$. This is the optimal value function for the linear program

$$\begin{aligned} \inf_{x} \quad & c^T x \\ \text{s.t.} \quad & Ax + y^* = b \\ & x + s^* \geq 0. \end{aligned}$$

The variables y^* and s^* represent perturbations of the constraints with the base problem occurring for $(y^*, s^*) = (0, 0)$. In the case of linear programming, the duality theory is developed by comparing $p(0, 0)$ with $p^{**}(0, 0)$, where p^{**} is the bi-conjugate of p. In particular, $p^*(0, 0) = f^*(0, y, s)$, where $f^* : \mathbb{R}^p \times \mathbb{R}^q \times \mathbb{R}^p \to \overline{\mathbb{R}}$ is the convex conjugate of f and is given by

$$\begin{aligned} f^*(x^*, y, s) &= \sup_{x,y^*,s^*} \langle (x, y^*, s^*),\, (x^*, y, s)\rangle - f(x, y^*, s^*) \\ &= \langle b,\, y\rangle + \delta(x^* - c - A^T + s|\{0\}) + \delta(s|\mathbb{R}^n_+), \end{aligned}$$

so that

$$\begin{aligned} p(0,0) = \inf_{x} \quad & c^T x \\ \text{s.t.} \quad & Ax = b \\ & x \geq 0 \end{aligned} \qquad\qquad \begin{aligned} p^{**}(0,0) = \inf_{y,s} \quad & b^T y \\ \text{s.t} \quad & A^T y + c = s \\ & s \geq 0. \end{aligned}$$

The optimization problem defining $p(0, 0)$ is referred to as the *primal* problem, and that for

$p^{**}(0,0)$ is called the *dual* problem.

Similarly, we can define a perturbation function $f : \mathcal{B}(\Omega) \times \mathcal{E} \to \overline{\mathbb{R}}$ for (1.1) by setting

$$f(\mu, z) := \psi(S\mu + z) + \delta(\mu|\mathcal{P}(\Omega)).$$

Importantly, the associated optimal value function $p(z) := \inf_{\mu \in \mathcal{B}(\Omega)} f(\mu, z)$ is a convex function with a finite-dimensional domain, as shown in Theorem 3.11. By focusing on $p(z)$, some of the difficulties in working with non-reflexive Banach spaces can be simplified, or avoided entirely. Theorem 3.11 uses the perturbation function $f(\mu, z)$ and associated optimal value function $p(z)$ to derive the dual problem

$$\min_{w \in \mathcal{E}} \delta^*_{\mathcal{P}}(S^* w) + \psi^*(-w). \tag{1.2}$$

Furthermore, it is shown that (1.1) and (1.2) satisfy strong duality if $\operatorname{ri} S\left[\mathcal{P}(\Omega)\right] \cap \operatorname{ri} \operatorname{dom} \psi \neq \emptyset$ with attainment in (1.2). Theorem 3.11 also uses the sequential compactness of $\mathcal{P}(\Omega)$ to demonstrate that the optimal value of (1.1), if finite, is attained by some measure $\mu \in \mathcal{P}(\Omega)$.

Properties of the solution set to (1.1) are also examined. The convex duality established in Theorem 3.11 allows the derivation in Theorem 3.18 of first-order necessary and sufficient optimality conditions for (1.1) and (1.2). These optimality conditions are

$$\begin{aligned} S^* w &\in \partial \delta_{\mathcal{P}(\Omega)}(\mu) \\ -w &\in \partial \psi(S\mu) \end{aligned} \tag{1.3}$$

where $S^* : \mathcal{E} \to C(\Omega)$ is given by $(S^* w)(\cdot) = \langle w,\, F(\cdot) \rangle$ for all $w \in \mathcal{E}$ and

$$\partial \delta_{\mathcal{P}(\Omega)}(\mu) = \left\{ \alpha \in C(\Omega) \mid \operatorname{supp} \nu \subset \operatorname{argmax}_{\beta \in \Omega} \alpha(\beta) \right\}$$

for $\mu \in \mathcal{P}(\Omega)$. It is also shown in Theorem 3.25 that the solution set $\mathscr{S}$ to (1.1) is

$$S_1^{-1}(\phi^{-1}(p(0))) \cap S_2^{-1}K \cap \mathcal{P}(\Omega).$$

Theorem 3.25 also establishes expressions for $T(\mu|\mathscr{S})$. In particular, if $\psi(z_1, z_2) = \phi(z_1) + \delta_K(z_2)$ with ϕ strictly convex, K polyhedral, and μ finitely supported, then

$$T\left(\mu_0 \,|\, \mathscr{S}\right) = \ker S_1 \cap T\left(\mu_0 \,\middle|\, S_2^{-1}K\right) \cap T\left(\mu_0 \,|\, \mathcal{P}(\Omega)\right). \tag{1.4}$$

It is shown in the proof of Lemma 3.10(2) that

$$T\left(\mu_0 \,|\, \mathcal{P}(\Omega)\right) = \left\{\mu \in \mathcal{B}(\Omega) \,\middle|\, \int_\Omega \mu(\beta) = 0,\ \mu|_{\Omega\setminus\operatorname{supp}\mu_0} \geq 0\right\}.$$

In addition, it is often useful to establish when the solution set is a singleton, as multiple solutions complicate our understanding of the variational behavior of the optimal value function. However, (1.4) suggests the solution set $\mathscr{S}$ will typically be "large" since it is the intersection of a three sets, each of which is defined by a finite number of equality and inequality constraints on $\mathcal{B}(\Omega)$ along with a non-negativity constraint.

It is well-known that the infinite dimensional optimization problem (1.1) has an embedding into an finite-dimensional problem. However, the finite-dimensional problem is not convex. The reduction to finite dimensions is achieved by applying the Krein-Milman theorem to the extreme points of $\mathcal{P}(\Omega)$, which are known to be the atomic measures on Ω. This shows that the image of the probability measures $\mathcal{P}(\Omega)$ under the linear operator S is the convex hull of the image of Ω under the function F. The Caratheodory theorem gives an upper-bound d on the number of points in $F(\Omega)$ required to represent an element of the convex hull of $F(\Omega)$, namely d must be greater than the dimension of $\mathcal{E}$. In terms of the original infinite-dimensional problem, this representation of $S[\mathcal{P}(\Omega)]$ and the Caratheodory theorem shows that if the optimal solution set to (1.1) is non-empty, then it contains a measure with finite support. Encoding the embedding of (1.1) into finite dimensions using the representation given by the Carathedory theorem yields

$$\begin{aligned} \min_{x,\lambda,z} \quad & \psi(z) \\ \text{s.t.} \quad & z = A_F(x)\lambda \\ & \lambda \in \Delta^{d-1} \\ & x \in \Omega^d, \end{aligned} \tag{1.5}$$

where

$$\begin{aligned} d &> \dim \mathcal{E} \\ x &= (\beta_1, \ldots, \beta_d) \in \Omega^d \\ A_F(x)\lambda &= \sum_{i=1}^{d} F(\beta_i)\lambda_i \\ \Delta^{d-1} &= \left\{ \lambda \in \mathbb{R}^d \,\middle|\, \lambda \geq 0,\ \sum_i \lambda_i = 1 \right\}. \end{aligned}$$

A pair $(x, \lambda) \in \Omega^d \times \Delta^{d-1}$ in (1.5) corresponds to the probability measure $\sum_{i=1}^d \lambda_i \mathsf{a}_{\beta_i}$, where a_β is the atomic probability measure with mass at β.

The finite-dimensional problem (1.5) is non-convex due to the constraint $z = A_F(x)\lambda$. More specifically, the problem is a convex-composite optimization problem [14]. Recall the objective function of a convex-composite optimization problem is the composition of a convex function with a smooth nonlinear mapping, e.g. non-linear least squares. Convex-composite problems have a local duality theory that, in general, does not satisfy strong duality. However, in this case the relationship between the convex optimization problem (1.1) and (1.5) gives hope that there exist stronger relationships between (1.5) and its convex-composite dual. Theorem 3.15 shows the convex-composite dual of (1.5) and the convex dual of (1.1) coincide. An immediate consequence is that if the constraint qualification $\operatorname{ri} S\left[\mathcal{P}(\Omega)\right] \cap \operatorname{ri}\operatorname{dom}\psi \neq \emptyset$ is satisfied, then strong duality exists between (1.5) and its convex-composite dual.

Recall the constraint qualification $\operatorname{ri} S\left[\mathcal{P}(\Omega)\right] \cap \operatorname{ri}\operatorname{dom}\psi \neq \emptyset$ is required to obtain strong duality between (1.1) and (1.2). This condition does not relate directly to the objects in (1.5). Instead, we propose the alternative constraint qualification

$$A_F(x)\operatorname{ri}\Delta^{d-1} \cap \operatorname{ri}\operatorname{dom}\psi \neq \emptyset, \quad \text{and}$$
$$\operatorname{Par}(S\mathcal{P}) \times \mathbb{R} \subset \operatorname{Ran}\left(\begin{bmatrix} A_F(x) \\ \mathsf{e}^T \end{bmatrix}\right).$$

In Theorem 3.21 it is shown this alternative constraint qualification implies $\operatorname{ri} S\left[\mathcal{P}(\Omega)\right] \cap \operatorname{ri}\operatorname{dom}\psi \neq \emptyset$, and hence strong duality between (1.1) and (1.2).

The relationship between (1.5) and (1.1) also allows the optimality conditions for (1.1) to be used as a test for global optimality of locally optimal solutions to (1.5). This is an important distinction, as the convex optimization problem (1.1) has no locally optimal solutions that are not globally optimal, while the non-convex optimization problem (1.5) could very well have locally optimal solutions that are not globally optimal. This distinction is explored in Example 3.17, and made explicit in Theorem 3.18. In particular, the (local) first-order stationary conditions for (1.5) are

$$\begin{aligned} \nabla F(\beta_i)^T w &\in N\left(\beta_i \,|\, \Omega\right), \quad i = 1, \dots, d \\ A_F(x)^T w &\in \partial\delta_{\Delta^{d-1}}(\lambda) \\ -w &\in \partial\psi(A_F(x)\lambda). \end{aligned} \tag{1.6}$$

The first-order stationary conditions in (1.6) relate to the first-order optimality conditions of (1.1) given in (1.3) in the following way: the condition $S^* w \in \partial\delta_{\mathcal{P}(\Omega)}(\mu)$, a global optimality condition on $\langle w,\, F(\cdot)\rangle \in C(\Omega)$, is replaced with the local optimality condition $\nabla F(\beta_i)^T w \in N\left(\beta_i \,|\, \Omega\right)$ for $i = 1, \dots, d$ and how those local optima relate to each other through $A_F(x)^T w \in \partial\delta_{\Delta^{d-1}}(\lambda)$.

The specific structure in (1.5) accommodates a decomposition-based approach based on the optimal value function $G : \Omega^d \to \overline{\mathbb{R}}$ given by

$$\begin{aligned} G(x) := \min_{\lambda,z} \quad & \psi(z) \\ \text{s.t.} \quad & z = A_F(x)\lambda \\ & \lambda \in \Delta^{d-1}. \end{aligned} \tag{1.7}$$

This decomposition isolates the convexity in (1.5) into the convex optimization problem defining $G(x)$. Ideally, $G(x)$ can then be optimized over Ω^d using standard nonlinear optimization techniques. Unfortunately, $G(x)$ may not be differentiable or finite-valued. In fact, when $\psi(z_1, z_2) = \phi(z_1) + \delta_K(z_2)$ with non-trivial cone constraints (i.e., $K \neq \mathcal{E}_2$), then Proposition 4.1 shows $G(x)$ is necessarily non-finite-valued on Ω^d.

The poor regularity properties of $G(x)$ motivate introducing a smoothing $G(x,t)$ of $G(x)$ for $t > 0$ when the objective ψ has the form $\psi(z_1, z_2) = \phi(z_1) + \delta_K(z_2)$ where $\phi : \mathbb{R}^n \to \overline{\mathbb{R}}$ is proper second-order differentiable strictly convex lsc and $K = \{0\}^{n_E} \times \mathbb{R}_-^{n_I} \times \mathbb{S}_-^{n_M}$. The resulting function $G(x,t)$ is shown to be second-order smooth in x and t for $t > 0$. In this case, the function F takes the form

$$F(\beta) = \begin{bmatrix} f_\phi(\beta) \\ f_E(\beta) \\ f_I(\beta) \\ f_M(\beta) \end{bmatrix}$$

and we can define a relaxation of (1.5) by

$$\begin{aligned} G(x,t) := \min_{\substack{(\lambda,s,S)\\(z,z_I,z_E,z_M)}} \quad & \phi(z) + t\,\mathrm{lb}(\lambda) + t\,\mathrm{lb}(-s) - t\,\mathrm{logdet}(-S) \\ & + \frac{1}{2t}\left\|z_E\right\|^2 + \frac{1}{2t}\left\|z_I - s\right\|^2 + \frac{1}{2t}\left\|z_M - S\right\|^2 \\ \text{s.t.} \quad & e^T\lambda = 1 \\ & z = A_\phi(x)\lambda \\ & z_E = A_E(x)\lambda \\ & z_I = A_I(x)\lambda \\ & z_M = A_M(x)\lambda, \end{aligned} \tag{1.8}$$

where $A_g(x) = [f_g(\beta_1)\cdots f_g(\beta_d)]$ and

$$\mathrm{lb}(x) = \begin{cases} -\sum_{i=1}^M \log\beta_i & \beta_i > 0 \text{ for each } i, \\ \infty & \text{else} \end{cases}. \tag{1.9}$$

An application of Theorem 2.42 to parametrized optimization problems with linear constraints shows that, as a function of x for fixed t, $G(x,t)$ has the differentiability properties of ϕ. Propositions 4.18 and 4.19 show that if ϕ is second-order differentiable, then so is $G(x,t)$ in x, and gives explicit formulas for the first and second derivatives of $G(x,t)$. In particular,

$$\begin{aligned} d_xG(x,t)[\Delta x] = {} & d\phi(A_\phi(x)\lambda(x,t))d_xA_\phi(x)[\Delta x]\lambda(x,t) \\ & + \frac{1}{t}\left(A_E(x)\lambda(x,t)\right)^T d_xA_E(x)[\Delta x]\lambda(x,t) \\ & + \frac{1}{t}(A_I(x)\lambda(x,t) - s(x,t))^T d_xA_I(x)[\Delta x]\lambda(x,t) \\ & + \frac{1}{t}\,\mathrm{tr}\left[\left(A_M(x)\lambda(x,t) - S(x,t)\right)\right)\left(d_xA_M(x)[\Delta x]\lambda(x,t)\right)\right], \end{aligned} \tag{1.10}$$

where $(\lambda(x,t), s(x,t), S(x,t))$ are the optimal solutions to the optimization problem defining $G(x,t)$, and $d_x F(x)[\Delta x]$ denotes the derivative of F with respect to x evaluated at Δx.

For $G(x,t)$ to be a useful relaxation of $G(x)$, it must be the case that the functions $G(x,t)$ converge in some sense to $G(x)$ as $t \searrow 0$, and, in addition, for the solutions to the problems $\min_{x\in\Omega^d} G(x,t)$ converge to solutions of $\min_{x\in\Omega^d} G(x)$. A notion of convergence that is suitable for these objectives is that of epi-convergence [55]. In Theorem 4.10 it is shown that $G(x,t)$ is epi-continuous for $t \geq 0$, where $G(x,0) = G(x)$ by definition. In particular, this implies that if the optimal solution sets of $\min_{x\in\Omega^d} G(x,t)$ converge as $t \searrow 0$, then they converge to solutions of $\min_{x\in\Omega^d} G(x)$.

However, in practice, it is typical to only know first-order critical points of $G(x,t)$ instead of global optima. So it is also very important to understand the limiting behavior of the critical points of $\min_{x\in\Omega^d} G(x,t)$ as $t \searrow 0$. Theorem 4.22 shows that, subject to certain regularity conditions, if $(x^\nu, t^\nu) \to (\overline{x}, 0)$ and $0 \in \nabla_x G(x^\nu, t^\nu) + N_{\Omega^d}(x^\nu)$ for all ν, then $\overline{x}$ is a Clarke first-order critical point of $G(x)$. Furthermore, if $G(x)$ has regular subgradients at $\overline{x}$, then $\overline{x}$ is in fact a regular critical point.

The proof of Theorem 4.22 requires several ingredients. The expression for $d_x G(x,t)$ in (1.10) shows the limiting behavior as $t \searrow 0$ is closely tied to the limiting behavior of $(\lambda(x,t), s(x,t), S(x,t))$, the optimal solutions to the optimization problem defining $G(x,t)$. Analysis of this behavior is complicated by the possibility of coalescence in x as $t \searrow 0$. More specifically, if $(x^\nu, t^\nu) \to (\overline{x}, 0)$ with $x^\nu = (\beta_1^\nu, \ldots, \beta_d^\nu)$, then elements of x^ν can "coalesce" in the sense that $\lim_{\nu\to\infty} \beta_i^\nu = \lim_{\nu\to\infty} \beta_j^\nu$ where $i \neq j$. This behavior is commonly seen in practice, and when present will cause all typical regularity conditions to fail.

In order to handle the coalescence problem, we define $\mathcal{I}(x,\beta) := \{i |\, x_i = \beta,\ 1 \leq i \leq d\}$, the set of indices of $x = (\beta_1, \ldots, \beta_d)$ such that $\beta_i = \beta$. Regularity conditions related to the ranks of $A_\phi(x)$, $A_E(x)$, $A_I(x)$, and $A_M(x)$ are given in Definition 4.6. Theorem 4.7 shows how these regularity conditions relate to uniqueness of the limiting solutions. The values of $\lambda(x^\nu, t^\nu)$, the (unique) optimal solution to the problem defined by $G(x^\nu, t^\nu)$, might not converge as $(x^\nu, t^\nu) \to (\overline{x}, 0)$. However, it is shown in Theorem 4.17 that the limit of

$\sum_{i\in\mathcal{I}(\overline{x},\beta)} \lambda_i(x^\nu, t^\nu)$ does exist and is unique as $(x^\nu, t^\nu) \to (\overline{x}, 0)$.

The formula (1.10) for $d_x G(x,t)$ also contains the expressions $d\phi(A_\phi(x)\lambda(x,t))$, $\frac{1}{t}\left(A_E(x)\lambda(x,t)\right)$, $\frac{1}{t}\left(A_I(x)\lambda(x,t) - s(x,t)\right)$, and $\frac{1}{t}\left(A_M(x)\lambda(x,t) - S(x,t)\right)$. Thus, their limiting behavior as $(x^\nu, t^\nu) \to (\overline{x}, 0)$ must be analyzed for the proof of Theorem 4.22. This is accomplished by considering the convex dual to the optimization problem defining $G(x,t)$. This convex dual is derived in Theorem 4.3. Written as a function $D(x,t)$ of (x,t), the convex dual is

$$
\begin{aligned}
D(x,t) := \min_{\substack{(w,w_E,w_I,w_M)\in\mathbb{F}\\ \gamma\in\mathbb{R}}} \quad & \gamma + \phi^*(-w) \\
& + t\,\mathrm{lb}\left(\gamma e - A_\phi(x)^T w - A_I(x)^T w_I - A_E(x)^T w_E - A_M(x)^T w_M\right) \\
& + t\,\mathrm{lb}(-w_I) - t\,\mathrm{logdet}(-w_M) \\
& + \frac{t}{2}\left(\|w_E\|^2 + \|w_I\|^2 + \|w_M\|^2\right) \\
& + t\left(\log(t) - 1\right)\left(d + n_I + n_M\right).
\end{aligned}
\tag{1.11}
$$

The dual variables w, w_E, w_I, and w_M correspond, respectively, to $d\phi(A_\phi(x)\lambda(x,t))$, $\frac{1}{t}\left(A_E(x)\lambda(x,t)\right)$, $\frac{1}{t}\left(A_I(x)\lambda(x,t) - s(x,t)\right)$, and $\frac{1}{t}\left(A_M(x)\lambda(x,t) - S(x,t)\right)$. The precise relationship is given in Proposition 4.4, which exhibits the first-order necessary and sufficient optimality conditions for the primal-dual pair of optimization problems defined by $G(x,t)$ and $D(x,t)$. Proposition 4.16 establishes that the optimal dual solutions $w(x^\nu, t^\nu)$, $w_E(x^\nu, t^\nu)$, $w_I(x^\nu, t^\nu)$, and $w_M(x^\nu, t^\nu)$ for $D(x^\nu, t^\nu)$ are, subject to regularity conditions, bounded as $(x^\nu, t^\nu) \to (\overline{x}, 0)$. In conjunction with non-degeneracy conditions in Definition 4.6, boundedness from Proposition 4.16 is used in Theorem 4.17 to show the limit of optimal dual variables w, w_E, w_I, and w_M exists.

Finally, proof of concept numerical comparisons are performed on simulated data. The problems examined are the non-parametric mixture likelihood estimator problem for mixtures of Poisson and Gaussian densities, and the D-optimal design problem. These experiments collectively show that the algorithm suggested by the above decomposition method, namely performing the optimization $\min_{x\in\Omega^d} G(x,t)$ for successively smaller values of t, takes longer, though less than an order of magnitude longer, than potential competing methods and typically yields objective values at least as good and often better than the potential alternatives. However, these comparisons are very preliminary, and further comparisons against a wider range of alternative algorithms and on a more comprehensive suite of problems is necessary.

The following outlines the topics covered by chapter.

Chapter 2: A review of convex analysis, variational analysis, and calculus notation, definitions, and results used in later chapters.

Chapter 3: The main problem of study is formally stated, and several instances of it are discussed. The problem is then explored both in infinite and finite dimensions. A reduction to finite dimensions and first-order optimal conditions are also described, and applications of the duality theory are demonstrated.

Chapter 4: A decomposition based on a partition of the variables of the non-convex finite dimensional embedding from the previous chapter is presented. This decomposition is then smoothed to provide a better-behaved function. The properties of this smoothing as the homotopy parameter defining the smoothing goes to 0 are derived and discussed.

Chapter 5: Proof of concept numerical results are given for the smoothed decomposition method discussed in Chapter 4.

The original contributions in the thesis are the duality theory and conditions given in Chapter 3, the relaxation analyzed in Chapter 4, and the convergence results given in Chapter 4. The numerical results presented in Chapter 5 are also novel. Much of the analysis in Chapter 4 is a generalization of results in [2], but that analysis used a different relaxation and only considered the case when $\psi(z_1, z_2) = -\sum_{i=1}^{n} \log((z_1)_i) + \delta_K(z_2)$, for some closed convex cone K. Additionally, the assumptions stated in several theorems were corrected.

Chapter 2

BACKGROUND MATERIAL

2.1 Introduction

This chapter introduces notation and reviews results used in later chapters. After introducing general notation, the main objects of study in convex analysis in infinite and finite dimensions are reviewed. Their generalizations to non-convex functions are then reviewed. Finally, results for taking derivatives of optimal value functions are given.

2.2 General Notation

This section briefly reviews notation for common sets, vector spaces, and functions.

The following sets, subsets, and supersets of the real line are used throughout:

$$\mathbb{Z} \quad \text{the integers} \tag{2.1}$$

$$\mathbb{Z}_+ \quad \text{the non-negative integers} \tag{2.2}$$

$$\mathbb{R} \quad \text{the real numbers} \tag{2.3}$$

$$\mathbb{R}_+ \quad \text{the non-negative real numbers} \tag{2.4}$$

$$\mathbb{R}_- \quad \text{the non-positive real numbers} \tag{2.5}$$

$$\mathbb{R}_{++} \quad \text{the strictly positive real numbers} \tag{2.6}$$

$$\mathbb{R}_{--} \quad \text{the strictly negative real numbers} \tag{2.7}$$

$$\overline{\mathbb{R}} \quad \text{the extended real numbers, } \overline{\mathbb{R}} := \mathbb{R} \cup \{+\infty\} \cup \{-\infty\} \tag{2.8}$$

Similar notation is also used for the symmetric matrices,

$$\mathbb{S}^m \quad \text{the } m \times m \text{ real symmetric matrices} \tag{2.9}$$

$$\mathbb{S}^m_+ \quad \text{the } m \times m \text{ positive semi-definite real symmetric matrices} \tag{2.10}$$

$$\mathbb{S}^m_{++} \quad \text{the } m \times m \text{ positive definite real symmetric matrices} \tag{2.11}$$

$$\mathbb{S}^m_- \quad \text{the } m \times m \text{ negative semi-definite real symmetric matrices} \tag{2.12}$$

$$\mathbb{S}^m_{--} \quad \text{the } m \times m \text{ negative definite real symmetric matrices} \tag{2.13}$$

Symmetric matrices are given an ordering $\preceq$ from the positive semi-definite cone $\mathbb{S}^m_+$. That is, if $A, B \in \mathbb{S}^m$, then $A \preceq B$ if and only if $B - A \in \mathbb{S}^m_+$ and, similarly, $A \succeq B$ if any only if $A - B \in \mathbb{S}^m_+$. Further, symmetric matrices are endowed with the inner product $\langle A,\, B\rangle = \operatorname{tr}(AB) = \sum_{i,j} A_{ij}B_{ij}$ for $A, B \in \mathbb{S}^m$, and the norm $\|A\| := \sqrt{\langle A,\, A\rangle}$.

If V, V_1, V_2 are vector spaces, not necessarily finite dimensional, and $C \subset V$, then we have the following sets

$$\operatorname{Lin}(V_1, V_2) \quad \text{the space of all linear functions from } V_1 \text{ to } V_2, \tag{2.14}$$

$$\operatorname{Bilin}(V_1, V_2) \quad \text{the space of all bilinear forms from } V_1 \times V_2 \text{ to } \mathbb{R}, \tag{2.15}$$

$$\operatorname{Bilin}(V_1, V_2, U) \quad \text{the space of all bilinear forms from } V_1 \times V_2 \text{ to } U, \tag{2.16}$$

$$\operatorname{Aff}(C) \quad \text{the smallest affine space in } V \text{ containing } C, \tag{2.17}$$

$$\operatorname{Par}(C) \quad \text{the subspace of } V \text{ given by } \operatorname{Aff}(C) - c \text{ for any } c \in C, \tag{2.18}$$

$$\operatorname{ri}(C) \quad \text{the relative interior of } C. \tag{2.19}$$

An alternative definition for $\operatorname{Par}(C)$ is $\operatorname{Par}(C) = \operatorname{Aff}(C) - v$ for any $v \in \operatorname{Aff}(C)$. The *relative interior* of C is the interior of C viewed as a subset of $\operatorname{Aff}(C)$. If $L \in \operatorname{Lin}(V_1, V_2)$, then the range $\operatorname{Ran} L$ is the image of V_1, so $\operatorname{Ran} L := LV_1$.

The Hadamard product $v_1 \circ v_2$ is defined for any vectors $v_1, v_2 \in \mathbb{R}^n$ for some n by

$$(v_1 \circ v_2)_i := (v_1)_i (v_2)_i \tag{2.20}$$

The vector $e \in \mathbb{R}^n$ is the vector of all ones, where the dimension of e is implicit from the context.

For any subset $C \subset V$ we associate two functions $\delta_C : V \to \overline{\mathbb{R}}$ and $\mathbf{1}_C : V \to \mathbb{R}$, which characterize the set. They are given by

$$\delta_C(v) := \begin{cases} 0 & v \in C \\ +\infty & v \notin C \end{cases} \tag{2.21}$$

$$\mathbf{1}_C(v) := \begin{cases} 1 & v \in C \\ 0 & v \notin C \end{cases} \tag{2.22}$$

The function δ_C is referred to as the convex indicator of C while the function $\mathbf{1}_C$ is referred to as the set indicator of C. The convex indicator is often used in convex duality theory, while the set indicator is useful in integration and summation formulas.

For any function $f : V \to \overline{\mathbb{R}}$ and $\alpha \in \mathbb{R}$, the following level sets are defined:

$$\mathrm{lev}_f^{\leq}(\alpha) := \{v \in V \,|\, f(v) \leq \alpha\} \tag{2.23}$$

$$\mathrm{lev}_f^{=}(\alpha) := \{v \in V \,|\, f(v) = \alpha\} \tag{2.24}$$

2.3 Convex Analysis

This section provides a brief review of convex analysis in vector spaces of arbitrary dimension. Several more specific results relevant to later chapters are also presented.

Convex analysis in infinite dimensions presents several technical hurdles. Although many basic separation theorems extend to infinite dimensions, certain biconjugation theorems do not extend if the spaces are non-reflexive. Nonetheless, a rich duality theory can be constructed using the notion of spaces paired in duality. Additionally, this section also briefly reviews convex duality in finite dimensions.

2.3.1 Paired Spaces

Some regularity is necessary for convex analysis in infinite dimensional vector spaces. Locally convex topological vector spaces provide sufficient regularity, though we shall only use Banach spaces, which are a more restrictive setting.

Definition 2.1. *A real vector space V endowed with a topology $\mathcal{T}$ is a* topological vector space *if the operations $(v_1, v_2) \mapsto v_1 + v_2$ and $(\alpha, v) \mapsto \alpha v$ for $v, v_1, v_2 \in V$ and $\alpha \in \mathbb{R}$ are continuous in the product topologies on $V \times V$ and $\mathbb{R} \times V$. A topological vector space is* locally convex *if $\mathcal{T}$ is Hausdorff and any (open) neighborhood of $0 \in V$ includes an open barrel set. (Recall that a barrel set is a balanced, absorbing convex set, where a subset $M \subset V$ is balanced if $-M = M$ and absorbing if $\cup_{k=1}^{\infty} kM = V$.)*

Note that any normed vector space endowed with the norm topology is a locally convex topological vector space. This is an immediate consequence of the norm topology being the topology generated by the open norm balls and that norm balls about the origin are barreled.

Obtaining duality results in arbitrary topological vector spaces requires a way to connect two spaces. This is done via a bilinear form placing the two topological vector spaces in symmetry, in a the sense defined below. This generalizes the notion of duality obtained when an inner product is present.

Definition 2.2 ([11], Definition 2.26)**.** *A* pairing *of two (real) linear spaces V and U is a (real-valued) bilinear form $\langle \cdot, \cdot \rangle$ on $V \times U$. A topology on V is* compatible *with this pairing if it is locally convex and a linear functional $f : V \to \mathbb{R}$ is continuous in this topology if and only if $f(\cdot) = \langle u, \cdot \rangle$ for some $u \in U$. Compatible topologies on U are likewise defined. Finally, V and U are* paired in duality *with respect to a particular bilinear form and topologies on V and U when the topologies are compatible with respect to the pairing.*

The following theorem gives an easy way to construct spaces paired in duality from a Banach space and its dual.

Theorem 2.3 ([11], Theorem 2.25)**.** *Let V be a Banach space and let $f \in \mathrm{Lin}(V^*, \mathbb{R})$ be a linear functional on V^* that is continuous in the weak* topology. Then there exists $v \in V$ such that $f(v^*) = \langle v^*, v\rangle$ for all $v^* \in V^*$. That is, the dual of V^* , endowed with with weak topology, is isomorphic to V.*

The following gives a more concrete example.

Example 2.4. For any topological space V, let $C(V)$ denote the continuous functions $f : V \to \mathbb{R}$ and $\mathcal{B}(V)$ denote the (regular) Borel measures on V. It is a standard result that if $\Omega \subset \mathbb{R}^p$ is compact then $C(\Omega)$ with the sup-norm topology is a Banach space, and $(C(\Omega))^* = \mathcal{B}(\Omega)$ [23, Proposition 4.13,Corollary 7.18]. We can immediately apply Theorem 2.3 to get that $C(\Omega)$ with the sup-norm and $\mathcal{B}(\Omega)$ with the weak* topology are paired in duality.

2.3.2 Cones in Finite Dimensions

Cones are a fundamental object in both convex and variational analysis. They are used as local approximations of the boundaries of sets at points where the boundary cannot be written as the image of a half-space under a smooth mapping with smooth inverse. A local approximation of $\mathbb{R}^n_+$ at the origin is a prototypical example.

Definition 2.5. *A set $K \subset \mathbb{R}^n$ is a* cone *if $\lambda K \subset K$ for all $\lambda \geq 0$.*

For any set, the horizon cone is a set representing the directions the set stretches out in infinitely. In other words, the horizon cone is the directions in which a set is unbounded. When C is a closed convex set, the horizon cone is also called the (global) recession cone.

Definition 2.6. *For a set $C \subset \mathbb{R}^n$, the* horizon cone *is the closed cone $C^\infty \subset \mathbb{R}^n$ given by*

$$C^\infty = \begin{cases} \{x \in \mathbb{R}^n |\, \textit{there exists } x^\nu \in C,\ \lambda^\nu \searrow 0,\ \textit{with } \lambda^\nu x^\nu \to x\} & \textit{when } C \neq \emptyset, \\ \{0\} & \textit{when } C = \emptyset. \end{cases}$$

The polar cone is the conical analogue of a perpendicular subspace. It is often useful when translating equivalent statements tangent cones and normal cones.

Definition 2.7. *If* $K \subset \mathbb{R}^n$ *is a cone, the* polar cone $K^\circ \subset \mathbb{R}^n$ *to* K *is*

$$K^\circ := \{\, v \in \mathbb{R}^n |\, \langle v,\, w\rangle \leq 0 \textit{ for all } w \in K \,\}\,. \tag{2.25}$$

In the same way it is sometimes convenient to consider the smallest affine space containing a set, it is also sometimes convenient to consider the smallest cone containing a set.

Definition 2.8. *If* $C \subset \mathbb{R}^n$*, then* cone(C) *is defined by*

$$\text{cone}(V) := \bigcup_{t>0} tC, \tag{2.26}$$

2.3.3 Convex Sets

Convex functions often do not possess many of the more standard notions of smoothness, particularly at the boundary of the set where their values are finite. Experience has shown it's often simpler to study convex functions through convex sets associated with the function.

Definition 2.9. *A set* C *is* convex *if* $c_1, c_2 \in C$ *and* $t \in [0, 1]$ *implies* $tc_1 + (1 - t)c_2 \in C$*.*

It is often convenient to work with "convexified" versions of sets. The following definition formalizes this notion.

Definition 2.10. *For any set* C*, the* convex hull*,* $\operatorname{co} C$*, of* C *is the smallest convex set containing* C*.*

Tangent cones are a local approximation to a convex set at a point. They are a conical generalization of tangent spaces. For example, consider the tangent space to $\mathbb{R}^n_+$ at the origin.

Definition 2.11. *[11, Proposition 2.55] If* $C \subset V$ *is a closed convex set and* $c \in C$*, then the* tangent cone *to* C *at* c *is the set*

$$T(c|C) := \mathrm{cl}\,(\mathrm{cone}(C - c))\,, \tag{2.27}$$

where cl *denotes the closure is taken in the topology of* V*.*

If V has the weak* topology, then we denote the closure in that topology by cl^* instead of cl.

2.3.4 Convex Functions

The definition and study of convex functions often reduces to studying the epigraph, a set associated with any function that uniquely characterizes the function.

Definition 2.12. *If* $f : V \to \overline{\mathbb{R}}$*, then the* epigraph $\mathrm{epi}\, f \subset V \times \mathbb{R}$ *is the set*

$$\mathrm{epi}\, f := \{\, (v,t) \in V \times \mathbb{R} |\, f(v) \le t \,\}\,.$$

Many theorems about convex functions require the mild regularity conditions defined below.

Definition 2.13. *The* (effective) domain $\mathrm{dom}\, f$ *of* f *is*

$$\mathrm{dom}\, f := \{\, v \in V |\, f(v) < +\infty \,\}$$

Definition 2.14. *A function* $f : V \to \overline{\mathbb{R}}$ *is proper if* $\mathrm{dom}\, f \neq \emptyset$ *and* $f(v) > -\infty$ *for all* $v \in V$

Convex functions are not differentiable everywhere, even on the interior of their domain. However, there is a useful generalization of derivatives of convex functions based on linear lower bounds for convex functions.

Definition 2.15. *For a convex function* $f : V \to \overline{\mathbb{R}}$ *with* $\overline{v} \in V$*, the* subdifferential $\partial f(\overline{v})$ *of* f *at* $\overline{v}$ *is*

$$\partial f(v) := \{ u \in U | \, f(v) \geq f(\overline{v}) + \langle u, \, v - \overline{v} \rangle \ \text{for all} \ v \in V \} \,. \tag{2.28}$$

Elements of $\partial f(\overline{v})$ *are called* subgradients.

Associated with any convex function $f : X \to \overline{\mathbb{R}}$ is its convex conjugate. Though not proven here, the convex conjugate of f is a "dual" representation of f obtained by writing the epigraph of f as the intersection of hyperplanes.

Definition 2.16. *If* $f : V \to \overline{\mathbb{R}}$ *is a convex function, then the* convex conjugate $f^* : U \to \overline{\mathbb{R}}$ *is defined by*

$$f^*(u) := \sup_{v \in V} \left[\langle u, \, v \rangle - f(v) \right] \tag{2.29}$$

for $u \in U$*.*

The following theorem is the fundamental underpinning much of convex analysis and duality theory.

Proposition 2.17. *[11, Proposition 2.112] If* V *and* U *are two spaces paired in duality and* $f : V \to \overline{\mathbb{R}}$ *is proper lsc convex, then* $f^{**} = f$*, where* $f^{**} := (f^*)^*$*.*

Theorem 2.18. *[11, Theorem 2.113] If* V *and* U *are two spaces paired in duality and* $f : V \to \overline{\mathbb{R}}$ *is proper lsc convex, then* f^* *is proper lsc convex.*

A useful strengthening of convexity is strict convexity.

Definition 2.19. *A function* $f : \mathbb{R}^n \to \overline{\mathbb{R}}$ *is* strictly convex *if for all* $v_1, v_2 \in \mathbb{R}^n$ *and* $\alpha \in (0, 1)$*,*

$$f(\alpha v_1 + (1 - \alpha) v_2) < \alpha f(v_1) + (1 - \alpha) f(v_2). \tag{2.30}$$

Strictly convex functions are useful because their optimal solution set is either empty or a singleton.

Proposition 2.20. *[7, Proposition 3.1.1] If* $f : \mathbb{R}^n \to \overline{\mathbb{R}}$ *is strictly convex and* $V \subset \mathbb{R}^n$ *is convex, then there exists at most one global minimum of* f *over* V*.*

Strict convexity also has a close relationship with differentiability of the convex conjugate.

Theorem 2.21. *[55, Theorem 11.13] The following properties are equivalent for a proper, lsc, convex function* $f : \mathbb{R}^n \to \overline{\mathbb{R}}$ *and its conjugate function* f^**:*

(a) f *is almost differentiable, in the sense that* f *is differentiable on the open, convex set* $\operatorname{int}(\operatorname{dom} f)$*, which is non-empty, but* $\partial f(x) = \emptyset$ *for all points* $x \in \operatorname{dom} f \setminus \operatorname{int}(\operatorname{dom} f)$*, if any;*

(b) f^* *is almost strictly convex, in the sense that* f^* *is strictly convex on every convex subset of* $\operatorname{dom} \partial f^*$ *(hence on* $\operatorname{ri}(\operatorname{dom} f^*)$*, in particular).*

Likewise, the function f^* *is almost differentiable if and only if* f *is almost strictly convex.*

2.3.5 Lagrangian Duality and First Order Optimality Conditions

One of the most important applications of convex analysis is proving results about convex optimization problems and their dual problems. The theorem below is a version of the celebrated Fenchel-Rockafellar duality between two convex optimization problems.

Theorem 2.22. *[55, Example 11.41] Assume* $k : \mathbb{R}^n \to \overline{\mathbb{R}}$ *and* $h : \mathbb{R}^m \to \overline{\mathbb{R}}$ *are proper, lsc, convex, and* $A \in \mathbb{R}^{m\times n}$*,* $b \in \mathbb{R}^m$*. Then the two problems*

$$\min_x k(x) + h(b - Ax)$$

and

$$\max_y \langle b,\, y\rangle - h^*(y) - k^*(A^T y)$$

have the same optimal value if

$$b \in \operatorname{int}(A \operatorname{dom} k + \operatorname{dom} h)$$

and optimal solutions are characterized by

$$\overline{y} \in \partial h(b - A\overline{x})$$
$$A^T\overline{y} \in \partial k(\overline{x}).$$

In practice, it is often easier to derive dual programs using Lagrangians. This formulation is equivalent to Fenchel-Rockafellar duality, but less convenient for theoretical computations.

Theorem 2.23. *[55, Example 11.47] The Lagrangian for the dual problems formulation in Theorem 2.22 is given by*

$$l(x, y) = k(x) + \langle y,\, b - Ax\rangle - h^*(y)$$

and the optimality conditions can be equivalently written as

$$0 \in \partial_x l(\overline{x}, \overline{y})$$
$$0 \in \partial_y[-l](\overline{x}, \overline{y}).$$

2.3.6 Convexity and Extreme Points

With some caveats, it's possible to represent non-boundary points of a convex set as a "weighted centroid" of points on the boundary of that set. This subsection reviews results making this intuition rigorous. Central to this idea is the notion of an extreme point, which are points that cannot be represented as lying between any other two points in a convex set.

Definition 2.24. *If $C \subset V$ is convex, then $c \in C$ is an* extreme point *of C if $c_1, c_2 \in C$ and $\lambda \in (0, 1)$ with $c = \lambda c_1 + (1-\lambda)c_2$ implies $c_1 = c_2 = c$. The set of all extreme points of C is denoted* $\operatorname{ext} C$.

The theorem below gives the extreme points for the probability measures on a compact set. It uses the atomic probability measures a_β, defined by

$$\int_C \mathsf{a}_\beta(\beta') = \begin{cases} 1 & \text{if } \beta \in C \\ 0 & \text{else.} \end{cases} \tag{2.31}$$

Theorem 2.25. *[3, Proposition III.8.4][1] If $\Omega \subset \mathbb{R}^p$ is compact and $\mathcal{P}(\Omega)$ is the set of (Borel) probability measures on Ω, then* $\operatorname{ext}\mathcal{P}(\Omega) = \{\mathsf{a}_\beta |\, \beta \in \Omega\}$.

Extreme points are exactly the notion needed to express how a convex set is related to its boundary points.

Theorem 2.26 (Krein-Milman). *[11, Theorem 2.19] Let C be a nonempty convex compact subset of a locally convex topological vector space V. Then $C = \operatorname{cl}(\operatorname{co}(\operatorname{ext} C))$.*

The above relationship can be refined in finite dimensions. To do so requires the notion of how to combine points in a convex set to obtain another point in the set.

Definition 2.27. *The $n-1$-dimensional unit simplex $\Delta^{n-1} \subset \mathbb{R}^n$ is the set*

$$\Delta^{n-1} := \left\{ \lambda \in \mathbb{R}^n \,\middle|\, \sum_{i=1}^{n} \lambda_i = 1 \text{ and } \lambda_i \geq 0,\ i = 1, \ldots, n \right\}. \tag{2.32}$$

Definition 2.28. *If V is a vector space and $v_1, \ldots, v_n \in V$, a* convex combination *of $v_1, \ldots, v_n$ is any sum $\sum_{i=1}^n \lambda_i v_i$, where $\lambda \in \Delta^{n-1}$*

Now we can state Caratheodory's Theorem, which puts an upper bound, depending only on the dimension of the convex set, for how many extreme points are required to represent any non-extreme point.

Theorem 2.29 (Caratheory). *[55, Theorem 2.29] For a set $C \subset \mathbb{R}^n$ with $C \neq \emptyset$, if $c \in \operatorname{co} C$, then c is a convex combination of at most $n+1$ points in C.*

---

[1] The reference proves this only for the case $\Omega = [0, 1] \subset \mathbb{R}^1$, but the proof generalizes to an arbitrary compact subset of $\mathbb{R}^p$.

2.4 Variational Analysis

Complimenting the previous section, this section reviews the variational analysis of non-convex functions. These constructions and results generalize much of convex analysis to a less restrictive setting, at the cost of less powerful results.

2.4.1 Set Convergence

As in convex analysis, variational analysis uses similar concepts to study both functions and the geometry of sets. Unlike in convex analysis, this requires studying limiting properties of sets and set-valued mappings.

For the definitions below we make use of the set of subsequences of $\mathbb{Z}_+$ which eventually includes all integers past some finite cut-off:

$$\mathcal{N}_\infty = \{\, N \subset \mathbb{Z}_+ \,|\, \mathbb{Z}_+ \setminus N \text{ is finite} \,\} \,.$$

Additionally, if $\{a^\nu\}_{\nu \in \mathbb{Z}_+}$ is a sequence, in some space, and indexed by the positive integers, then the notation $a^\nu \underset{N}{\to} \overline{a}$ for $N \subset \mathbb{Z}_+$ means the subsequence $\{a^\nu\}_{\nu \in N}$ converges to $\overline{a}$.

The definitions below rigorously establish what it means to take a limit of set-valued mappings.

Definition 2.30. *The* outer limit *of a set-valued mapping $S : \mathbb{R}^n \rightrightarrows \mathbb{R}^m$ at $\overline{x}$ is the set*

$$\limsup_{x \to \overline{x}} S(x) = \{\, u \,|\, \exists x^\nu \to \overline{x},\ u^\nu \in S(x^\nu) \ \textit{with}\ u^\nu \to u \,\} \,.$$

Similarly the inner limit *is the set*

$$\liminf_{x \to \overline{x}} S(x) = \left\{ u \,\middle|\, \forall x^\nu \to \overline{x},\ \exists N \in \mathcal{N}_\infty,\ u^\nu \in S(x^\nu),\ \textit{with}\ u^\nu \underset{N}{\to} u \right\}$$

The limit of $S(x^\nu) \to S(\overline{x})$ is said to exist if both of the above limits are equal, and the limit is that common set.

The above definitions can be used to define generalizations of upper-semicontinuity and lower-semicontinuity of real-valued functions to set-valued mappings.

Definition 2.31. *The mapping S is* outer semicontinuous (osc) *at $\overline{x}$ if*

$$\limsup_{x \to \overline{x}} S(x) \subset S(\overline{x}).$$

Similarly, S is inner semicontinuous (isc) *at $\overline{x}$ if*

$$\liminf_{x \to \overline{x}} S(x) \supset S(\overline{x}),$$

and S is continuous *at $\overline{x}$ if it is osc and isc at $\overline{x}$.*

2.4.2 Tangent and Normal Cones

The set limits in the previous subsection allow for somewhat natural definitions of tangent and normal cones for non-convex sets.

Definition 2.32. *[55, Proposition 6.2] At any point $\overline{x}$ of a set $C \subset \mathbb{R}^n$, the set $T_C(\overline{x})$ of all tangent vectors is the closed cone*

$$T_C(\overline{x}) = \limsup_{\tau \searrow 0} \frac{1}{\tau}(C - \overline{x}).$$

Definition 2.33. *[55, Definition 6.3 and Proposition 6.5] Let $C \subset \mathbb{R}^n$ and $\overline{x} \in C$. The regular normal vectors $\widehat{N}_C(\overline{x})$ is the close cone of vectors v such that*

$$\langle v, x - \overline{x} \rangle \leq o(|x - \overline{x}|) \quad \text{for } x \in C.$$

The normal vectors $N_C(\overline{x})$ is the closed cone

$$N_C(\overline{x}) = \limsup_{x \underset{C}{\to} \overline{x}} \widehat{N}_C(x),$$

where $x \underset{C}{\to} \overline{x}$ indicates the limit is taken only over $x \in C$.

2.4.3 Subdiffential Calculus

The definitions of subdifferential below generalize the subdifferential for convex functions. Though not proven here, the two definitions agree for convex functions.

The definition below uses the concept of *f-attentive* convergence, where

$$x \underset{f}{\to} \overline{x} \tag{2.33}$$

denotes

$$x \to \overline{x} \text{ and } f(x) \to f(\overline{x}). \tag{2.34}$$

This notation is necessary for non-continuous functions.

Definition 2.34. *[55, Definition 8.3] Consider a function $f : \mathbb{R}^n \to \overline{\mathbb{R}}$ and a point $\overline{x} \in \mathbb{R}^n$ with $f(x)$ finite. For a vector $v \in \mathbb{R}^n$, one says that*

(a) v is a regular subgradient *of f at x, written $v \in \widehat{\partial} f(\overline{x})$, if*

$$f(x) \geq f(\overline{x}) + \langle v, x - \overline{x} \rangle + o(|x - \overline{x}|); \tag{2.35}$$

(b) v is a subgradient *of f at x, written $v \in \partial f(\overline{x})$, if there are sequences*

$$x^\nu \underset{f}{\to} \overline{x} \text{ and } v^\nu \in \widehat{D} f(x^\nu) \text{ with } v^\nu \to v; \tag{2.36}$$

(c) v is a horizon subgradient *of f at x, written $v \in \partial^\infty f(\overline{x})$, if there exist sequences*

$$x^\nu \underset{f}{\to} \overline{x},\ \lambda^\nu \searrow 0, \text{ and } v^\nu \in \widehat{\partial} f(x^\nu) \text{ with } \lambda^\nu v^\nu \to v. \tag{2.37}$$

The set $\widehat{\partial} f(\overline{x})$ *is the* regular subdifferential. *Similarly,* $\widehat{\partial} f(\overline{x})$ *is the* subdifferential *and* $\partial^\infty f(\overline{x})$ *is the* horizon subdifferential.

The Clarke subdifferential is another generalization of the subdifferential to non-convex functions.

Definition 2.35. *The* Clarke normal cone $\overline{N}_{\mathrm{epi}\, f}(\overline{x}, f(\overline{x}))$ *is the set*

$$\overline{N}_{\mathrm{epi}\, f}(\overline{x}, f(\overline{x})) := \mathrm{cl}\left(\mathrm{co}\left(N_{\mathrm{epi}\, f}\left(\overline{x}, f(\overline{x})\right)\right)\right) \tag{2.38}$$

and the Clarke subdifferential $\overline{\partial} f(\overline{x})$ *is the set*

$$\overline{\partial} f(\overline{x}) := \left\{ v \in \mathbb{R}^n \middle| \, (v, -1) \in \overline{N}_{\mathrm{epi}\, f}(\overline{x}, f(\overline{x})) \right\}. \tag{2.39}$$

The Clarke subdifferential is always a closed convex set. Under mild regularity conditions the Clarke subdifferential is also the convex hull of the subdifferential.

Theorem 2.36. *[55, Theorem 8.49] Let* $f : \mathbb{R}^n \to \overline{\mathbb{R}}$ *be locally lsc and finite at* $\overline{x} \in \mathbb{R}^n$*. Then* $\overline{\partial} f(\overline{x})$ *is a closed convex set. If, in addition,* $\partial^\infty f(\overline{x}) = \{0\}$*, then*

$$\overline{\partial} f(\overline{x}) = \mathrm{co}\left(\partial f(\overline{x})\right). \tag{2.40}$$

2.4.4 Epi-Convergence

Many notions of a convergent sequence of functions are not useful for studying a sequence of optimization problems. For example, neither pointwise or mean squared convergence imply convergence of the optimal values or the optimal solutions. Epi-continuity, defined below, is a method for taking limits of functions that implies optimal values and optimal solution sets converge as well. Epi-continuity is particularly useful for characterizing how a parameterized family of optimization problems behaves as the parameters converge to a limit.

Definition 2.37. *For* $f : \mathbb{R}^n \times \mathbb{R}^m \to \overline{\mathbb{R}}$, *the function-valued mapping* $u \mapsto f(\cdot, u)$ *is epi-continuous at* $\overline{u}$ *if*

$$f(\cdot, u) \xrightarrow{e} f(\cdot, \overline{u}) \text{ as } u \to \overline{u}.$$

That is, the set-valued mapping $u \mapsto \operatorname{epi} f(\cdot, u)$ *is continuous at* $\overline{u}$.

2.4.5 Amenability and Convex-Composite Duality

Amenable functions are compositions of a convex function with a smooth nonlinear change of variables with sufficient regularity for the subdifferential calculus chain rule to hold. The definition below makes this rigorous.

Definition 2.38. *A function* $f : \mathbb{R}^n \to \overline{\mathbb{R}}$ *is* amenable *at* $\overline{x} \in \mathbb{R}^n$ *if* $f(\overline{x})$ *is finite and there is an open neighborhood* V *of* $\overline{x}$ *on which* f *can be represented in the form* $f = g \circ F$ *for a* C^1 *mapping* $F : V \to \mathbb{R}^m$ *and a proper, lsc, convex function* $g : \mathbb{R}^m \to \overline{\mathbb{R}}$ *such that, in terms of* $D = \operatorname{cl}(\operatorname{dom} g)$, *the only vector* $y \in N_D(F(\overline{x}))$ *with* $DF(\overline{x})^T y = 0$ *is* $y = 0$.

Amenable functions have many applications, such as providing a natural extension of the Gauss-Newton method for least-squares. Optimization problems with amenable functions also have a duality theory that can be associated with them [14, 15]. Such problems are often called *convex composite*.

Definition 2.39. *Suppose* $V \subset \mathbb{R}^n$, *the function* $g : \mathbb{R}^m \to \overline{\mathbb{R}}$ *is convex, and* $F : \mathbb{R}^n \to \mathbb{R}^m$ *is* C^1. *Then the* convex composite Lagrangian $L : \mathbb{R}^n \times \mathbb{R}^m \to \overline{\mathbb{R}}$ *associated with the (primal) problem*

$$\min_{v \in V} g(F(v)) \tag{2.41}$$

is

$$L(v, u) := \langle u,\, F(v) \rangle - g^*(u) + \delta_V(v) \tag{2.42}$$

The dual problem associated with the above primal problem is

$$\sup_{u \in \mathbb{R}^m} \left(\inf_{v \in \mathbb{R}^n} L(v, u) \right) \tag{2.43}$$

It is straightforward to verify that switching the order of the supremum and infimum above recovers the primal problem, so

$$\min_{v \in V} g(F(v)) = \inf_{v \in \mathbb{R}^n} \left(\sup_{u \in \mathbb{R}^m} L(v, u) \right).$$

In general, the convex composite dual problem provides only a local dual, in the sense that only local optimality is established. Thus strong duality does not hold in general between convex composite primal and dual problems.

2.5 Derivatives of Optimal Value Functions

This section reviews the notation used for derivatives and the rules for taking derivatives of optimal value functions.

2.5.1 Calculus

Suppose V, U are normed vector spaces and $V_1 \subset V$ is open. For a mapping $f : V_1 \to U$, the Frchet derivative $df : V_1 \to \mathrm{Lin}(V, U)$ is the mapping satisfying

$$f(v + \epsilon) = f(v) + df(v)\epsilon + o(\|\epsilon\|)$$

for $v, \epsilon \in V$. In some cases it is simpler to write $df(v)$ applied to an element $\Delta v \in V$, which is denoted by $df(v)[\Delta v] \in U$. If $V = V_1 \times V_2$, the derivative of f with respect to only its first argument v_1 is typically denoted by $d_{v_1} f$. In cases when confusion may arise, such as applications involving the implicit function theorem, the derivative of f with respect to its first argument, $d_{v_1} f$, may also be denoted simply by $d_1 f$. This notation extends arguments other than the first, as well as when V is the cartesian product of more than 2 spaces. When $U = \mathbb{R}$ and V is a Euclidian space, then the gradient $\nabla f(v) \in V$ is the element of V identified with $df(v) \in \mathrm{Lin}(V, \mathbb{R})$.

Similar notation is used for second derivatives. If f is as in the previous paragraph, the second derivative $d^2 f : V_1 \to \mathrm{Bilin}(V, V, U)$ is the symmetric bilinear form such that

$$f(v+\epsilon) = f(v) + df(v)\epsilon + \frac{1}{2}d^2 f(v)[\epsilon,\epsilon] + O(\|\epsilon\|^2),$$

for $v, \epsilon \in V$. As with the first derivative, in some cases it is simpler to express the second derivative $d^2 f(v)$ when applied to $\Delta v \in V$, which is denoted by $d^2 f(v)[\Delta v, \Delta v]$. Also as with the first derivative, if $V = V_1 \times V_2$, then the second derivative with respect to only the first argument is denoted by $d_{v_1 v_1} f$ or, in cases when confusion may arise, by $d_{11} f$. Mixed second derivatives can also occur, such as $d_{v_1 v_2} f : V \to \mathrm{Bilin}(V_1, V_2, U)$ or, equivalently, $d_{12} f$. Sometimes the second derivatives will be associated with their matrix representation in the standard basis. For mixed second derivatives, if $U = \mathbb{R}$ the convention followed is if $d_{v_1 v_2} f(v_1 v_2) \in \mathrm{Bilin}(V_1, V_2)$, then $d_{v_1 v_2} f(v_1, v_2)$ is associated with a linear map in $\mathrm{Lin}(V_2, V_1)$.

2.5.2 Derivatives of Optimal Value Functions

The computational methods discussed in Chapter 4 rely on a relaxed optimization problem with an optimal value function that is smooth with respect to a parameterized set of constraints. To prove it is smooth and compute its derivatives requires the implicit function theorem and a computation based on it.

Two definitions are necessary to state the version of the implicit function theorem that will be used.

Definition 2.40. *The point $x_0 \in M$ is a* strong local minimizer *of $f : M \to \overline{\mathbb{R}}$, where $M \subset \mathbb{R}^p$, if there exists some $\delta > 0$ such that $f(x) \geq f(x_0) + \delta \|x - x_0\|^2$ for all x sufficiently near x_0.*

Definition 2.41. *A manifold $Q \subset \mathbb{R}^p \times \mathbb{R}^q$ is a* transversal embedding *at $(y_0, z_0) \in Q$ if*

$$(w, 0) \in N_Q(y_0, z_0) \Rightarrow w = 0.$$

The variant of the implicit function theorem stated below is well-suited for studying optimization problems.

Theorem 2.42 ([36, Theorem 5.5]). *Let $Q \subset \mathbb{R}^q \times \mathbb{R}^q$ be a smooth manifold and for each $y \in \mathbb{R}^p$ define $Q_y = \{z \in \mathbb{R}^q | (y, z) \in Q\}$. Suppose $f|_Q$ is twice continuously differentiable about $(y_0, z_0) \in Q$, that Q is a transversal embedding at (y_0, z_0), and z_0 is a strong local minimizer of $f_{y_0}|_{Q_{y_0}}$, defined by $f_{y_0}|_{Q_{y_0}}(z) = f|_Q(y_0, z)$. Then there are open neighborhoods $U \subset \mathbb{R}^p$ and $V \subset \mathbb{R}^q$ and a continuously differentiable function $\Psi : U \to V$ such that $\Psi(y_0) = z_0$ and for all $y \in V$ the function $f_y|_{Q_y \cap U}$ has a unique critical point $\Psi(y)$, which is furthermore a strong local minimizer.*

The above theorem can be used to show the optimal value function of a strictly convex program with linear equality constraints is differentiable. The chain rule and KKT conditions can be used to derive the first and second derivatives of the optimal value function.

Theorem 2.43. *Assume that $f : \mathbb{R}^p \times \mathbb{R}^q \to \overline{\mathbb{R}}$ is twice continuously differentiable, $A : \mathbb{R}^q \to \mathbb{R}^s$ is linear, $b \in \mathbb{R}^s$, and define $g : \mathbb{R}^p \to \overline{\mathbb{R}}$ by*

$$g(y) = \min_{z:Az=b} f(y, z).$$

Furthermore, assume that $f(y, z)$ is strictly convex in $z \in \mathbb{R}^q$ for fixed $y \in \mathbb{R}^p$.

1. *The set of optimal solutions to $\min_{z:Az=b} f(y, z)$ is either empty or a singleton. The optimal solution is denoted $z(y)$ when it exists.*

2. *If $(y_0, z(y_0)) \in \operatorname{int}(\operatorname{dom} f)$ and $z(y_0)$ is a strong local minimizer for $f(y_0, z)$, then the optimal solution $z(y)$ and the function $g(y)$ are continuously differentiable functions of y.*

3. *If, in addition, A is surjective, then g is twice continuously differentiable at y_0 with*

first and second derivatives

$$
\begin{aligned}
dg(y_0) &= d_1 f(y_0, z(y_0)),\\
d^2 g(y_0) &= d_{11} f(y_0, z(y_0)) + d_{12} f(y_0, z(y_0)) dz(y_0),
\end{aligned}
$$

where $dz(y_0)$ *is* v_1 *in the solution to*

$$
\begin{bmatrix} d_{22} f(y_0, z(y_0)) & -A^T \\ A & 0 \end{bmatrix} \begin{bmatrix} v_1 \\ v_2 \end{bmatrix} = \begin{bmatrix} -d_{21} f(y_0, z(y_0)) \\ 0 \end{bmatrix}.
$$

Proof. (1) This follows directly from strict convexity of $f(y, z)$ in z for fixed y.

(2) We obviously have $g(y) = f(y, z(y))$ when $g(y) < +\infty$. So g is smooth at y_0 if $z(y)$ is smooth at y_0. This follows by applying Theorem 2.42, using

$$
Q = \{ (y, z) | \, (y, z) \in \operatorname{dom} f, \; Az = b \}.
$$

The transversality condition of Theorem 2.42 is trivially satisfied for this Q.

(3) The chain rule applied to $f(y, z(y))$ gives

$$
dg(y) = d_1 f(y, z(y)) + d_2 f(y, z(y)) dz(y). \tag{2.44}
$$

Surjectivity of A implies the constraint qualification $b \in \operatorname{int}(A \{ z | \, z \in \operatorname{dom} f(y, z) \})$ is satisfied for y in an open neighborhood of y_0. For each y in that neighborhood, Theorem 2.22 shows the first-order conditions for $\min_{z:\ Az=b} f(y, z)$ are

$$
\begin{aligned}
d_2 f(y, z) - \alpha^T A &= 0\\
Az &= b
\end{aligned}
$$

for some $\alpha \in \mathbb{R}^s$. Inserting $d_2 f(y, z(y)) = \alpha^T A$ into (2.44) gives

$$dg(y) = d_1 f(y, z(y)) + \alpha^T A(dz(y)).$$

But $Az(y) = b$ for all y, so $dz(y) \in \ker A$. Therefore

$$dg(y) = d_1 f(y, z(y)),$$

proving the first formula.

Since A is surjective, the condition $\alpha^T A = d_2 f(y, z)$ implies there exists a unique α for each y. Thus for a neighborhood about y_0 there exists a function $\alpha(y)$ taking values in $\mathbb{R}^s$ and satisfying the first-order optimality conditions stated above. In particular, $\alpha(y)$ satisfies $\alpha(y)^T A = d_2 f(y, z(y))$. So $d_2 f(y, z(y))^T \in \operatorname{Ran}(A^T)$ for each y in a neighborhood of y_0, implying $\alpha(y)$ is the linear image of a continuously differentiable function, and hence is itself continuously differentiable.

Differentiating both $dg(y)$ and the first-order optimality conditions for $g(y)$ with respect to y yields

$$d^2 g(y) = d_{11} f(y, z(y)) + d_{12}(y, z(y)) dz(y)$$
$$\begin{bmatrix} d_{22} f(y, z(y)) & -A^T \\ A & 0 \end{bmatrix} \begin{bmatrix} dz(y) \\ d\alpha(y) \end{bmatrix} = \begin{bmatrix} -d_{21} f(y, z(y)) \\ 0 \end{bmatrix},$$

proving the second formula. □

The following example to demonstrates Theorem 2.43.

Example 2.44. Given non-negative integers $w_1, \ldots, w_n$ define $P : \mathbb{R}^q_{++} \to \mathbb{R}^{n \times q}$ given by

$$P(y)_{ij} = \frac{y_j^{w_i}}{w_i!} e^{-y_j}.$$

For each $y \in \mathbb{R}^q_{++}$, $P(y)$ is the matrix of the Poisson distribution density for all combinations of observed values w_i and rates y_j. Define $\text{lb} : \mathbb{R}^q \to \overline{\mathbb{R}}$ by

$$\text{lb}(x) = \begin{cases} -\sum_{i=1}^{M} \log \beta_i & \beta_i > 0 \text{ for each } i, \\ \infty & \text{else} \end{cases}, \tag{2.45}$$

and define $f : \mathbb{R}^q \times \mathbb{R}^q \to \overline{\mathbb{R}}$ by

$$f(y, z) = \text{lb}(P(y)z) + t\,\text{lb}(z),$$

Then for z with $z_i \geq 0$ for $i = 1, \ldots, q$ and $\sum_{i=1}^{q} z_i = 1$, the function $f(y, z)$ is $t\,\text{lb}(z)$ plus the log-likelihood of a Poisson mixture model whose components have rates $y_1, \ldots, y_q$.

The function $g : \mathbb{R}^q \to \overline{\mathbb{R}}$ is defined by

$$g(y) = \min_{z : e^T z = 1} f(y, z),$$

where $e \in \mathbb{R}^q$ is the vector of all ones. This function g is an specific example of a class of problems that will be examined in detail in Chapter 4. The following notation is introduced for use in computing the derivatives of g. Use $\circ$ to mean the Hadamard product and $\text{diag}(z)$ for $z \in \mathbb{R}^q$ to mean diagonal $q \times q$ matrix with $\text{diag}(z)_{ii} = z_i$. Then $z_0 \in \mathbb{R}^q$, $v_0 \in \mathbb{R}^n$, $V_0 \in \mathbb{R}^{n \times n}$, and $B \in \mathbb{R}^{n \times p}$ are defined to be

$$\begin{aligned} z_0 &= z(y_0) \\ Z_0 &= \text{diag}(z_0) \\ v_0 \circ (P(y_0) z_0) &= 1 \\ V_0 &= \text{diag}(v_0) \\ B_{ij} &= \frac{dP_{ij}}{d\beta_j}(y_0) \\ C_{ij} &= \frac{d^2 P_{ij}}{d^2 \beta_j}(y_0). \end{aligned}$$

With those definitions,

$$dg(y_0) = (v_0^T B) \circ z_0^T$$

and the hessian $d^2 g(y_0)$ can be calculated using the derivatives

$$\begin{aligned}
d_{11} f(y,z) &= Z_0 B^T V_0^2 B Z_0 - V_0 C Z_0, \\
d_{12} f(y,z) &= Z_0 B^T V_0^2 P(y) + V_0 B, \\
d_{22} f(y,z) &= P(y_0)^T V_0^2 P(y_0) + t Z_0^2.
\end{aligned}$$

For completeness, note

$$\begin{aligned}
\frac{d}{dy_j}\left(P_{ij}(y)\right) &= \left(\frac{w_i}{y_j} - 1\right) P_{ij}(y) \\
\frac{d^2}{d^2 y_j}\left(P_{ij}(y)\right) &= \left(\left(\frac{w_i}{y_j} - 1\right)^2 - \frac{w_i}{y_j^2}\right) P_{ij}(y).
\end{aligned}$$

Chapter 3
DUALITY THEORY

3.1 Introduction

In this chapter we define the optimization problem under study. A duality theory in finite and infinite dimensions, and the relationship between these problems is explored. Finally, we examine the consequences for the applications described in Section 3.3.

3.2 Problem Statement

We consider optimization problems of the form

$$\min_{\mu \in \mathcal{B}(\Omega)} \psi(S\mu) + \delta_{\mathcal{P}(\Omega)}(\mu). \tag{Primal}$$

where

$$\Omega \subset \mathbb{R}^p \text{ is compact,} \tag{3.1}$$

$$\mathcal{B}(\Omega) \text{ are the Borel measures over } \Omega, \tag{3.2}$$

$$\mathcal{P}(\Omega) \subset \mathcal{B}(\Omega) \text{ are the probability measures over } \Omega, \tag{3.3}$$

$$\mathcal{E} \text{ is a finite dimensional real inner product space,} \tag{3.4}$$

$$\psi : \mathcal{E} \to \overline{\mathbb{R}} \text{ is proper convex lsc,} \tag{3.5}$$

$$F : \Omega \to \mathcal{E} \text{ is differentiable,} \tag{3.6}$$

$$S : \mathcal{B}(\Omega) \to \mathcal{E} \text{ is the continuous linear transformation} \tag{3.7}$$

$$S\mu := \int_\Omega F(\beta)\,\mu(d\beta).$$

To be consistent with later applications and theorems a convex indicator function is used for the constraint that $\mu \in \mathcal{P}(\Omega)$. Further restrictions on the set Ω can arise in specific applications, but, in general, it is only assumed to be compact. If Ω is finite, then the continuity and differentiability of functions on Ω are defined in terms of the discrete topology, and are trivially satisfied. In the sequel, we let $\mathscr{S}$ denote the set of optimal solutions to `Primal`, which may be empty.

In applications, ψ often takes the form

$$\psi(z_1, z_2) = \phi(z_1) + \delta_K(z_2), \tag{3.8}$$

where $\mathcal{E} = \mathcal{E}_1 \times \mathcal{E}_2$, the function $\phi : \mathcal{E}_1 \to \overline{\mathbb{R}}$ is convex and $K \subset \mathcal{E}_2$ is a non-empty closed convex cone. General convex set constraints can easily be written as convex cone constraints by exploiting $\int_\Omega 1\, \mu(d\beta) = 1$. In particular, the constraint

$$\int_\Omega F(\beta)\, \mu(d\beta) \in C$$

can be replaced by the constraint

$$\int_\Omega \begin{bmatrix} F(\beta) \\ 1 \end{bmatrix} \mu(d\beta) \in \operatorname{cl}\left(\operatorname{cone}(C \times \{1\})\right).$$

This implies that any functional constraint of the form $\kappa(S\mu) \leq \epsilon$ for κ convex can be expressed as a cone constraint using $C = \operatorname{lev}_\kappa^{\leq}(\epsilon)$. In this case, the first-order conditions for `Primal` that use normal cone conditions on $\operatorname{cl}\left(\operatorname{cone}(C \times \{1\})\right)$ easily translate to normal cone conditions on C, as shown below.

Proposition 3.1. *Suppose that $C \subset \mathcal{E}_2$ is a closed convex set and let $K = \operatorname{cl}\left(\operatorname{cone}(C \times \{1\})\right)$. Then, for any $\bar{c} \in C$,*

$$N\left((\bar{c}, 1) \,|\, K\right) = \left\{ (v, -\langle v, \bar{c}\rangle_{\mathcal{E}_2}) \,\middle|\, v \in N\left(\bar{c} \,|\, C\right) \right\}.$$

Proof. For ease of notation, $\langle \cdot, \cdot \rangle$ means $\langle \cdot, \cdot \rangle_{\mathcal{E}_2}$ for the remainder of this proof. By the definition of the normal cone, $(v, \alpha) \in N((\bar{c}, 1) \,|\, K)$ if and only if

$$\langle v, c - \bar{c} \rangle + \alpha(t - 1) \leq 0 \text{ for all } (c, t) \in K. \tag{3.9}$$

Taking $(c, t) = (0, 0)$ gives $\langle v, \bar{c} \rangle + \alpha \geq 0$, while taking $c = t\bar{c}$ and dividing by $t - 1$ shows $\langle v, \bar{c} \rangle + \alpha \leq 0$. So $\alpha = -\langle v, \bar{c} \rangle$, and (3.9) becomes

$$\langle v, c - t\bar{c} \rangle \leq 0 \text{ for all } (c, t) \in K. \tag{3.10}$$

For $t = 1$, (3.10) is equivalent to $\langle v, c - \bar{c} \rangle \leq 0$ for all $c \in C$. Thus if $(v, -\langle v, \bar{c} \rangle) \in N((\bar{c}, 1) \,|\, K)$, then $v \in N(\bar{c} \,|\, C)$. So $N((\bar{c}, 1) \,|\, K) \subset \{(v, -\langle v, \bar{c} \rangle) | \, v \in N(\bar{c} \,|\, C)\}$.

Conversely, suppose $v \in N(\bar{c} \,|\, C)$. Then $\langle v, \tilde{c} - \bar{c} \rangle \leq 0$ for all $\tilde{c} \in C$. Multiplying through by t shows $\langle v, c - t\bar{c} \rangle \leq 0$ for all $(c, t) \in K$ with $t > 0$, since $(c, t) \in K$ with $t > 0$ if and only if $(c, t) = (t\tilde{c}, t)$ for some $\tilde{c} \in C$. If $t = 0$ and $(c, t) \in K$, then $c \in C^\infty$. But $\langle v, c \rangle \leq 0$ for all $v \in N(\bar{c} \,|\, C)$ and $c \in C^\infty$ [55, Proposition 6.35]. Thus $\langle v, c - t\bar{c} \rangle \leq 0$ for all $(c, t) \in K$, showing $(v, -\langle v, \bar{c} \rangle) \in N((\bar{c}, 1) \,|\, K)$ and completing the proof. □

When $\psi(v, w) = \phi(v) + \delta_K(w)$, we decompose $S : \mathcal{P}(\Omega) \to \mathcal{E}_1 \times \mathcal{E}_2$ as

$$S\mu = \begin{bmatrix} S_1\mu \\ S_2\mu \end{bmatrix} = \begin{bmatrix} \int_\Omega F_1(\beta)\, \mu(d\beta) \\ \\ \int_\Omega F_2(\beta)\, \mu(d\beta) \end{bmatrix}, \tag{3.11}$$

with $F_1 : \Omega \to \mathcal{E}_1$ and $F_2 : \Omega \to \mathcal{E}_2$ both continuously differentiable. The product space notation chosen in (3.11) allows a formal manipulation of operators as matrices. We interpret the real vector spaces $\mathcal{E}_i$, the cone $K \subset \mathcal{E}_2$, the operators S_i, and the functions $F_i : \Omega \to \mathcal{E}_i$ for $i = 1, 2$ as above.

Several of the most common constraints are listed below.

(i) Mean constraints:

$$L \leq \int_\Omega \beta\, \mu(d\beta) \leq U \tag{3.12}$$

with lower and upper bounds L and U given, and $\leq$ taken component-wise.

(ii) Mean-Variance constraints:

$$\int_\Omega (\beta - \theta)\mu(d\beta) = 0 \quad \text{and} \quad \Sigma_L \preceq \int_\Omega (\beta - \theta)(\beta - \theta)^T \mu(d\beta) \preceq \Sigma_U$$

with $0 \prec \Sigma_L \preceq \Sigma_U$ given.

(iii) Approximate (univariate) quantile constraints:

$$\int_\Omega \tau^-_{(\gamma,\theta)}(\beta)\mu(d\beta) \geq \alpha - \epsilon \qquad \int_\Omega \tau^+_{(\gamma,\theta)}(\beta)\mu(d\beta) \geq (1-\alpha) - \epsilon$$

with $\gamma > 0$, $\alpha \in (0,1)$ is the constrained quantile and

$$\tau^-_{(\gamma,\theta)}(\beta) = \operatorname{expit}(-\gamma(\beta - \theta)) \text{ and } \tau^+_{(\gamma,\theta)}(\beta) = \operatorname{expit}(\gamma(\beta - \theta))$$

where $\operatorname{expit} : \mathbb{R} \to \mathbb{R}$ is given by

$$\operatorname{expit}(\alpha) = (1 + \exp(-\alpha))^{-1}. \tag{3.13}$$

See Figure 3.1 examples of $\tau^+_{(\gamma,0)}$. The functions τ^- and τ^+ are used rather than $\mathbf{1}_{(-\infty,\theta]}$ and $\mathbf{1}_{[\theta,+\infty)}$ because the theory requires integration against continuous functions on Ω.

(iv) Approximate chance constraints:

$$\int_\Omega \widetilde{\mathbf{1}_C}(\beta)\, \mu(d\beta) \geq 1 - \epsilon$$

for some $\epsilon \in (0,1)$ and any continuous approximation $\widetilde{\mathbf{1}_C} : \mathbb{R}^p \to [0,1]$ of the set

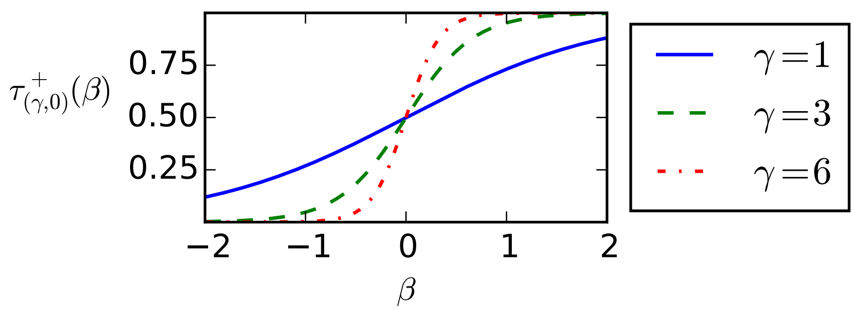

Figure 3.1: A graph of the functions $\tau^{+}_{\gamma,0}(\beta)$ used for approximate quantile constraints for various values of γ.

indicator function $\mathbf{1}_C(\beta)$ for $C \subset \mathbb{R}^p$. This approximates chance constraints [51] of the form

$$P(\beta \in C) \geq 1 - \epsilon.$$

(v) ϕ-divergence constraints: Given observed data $y_1, \ldots, y_D$ we can require a statistic related to the data, such as the log-likelihood of a related model, meets a minimum threshold. For example,

$$\sum_{i=1}^{D} -\log\left(\int_{\Omega} f(y_i|\beta)\,\mu(d\beta)\right) \leq -\epsilon$$

where $\int_{\Omega} f(y|\beta)\,\mu(d\beta)$ is the marginal density of the data and the value of ϵ can be chosen by appealing to the asymptotic distribution of the statistic. The marginal likelihood will be discussed further in the next section. This is a convex set constraint, where the convex set is the lower level-set of the negative log-likelihood. It can be transformed to a convex cone constraint using Proposition 3.1. This form of constraint is referred to as *likelihood robust optimization* [60]. It can be generalized to ϕ-divergence constraints, such as the Kullback-Leibler divergence, and to penalizing the distance from an arbitrary reference density [4, 17, 40].

3.3 Applications

3.3.1 Non-parametric Mixture Models

A mixture model is a hierarchical statistical model

$$\begin{aligned} Y|\beta &\sim f(y|\beta) \\ \beta &\sim \mu \in \mathcal{P}(\Omega), \end{aligned}$$

where $f(\cdot|\beta)$ is a parameterized family of distributions and $\mu \in \mathcal{P}(\Omega)$. The density $f(y|\mu) = \int_\Omega f(y|\beta)\,\mu(d\beta)$ is the marginal density of Y for the mixing measure μ. For fixed y the non-parametric likelihood $L : \mathcal{P}(\Omega) \to \mathbb{R}$ is the function $L(\mu) = f(y|\mu)$. A mixing measure is finite if its support, $\operatorname{supp}(\mu)$, is finite and a finite mixture model is a mixture model with a finite mixing measure. Finite mixture models are widely used to model data from heterogenous populations, or to approximate unknown densities as a mixture of simpler densities, e.g., see [9, 20, 21, 25, 27, 45]. It has long been know that the non-parametric likelihood can be optimized using finite mixing measures, e.g. [37, Theorem 3.1]. We review this result in Section 3.4.2 using modern techniques of convex analysis.

Suppose $(y_1, d_1), \ldots, (y_N, d_N)$ are the (value, count) pairs of the number of times d_i value y_i has been observed from $D = \sum_{i=1}^N d_i$ independent observations of a finite mixture model. Note $d_i = 1$ for all i almost surely if, for fixed β, the density $f(y|\beta)$ has no point masses. For ease of exposition, $f(y|\beta)$ will, depending on the context, denote either the probability density function of a continuous-valued random variable or the probability mass function of a discrete-valued random variable.

The mixing measure μ can be estimated using only observed data by maximizing the non-parametric log-likelihood $\ell(\mu) = \log L(\mu)$. In particular, $\operatorname{supp}\mu \subset \Omega$ is the only assumption made on μ. The likelihood factors as a product of terms corresponding to the independent observations, giving the following maximum likelihood problem:

$$\min_{\mu \in \mathcal{B}(\Omega)} -\sum_{i=1}^{N} d_i \log \left(\int_\Omega f(y_i|\beta)\, \mu(d\beta) \right) + \delta_{\mathcal{P}(\Omega)}(\mu).$$

If one adds constraints of the form $\delta_K(S_2\mu)$, as discussed in Section 3.2, the *non-parametric maximum likelihood* (`NPMLE`) problem becomes

$$\min_{\mu \in \mathcal{B}(\Omega)} -\sum_{i=1}^{N} d_i \log \left(\int_\Omega f(y_i|\beta)\, \mu(d\beta) \right) + \delta_K(S_2\mu) + \delta_{\mathcal{P}(\Omega)}(\mu). \tag{NPMLE}$$

The `NPMLE` problem is an instance of `Primal` with $\mathcal{E}_1 = \mathbb{R}^N$,

$$\psi(z_1, z_2) = \begin{cases} -\left(\sum_{i=1}^{N} d_i \log(z_1)_i\right) + \delta_K(z_2) & (z_1)_i > 0 \ \text{ for each } i, \\ \infty & \text{else,} \end{cases}$$

and F, the integral kernel of S, is

$$F(\beta) = \begin{bmatrix} f(\beta) \\ F_2(\beta) \end{bmatrix}, \quad f(\beta) = \begin{bmatrix} f(y_1|\beta) \\ \vdots \\ f(y_N|\beta) \end{bmatrix}.$$

The following two examples illustrate these ideas.

Example 3.2. [42] Let $n(y|\gamma, \sigma^2)$ denote the normal density with mean γ and variance σ^2:

$$n(y|\gamma, \sigma^2) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{(y-\gamma)^2}{2\sigma^2} \right).$$

If

$$\overline{y}, \sigma_L^2, \sigma_U^2 \in \mathbb{R}_{++} \text{ with } \sigma_L^2 < \sigma_U^2,$$
$$\Omega = [-\overline{y}, \overline{y}] \times [\sigma_L^2, \sigma_U^2],$$
$$f(y|\beta) = n(y|\gamma, \sigma^2) \text{ with } \beta = (\gamma, \sigma^2),$$

then the NPMLE problem is finding a non-parametric mixture of Gaussians with upper and lower bounds on the variance of each component.

Example 3.3. [43]. Suppose $s(z)$ is a smooth function, possibly implicitly depending on a covariate. Then the model $y|\beta \sim n(s(\beta_1), \beta_2^2)$, i.e. $f(y|\beta) = n(y|s(\beta_1), \beta_2^2)$, can be extended to allow z to be a random variable whose density comes from a known parameterized family of densities. For example, this is often done in hierarchical or mixed effects models. Doing so yields a non-parametric maximum likelihood problem.

3.3.2 Optimal Design

The problem of optimal design is to specify values of design parameters β that maximize a given function of the Fisher information matrix [10, 13, 16, 22, 41, 47, 48]. In the general problem, we are given a family of probability densities $f(y|\theta, \beta)$ associated with an outcome y, where the parameters θ are to be estimated through repeated experiments while the parameters β can be set by the experimenter. For a single observation y the Fisher information matrix for θ given β is

$$I_\theta(\beta) = E[gg^T], \quad g = \nabla_\theta f(Y|\theta, \beta),$$

with the expectation taken over Y. Under mild conditions, if β is given and $\widehat{\theta}_N$ is maximum likelihood estimate for the true value of θ based on N independent samples of Y, then $\sqrt{N}\left(\widehat{\theta}_N - \theta\right)$ is asymptotically normal with mean 0 and covariance $I_\theta(\beta)^{-1}$ [58, Theorem 5.39]. So the quality of the estimate $\widehat{\theta}_N$ is improved by making the inverse Fisher information matrix smaller, in some sense. The parameter β yielding the *best* optimal design for a single observed value y is taken to be a solution to an optimization problem of the form

$$\min_{\beta\in\Omega} \Phi\left(I_\theta(\beta)\right),$$

where the function $\Phi : \mathbb{S}^n_+ \to \overline{\mathbb{R}}$ is a convex loss function on the symmetric semi-positive definite matrices $\mathbb{S}^n_+$. This problem can be relaxed to a constrained optimization problem on $\mathcal{P}(\Omega)$ where rather than picking a single $\beta \in \Omega$ we choose a measure over Ω, potentially satisfying a finite number of moment constraints, that optimizes the experimental design:

$$\min_{\mu\in\mathcal{B}(\Omega)} \Phi(I_\theta(\mu)) + \delta_K\left(S_2\mu\right) + \delta_{\mathcal{P}(\Omega)}(\mu), \tag{OptD}$$

where $I_\theta(\mu) = \int_\Omega I_\theta(\beta)\,\mu(d\beta)$. This is an instance of `Primal`, with $\mathcal{E}_1 = \mathbb{S}^n$, $\langle X, Y\rangle_{\mathcal{E}_1} = \operatorname{tr}(X^TY)$, and

$$\psi(z_1, z_2) = \Phi(z_1) + \delta_K(z_2), \quad F(\beta) = \begin{bmatrix} I_\theta(\beta) \\ F_2(\beta) \end{bmatrix}.$$

Usually Φ takes the form $\Phi = f \circ \lambda$, where $f : \mathbb{R}^n \to \overline{\mathbb{R}}$ is a permutation invariant (i.e., symmetric) proper convex function and $\lambda : \mathbb{S}^n \to \mathbb{R}^n$ is the vector of eigenvalues in non-decreasing order. Common choices for the function f are listed in Table 3.1 [13, 48, 52]. In the table, the $\log(\cdot)$ and $\sqrt{\cdot}$ functions are assumed to be $+\infty$ off their standard domains.

Example 3.4. A simple example is the linear model with Gaussian errors. Let $q(\beta) = (q_1(\beta), \ldots, q_n(\beta))$, where each $q_i : \Omega \to \mathbb{R}$ is differentiable, and let $w : \Omega \to \mathbb{R}_{++}$ also be differentiable. Consider the model

$$y = \theta^T q(\beta) + \epsilon, \quad \epsilon \sim N\left(0, \frac{1}{w(\beta)}\right).$$

A straight-forward computation shows the Fisher Information matrix for such a model ob-

Table 3.1: Names of common optimal design criteria and the eigenvalue functions $\Phi = f \circ \lambda$ they optimize, along with their conjugates.

Scalarization	$f(\lambda)$	$f^*(\lambda^*)$	$\Phi(I_\theta)$
D-optimal	$-\sum_{i=1}^n \log \lambda_i$	$-n - \sum_{i=1}^n \log(-\lambda_i^*)$	$-\log\det I_\theta$
E-optimal	$\max_i \lambda_i$	$\delta_{\Delta^{n-1}}(\lambda^*)$	$\lambda_{\max}(I_\theta)$
A-optimal	$\sum_{i=1}^n \lambda_i^{-1}$	$-2\sum_{i=1}^n \sqrt{-\lambda_i^*}$	$\mathrm{tr}(I_\theta^{-1})$
pth Mean-optimal	$\sum_{i=1}^n \lambda_i^p$, $p < 0$	$\frac{2}{p}\sum_{i=1}^n (-\lambda_i^*)^{p/(p-1)}$	$\mathrm{tr}(I_\theta^p)$
Φ_p-optimal	$\lVert\lambda\rVert_p$, $1 \le p$	$\delta_{\mathbb{B}_q}(\lambda^*)$, $p^{-1} + q^{-1} = 1$	$\mathrm{tr}(I_\theta^p)^{1/p}$

served at $\beta_1, \ldots, \beta_M$ is

$$I_\theta\left(\frac{1}{M}\sum_{i=1}^M \mathsf{a}_{\beta_i}\right) = \frac{1}{M}\sum_{i=1}^M w(\beta_i)q(\beta_i)q(\beta_i)^T$$

In this case there is no dependence of the Fisher information on the model parameters θ.

Example 3.5. Another common model is logistic regression,

$$\mathrm{logit}(E[y]) = \theta^T q(\beta),$$

where $\mathrm{logit} : (0,1) \to \mathbb{R}$ is defined by

$$\mathrm{logit}(p) = \log\left(\frac{p}{1-p}\right). \tag{3.14}$$

The Fisher information matrix takes a similar form [44],

$$I_\theta\left(\frac{1}{M}\sum_{i=1}^M \mathsf{a}_{\beta_i}\right) = \frac{1}{M}\sum_{i=1}^M w(\theta,\beta_i)q(\beta_i)q(\beta_i)^T$$

with $w(\theta, \beta_i) = \mathrm{expit}(\theta^T\beta_i)(1 - \mathrm{expit}(\theta^T\beta_i))$.

3.3.3 Distributionally Robust Stochastic Programming

Many common stochastic programming problems have the form

$$\min_{\xi \in X} E_{U\sim\nu}[g(\xi, U)] \tag{3.15}$$

where $g : X \times \widetilde{\mathcal{E}} \to \overline{\mathbb{R}}$ is a given loss function and U is a random variable takings values in a Euclidian space $\widetilde{\mathcal{E}}$ and distributed according to $\nu \in \mathcal{P}(\widetilde{\mathcal{E}})$. It is assumed U has density $f(u|\beta_0)$, where $f(u|\beta)$ is a family of densities with parameters in Ω and $\beta_0 \in \Omega$ is fixed.

Suppose, on the other hand, β_0 is unknown but it is known U is distributed as $\nu = \int_\Omega f(u|\beta)\,\mu(d\beta)$ for some $\mu \in \mathcal{F} \subset \mathcal{P}(\Omega)$. In this case we can consider minimizing the function

$$\phi_{\mathcal{F}}(\xi) := \min_{\nu = \int_\Omega f(u|\beta)\,\mu(d\beta)} E_{U\sim\nu}\left[g(\xi, U)\right] + \delta_{\mathcal{F}}(\mu). \tag{\texttt{DRSP}}$$

Stochastic programming problems of this type are robust to our ignorance of the true distribution of U, and so fall into the framework of robust optimization [5]. Distributionally robust stochastic programming has been intensively researched [18, 29, 50, 62, 57], and a wide variety of both constraints and loss functions have been considered.

The set $\mathcal{F}$ can be specified in a number of ways. For example, the elements of $\mathcal{F}$ may simply satisfy a finite collection of moment constraints and so take the form

$$\mathcal{F} := \left\{ \mu \in \mathcal{P}(\Omega) \,\middle|\, \int F_2(\beta)\,\mu(d\beta) \in K \right\}.$$

If the Fubini theorem can be applied so that

$$E_{U\sim\nu}\left[g(\xi, U)\right] = \int_\Omega \left(\int g(\xi, u) f(u|\beta)\,du \right) \mu(d\beta),$$

then DRSP problem is an instance of Primal where $\mathcal{E}_1 = \mathbb{R}$,

$$\psi(z_1, z_2) = z_1 + \delta_K(z_2) \text{ and } F(\beta) = \begin{bmatrix} \int g(\xi, u) f(u|\beta)\, du \\ F_2(\beta) \end{bmatrix}. \tag{3.16}$$

It is important to note that the above requires the expression

$$\int g(\xi, u) f(u|\beta)\, du$$

to be continuous as a function of β, which does not require that $f(u|\beta)$ is continuous as a function of β. This can be generalized to expressions of the form

$$\int g(\xi, u) h(\beta)(du),$$

where $h : \Omega \to \mathcal{P}(\mathcal{E})$, as long as the expression is continuous as a function of β. For example, if $g(\xi, u)$ is continuous, $\Omega \subset \mathcal{E}$, and $h(\beta) = \mathsf{a}_\beta$, then $\int g(\xi, u) h(\beta)(du) = g(\xi, \beta)$ is continuous in β.

Example 3.6. [8, Section 1.1e]

The classic News Vendor problem is a simple example of distributionally robust stochastic programming. Suppose that a news vendor buys ξ newspapers from a publisher at cost c. The news vendor cannot buy more than u newspapers, so $\xi \in [0, u]$, and sells newspapers to customers at price q. Any unsold newspapers can be returned to the publisher at price r, with $r < c$. The demand $U > 0$ for newspapers is random, and follows distribution $\nu = \int_\Omega f(u|\beta)\, \mu(\beta)$. The loss the news vendor suffers given demand U and newspapers x is

$$g(\xi, U) = cx - q \min(U, \xi) - r \max(\xi - U, 0),$$

i.e. cost of the newspapers minus the revenue from selling them. In this example, $\mathcal{F}$ is commonly taken to be all densities on Ω that specify a given mean $\overline{u}$ and variance σ_U^2 for U, i.e.

$$F_2(\beta) = \begin{bmatrix} \int_{\widetilde{\mathcal{E}}} u f(u|\beta)\,du - \overline{u} \\ \int_{\widetilde{\mathcal{E}}} (u-\overline{u})^2 f(u|\beta)\,du - \sigma_U^2 \end{bmatrix} \text{ and } K = (0,0) \in \mathbb{R}^2.$$

Example 3.7. [60] The news vendor example above can be extended to include historical data. Suppose $U \in \mathbb{Z}$ and demands $u_i \geq 0$, have each occurred N_i times previously, for $i = 1, \ldots, n$, and fix $\gamma \in \mathbb{R}$. We can require $\nu = \int_\Omega f(u|\beta)\,\mu(d\beta)$, the distribution of U, to be close to the observed data with the constraint

$$-\sum_{i=1}^{n} N_i \log \nu(u) \leq \gamma,$$

where $\nu(u)$ is the probability density on $u \in \mathbb{Z}$.

3.3.4 Maximum Entropy

Let

$$g(x) = \begin{cases} x \log x & , \ x > 0, \\ 0 & , \ x = 0, \\ +\infty & , \ x \leq 0 \ . \end{cases}$$

The maximum entropy problem for a parametric family of distributions $f(y|\beta)$ for a random variable Y taking values in $\mathscr{Y}$ is

$$\min_{\beta \in \Omega} \int_{\mathscr{Y}} g(f(y|\beta))\,dy.$$

This problem has many applications [6, 13, 26]. This problem can be generalized to mixture densities with cone constraints by using the marginal density $f(y|\mu) = \int_\Omega f(y|\beta)\ \mu(d\beta)$. The maximum entropy problem then takes the form

$$\min_{\mu \in \mathcal{B}(\Omega)} \left(\int_{\mathscr{Y}} g(f(y|\mu))\,dy \right) + \delta_K \left(\int_\Omega F_2(\beta)\,\mu(d\beta) \right) + \delta_{\mathcal{P}(\Omega)}(\mu).$$

The goal is to maximize the entropy over mixtures of the parametric family $f(y|\beta)$. By applying Jensen's inequality, observe that

$$g\left(f(y|\mu)\right) = g\left(E\left[f(y|\beta)\right]\right) \le E\left[g(f(y|\beta))\right] = \int_\Omega g(f(y|\beta))\,\mu(d\beta).$$

This suggests the following relaxation of the maximum entropy problem:

$$\min_{\mu\in\mathcal{B}(\Omega)} \int_\Omega \left(\int_{\mathscr{Y}} g(f(y|\beta))\,dy\right)\mu(d\beta) + \delta_K\left(\int_\Omega F_2(\beta)\,\mu(d\beta)\right) + \delta_{\mathcal{P}(\Omega)}(\mu). \tag{MER}$$

This relaxation can also be derived as a convexification of the original problem since $\operatorname{cl}\left(\operatorname{co}\{\mathsf{a}_\beta|\,\beta\in\Omega\}\right) = \mathcal{P}(\Omega)$. `MER` is an instance of `Primal` with a similar form to `DRSP`, namely $\mathcal{E}_1 = \mathbb{R}$,

$$\psi(z_1, z_2) = z_1 + \delta_K(z_2) \text{ and } F(\beta) = \begin{bmatrix} \int_{\mathscr{Y}} g(f(y|\beta))\,dy \\ F_2(\beta) \end{bmatrix}. \tag{3.17}$$

The integral $\int_{\mathscr{Y}} g(f(y|\beta))\,dy$ is the negative entropy of Y, typically denoted $-H(Y|\beta)$. Closed form expressions for $H(Y|\beta)$ exist for several well-known distributions.

3.4 Duality Theory

3.4.1 Infinite Dimensional Convex Duality

We now derive a general duality theorem for `Primal` using the theory of convex perturbation functions. This result parallels those in [55, Theorem 11.39] and [11, Theorem 2.142]. However, we provide a somewhat different statement and proof that takes advantage of the special structure of `Primal`. Specifically, `Primal` is only partially infinite dimensional and the set of probability measures $\mathcal{P}(\Omega)$ is weak* compact.

Deriving the dual problem to `Primal` requires a space paired in duality with $\mathcal{B}(\Omega)$. As in example 2.4, $\mathcal{B}(\Omega)$ is paired in duality with $C(\Omega)$, the space of continuous functions on Ω. The Borel space $C(\Omega)$ is given the strong topology induced by the norm $\|f\| = \sup_{\beta\in\Omega}|f(\beta)|$ and

$\mathcal{B}(\Omega)$ is given the associated weak* topology. The duality pairing is $\langle f, \mu\rangle = \int_\Omega f(\beta)\mu(d\beta)$.

Theorem 3.8. *Let* $f : \mathcal{B}(\Omega) \times \mathcal{E} \to \overline{\mathbb{R}}$ *be proper, lower semicontinuous, and convex, with* $\operatorname{dom} f \subset \mathcal{P}(\Omega) \times \mathcal{E}$*, and define*

$$\begin{aligned} g(\mu) &:= f(\mu, 0) \\ h(z^*) &:= -f^*(0, z^*) \\ p(z) &:= \inf_\mu f(\mu, z) \\ U &:= \operatorname{dom} p. \end{aligned}$$

Consider the primal problem

$$\inf_\mu g(\mu),$$

along with the dual problem

$$\sup_{z^* \in \mathbb{R}^d} h(z^*).$$

Then g *is lsc convex,* h *is usc concave,* p *is lsc convex, and* U *is convex.*

1. *The function* p *is proper. If* $z \in U$*, then there exists* $\mu \in \mathcal{P}(\Omega)$ *such that* $p(z) = f(\mu, z)$.

2. *The weak duality inequality* $\inf_\mu g(\mu) \geq \sup_{z^*} h(z^*)$ *always holds, and*

   $$\inf_\mu g(\mu) < +\infty \iff 0 \in U.$$

   Moreover, if $0 \in U$*, then the duality gap is zero, that is*

   $$\inf_\mu g(\mu) = \sup_{z^*} h(z^*).$$

3. *If* $0 \in \operatorname{ri} U$*, then* $\operatorname{argmax}_{z^*} h(z^*) = \partial p(0) \neq \emptyset$.

4. *The set* $\operatorname{argmax}_{z^*} h(z^*)$ *is non-empty and bounded if and only if* $0 \in \operatorname{int} U$.

5. *Optimal solutions are characterized jointly through primal and dual forms of Fermat's rule:*

$$\left.\begin{array}{c}\overline{\mu} \in \operatorname{argmin}_{\mu} g(\mu) \\ \overline{z^*} \in \operatorname{argmax}_{z^*} h(z^*) \\ \inf_{\mu} g(\mu) = \max_{z^*} h(z^*)\end{array}\right\} \iff (0, \overline{z^*}) \in \partial f(\overline{\mu}, 0) \iff (\overline{\mu}, 0) \in \partial f^*(0, \overline{z^*}).$$

Proof. By their definitions, $p^*(z^*) = -h(z^*)$ and $p^{**}(0) = \sup_{z^*} h(z^*)$. Likewise, convexity of p and U follow because p is the inf-projection in the μ component of the convex function f [53, Theorem 1]. The dual objective h is concave and usc because f^* is convex and lsc.

The probability measures $\mathcal{P}(\Omega)$ are sequentially compact by Prohorov's theorem [58, Theorem 2.4]. Thus $p(z) = f(\mu_z, z)$ for some $\mu_z \in \mathcal{P}(\Omega)$ for all $z \in \operatorname{dom} p$. Furthermore, if $z_k \to \overline{z}$, then, passing to a subsequence if necessary, $(\mu_{z_k}, z_k) \to (\overline{\mu}, \overline{z})$ for some $\overline{\mu} \in \mathcal{P}(\Omega)$, so

$$p(\overline{z}) \leq f(\overline{\mu}, \overline{z}) \leq \liminf_k f(\mu_{z_k}, z_k) = \liminf_k p(z_k)$$

Thus p is lsc. We now prove the remaining claims in sequence .

(1) Attainment was shown in the previous paragraph. If $z \in \mathcal{E}$, then sequential compactness of $\mathcal{P}(\Omega)$ implies there exists a convergence sequence $\{\mu_n\} \subset \mathcal{P}(\Omega)$, $\mu_n \to \overline{\mu}$, with

$$p(z) = \lim_{n\to\infty} f(\mu_n, z).$$

But f is proper lsc, so

$$p(z) = \lim_{n\to\infty} f(\mu_n, z) \geq f(\overline{\mu}, z) > -\infty.$$

Since f is proper, there exists $(\overline{\mu}, \overline{z})$ with $f(\overline{\mu}, \overline{z}) < +\infty$, and hence $p(\overline{z}) < +\infty$. Therefore p is proper.

(2) This follows by the finite-dimensional biconjugacy theorem [55, Theorem 11.1]. In par-

ticular, p is proper convex, so

$$\sup_{z^*} h(z^*) = p^{**}(0) \leq p(0) = \inf_{\mu} g(\mu),$$

proving weak duality. Equality holds if $0 \in U$ since p is lsc.

(3) Since p is proper, convex, lsc, if $0 \in \operatorname{ri}\operatorname{dom} p$, then $\partial p(0) \neq \emptyset$ [54, Theorem 23.4] and is given by $\operatorname{argmin}_{z^*} p^*(z^*)$ [55, Proposition 11.3]. Hence $\partial p(0) = \operatorname{argmax}_{z^*} h(z^*)$ since $h(z^*) = -p^*(z^*)$.

(4) The function p is lsc, convex, and proper, hence so is p^* and $p^{**} = p$ [55, Theorem 11.1]. So $U = \operatorname{dom} p^{**}$. Then $0 \in \operatorname{int}(U)$ if and only if p^* is level bounded [55, Theorem 11.8c]. If p^* is level bounded, then $\operatorname{argmin}_{z^*} p^*(z^*) = \operatorname{argmax}_{z^*} h(z^*)$ is non-empty and bounded [55, Theorem 1.9], while if $\operatorname{argmax}_{z^*} h(z^*)$ is non-empty and bounded, then p^* is level bounded [55, Proposition 3.23].

(5) This is an immediate consequence of Fenchel's inequality and the case where equality holds [53, Corollary 12A]. In this case Fenchel's inequality states

$$g(\mu) - h(z^*) = f(\mu, 0) + f^*(0, z^*) \geq \langle \mu,\, 0\rangle + \langle 0,\, z^*\rangle = 0,$$

and equality holds if and only if $(0, z^*) \in \partial f(\mu, 0)$ or, equivalently, $(\mu, 0) \in \partial f^*(0, z^*)$. If equality holds for $\overline{\mu}$ and $\overline{z^*}$, then $g(\overline{\mu}) = h(\overline{z^*})$. But $g(\mu) \geq h(z^*)$ for all μ, z^*, so equality holding implies $g(\overline{\mu}) = \min_\mu g(\mu)$ and $h(\overline{z^*}) = \max_{z^*} h(z^*)$. Conversely, if $g(\overline{\mu}) = h(\overline{z^*})$, then equality holds in Fenchel's inequality.

□

Recall the optimization problem `Primal` is

$$\min_{\mu \in \mathcal{B}(\Omega)} \psi(S\mu) + \delta_{\mathcal{P}(\Omega)}(\mu).$$

Applying the previous theorem to `Primal` gives a duality theory specific to our problem.

Before stating this result, we discuss a constraint qualification for `Primal`, by which we mean a testable condition under which $0 \in \operatorname{ri} \operatorname{dom} p$.

Definition 3.9. *The* Basic CQ *for* `Primal` *is satisfied if*

$$\operatorname{ri} S\left[\mathcal{P}(\Omega)\right] \cap \operatorname{ri} \operatorname{dom} \psi \neq \emptyset.$$

When ψ takes the form $\psi(z_1, z_2) = \phi(z_1) + \delta_K(z_2)$ and $S\mu = (S_1\mu, S_2\mu)$ then the Basic CQ requires there exists $\mu \in \mathcal{P}(\Omega)$ such that

$$\begin{aligned} S_1\mu &\in \operatorname{ri} S_1\left[\mathcal{P}(\Omega)\right] \cap \operatorname{ri} \operatorname{dom} \phi, \\ S_2\mu &\in \operatorname{ri} S_2\left[\mathcal{P}(\Omega)\right] \cap \operatorname{ri} K. \end{aligned}$$

In the following lemma we collect several straight-forward or well-known results for future reference.

Lemma 3.10. *The following formulas hold.*

1. *[11, Example 2.122]*

$$\partial \delta_{\mathcal{P}}(\nu) = \begin{cases} \left\{\alpha \in C(\Omega) \;\middle|\; \operatorname{supp} \nu \subset \operatorname{argmax}_{\beta \in \Omega} \alpha(\beta)\right\} & \text{if } \nu \in \mathcal{P}(\Omega) \\ \emptyset & \text{else} \end{cases}$$

2. *If $\mu_0 \in \mathcal{P}(\Omega)$ is a finite sum of atomic measures, then*

$$T\left(\mu_0 \,|\, \mathcal{P}(\Omega)\right) = \bigcup_{t>0} \frac{\mathcal{P}(\Omega) - \mu_0}{t}.$$

3. *For $\alpha \in C(\Omega)$,*

$$\delta^*_{\mathcal{P}(\Omega)}(\alpha) = \max_{\beta \in \Omega} \alpha(\beta).$$

4. *Let $S : \mathcal{B}(\Omega) \to \mathcal{E}$ be the linear operator given in Definition 3.7. Then $S^* : \mathcal{E} \to C(\Omega)$ is given by $(S^*w)(\cdot) = \langle w,\, F(\cdot)\rangle$ for all $w \in \mathcal{E}$, where the inner product is that of $\mathcal{E}$.*

5. *Define $p(z) := \inf_{\mu\in\mathcal{P}(\Omega)} \psi(S\mu + z)$. If the Basic CQ is satisfied, then $0 \in \operatorname{ri}\operatorname{dom} p$.*

Proof. (2) By [11, Proposition 2.55], $T\left(\mu_0 \,|\, \mathcal{P}(\Omega)\right) = \mathrm{cl}^*\left[\bigcup_{t>0} \frac{\mathcal{P}(\Omega)-\mu_0}{t}\right]$. Assume $\mu_0 \in \mathcal{P}(\Omega)$ is a finite discrete measure with $\operatorname{supp}\mu_0 = \{\beta_i : i = 1,\ldots,k\}$, and suppose that $\nu \in T\left(\mu_0 \,|\, \mathcal{P}(\Omega)\right)$.

If $A \subset \Omega \setminus \operatorname{supp}\mu_0$ is closed, then Urysohn's lemma [23, 4.15] implies there exists a continuous function $g : \Omega \to [0,1]$ with $g|_A \equiv 1$ and $g|_{\operatorname{supp}\mu_0} \equiv 0$. If $\{\nu_n\} \subset \mathcal{P}(\Omega)$ and $t_n > 0$ are a sequence such that $t_n^{-1}(\nu_n - \mu_0)$ weakly converges to ν, then

$$\frac{1}{t_n}\int_\Omega g(\beta)(\nu_n(d\beta) - \mu_0(d\beta)) \to \int_\Omega g(\beta)\nu(d\beta)$$

and

$$\frac{1}{t_n}\int_\Omega g(\beta)(\nu_n(d\beta) - \mu_0(d\beta)) = \frac{1}{t_n}\int_\Omega g(\beta)\nu_n(d\beta) \geq \frac{1}{t_n}\nu_n(A) \geq 0,$$

so

$$\int_\Omega g(\beta)\nu(d\beta) \geq 0.$$

Since A was arbitrary, this proves $\nu|_{\Omega\setminus\operatorname{supp}\mu_0} \geq 0$. Additionally, since $\nu_n \in \mathcal{P}(\Omega)$ for each n,

$$\int_\Omega (\nu_n(d\beta) - \mu_0(d\beta)) = 1 - 1 = 0,$$

so $\int_\Omega \nu(d\beta) = 0$.

The Lesbesgue-Radon-Nikodym Theorem [23, 3.8] implies there exists unique measures $\nu^+, \rho \in \mathcal{B}(\Omega)(\Omega)$ satisfying

$$\nu = \nu^+ + \rho, \quad \nu^+|_{\operatorname{supp}\mu_0} \equiv 0, \quad \text{and } \operatorname{supp}\rho \subset \operatorname{supp}\mu_0.$$

In particular this implies $\nu^+|_{\Omega\setminus\operatorname{supp}\mu_0} = \nu|_{\Omega\setminus\operatorname{supp}\mu_0}$ so that ν^+ is a non-negative measure.

Define $(\mu_0)_i$ and ρ_i for $i = 1, \ldots, k$ by

$$\mu_0 = \sum_{i=1}^{k} (\mu_0)_i \mathsf{a}_{\beta_i} \text{ and } \rho = \sum_{i=1}^{k} \rho_i \mathsf{a}_{\beta_i}.$$

Let $\widehat{t} > 0$ be chosen such that $\widehat{t}\rho_i + (\mu_0)_i \geq 0$ for $i = 1, \ldots, k$. Then $\widehat{t}\nu + \mu_0$ is a non-negative measure with

$$\int_\Omega \big(\widehat{t}\nu(d\beta) + \mu_0(d\beta)\big) = \widehat{t}\left(\int_\Omega \nu(d\beta)\right) + \int_\Omega \mu_0(d\beta) = 0 + 1 = 1.$$

So $\widehat{\nu} = \widehat{t}\nu + \mu_0$ is a probability measure such that $\widehat{t}^{-1}(\widehat{\nu} - \mu_0) = \nu$, proving the statement.

(3) Observe that for every continuous function $\alpha \in C(\Omega)$,

$$\delta^*_{\mathcal{P}}(\alpha) = \sup_{\mu \in \mathcal{P}} \int_\Omega \alpha(\beta)\, \mu(d\beta) = \max_{\beta \in \Omega} \alpha(\beta),$$

since Ω is compact.

(4) We temporarily use the notation $\langle \mu,\, g \rangle_{\mathcal{B}(\Omega)} = \int g(\beta)\, \mu(d\beta)$ for integrating $g \in C(\Omega)$ with respect to $\mu \in \mathcal{P}(\Omega)$, and $\langle \cdot,\, \cdot \rangle_{\mathcal{E}}$ for the usual inner product in $\mathcal{E}$. Then

$$\begin{aligned}
\langle w,\, S\mu \rangle_{\mathcal{E}} &= \left\langle w,\, \int_\Omega F(\beta)\, \mu(d\beta) \right\rangle_{\mathcal{E}} \\
&= \int_\Omega \langle w,\, F(\beta) \rangle_{\mathcal{E}}\, \mu(d\beta) \\
&= \left\langle \langle w,\, F(\beta) \rangle_{\mathcal{E}},\, \mu \right\rangle_{\mathcal{B}(\Omega)},
\end{aligned}$$

showing that $S^* w(\beta) = \langle w,\, F(\beta) \rangle_{\mathcal{E}} \in C(\Omega)$.

(5) Suppose $v \in \operatorname{dom} p$ so that, by the definition of p, there exists $w \in S\mathcal{P}$ such that $w + v \in \operatorname{dom} \psi$. By the Basic CQ, there exists $u_0 \in\ \operatorname{ri} S\mathcal{P}\ \cap\ \operatorname{ri} \operatorname{dom} \psi$, and $\epsilon_1, \epsilon_2 > 0$ such that

$$u_0 + \epsilon_1(u_0 - (w + v)) \in \operatorname{dom} \psi \text{ and } u_0 + \epsilon_2(u_0 - w) \in S\mathcal{P}.$$

Choose $\epsilon = \min\{\epsilon_1, \epsilon_2\}$. Since $S\mathcal{P}$ is convex and

$$u_0 + \epsilon(u_0 - w) = \left(1 - \frac{\epsilon}{\epsilon_2}\right) u_0 + \frac{\epsilon}{\epsilon_2}\left(u_0 + \epsilon_2(u_0 - w)\right),$$

there exists $\mu_0 \in \mathcal{P}(\Omega)$ such that $S\mu_0 = u_0 + \epsilon(u_0 - w)$ and

$$p(-\epsilon v) \leq \psi(S\mu_0 - \epsilon v) = \psi\left[u_0 + \epsilon_1(u_0 - (w + v))\right] < +\infty.$$

Since $v \in \operatorname{dom} p$ was arbitrary this shows $0 \in \operatorname{ri} \operatorname{dom} p$. □

Theorem 3.11. *Assume ψ is lsc proper convex. The (pre-)dual problem to*

$$\min_{\mu \in \mathcal{B}(\Omega)} \psi(S\mu) + \delta_{\mathcal{P}}(\mu). \tag{Primal}$$

is

$$\min_{w \in \mathcal{E}} \delta^*_{\mathcal{P}}(S^* w) + \psi^*(-w), \tag{Dual}$$

or, equivalently,

$$\begin{aligned} &\min_{w \in \mathcal{E}, \gamma \in \mathbb{R}} \gamma + \psi^*(-w) \\ &\langle w,\, F(\beta)\rangle \leq \gamma \quad \forall \beta \in \Omega, \end{aligned} \tag{Dual}$$

where the inner product is the inner product on $\mathcal{E}$.

1. *If $S\left[\mathcal{P}(\Omega)\right] \cap \operatorname{dom} \psi \neq \emptyset$, then there exists an optimal solution to* `Primal`*.*

2. *(Weak Duality) For all $\mu \in \mathcal{B}(\Omega)$ and $y \in \mathcal{E}$,*

$$\left[\psi(S\mu) + \delta_{\mathcal{P}(\Omega)}(\mu)\right] + \left[\delta^*_{\mathcal{P}(\Omega)}(S^* w) + \psi^*(-w)\right] \geq 0. \tag{3.18}$$

3. *If $\mu \in \mathcal{B}(\Omega)$ and $w \in \mathcal{E}$ are such that the equality is achieved in (3.18), then μ solves*

`Primal` *and* w *solves* `Dual`.

4. *(Strong Duality) If the Basic CQ is satisfied, then the optimal values of* `Primal` *and* `Dual` *are finite, attained at some* μ *and* w *respectively, and equality is achieved in (3.18) at* (μ, w).

5. *(Coercivity) If* $\operatorname{int}(S[\mathcal{P}(\Omega)] \cap \operatorname{dom}\psi) \neq \emptyset$*, then the Basic CQ is satisfied and the optimal solution set of* `Dual` *is compact.*

Proof. The first form of `Dual` is derived from Theorem 3.8 using the dualizing parameterization $f(\mu, z) = \psi(S\mu + z) + \delta(\mu|\mathcal{P}(\Omega))$. By Lemma 3.10 parts (3) and (4),

$$\delta^*_{\mathcal{P}(\Omega)}(S^*w) = \max_{\beta\in\Omega} \langle w,\ F(\beta)\rangle\,,$$

so the dual problem can be rephrased as

$$\begin{aligned}&\min_{w\in\mathcal{E},\gamma\in\mathbb{R}} \gamma + \psi^*(-w) \\ &\langle w,\ F(\beta)\rangle \leq \gamma \quad \forall\beta\in\Omega.\end{aligned} \tag{Dual}$$

(1) The given condition is equivalent to $0 \in \operatorname{dom} p$, so the result follows from Theorem 3.8(1).

(2) By the definition of the convex conjugate,

$$\langle S^*w,\ \mu\rangle \leq \delta^*_{\mathcal{P}(\Omega)}(S^*w) + \delta_{\mathcal{P}(\Omega)}(\mu), \tag{3.19}$$

$$\langle -w,\ S\mu\rangle \leq \psi^*(-w) + \psi(S\mu), \tag{3.20}$$

where $\mu \in \mathcal{P}(\Omega)$ and $w \in \mathcal{E}$. Adding the above inequalities proves (3.18), weak duality.

(3) Inequality (3.18) can be rearranged to

$$\psi(S\mu) + \delta_{\mathcal{P}(\Omega)}(\mu) \geq -\left[\delta^*_{\mathcal{P}(\Omega)}(S^*w) + \psi^*(-w)\right] \quad \text{for all } \mu\in\mathcal{B}(\Omega) \text{ and } w\in\mathcal{E}.$$

In that form it is clear that if $\mu \in \mathcal{P}(\Omega)$ and $w \in \mathcal{E}$ exist such that equality is achieved in (3.18), then μ is optimal for `Primal` and w is optimal for `Dual`.

(4) The optimal value function for our dualizing parameterization is $p(z) = \inf_\mu f(\mu, z)$. By Lemma 3.10(5), $0 \in \operatorname{ri} \operatorname{dom} p$. Hence, by Theorem 3.8(3,5), the optimal values of `Primal` and `Dual` are finite, equal, and attained in the dual. The existence of primal solution is given in Theorem 3.8(1).

(5) If $\operatorname{int}(S\mathcal{P} \cap \operatorname{dom} \psi) \neq \emptyset$, then $0 \in \operatorname{int}(\operatorname{dom} p)$, and so the result follows by from Theorem 3.8(4). □

Remark 3.12. Note `Dual` is a semi-infinite program. Therefore, one approach to understanding the duality is through results on semi-infinite programming, e.g. [11, Section 5.4]. However, the direct approach taken here is significantly more transparent and expedient.

3.4.2 Finite Dimensional Reduction

Since S is a continuous linear transformation and $\mathcal{P}(\Omega)$ is a weak* compact set, the set $S(\mathcal{P}(\Omega)) \subset \mathcal{E}$ is a compact convex set. Hence the infinite dimensional convex problem `Primal` has the same optimal value as the finite-dimensional convex problem

$$\min_{z \in S(\mathcal{P}(\Omega))} \psi(z). \tag{F-Primal}$$

This reformulation of `Primal` has two obvious drawbacks. First, we do not have a manageable representation of the set $S(\mathcal{P}(\Omega))$ for the purposes of computation, and second, even if we are given the solution w to `F-Primal`, we have no method for recovering a solution μ to `Primal`.

A remedy for both drawbacks can be found through an application of the Krein-Milman and Caratheodory Theorems. For $x = (\beta_1, \ldots, \beta_{\widehat{n}}) \in \Omega^{\widehat{n}}$, define

$$A_F(x) \in \operatorname{Lin}(\mathbb{R}^{\widehat{n}}, \mathcal{E}), \quad A_F(x)\lambda = \sum_{i=1}^{\widehat{n}} \lambda_i F(\beta_i). \tag{3.21}$$

In this notation, the dimension of the domain of $A_F(x)$ implicitly depends on the dimension of the argument x, i.e. A_F is a function $A_F : \Omega^{\widehat{n}} \to \mathrm{Lin}(\mathbb{R}^{\widehat{n}}, \mathcal{E})$.

The following theorem is a generalization from [2].

Theorem 3.13. *Let ψ, S, and Ω be as in* `Primal`*. Then for any integer $\widehat{n} > \dim \mathcal{E}$,*

$$S\left(\mathcal{P}(\Omega)\right) = \left\{ A_F(x)\lambda \middle| \, x \in \Omega^{\widehat{n}} \text{ and } \lambda \in \Delta^{\widehat{n}-1} \right\}. \tag{3.22}$$

Hence, for any integer $\widehat{n} > \dim \mathcal{E}$, the problem `Primal` *has the same optimal value as the finite dimensional problem*

$$\begin{aligned} \min \quad & \psi(z) \\ s.t. \quad & z = A_F(x)\lambda, \\ & \lambda \in \Delta^{\widehat{n}-1} \text{ and } x \in \Omega^{\widehat{n}}. \end{aligned} \tag{F-Primal}$$

Futhermore, a solution $(\lambda, x) \in \Delta^{\widehat{n}-1} \times \Omega^{\widehat{n}}$ to `F-Primal` *exists, and for each such solution the discrete measure $\mu := \sum_{i=1}^{\widehat{n}} \lambda_i \mathsf{a}_{\beta_i} \in \mathcal{P}(\Omega)$ is a solution to* `Primal`*.*

Proof. First we establish the representation (3.22). For convenience let $\mathcal{P}$ denote $\mathcal{P}(\Omega)$. By the Krein-Milman Theorem, $\mathcal{P} = \mathrm{cl}^*\mathrm{co}(\mathrm{ext}\,\mathcal{P})$, where cl^* denotes the weak* closure and co denotes the convex hull. Since S is a continuous linear operator and $\mathcal{P}$ is weak* compact,

$$S\mathcal{P} \subset \mathrm{cl}(S(\mathrm{co}(\mathrm{ext}\,\mathcal{P}))) \subset \mathrm{cl}(S\mathcal{P}) = S\mathcal{P}.$$

So $S\mathcal{P} = \mathrm{cl}(S(\mathrm{co}(\mathrm{ext}\,\mathcal{P})))$.

The extreme points of $\mathcal{P}$ are $\mathrm{ext}\,\mathcal{P} = \left\{ \mathsf{a}_\beta \middle| \, \beta \in \Omega \right\}$ [3, Proposition 8.4], so

$$S(\mathrm{ext}\,\mathcal{P}) = \left\{ \int_\Omega F(\gamma)\, a_\beta(d\gamma) \middle| \, \beta \in \Omega \right\} = F(\Omega).$$

Since $F(\Omega)$ is compact, $\mathrm{cl}(\mathrm{co}\, S(\mathrm{ext}\,\mathcal{P})) = \mathrm{co}\, S(\mathrm{ext}\,\mathcal{P})$ and so

$$S\mathcal{P} = \mathrm{co}\,[S(\mathrm{ext}\,\mathcal{P})]\,,$$

Using Caratheodory's Theorem, we may represent $\mathrm{co}\,[S(\mathrm{ext}\,\mathcal{P})]$ as in (3.22) for any fixed integer with $\widehat{n} > \dim \mathcal{E}$. Finally, replacing $S\mathcal{P}$ with this representation of $S\mathcal{P}$ gives the equivalent finite-dimensional optimization problem `F-Primal`.

The optimal value in `F-Primal` is finite and attained by Weierstrass' Theorem [7, Proposition 3.2.1], since the objective is proper lower semicontinuous and the constraint region is compact. The measure $\widehat{\mu} = \sum_{i=1}^{\widehat{n}} \widehat{\lambda}_i \mathsf{a}_{\widehat{\beta}_i}$ is optimal for `Primal` if $(\widehat{\lambda}, \widehat{x})$ is optimal for `F-Primal` because $\psi(S\widehat{\mu}) = \psi(A_F(\widehat{x})\widehat{\lambda})$, so the optimal values coincide . □

Remark 3.14. In the proof of Theorem 3.13 only $\widehat{n} > \dim S\,[\mathcal{P}(\Omega)]$ is required. This reduction in the number of required support points required is useful if F maps Ω to a lower dimensional affine set in $\mathcal{E}$.

An immediate consequence of the above theorem is that the primal infinite dimensional problem has a finite optimal value, and this optimal value is attained by a finite sum of atomic probability measures. The reduction using Caratheory's Theorem has appeared many times in the literature related to the applications given, for example see [22, Theorem 2.1.2]. The price paid to achieve this finite dimensional reformulation of `Primal` is that `F-Primal` is not convex. It is also interesting to note the addition of cone constraints to objectives, as in $\psi(z_1, z_2) = \phi(z_1) + \delta_K(z_2)$, increases the maximum number of support points required for a finitely supported optimal measure.

3.4.3 Convex-Composite Duality

Writing the objective in `F-Primal` as

$$\min_{x\in\Omega^{\widehat{n}},\ \lambda\in\Delta^{\widehat{n}-1}} \psi(A_F(x)\lambda),$$

we see that F-Primal is convex-composite.

Theorem 3.15. *The convex-composite dual* F-Dual *to* F-Primal *is equivalent to the infinite dimensional dual problem* Dual*. Consequently strong duality holds for the (non-convex dual) problems* F-Primal *and* F-Dual *when strong duality holds for the (infinte dimensional) convex problems* Primal *and* Dual*.*

Proof. The convex-composite dual problem takes the form

$$\sup_{w\in\mathcal{E}} \inf_{\substack{\lambda\in\Delta^{\widehat{n}-1}\\ x\in\Omega^{\widehat{n}}}} \langle w,\, A_F(x)\lambda\rangle - \psi^*(w) = \sup_{w\in\mathcal{E}} \left\{ -\psi^*(w) + \inf_{\substack{\lambda\in\Delta^{\widehat{n}-1}\\ x\in\Omega^{\widehat{n}}}} \langle w,\, A_F(x)\lambda\rangle \right\}. \tag{F-Dual}$$

We can further simplify the inner term:

$$\begin{aligned}
\inf_{\substack{\lambda\in\Delta^{\widehat{n}-1}\\ x\in\Omega^{\widehat{n}}}} \langle w,\, A_F(x)\lambda\rangle &= \inf_{x\in\Omega^{\widehat{n}}} \left(\inf_{\lambda\in\Delta^{\widehat{n}-1}} \langle w,\, A_F(x)\lambda\rangle \right) \\
&= \inf_{x\in\Omega^{\widehat{n}}} \left(\inf_{\lambda\in\Delta^{\widehat{n}-1}} \sum_{i=1}^{\widehat{n}} \lambda_i \langle w,\, F(\beta_i)\rangle \right) \\
&= \inf_{x\in\Omega^{\widehat{n}}} \min_{i=1,\dots,\widehat{n}} \langle w,\, F(\beta_i)\rangle \\
&= \inf_{\beta\in\Omega} \langle w,\, F(\beta)\rangle
\end{aligned}$$

So F-Dual can be restated as

$$\begin{aligned}
&\sup_{w\in\mathcal{E}} \left\{ -\psi^*(w) + \inf_{\beta\in\Omega} \langle w,\, F(\beta)\rangle \right\} \qquad \text{(F-Dual)}\\
&= -\left(\inf_{w\in\mathcal{E}} \left\{ \psi^*(w) + \sup_{\beta\in\Omega} \langle -w,\, F(\beta)\rangle \right\} \right)
\end{aligned}$$

By removing the outer minus sign and replacing w by $-w$ gives the equivalent problem

$$\begin{aligned}
\inf_{w\in\mathcal{E}} \left[\psi^*(-w) + \sup_{\beta\in\Omega} \langle w,\, F(\beta)\rangle \right] = \quad & \inf_{w\in\mathcal{E}} \psi^*(-w) + \gamma \\
& \langle w,\, F(\beta)\rangle \le \gamma \quad \forall \beta\in\Omega,
\end{aligned}$$

which is the convex dual of `Primal`. □

3.4.4 First-order Stationarity Conditions

Theorem 3.15 tells us that the problems `Primal` and `F-Primal` have the same dual even though `F-Primal` is non-convex. In this section we investigate the implications of this relationship on the first-order stationarity conditions for `Primal` and `F-Primal`. We begin by defining first-order conditions and then prove they are necessary conditions for their respective problems. Recall that $\partial\delta_{\mathcal{P}(\Omega)}$ was computed in Lemma 3.10(1) and that, as in the notation section, the normal cone $N\left(\beta\,|\,\Omega\right)$ is defined even when Ω is non-convex.

Definition 3.16. *1. We say that the first-order stationarity conditions for* `Primal` *are satisfied at* $\mu \in \mathcal{P}(\Omega)$ *if*

$$S^*w \in \partial\delta_{\mathcal{P}(\Omega)}(\mu) \tag{3.23}$$

$$-w \in \partial\psi(S\mu). \tag{3.24}$$

Any pair $(\mu,\, w)$ *satisfying these conditions is called a first-order (stationary) point for* `Primal`*.*

2. We say that the first-order stationarity conditions for `F-Primal` *are satisfied at* $(\lambda,\ (\beta_1, \ldots,\ \beta_{\widehat{n}}), w) \in \Delta^{\widehat{n}-1} \times \Omega^{\widehat{n}} \times \mathcal{E}$ *if*

$$\nabla F(\beta_i)^T w \in N\left(\beta_i\,|\,\Omega\right), \quad i = 1, \ldots, \widehat{n} \tag{3.25}$$

$$A_F(x)^T w \in \partial\delta_{\Delta^{\widehat{n}-1}}(\lambda) \tag{3.26}$$

$$-w \in \partial\psi(A_F(x)\lambda) \tag{3.27}$$

where $A_F(x)$ *is as defined in 3.21. Any triple* $(\lambda,\ (\beta_1, \ldots,\ \beta_{\widehat{n}}), w)$ *satisfying these conditions is called a first-order stationary point for* `F-Primal`*.*

For brevity, we will say *first-order conditions* to mean *first-order stationary conditions.*

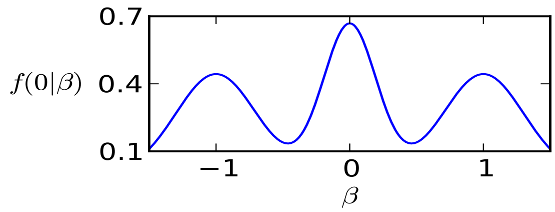

Figure 3.2: The graph of the function used in Example 3.17

Before showing that these conditions are necessary first-order optimality conditions for the problems (`Primal`) and (`F-Primal`), we examine the relationships between the conditions in Definition 3.16. Consider the discrete measure $\mu = \sum_{i=1}^{\widehat{n}} \lambda_i \mathsf{a}_{\beta_i}$ where $(\lambda,\ (\beta_1, \ldots,\ \beta_{\widehat{n}})) \in \Delta^{\widehat{n}-1} \times \Omega^{\widehat{n}}$, and set $x = (\beta_1, \ldots, \beta_{\widehat{n}})$. Then $S\mu = A_F(x)\lambda$. Hence, for $w \in \mathcal{E}$, the first-order condition (3.24) applied to the pair (μ, w) is exactly the same as the condition (3.27). On the other hand, the condition (3.23) corresponds to the two conditions (3.25) and (3.26) combined. In order to see this, note that condition (3.23) applied to $\mu = \sum_{i=1}^{\widehat{n}} \lambda_i \mathsf{a}_{\beta_i}$ is equivalent to

$$\{\beta_j\}_{\lambda_j \neq 0} \subset \operatorname{argmax}_{\beta \in \Omega} \langle w,\ F(\beta) \rangle\,, \tag{3.28}$$

by the representations for S^* and $\partial \delta_{\mathcal{P}}$ given in Lemma 3.10 parts (4) and (1), respectively. By comparison, for each $i = 1, \ldots, \widehat{n}$, the inclusion in (3.25) is the first-order necessary optimality condition for the optimization problem on the right-hand side of (3.29), and one can show that (3.26) is equivalent to

$$\{\beta_j\}_{\lambda_j \neq 0} \subset \operatorname{argmax}_{\beta \in \{\beta_1, \ldots, \beta_{\widehat{n}}\}} \langle w,\ F(\beta) \rangle\,.$$

While the first-order conditions for `Primal` and `F-Primal` have strong similarities, there can exist local minima to `F-Primal` which satisfy (3.25) – (3.27) but are nonetheless not global minima. The following simple example illustrates how this can occur.

Example 3.17. Consider the `NPMLE` mixture model problem. As in Example 3.2, let $n(w|\gamma, \sigma^2)$

denote the normal density with mean γ and variance σ^2. Define

$$f(w|\beta) = \frac{1}{3}n(w|-1+\beta, 0.3) + \frac{1}{3}n(w|\beta, 0.2) + \frac{1}{3}n(w|1+\beta, 0.3),$$

and $\Omega = [-10, 10]$. We suppose there is a single observation, $w = 0$, set $\widehat{n} = 2$, and add no cone constraints. The graph of $f(0|\beta)$ as a function of β is depicted in Figure 3.2.

Then $\partial\psi(z) = \{-z^{-1}\}$ where the inverse is element-wise. The conditions $-w \in \partial\psi(S\mu)$ and $-w \in \partial\psi(A_F(x)\lambda)$ both simplify to $w_i = \frac{1}{\int_\Omega f(x_i|\beta)\,\mu(d\beta)}$. In the later case μ is the discrete density $\mu = \lambda_1 \mathsf{a}_{\beta_1} + \lambda_2 \mathsf{a}_{\beta_2}$.

The condition $S^*w \in \partial\delta_{\mathcal{P}(\Omega)}(\mu)$ implies the support points of μ belong to the *global* maximum set of $f(0|\beta)$, as a function of β. Thus $\mu = \mathsf{a}_0$ is the only $\mu \in \mathcal{P}(\Omega)$ satisfying the first-order conditions.

Next consider the finite dimensional problem `F-Primal`. The globally optimal solution to `Primal` is a_0, and $(\lambda, (\beta_1, \beta_2))$ is a globally optimal solution to `F-Primal` if $\sum_{i=1,2} \lambda_i \mathsf{a}_{\beta_i} = \mathsf{a}_0$. But these aren't the only first-order stationary points. Take $\widehat{\beta_1}$ and $\widehat{\beta_2}$ to be the two modes of $f(0|\beta)$ near ± 1. Then for any $\lambda \in \Delta^{\widehat{n}-1}$, the point $(\lambda, (\widehat{\beta_1}, \widehat{\beta_2}))$ is a first-order stationary point for `F-Primal`, but $(\lambda, (\widehat{\beta_1}, \widehat{\beta_2}))$ is not globally optimal and $\sum_{i=1,2} \lambda_i \mathsf{a}_{\widehat{\beta_i}}$ is not a first-order stationary point for `Primal`.

We conclude this section by showing that the first-order conditions given in Definition 3.16 are, respectively, the first-order optimality conditions for `Primal` and `F-Primal`.

Theorem 3.18. *1. If (μ, w) is a first-order stationary point for* `Primal`*, then μ solves* `Primal` *and w solves* `Dual`*.*

2. If $S[\mathcal{P}(\Omega)] \cap \operatorname{dom}\psi \neq \emptyset$, then a solution to `Primal` *always exists. If, in addition, the Basic CQ is satisfied, then for every optimal solution μ to* `Primal` *there exists $w \in \mathcal{E}$ such that (μ, w) is a first-order stationary point and w solves* `Dual`*.*

3. A (global) optimal solution to `F-Primal` *always exists. If the Basic CQ is satisfied and Ω is convex, then for every globally optimal solution $(\lambda, (\beta_1, \ldots, \beta_{\widehat{n}}))$ to* `F-Primal` *there exists a* $w \in \mathcal{E}$ *such that* $(\lambda, (\beta_1, \ldots, \beta_{\widehat{n}}), w)$ *is a first-order stationary point for* `F-Primal` *and* $(\sum_{i=1}^{\widehat{n}} \lambda_i \mathsf{a}_{\beta_i}, w)$ *is a first-order stationary point for* `Primal`*.*

Proof. (1) If (μ, w) is a first-order stationary point for `Primal`, then equality is achieved in (3.19) and (3.20), so by Theorem 3.11(3) μ is an optimal solution for `Primal` and w is an optimal solution for `Dual`.

(2) A solution to `Primal` always exists by Theorem 3.11(1). If the Basic CQ is satisfied, then by Theorem 3.11(4) there exist $\mu \in \mathcal{P}(\Omega)$ and $w \in \mathcal{E}$ satisfying (3.18), so $-w \in \partial\psi(S\mu)$ and $S^*w \in \partial\delta_{\mathcal{P}(\Omega)}(\mu)$, i.e. (μ, w) is a first-order stationary point.

(3) By Theorem 3.13 a global optimal solution $(\lambda, (\beta_1, \ldots, \beta_{\widehat{n}}))$ to `F-Primal` always exists, and $\mu := \sum_{i=1}^{\widehat{n}} \lambda_i \mathsf{a}_{\beta_i}$ solves `Primal`. So by part (2) there exists an optimal solution w for `Dual`. Thus (μ, w) is a first-order stationary point for `Primal`. Set $x = (\beta_1, \ldots, \beta_{\widehat{n}})$ so that $S\mu = A_F(x)\lambda$. In particular, this shows that equations (3.24) and (3.27) are equivalent for $\mu = \sum_{i=1}^{\widehat{n}} \lambda_i \mathsf{a}_{\beta_i}$. Condition (3.23) applied to $\mu = \sum_{i=1}^{\widehat{n}} \lambda_i \mathsf{a}_{\beta_i}$ is equivalent to

$$\{\beta_j\}_{\lambda_j \neq 0} \subset \operatorname{argmax}_{\beta\in\Omega} \langle w,\, F(\beta)\rangle\,, \tag{3.29}$$

by the representations for S^* and $\partial\delta_{\mathcal{P}}$ given in Lemma 3.10 parts (4) and (1), respectively. By [55, Corollary 10.9], every $\overline{\beta} \in \operatorname{argmin}_{\beta\in\Omega} \langle -w,\, F(\beta)\rangle$ satisfies

$$\nabla F\left(\overline{\beta}\right)^T w \in N\left(\overline{\beta}\,\middle|\,\Omega\right),$$

from which (3.25) follows. □

3.4.5 Finite Mixture CQ for `Primal`

In this section we exploit the close connection between the primal and dual problems in the infinite and finite dimensional representations to obtain a constraint qualification in terms of the structures associated with the finite dimensional problem `F-Primal`. We show that this constraint qualification implies the Basic CQ for `Primal` given in Definition 3.9.

Definition 3.19. Finite Mixture Constraint Qualification *We say that the* Finite Mixture CQ *for* `Primal` *is satisfied if there exists $d > 0$ and $x = (\beta_1, \ldots, \beta_d) \in \Omega^d$ such that*

$$A_F(x)\,\mathrm{ri}\,\Delta^{d-1} \cap \mathrm{ri}\,\mathrm{dom}\,\psi \neq \emptyset, \quad \textit{and} \tag{3.30}$$

$$\mathrm{Par}(S\mathcal{P}) \times \mathbb{R} \subset \mathrm{Ran}\left(\begin{bmatrix} A_F(x) \\ \mathsf{e}^T \end{bmatrix}\right). \tag{3.31}$$

The dimension d in the above definition is unrelated to the maximum number of support points required for a finitely supported solution in Theorem 3.13. That theorem concerns solutions of `Primal` while the above constraint qualification concerns the finite dimensional problem `F-Primal`. Conditions similar to (3.31) have appeared elsewhere for moment constrained probability measures [34, 61].

The connection between the (3.30) and (3.31) and the Basic CQ is revealed by the following lemma and theorem.

Lemma 3.20. *Let $\lambda \in \mathrm{ri}\,\Delta^{d-1}$ and $x = (\beta_1, \ldots, \beta_d) \in \Omega^d$. If (3.31) is satisfied, then $A_F(x)\lambda \in \mathrm{ri}\,S\mathcal{P}(\Omega)$.*

Proof. Set $\widehat{\mu} = \sum_{i=1}^d \lambda_i \mathsf{a}_{\beta_i}$ and $\widehat{v} = S\widehat{\mu}$, and let $v \in S\mathcal{P}$. By [54, Theorem 6.4], the result follows once we show there exists $\alpha_v > 0$ such that $\widehat{v} + \alpha_v(\widehat{v} - v) \in S\mathcal{P}$. By assumption there exists $z \in \mathbb{R}^d$ such that $(\widehat{v} - v) = A_F(x)z$ and $e^T z = 0$. So for all $\alpha \in \mathbb{R}$,

$$e^T(\lambda + \alpha z) = 1 \quad \text{and} \quad \widehat{v} + \alpha(\widehat{v} - v) = A_F(x)(\lambda + \alpha z).$$

Since $\lambda > 0$ there exists $\overline{\alpha}_v > 0$ with $\lambda + \overline{\alpha}_v z > 0$. Hence $\mu = \sum_{i=1}^d (\lambda_i + \overline{\alpha}_v z_i)\mathsf{a}_{\beta_i} \in \mathcal{P}(\Omega)$ and

$$\hat{v} + \overline{\alpha}_v(\hat{v} - v) = A_F(x)(\lambda + \overline{\alpha}_v z) = S\mu \in S\mathcal{P}\ .$$

□

Theorem 3.21. *The Finite Mixture CQ implies the Basic CQ.*

Proof. If the Finite Mixture CQ for `Primal` is satisfied, then there exists $\lambda \in \operatorname{ri}\Delta^{d-1}$ and $x = (\beta_1, \ldots, \beta_d) \in \Omega^d$ such that (3.30) and (3.31) are satisfied. Consequently, $A_F(x)\lambda \in S\mathcal{P}(\Omega) \cap \operatorname{ri}\operatorname{dom}\psi$, and the previous lemma shows $A_F(x)\lambda \in \operatorname{ri} S\mathcal{P}(\Omega)$. Set $\mu = \sum_{i=1}^{\widehat{n}} \lambda_i \mathsf{a}_{\beta_i}$ so that $A_F(x)\lambda = S\mu$. This gives $S\mu \in \operatorname{ri} S\mathcal{P}(\Omega) \cap \operatorname{ri}\operatorname{dom}\psi$, showing the Basic CQ is satisfied. □

In the applications considered in Section 3.3 the set $\operatorname{dom}\psi$ is a convex cone. In this case there is a simple way to check (3.30).

Proposition 3.22. *Let $x \in \Omega^{\widehat{n}}$. If $\operatorname{dom}\psi$ is a convex cone with*

$$\left[A_F(x)^T w \geq 0 \textit{ and } w \in (\operatorname{dom}\psi)^\circ\right] \Rightarrow w = 0,$$

then (3.30) *is satisfied.*

Proof. The hypothesis can be restated as

$$\ker\left[A_F(x)^T \quad I_{d\times d}\right] \bigcap (\operatorname{dom}\psi)^\circ \times \mathbb{R}^d_- = \{0\}. \tag{3.32}$$

Define

$$L = \operatorname{Ran}\begin{bmatrix} A_F(x) \\ I_{d\times d} \end{bmatrix} \text{ and } \widehat{K} = \operatorname{cl}(\operatorname{dom}\psi) \times \mathbb{R}^d_+.$$

A standard theorem [54, Corollary 16.4.2] shows taking the polar of both sides of (3.32) yields

$$\operatorname{cl}\left(L + \widehat{K}\right) = \mathcal{E} \times \mathbb{R}^d.$$

If

$$L \bigcap \operatorname{ri}\widehat{K} = \emptyset,$$

then a standard separation theorem [54, Theorem 11.2] implies there exists $v \in \mathcal{E} \times \mathbb{R}^d$ such that $\langle v, x\rangle = 0$ for all $x \in L$ and $\langle v, y\rangle \geq 0$ for all $y \in \widehat{K}$. This obviously implies $\langle v, z\rangle \geq 0$ for all $z \in L + \widehat{K}$, contradicting $\operatorname{cl}\left(L + \widehat{K}\right) = \mathcal{E} \times \mathbb{R}^d$. Therefore $L \cap \operatorname{ri}\widehat{K} \neq \emptyset$. In particular, this immediately implies there exists $(z, \lambda) \in \operatorname{ri}\operatorname{dom}\psi \times \operatorname{ri}\mathbb{R}^d_+$ such that $z = A_F(x)\lambda$. So $A_F(x)\lambda \in \operatorname{ri}\operatorname{dom}\psi$. Since $\operatorname{dom}\psi$ is a convex cone we may scale λ so that $\lambda \in \Delta^{d-1}$ and $A_F(x)\lambda \in \operatorname{ri}\operatorname{dom}\psi$ remains true. □

The condition (3.31) is often established by showing the matrix on the right-hand side is surjective. We return to this issue in the final section of the paper, where we revisit the application in Section 3.3.

3.5 Structure of the Solution Set

Since the set of optimal solutions to a convex optimization problem is itself convex, the optimal solution set is infinite whenever it is not a singleton. When the optimal solution is not unique, understanding the properties of the set of optimal solutions is important for both a deeper understanding of the application and the performance of a numerical solution method. In this section, we focus on the problem structure specified by (3.8) and (3.11) and characterize the optimal solution set $\mathscr{S}$ for `Primal`. This characterization shows that unique solutions are unlikely. This should be no surprise since the linear operator S necessarily contains a non-trivial kernel.

There are two notions of non-uniqueness that appear for solutions of mixture models: identifiability vs. uniqueness of optimal measures. We focus on the uniqueness of the optimal measure, but to avoid confusion we briefly review identifiability.

Definition 3.23. *A family of mixture densities*

$$\mathcal{F} = \{p_\mu(y) : \; p_\mu(y) = \int_\Omega p(y, \omega)\, \mu(d\omega), \mu \in \mathcal{C} \subset \mathcal{P}(\Omega)\}$$

is identifiable *if*

$$p_{\mu_1} = p_{\mu_2} \Rightarrow \mu_1 = \mu_2.$$

The following concrete example demonstrates how a non-identifiable family can lead to non-uniqueness of mixing measures.

Example 3.24. Consider a family $\mathcal{F}$ of mixtures of uniform densities given by

$$
\begin{aligned}
L &\in \mathbb{R}_{++} \\
\Omega &= [-L, L]^2 \subset \mathbb{R}^2 \\
\mathcal{C} &= \{\, \mu \in \mathcal{P}(\Omega) |\, \operatorname{supp} \mu \subset \{\, (a,b) \in \Omega |\, a < b \,\}\,\} \\
p(y, \omega) &= \frac{1}{b-a} \mathbf{1}_{[a,b]}(y), \text{ where } \omega = (a,b)
\end{aligned}
$$

Then $\mathcal{F}$ is not identifiable since, for example, both of the mixing measures $\mathsf{a}_{(0,1)}$ and $\frac{1}{2}\mathsf{a}_{(0,\frac{1}{2})} + \frac{1}{2}\mathsf{a}_{(\frac{1}{2},1)}$ give the same mixture density, $\mathbf{1}_{[0,1]}$, almost everywhere.

Identifiability is principally a concern for the statistical model rather than the fitting procedure. We refer the reader to [24, 30, 31, 63] and their references for more information on identifiability. We merely note its distinction from whether optimal measures are unique.

Given the optimal value we can characterize the entirety of $\mathscr{S}$, the optimal solution set for `Primal`.

Proposition 3.25. *Assume that $\psi(z_1, z_2) = \phi(z_1) + \delta_K(z_2)$ and $S = (S_1, S_2)$. If α^* is the optimal value of* `Primal` *and $\mu_0 \in \mathcal{B}(\Omega)$ is an optimal solution, then*

$$
\mathscr{S} = S_1^{-1}(\operatorname{lev}_{\phi}^{=}(\alpha^*)) \cap S_2^{-1}K \cap \mathcal{P}(\Omega), \tag{3.33}
$$

and

$$
T\left(\mu_0 \,|\, \mathscr{S}\right) \subset T\left(\mu_0 \,\middle|\, S_1^{-1}(\operatorname{lev}_{\phi}^{\leq}(\alpha^*))\right) \cap T\left(\mu_0 \,\middle|\, S_2^{-1}K\right) \cap T\left(\mu_0 \,|\, \mathcal{P}(\Omega)\right). \tag{3.34}
$$

If ϕ is strictly convex and $\mu_0 \in \mathscr{S}$, then $S_1^{-1}(\operatorname{lev}_{\phi}^{=}(\alpha^*))$ may be replaced with $\mu_0 + \ker S_1$. If ϕ is strictly convex, K is polyhedral, and $\mu_0 \in \mathscr{S}$ is finitely supported, then (3.34) *can be strengthened to*

$$
T\left(\mu_0 \,|\, \mathscr{S}\right) = \ker S_1 \cap T\left(\mu_0 \,\middle|\, S_2^{-1}K\right) \cap T\left(\mu_0 \,|\, \mathcal{P}(\Omega)\right). \tag{3.35}
$$

Proof. Formula (3.33) is straightforward. If $\mu \in \mathscr{S}$, then we must have $\psi(S_1\mu) = \alpha^*$, $S_2\mu \in K$, and $\mu \in \mathcal{P}(\Omega)$. Conversely if $\mu \in \mathcal{P}(\Omega)$, $S_2\mu \in K$, and $\psi(S_1\mu) = \alpha^*$ then $\mu \in \mathscr{S}$.

The inclusion (3.34) follows immediately from the definitions and (3.33).

If ϕ is strictly convex, then $S_1\mu_0$ is unique, so $\operatorname{lev}^{=}_{\phi}(\alpha^*)$ may be replaced with $\mu_0 + \ker S_1$.

If K is polyhedral, then so is $S_2^{-1}K$. Thus, in particular, the tangent cone to $S_2^{-1}K$ does not require taking the closure, so that $T(\nu|S_2^{-1}K) = \cup_{t>0} t^{-1}(S_2^{-1}K - \nu)$ for all $\nu \in S_2^{-1}K$. Lemma 3.10(2) shows the same is true of $\mathcal{P}(\Omega)$ at finitely supported measures. Since the tangent cone does not require taking a closure for each of the sets which intersect to form $\mathscr{S}$, the tangent cone to the intersection is the same as the intersection of the tangent cones. $\square$

By Lemma 3.10(2), if $\mu = \sum_i \lambda_i \mathsf{a}_{\beta_i} \in \mathcal{P}(\Omega)$ with $\lambda_i > 0$, then

$$T\left(\mu_0 \,|\, \mathcal{P}(\Omega)\right) = \left\{ \mu \in \mathcal{B}(\Omega) \,\middle|\, \int_\Omega \mu(\beta) = 0,\ \mu|_{\Omega \setminus \operatorname{supp} \mu_0} \geq 0 \right\},$$

consequently the tangent cone for finite discrete measures is very large. This makes it very difficult for the intersection in (3.34) to be trivial. However it can still occur, as the following example shows.

Example 3.26. In Example 3.17 the solution a_0 was unique. The uniqueness is easy to verify because the first-order condition (3.23) is

$$\operatorname{supp} \mu \subset \operatorname{argmax}_{\beta \in \Omega} (S^* w)(\beta) = \operatorname{argmax}_{\beta \in \Omega} f(0|\beta) w = \{0\},$$

where the last equality follows because $w > 0$ by (3.24). Since the solution is unique, ψ is strictly convex and there are no constraints, the previous theorem tells us that $T\left(\mathsf{a}_0 \,|\, \mathscr{S}\right) = \{0\}$. In particular,

$$T\left(\mathsf{a}_0 \,|\, \mathcal{P}(\Omega)\right) = \left\{ \nu \in \mathcal{P}(\Omega) \,\middle|\, \langle \nu,\, 1_\Omega \rangle = 0,\ \nu|_{\Omega \setminus \{0\}} \geq 0 \right\},$$

while $\ker S_1 = \{ \nu \in \mathcal{P}(\Omega) | \, \langle \nu,\, f(0|\beta) \rangle = 0 \}$, where $f(0|\beta)$ is given in Example 3.17. Since the global maximum of $f(0|\beta)$ occurs at 0, no measure except 0 can assign equal but opposite weight to $\Omega \setminus \{0\}$ and have $\langle \nu,\ f(0|\beta)\rangle = 0$. Hence $T\left(\mathsf{a}_0 \,|\, \mathscr{S}\right) = \{0\}$.

A more restricted question is uniqueness over all finite discrete measures,

$$\mathcal{M}_{\text{finite}} := \left\{ \sum_{k=1}^{d} p_k \mathsf{a}_{\beta_k} \,\middle|\, \beta_k \in \Omega,\ p \in \Delta^{d-1},\ d > 0 \right\},$$

or all finite discrete measures with d or fewer components,

$$\mathcal{M}_d := \left\{ \sum_{k=1}^{d} p_k \mathsf{a}_{\beta_k} \,\middle|\, \beta_k \in \Omega,\ p \in \Delta^{d-1} \right\}.$$

See [39] for theorems on uniqueness of the non-parametric MLE. Below, we consider the uniqueness of the weights when the support points are given and fixed.

If ϕ is strictly convex and we know the optimal value, then determining uniqnuess of an optimal solution over a fixed support set reduces to computing the intersection of three tangent cones. To show this the following notation is required. Suppose $\widetilde{\Omega} = \{\beta_1, \ldots, \beta_d\} \subset \Omega$ is fixed. We consider $\mathcal{B}(\widetilde{\Omega})$ as a subset of $\mathcal{B}(\Omega)$ by identifying measures $\sum_{i=1}^{d} \lambda_i \mathsf{a}_{\beta_i} \in \mathcal{B}(\widetilde{\Omega})$ with the same measure in $\mathcal{B}(\Omega)$.

Corollary 3.27. *Assume that $\psi(z_1, z_2) = \phi(z_1) + \delta_K(z_2)$ and $S = (S_1, S_2)$. The optimal value α^* of* `Primal` *over $\mathcal{B}(\widetilde{\Omega})$ and* `Primal` *over $\mathcal{B}(\Omega)$ agree if and only if $\mathscr{S} \cap \mathcal{B}(\widetilde{\Omega}) \neq \emptyset$. Further, an optimal solution $\overline{\mu} = \sum_{i=1}^{d} \overline{\lambda}_i \mathsf{a}_{\beta_i}$ is unique over $\mathcal{B}(\widetilde{\Omega})$ if*

$$T\left(\overline{\lambda} \,\middle|\, A_1^{-1} \operatorname{lev}_{\phi}^{\leq}(\alpha^*)\right) \cap T\left(\overline{\lambda} \,\middle|\, A_2^{-1} K\right) \cap T\left(\overline{\lambda} \,\middle|\, \Delta^d\right) = \{0\}, \tag{3.36}$$

where $A_1 \in \operatorname{Lin}(\mathbb{R}^{\widehat{n}}, \mathcal{E}_1)$ and $A_2 \in \operatorname{Lin}(\mathbb{R}^{\widehat{n}}, \mathcal{E}_2)$ are defined by

$$A_1\lambda = \sum_{i=1}^{\widehat{n}} \lambda_i F_1(\beta_i) \text{ and } A_2\lambda = \sum_{i=1}^{\widehat{n}} \lambda_i F_2(\beta_i).$$

Moreover, if ϕ is strictly convex and K is polyhedral, then μ_0 is unique over $\mathcal{B}(\widetilde{\Omega})$ only if

(3.36) *holds.*

Proof. It is easy to see the optimal values of `Primal` over $\mathcal{B}(\Omega)$ and $\mathcal{B}(\widetilde{\Omega})$ agree if and only if they share a common solution. To obtain the tangent cone formulas we simply apply Proposition 3.25 to $\mathcal{B}(\widetilde{\Omega})$. □

3.6 Applications Revisted

We apply the theory outlined in the previous sections to each of the applications in Section 3.3. Unless otherwise noted, the notation in Section 3.3 is employed.

3.6.1 Non-parametric Maximum Likelihood

In the case of `NPMLE`, the dual problem (`Dual`) in Theorem 3.11 is given by

$$\begin{array}{lll} \min\limits_{z\in\mathbb{R}^N, w\in\mathcal{E}_2, \gamma\in\mathbb{R}} & -\sum_i d_i \log(z_i/d_i) - D + \gamma & \\ \text{s.t.} & z^T f(\beta) + \langle w,\, F_2(\beta)\rangle_{\mathcal{E}_2} \leq \gamma \quad \forall\, \beta\in\Omega & \qquad (\texttt{NPMLEDual}) \\ & w \in -K^\circ, & \end{array}$$

where $D = \sum_i^N d_i$. In this application it is important to note the dual variable γ can be eliminated.

Proposition 3.28. *The pair* $(\overline{z}, \overline{w})$ *is an optimal solution to*

$$\begin{array}{ll} \min\limits_{z\in\mathbb{R}^N, w\in\mathcal{E}_2} & -\sum_i d_i \log(z_i/d_i) \\ s.t. & z^T f(\beta) + \langle w,\, F_2(\beta)\rangle_{\mathcal{E}_2} \leq D \quad \forall\, \beta\in\Omega \\ & w \in -K^\circ. \end{array}$$

if and only if $(\overline{z}, \overline{w}, D)$ *is a solution to* `NPMLEDual`*, and their optimal values coincide.*

Proof. Replace (z, w, γ) with $(tz', tw', t\gamma')$ for $t \geq 0$. Then NPMLEDual becomes

$$\begin{aligned} \min_{z' \in \mathbb{R}^N, w' \in \mathcal{E}_2, \gamma' \in \mathbb{R}, t \geq 0} \quad & -\textstyle\sum_i d_i \log(z_i'/d_i) - D\log(t) - D + t\gamma' \\ \text{s.t.} \quad & t(z')^T f(\beta) + t\,\langle w',\, F_2(\beta)\rangle_{\mathcal{E}_2} \leq t\gamma' \quad \forall\, \beta \in \Omega \\ & tw' \in -K^\circ \\ & t \geq 0. \end{aligned}$$

The constraints are homogenous in t, and so they may be ignored when optimizing over $t \geq 0$ for fixed (z', w', γ'). Doing so gives $\bar{t} = D/\gamma'$. Substituting this into the objective gives the problem

$$\begin{aligned} \min_{z' \in \mathbb{R}^N, w' \in \mathcal{E}_2} \quad & -\textstyle\sum_i d_i \log\left(\bar{t} z_i'/d_i\right) \\ \text{s.t.} \quad & \bar{t}(z')^T f(\beta) + \bar{t}\,\langle w',\, F_2(\beta)\rangle_{\mathcal{E}_2} \leq D \quad \forall\, \beta \in \Omega \\ & \bar{t}w' \in -K^\circ \end{aligned}$$

Converting $(z', w', \bar{t})$ back to the original variables (z, w) with this optimal value for t completes the proof. □

Additionally, a straight-forward computation shows, by Theorem 3.18, that the first-order conditions (3.23) and (3.24) for NPMLE at $(\mu, (z, w))$ are

$$\begin{aligned} \operatorname{supp}(\mu) \subset \operatorname{argmax}_{\beta \in \Omega} z^T f(\beta) + \langle w,\, F_2(\beta)\rangle_{\mathcal{E}_2} \\ z_i = \frac{d_i}{\int_\Omega f_i(\beta)\,\mu(d\beta)}, \quad i = 1, \ldots, N \\ -w \in N\left(\int_\Omega F_2(\beta)\,\mu(d\beta) \,\middle|\, K\right). \end{aligned}$$

The Finite Mixture CQ can be employed to yield a strong duality result between F-Primal and F-Dual based on Theorems 3.11 and 3.21. To see how this is done let us first consider the case where there are no cone constraints, that is, $F = F_1$ in (3.11). Recall that N is the number of distinct observations.

Proposition 3.29. *Consider problem* NPMLE *where the term* $\delta_K(S_2\mu)$ *is absent.*

1. *If there exists* $\widehat{x} = (\beta_1, \ldots, \beta_N)$ *such that*

$$f(y_i|\beta_i) > \sum_{\substack{1 \le j \le N \\ i \ne j}} f(y_i|\beta_j), \tag{3.37}$$

then $A_F(\widehat{x}) = [f(y_i|\beta_j)]_{ij} \in \mathbb{R}^{N \times N}$ *is invertible.*

2. *The Finite Mixture CQ for* `NPMLE` *is satisfied if there exists* $x = (\beta_0, \widehat{x})$ *with* $\widehat{x} = (\beta_1, \ldots, \beta_N)$ *satisfying* (3.37)*, and* β_0 *such that* $1 - e^T A_F(\widehat{x})^{-1} f(\beta_0) \ne 0$.

Proof. The first statement follows immediately from the Gershgorin Circle theorem [32, Theorem 6.1.1]. For the second, condition (3.31) follows because

$$\begin{bmatrix} f(\beta_0) & A_F(\widehat{x}) \\ 1 & e^T \end{bmatrix}$$

is invertible. This can be seen by doing a block row reduction:

$$\begin{bmatrix} A_F(\widehat{x})^{-1} & 0 \\ -e^T A_F(\widehat{x})^{-1} & 1 \end{bmatrix} \begin{bmatrix} f(\beta_0) & A_F(\widehat{x}) \\ 1 & e^T \end{bmatrix} = \begin{bmatrix} A_F(\widehat{x})^{-1} f(\beta_0) & I \\ 1 - e^T A_F(\widehat{x})^{-1} f(\beta_0) & 0 \end{bmatrix}.$$

The later matrix is invertible if and only if $1 - e^T A_F(\widehat{x})^{-1} f(\beta_0) \ne 0$. Condition (3.30) follows because the matrix $A_F(\widehat{x})$ can have no row of zeros, as it is invertible, so $A_F(x) \left(\frac{1}{1+N}\right) e > 0$ since all entries of $A_F(x)$ are non-negative. ☐

Informally, Proposition 3.29 says $A_F(\widehat{x})$ can be made invertible if the family of densities $f(y|\beta)$ is sufficiently rich. That is, if the family can separate points in the sense that is there exist $\{\beta_j\}_{j=1}^N$ for which the i^{th} observation y_i has high probability under β_i and low probability under β_j with $i \ne j$. If, in addition, there exists a parameter β_0 where all observations have sufficiently low probability, then the Finite Mixture CQ is satisfied.

The interpretation of the Finite Mixture CQ in the presence of cone constraints depends on the particular structure of these constraints. The following example shows the application of the Finite Mixture CQ, either directly or using Proposition 3.22, can be straight-forward in practice.

Example 3.30. If `NPMLE` is given the mean constraint

$$L \leq E[\beta] \leq U,$$

where $L, U \in \Omega \subset \mathbb{R}^p$ and $L < U$ component-wise, then

$$F_2(\beta) = \begin{bmatrix} U - \beta \\ \beta - L \end{bmatrix}$$

and $K_2 = \mathbb{R}_+^{2p}$. In this case, the relative interior condition (3.30) is satisfied if there exists $x = (\beta_1, \ldots, \beta_d)$ such that $A_F(x)$ has no row of zeros and $L < \beta_i < U$ componentwise for $i = 1, \ldots, d$. This condition can also be reached by applying Proposition 3.22, which specialized to this problem requires that if $(z_1, z_2, z_3) \in \mathbb{R}_-^N \times \mathbb{R}_-^p \times \mathbb{R}_-^p$ with

$$f(\beta_j)^T z_1 + (U - \beta_j)^T z_2 + (\beta_j - L)^T z_3 \geq 0 \text{ for all } j = 1, \ldots, \widehat{n},$$

then $(z_1, z_2, z_3) = (0, 0, 0)$. So it follows easily that if for each $j = 1, \ldots, \widehat{n}$ we have

$$\begin{bmatrix} U - \beta_j \\ \beta_j - L \end{bmatrix} > \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

and there exists $i \in \{1, \ldots, N\}$ with $f(y_i|\beta_j) > 0$, then the hypotheses of Proposition 3.22 are satisfied.

Note that with constraints the conditions of Proposition 3.29 no longer applies to show (3.31), so this must be verified directly.

3.6.2 Optimal Design

Following Section 3.3.2, assume $\Phi : \mathbb{S}^n_+ \to \overline{\mathbb{R}}$ is given by $\Phi = f \circ \lambda$, where $f : \mathbb{R}^n \to \overline{\mathbb{R}}$ is permutation invariant (i.e., symmetric) proper convex and $\lambda : \mathbb{S}^n \to \mathbb{R}^n$ is the eigenvalue mapping. In [35, Theorems 2.3 and 3.1] it is shown $(f \circ \lambda)^* = f^* \circ \lambda$ and

$$\partial(f \circ \lambda)(X) = \left\{ Y \in \mathbb{S}^n_+ \,\middle|\, \lambda(Y) \in \partial f(\lambda(X)),\ Y \sim_O X \right\}$$

where $Y \sim_O X$ denotes X and Y can be simultaneously orthogonally diagonalized.

The dual problem to `OptD` is

$$\begin{aligned} &\min_{\gamma\in\mathbb{R},\ V\in\mathbb{S}^n, w\in\mathcal{E}_2} \gamma + f^*(-\lambda(V)) \\ &\langle V,\, I_\theta(\beta)\rangle_{\mathbb{S}^n} + \langle w,\, F_2(\beta)\rangle_{\mathcal{E}_2} \le \gamma \quad \forall \beta \in \Omega \\ &w \in -K^\circ \end{aligned} \qquad (\texttt{OptDDual})$$

and the first-order conditions (3.23) and (3.24) are

$$\begin{aligned} \langle V,\, I_\theta(\beta)\rangle_{\mathbb{S}^n} + \langle w,\, F_2(\beta)\rangle_{\mathcal{E}_2} &\in \partial\delta_{\mathcal{P}}(\mu) \\ -V &\in \partial(f \circ \lambda)\,(I_\theta(\mu)) \\ -w &\in N\left(\int_\Omega F_2(\beta)\,\mu(d\beta) \,\middle|\, K \right) \end{aligned}$$

The following proposition is analogous to Proposition 3.29.

Proposition 3.31. *Consider the* `OptD` *problem with the term $\delta_K(S_2\mu)$ absent.*

1. *If there exist $\{\beta_i\}_{i=1}^{\binom{n+1}{2}}$ such that $\{q(\beta_i)q(\beta_i)^T\}_{i=1}^{\binom{n+1}{2}}$ are linearly independent, then* $\mathrm{Ran}(A_F(\widehat{x})) = \mathbb{S}^n$*, where $\widehat{x} = (\beta_1, \ldots, \beta_{\binom{n+1}{2}})$.*

2. *Let $x = (\beta_0, \widehat{x})$ with $\widehat{x}$ as above so that $A_F(\widehat{x})$ is invertible on $\mathbb{S}^n_+$. If, in addition, $e^T A_F(\widehat{x})^{-1}(q(\beta_0)q(\beta_0)^T) \neq 1$ then the Finite Mixture CQ is satisfied at x.*

Proof. The first statement follows directly from comparing dimensions. As

$$A_F(\widehat{x})\lambda = \sum_{i=1}^{\binom{n+1}{2}} \lambda_i q(\beta_i)q(\beta_i)^T,$$

if $\{q(\beta_i)q(\beta_i)^T\}_{i=1}^{\binom{n+1}{2}}$ are linearly independent, then $\mathrm{Ran}(A_F(\widehat{x})) \subset \mathbb{S}^n$ and

$$\dim \mathrm{Ran}(A_F(\widehat{x})) = \binom{n+1}{2} = \dim(\mathbb{S}^n).$$

We show the second statement in two steps. Condition (3.31) of the Finite Mixture CQ follows from a block reduction argument similar to Proposition 3.29. To show the relative interior condition, let $\widehat{\lambda} = A_F(\widehat{x})^{-1} I_{n\times n}$. If $\widehat{\lambda} \leq 0$, then $A_F(\widehat{x})\widehat{\lambda}$ is negative semi-definite, a contradiction. Thus $\widehat{\lambda} \not\leq 0$, so we may define $\widehat{\lambda} = \widehat{\lambda}_1 - \widehat{\lambda}_2$ where $\widehat{\lambda}_1 > 0$ and $\widehat{\lambda}_2 \geq 0$ and $A_i = A_F(\widehat{x})\widehat{\lambda}_i$. Since $q(\beta_j)q(\beta_j)^T$ for $j = 1, \ldots, \binom{n+1}{2}$ is a rank-1 positive semidefinite matrix for all β with $q(\beta) \neq 0$, we have A_i is positive semi-definite for $i = 1, 2$. Therefore A_1 is positive definite, as $I_{n\times n}$ is positive definite and $I_{n\times n} = A_1 - A_2$. Since A_1 is positive definite, so is $(1-\epsilon)A_1 + \epsilon A_F(x)\left(\frac{1}{\binom{n+1}{2}+1}e\right)$ for small enough $\epsilon > 0$. If,

$$\widetilde{\lambda} = (1-\epsilon)\begin{bmatrix}\widehat{\lambda}_1 \\ 0\end{bmatrix} + \epsilon\left(\frac{1}{\binom{n+1}{2}+1}e\right),$$

then

$$A_F(x)\widetilde{\lambda} = (1-\epsilon)A_1 + \epsilon A_F(x)\left(\frac{1}{\binom{n+1}{2}+1}e\right),$$

which shows condition (3.30) is satisfied. □

Intuitively, the above theorem is similar to Proposition 3.29, but using $\{\beta_i\}$ such that $\{q(\beta_i)q(\beta_i)^T\}$ approximates $\{(e_i+e_j)(e_i+e_j)^T\}$.

3.6.3 Distributionally Robust Stochastic Programming

As in Section 3.3.3, set

$$F_1(\beta) = \int g(\xi, u) f(u|\beta) du.$$

Then the dual to `DRSP` is

$$\begin{aligned} \min_{\gamma \in \mathbb{R}, w \in \mathcal{E}_2} \quad & \gamma \\ \text{s.t.} \quad & -F_1(\beta) + \langle w,\, F_2(\beta)\rangle_{\mathcal{E}_2} \leq \gamma \text{ for all } \beta \in \Omega, \\ & -w \in K^\circ \end{aligned} \tag{DRPSDual}$$

and the first-order conditions (3.23) and (3.24) for `DRSP` are

$$\begin{aligned} -F_1(\beta) + \langle w,\, F_2(\beta)\rangle_{\mathcal{E}_2} &\in \partial \delta_{\mathcal{P}(\Omega)}(\mu), \\ -w &\in N\left(\int_\Omega F_2(\beta)\, \mu(d\beta) \,\middle|\, K \right). \end{aligned}$$

The structure of the objective in `DRSP` is much simpler because the outer convex function is linear. However, the constraints are no simpler, so satisfying the Finite Mixture is not trivial. A similar condition to Propositions 3.29 and 3.31 can formulated, as below.

Proposition 3.32. *Suppose $\widehat{x} \in \Omega^k$, where $k = \dim \mathcal{E}_2$, and $\beta_1, \beta_2 \in \Omega$. Let*

$$\begin{aligned} B &= \begin{bmatrix} F_2(\beta_1) & F_2(\beta_2) \end{bmatrix} \\ C &= \begin{bmatrix} A_{F_1}(\widehat{x}) \\ e^T \end{bmatrix} \\ D &= \begin{bmatrix} F_1(\beta_1) & F_1(\beta_2) \\ 1 & 1 \end{bmatrix}. \end{aligned}$$

Then condition (3.31) of the Finite Mixture CQ is satisfied if $A_{F_2}(\widehat{x})$ and $D - CA_{F_2}(\widehat{x})^{-1}B$ are invertible. The Finite Mixture CQ is satisfied if, in addition, there exists $\lambda \in \operatorname{ri} \Delta^{k+1}$

such that $A_{F_2}(\widehat{x}, \beta_1, \beta_2)\lambda \in \operatorname{ri} K$.

Proof. Observe that by exchanging the first and second block rows of

$$\begin{bmatrix} A_F(\widehat{x}, \beta_1, \beta_2) \\ e^T \end{bmatrix} := \begin{bmatrix} A_{F_1}(\widehat{x}) & F_1(\beta_1) & F_1(\beta_2) \\ A_{F_2}(\widehat{x}) & F_2(\beta_1) & F_2(\beta_2) \\ e^T & 1 & 1 \end{bmatrix} \tag{3.38}$$

we obtain

$$\begin{bmatrix} A_{F_2}(\widehat{x}) & B \\ C & D \end{bmatrix}.$$

Then

$$\begin{bmatrix} I & 0 \\ -CA_{F_2}(\widehat{x})^{-1} & I \end{bmatrix} \begin{bmatrix} A_{F_2}(\widehat{x}) & B \\ C & D \end{bmatrix} = \begin{bmatrix} A_{F_2}(\widehat{x}) & B \\ 0 & D - CA_{F_2}(\widehat{x})^{-1}B \end{bmatrix}.$$

Hence the hypotheses imply the matrix in (3.38) is invertible, and the result follows from the definition of the Finite Mixture CQ. □

As observed in subsection 3.6.1, this result can be refined by specifying F_2 and K. (See Example 3.30.)

3.6.4 Maximum Entropy

Comparing the structure of `DRSP` in (3.16) and `MER` in (3.17), the convex function ψ and the linear map S defining `DRSP` and `MER` take the same form. Therefore the results of the previous section can be applied to the `MER` problem.

3.7 Conclusion

This chapter has introduced and discussed a framework for convex optimization over $\mathcal{P}(\Omega)$ and a finite dimensional convex-composite embedding of the problem amenable to computational algorithms. We have also developed a duality theory for these optimization problems, given first-order conditions, investigated the solution set, and discussed several specific applications of the framework.

Chapter 4

INTERIOR POINT RELAXATIONS OF BENDERS' DECOMPOSITION

4.1 Introduction

In applications, the function ψ in the problem

$$\min_{\mu \in \mathcal{B}(\Omega)} \psi\left(\int_\Omega F(\beta)\,\mu(d\beta)\right) + \delta_{\mathcal{P}(\Omega)}(\mu) \tag{Primal}$$

often takes the form

$$\psi(z_1, z_2) = \phi(z_1) + \delta_K(z_2),$$

where $\phi : \mathbb{R}^n \to \overline{\mathbb{R}}$ is proper convex lsc and $K \subset \mathbb{R}^k$ is a closed convex cone. In this chapter we consider solving such problems in the special case when $K = \{0\}^{n_E} \times \mathbb{R}_-^{n_I} \times \mathbb{S}_-^{n_M}$ and ϕ satisfies regularity conditions described below.

For ψ of this form, $F(\beta)$ is

$$F(\beta) = \begin{bmatrix} f_\phi(\beta) \\ f_E(\beta) \\ f_I(\beta) \\ f_M(\beta) \end{bmatrix},$$

where

$$f_\phi : \Omega \to \mathbb{R}^n \tag{4.1}$$

$$f_E : \Omega \to \mathbb{R}^{n_E} \tag{4.2}$$

$$f_I : \Omega \to \mathbb{R}^{n_I} \tag{4.3}$$

$$f_M : \Omega \to \mathbb{S}^{n_M} \tag{4.4}$$

are smooth. Recall for a Euclidian space $\mathcal{E}$, an arbitrary smooth function $g : \Omega \to \mathcal{E}$, and $x = (\beta_1, \ldots, \beta_d) \in \Omega^d$, the linear operator $A_g(x) \in \mathrm{Lin}(\mathbb{R}^d, \mathcal{E})$ is defined by

$$A_g(x)\lambda = \sum_{i=1}^{d} \lambda_i g(\beta_i) \tag{4.5}$$

for $\lambda \in \mathbb{R}^d$. Furthermore, a convenient abuse of notation is to treat $A_g(x)$ as a matrix, for instance by writing

$$A_g(x) = \begin{bmatrix} g(\beta_1) & \cdots & g(\beta_d) \end{bmatrix}.$$

Then for $x \in \Omega^d$,

$$A_F(x)\lambda = \begin{bmatrix} A_{f_\phi}(x)\lambda \\ A_{f_E}(x)\lambda \\ A_{f_I}(x)\lambda \\ A_{f_M}(x)\lambda \end{bmatrix} = \begin{bmatrix} \sum_{i=1}^d \lambda_i f_\phi(\beta_i) \\ \sum_{i=1}^d \lambda_i f_E(\beta_i) \\ \sum_{i=1}^d \lambda_i f_I(\beta_i) \\ \sum_{i=1}^d \lambda_i f_M(\beta_i) \end{bmatrix}. \tag{4.6}$$

To simplify notation, write $A_\phi(x)$ instead of $A_{f_\phi}(x)$, and similarly $A_E(x)$ for A_{f_E}, $A_I(x)$ for

A_{f_I}, and $A_{\scriptscriptstyle M}(x)$ for A_{f_M}, so that

$$A_F(x) = \begin{bmatrix} A_{f_\phi}(x) \\ A_{f_E}(x) \\ A_{f_I}(x) \\ A_{f_M}(x) \end{bmatrix} = \begin{bmatrix} A_\phi(x) \\ A_{\scriptscriptstyle E}(x) \\ A_{\scriptscriptstyle I}(x) \\ A_{\scriptscriptstyle M}(x) \end{bmatrix}.$$

The argument x is sometimes omitted in contexts where it is fixed to a specific value.

The object $A_{\scriptscriptstyle M}(x)$ is the linear operator whose action on λ is given by $A_{\scriptscriptstyle M}(x)\lambda = \sum_{i=1}^d \lambda_i f_{\scriptscriptstyle M}(\beta_i) \in \mathbb{S}^{n_M}$. Furthermore, $A_{\scriptscriptstyle M}(x)^T$ is taken to mean the transpose of $A_{\scriptscriptstyle M}(x)$ as a linear operator between inner product spaces rather than the transpose of a matrix. Thus $A_{\scriptscriptstyle M}(x)^T$ is a mapping $A_{\scriptscriptstyle M}(x)^T : \mathbb{S}^{n_M} \to \mathbb{R}^d$ given by

$$\left\langle e_i,\, A_{\scriptscriptstyle M}(x)^T X \right\rangle = \langle f_{\scriptscriptstyle M}(\beta_i),\, X \rangle$$

for $X \in \mathbb{S}^{n_M}$.

For later user, the following spaces are also introduced to simplify notation:

$$\mathbb{E} = \mathbb{R}^d \times \mathbb{R}^{n_I} \times \mathbb{S}^{n_M} \tag{4.7}$$

$$\mathbb{F} = \mathbb{R}^n \times \{0\}^{n_E} \times \mathbb{R}_-^{n_I} \times \mathbb{S}_-^{n_M} \tag{4.8}$$

$$\mathbb{K} = \{0\}^{n_E} \times \mathbb{R}_-^{n_I} \times \mathbb{S}_-^{n_M}. \tag{4.9}$$

Then $(\lambda, s, S) \in \mathbb{E}$ denotes $\lambda \in \mathbb{R}^d$, $s \in \mathbb{R}^{n_I}$, and $S \in \mathbb{S}^{n_M}$, and similarly for $(z, z_{\scriptscriptstyle E}, z_{\scriptscriptstyle I}, z_{\scriptscriptstyle M}) \in \mathbb{F}$ or $(z_{\scriptscriptstyle E}, z_{\scriptscriptstyle I}, z_{\scriptscriptstyle M}) \in \mathbb{K}$.

For the entirety of this chapter we assume ϕ satisfies the following assumptions:

1. The function $\phi : \mathbb{R}^n \to \overline{\mathbb{R}}$ is proper, convex, lsc

2. That $\operatorname{dom} \phi \cap A_\phi(x)\Delta^{d-1} \neq \emptyset$ for all $x \in \Omega^d$

3. The domain $\operatorname{dom}\phi$ has non-empty interior, i.e. $\operatorname{int}\operatorname{dom}\phi \neq \emptyset$

4. The function ϕ is differentiable on the interior of its domain.

4.2 Benders Decomposition

The special case of `Primal` analyzed in this chapter is

$$\begin{aligned} \min_{\mu} \quad & \phi\left(\int_\Omega f_\phi(\beta)\mu(d\beta)\right) \\ \text{s.t.} \quad & \mu \in \mathcal{P}(\Omega) \\ & \int_\Omega f_E(\beta)\mu(d\beta) \in \{0\}^{n_E} \\ & \int_\Omega f_I(\beta)\mu(d\beta) \in \mathbb{R}_-^{n_I} \\ & \int_\Omega f_M(\beta)\mu(d\beta) \in \mathbb{S}_-^{n_M}. \end{aligned} \tag{4.10}$$

Recall from Theorem 3.13 that the above infinite-dimensional convex problem is, for sufficiently large d, equivalent to the following finite-dimensional non-convex instance of `F-Primal`,

$$\begin{aligned} \min_{x,\lambda} \quad & \phi(A_\phi(x)\lambda) \\ \text{s.t.} \quad & (x,\lambda) \in \Omega^d \times \Delta^{d-1} \\ & A_E(x)\lambda \in \{0\}^{n_E} \\ & A_I(x)\lambda \in \mathbb{R}_-^{n_I} \\ & A_M(x)\lambda \in \mathbb{S}_-^{n_M}. \end{aligned} \tag{4.11}$$

The above optimization problem is non-convex. However, for fixed $x \in \Omega$ the problem is a nonlinear convex optimization problem over the finite-dimensional probability simplex Δ^{d-1}. This is an example of a more general class of optimization problems which possess *complicating variables* that, when held constant, yield optimization problems that are cheap or easy to solve. This structure was studied in [28], where it was suggested that the overall optimization problem be partitioned out so that the complicating variables were optimized separately. Such a partition of the variables was called a *Generalized Benders Decomposition*, in reference to the Benders Decomposition in linear programming. To apply the Generalized Benders Decomposition to `F-Primal`, define $G : \Omega^d \to \overline{\mathbb{R}}$ by

$$\begin{aligned} G(x) := \min_{(\lambda,s,S)\in\mathbb{E}} \quad & \phi(A_\phi(x)\lambda) \\ \text{s.t.} \quad & e^T\lambda = 1 \\ & A_E(x)\lambda = 0 \\ & A_I(x)\lambda = s \\ & A_M(x)\lambda = S \\ & \lambda \geq 0 \\ & s \leq 0 \\ & S \preceq 0, \end{aligned} \qquad (\texttt{BendersPrimal}(\mathrm{x}))$$

where $G(x)$ is the optimal value while $\texttt{BendersPrimal}(x)$ is the optimization problem. Then finding the optimal solution to `F-Primal` reduces to optimizing G on Ω^d. Unfortunately, as shown in the next proposition, the function G is not well-behaved. In particular, if $G(x) < +\infty$ for all $x \in \Omega^d$, then the constraints are trivially satisfied.

Proposition 4.1. *If $G(x)$ is finite for all $x \in \Omega^d$, then*

$$\begin{aligned} G(x) = \min_{(\lambda,s,S)\in\mathbb{E}} \quad & \phi(A_\phi(x)\lambda) \\ s.t. \quad & e^T\lambda = 1 \\ & \lambda \geq 0 \end{aligned} \qquad (\texttt{BendersPrimal}(\mathrm{x}))$$

Proof. Fix $\beta \in \Omega$ and consider $\overline{x} = (\beta, \ldots, \beta)$. If $G(\overline{x})$ is finite, then $f_E(\beta) = 0$, $f_I(\beta) \leq 0$, and $f_M(\beta) \preceq 0$. Since β was arbitrary, this holds for all $\beta \in \Omega$. Then $A_E(x)\lambda = 0$ for all $x \in \Omega^d$ and $\lambda \in \Delta^{d-1}$, and similarly for $A_I(x)\lambda$ and $A_M(x)\lambda$. □

In addition, there is no guarantee G is convex or smooth, as the following example shows.

Example 4.2. Let

$$\begin{aligned} \Omega &= [-1, 1] \subset \mathbb{R}, \\ d &= 2, \\ f_\phi(\beta) &= 1, \\ \phi(z) &= \frac{1}{2} z^2, \end{aligned}$$

with no equality, inequality, or semidefinite constraints. Then

$$G(\beta_1, \beta_2) = \min_{\lambda \in \Delta^{d-1}} \frac{1}{2} \left(\lambda_1 \beta_1 + \lambda_2 \beta_2\right)^2$$

It is simple to show

$$G(\beta_1, \beta_2) = \begin{cases} 0 & \beta_1 \leq 0 \leq \beta_2 \\ 0 & \beta_2 \leq 0 \leq \beta_1 \\ \min(\frac{1}{2}\beta_1^2, \frac{1}{2}\beta_2^2) & \text{else} \end{cases}$$

In particular, G is non-convex and is not second-order differentiable on the interior of its domain.

Proposition 4.1 and Example 4.2 show G is a poorly behaved function. There is no guarantee many standard constrained optimization algorithms would find good solutions when applied to G. The next section explores a relaxation of `BendersPrimal`(x) that yields an optimal value function with improved behavior.

4.3 *Interior Point Relaxation*

As shown in the previous section, the optimal value function $G(x)$ for the optimization problem $\texttt{BendersPrimal}(x)$ is non-convex, has limited smoothness properties, and is non-finite-valued. This section introduces an relaxation of $\texttt{BendersPrimal}(x)$ whose corresponding optimal value function is smooth and finite-valued. This relaxation will be used to introduce a practical algorithm for optimizing $G(x)$ at the end of the chapter.

For $t \in \mathbb{R}$, $t > 0$, consider the relaxation of $\texttt{BendersPrimal}(x)$

$$
\begin{aligned}
G(x,t) := \min_{\substack{(\lambda,s,S)\in\mathbb{E} \\ (z,z_I,z_E,z_M)\in\mathbb{F}}} \quad & \phi(z) + t\,\mathrm{lb}(\lambda) + t\,\mathrm{lb}(-s) - t\,\mathrm{logdet}(-S) \qquad (\texttt{BendersPrimal}(\mathrm{x,t})) \\
& + \frac{1}{2t}\left\|z_E\right\|^2 + \frac{1}{2t}\left\|z_I - s\right\|^2 + \frac{1}{2t}\left\|z_M - S\right\|^2 \\
\text{s.t.} \quad & e^T\lambda = 1 \\
& z = A_\phi(x)\lambda \\
& z_E = A_E(x)\lambda \\
& z_I = A_I(x)\lambda \\
& z_M = A_M(x)\lambda,
\end{aligned}
$$

where the function $\mathrm{lb} : \mathbb{R}^M \to \overline{\mathbb{R}}$ is, for any M, given by

$$
\mathrm{lb}(x) = \begin{cases} -\sum_{i=1}^M \log \beta_i & \beta_i > 0 \quad \text{for each } i, \\ \infty & \text{else} \end{cases}, \tag{4.12}
$$

and, as before, $G(x,t)$ refers to the optimal value while $\texttt{BendersPrimal}(x,t)$ is the optimization problem. As shown later, for fixed $t > 0$ the constraint relaxations imply $G(x,t)$ is a smooth function of x on Ω. Thus optimization of $G(x,t)$ over x for fixed $t > 0$ is possible using standard optimization algorithms. Understanding the behavior of $\min_x G(x,t)$ and its optimal solutions as $t \searrow 0$ is the major focus of this chapter.

Both $\texttt{BendersPrimal}(x)$ and $\texttt{BendersPrimal}(x,t)$ are convex programs for fixed x and t. So it's natural to consider their dual programs.

Theorem 4.3. *1. For fixed $x \in \Omega^d$, the dual program to* $\texttt{BendersPrimal}(x)$ *is*

$$D(x) := \min_{\substack{(w,w_E,w_I,w_M)\in\mathbb{F}\\ \gamma\in\mathbb{R}}} \quad \phi^*(-w) + \gamma \qquad (\texttt{BendersDual(x)})$$

$$\text{s.t.} \quad A_\phi(x)^T w + A_E(x)^T w_E + A_I(x)^T w_I + A_M(x)^T w_M \leq \gamma e$$

$$w_I \leq 0$$

$$w_M \preceq 0$$

where $D(x)$ is the optimal value and $\texttt{BendersDual}(x)$ *is the dual optimization problem taken as a whole.*

2. For fixed $(x,t) \in \Omega^d \times \mathbb{R}_+$, the dual program to $\texttt{BendersPrimal}(x,t)$ *is*

$$\begin{aligned} D(x,t) := \min_{\substack{(w,w_E,w_I,w_M)\in\mathbb{F}\\ \gamma\in\mathbb{R}}} \quad & \gamma + \phi^*(-w) \qquad (\texttt{BendersDual(x,t)}) \\ & + t\,\mathrm{lb}\left(\gamma e - A_\phi(x)^T w - A_I(x)^T w_I - A_E(x)^T w_E - A_M(x)^T w_M\right) \\ & + t\,\mathrm{lb}(-w_I) - t\,\mathrm{logdet}(-w_M) \\ & + \frac{t}{2}\left(\|w_E\|^2 + \|w_I\|^2 + \|w_M\|^2\right) \\ & + t\left(\log(t) - 1\right)\left(d + n_I + n_M\right), \end{aligned}$$

where $D(x,t)$ is the optimal value and $\texttt{BendersDual}(x,t)$ *is the optimization problem.*

Proof. We use the Lagrangian duality in Theorem 2.22 to derive the dual problems.

(1) Eliminating the variables s and S, the Lagrangian $L : \mathbb{R}^d \times \mathbb{F} \times \mathbb{F} \times \mathbb{R} \times \mathbb{R}^d \to \overline{\mathbb{R}}$ is

$$
\begin{aligned}
L(\lambda, z, &z_E, z_I, z_M, w, w_E, w_I, w_M, \gamma, \sigma) \\
&:= \phi(z) + \delta(z_E, z_I, z_M | \mathbb{K}) \\
&+ \langle w,\, z - A_\phi(x)\lambda \rangle + \langle w_E,\, z_E - A_E(x)\lambda \rangle \\
&+ \langle w_I,\, z_I - A_I(x)\lambda \rangle + \langle w_M,\, z_M - A_M(x)\lambda \rangle \\
&+ \gamma(e^T\lambda - 1) + \sigma^T\lambda - \delta(\sigma | \mathbb{R}^d_-) \\
&= -\left(\langle z,\, -w \rangle - \phi(z)\right) \\
&- \left(\langle z_E,\, -w_E \rangle - \delta(z_E | \{0\}^{n_E})\right) \\
&- \left(\langle z_I,\, -w_I \rangle - \delta(z_I | \mathbb{R}^{n_I}_-)\right) \\
&- \left(\langle z_M,\, -w_M \rangle - \delta(z_M | \mathbb{S}^{n_M}_-)\right) \\
&+ \left\langle \lambda,\, \gamma e + \sigma - A_\phi(x)^T w - A_E(x)^T w_E - A_I(x)^T w_I - A_M(x)^T w_M \right\rangle \\
&- \gamma - \delta(\sigma | \mathbb{R}^d_-)
\end{aligned}
$$

where the later expression shows the Lagrangian decouples over the primal variables $(\lambda, z, z_E, z_I\ z_M)$, simplifying minimization over them. Maximizing $L(\lambda, z, z_E, z_I, z_M, w, w_E, w_I, w_M, \gamma,)$ over $(w, w_E, w_I, w_M, \gamma, \sigma)$ recovers the primal problem, while minimizing jointly over $(\lambda, z, z_E, z_I\ z_M)$ gives the dual program in the theorem statement.

(2) The dual program $\texttt{BendersDual}(x, t)$ can be derived directly from the Lagrangian

$$L : \mathbb{E} \times \mathbb{F} \times \mathbb{F} \times \mathbb{R} \to \overline{\mathbb{R}},$$

given by

$$
\begin{aligned}
L(\lambda, s, &S, z, z_{I}, z_{E}, z_{M}, w, w_{E}, w_{I}, w_{M}, \gamma) \\
&:= \phi(z) + t\,\mathrm{lb}(\lambda) + t\,\mathrm{lb}(-s) - t\,\mathrm{logdet}(-S) \\
&+ \frac{1}{2t}\left\|z_{E}\right\|^2 + \frac{1}{2t}\left\|z_{I} - s\right\|^2 + \frac{1}{2t}\left\|z_{M} - S\right\|^2 \\
&+ \langle w,\, z - A_{\phi}(x)\lambda\rangle + \langle w_{E},\, z_{E} - A_{E}(x)\lambda\rangle \\
&+ \langle w_{I},\, z_{I} - A_{I}(x)\lambda\rangle + \langle w_{M},\, z_{M} - A_{M}(x)\lambda\rangle \\
&+ \gamma(e^T\lambda - 1) \\
&= -\left(\langle z,\, -w\rangle - \phi(z)\right) \\
&- \left(\langle z_{E},\, -w_{E}\rangle - \frac{1}{2t}\left\|z_{E}\right\|^2\right) \\
&- \left(\langle z_{I},\, -w_{I}\rangle - \frac{1}{2t}\left\|z_{I} - s\right\|^2 - t\,\mathrm{lb}(-s)\right) \\
&- \left(\langle z_{M},\, -w_{M}\rangle - \frac{1}{2t}\left\|z_{M} - S\right\|^2 + \mathrm{logdet}(-S)\right) \\
&+ \left(\left\langle \lambda,\, \gamma e - A_{\phi}(x)^T w - A_{E}(x)^T w_{E} - A_{I}(x)^T w_{I} - A_{M}(x)^T w_{M}\right\rangle\right) \\
&- \gamma,
\end{aligned}
$$

where the second expression for L shows how the Lagrangian decouples over the primal variables $(\lambda, s, S, z, z_{E}, z_{I}, z_{M})$. □

We also define $G(x, 0) = G(x)$ and similarly for $D(x, t)$, $\texttt{BendersPrimal}(x, t)$, and $\texttt{BendersDual}(x, t)$. For notational simplicity in later arguments, we let $p : \mathbb{E} \times \Omega^d \times \mathbb{R}_+ \to \overline{\mathbb{R}}$ be the objective of $\texttt{BendersPrimal}(x, t)$,

$$
\begin{aligned}
p(\lambda, s, S; x, t) := {} & \phi(A_{\phi}(x)\lambda) + t\,\mathrm{lb}(\lambda) + t\,\mathrm{lb}(-s) - t\,\mathrm{logdet}(-S) \\
& + \frac{1}{2t}\left\|A_{E}(x)\lambda\right\|^2 + \frac{1}{2t}\left\|A_{I}(x)\lambda - s\right\|^2 + \frac{1}{2t}\left\|A_{M}(x)\lambda - S\right\|^2
\end{aligned} \tag{4.13}
$$

and let $d : \mathbb{F} \times \mathbb{R} \times \Omega^d \times \mathbb{R}_+ \to \overline{\mathbb{R}}$ be the objective for $\texttt{BendersDual}(x,t)$,

$$\begin{aligned} d(w, w_E, w_I, w_M, \gamma; x, t) := {} & \gamma + \phi^*(-w) \\ & + t\,\mathrm{lb}\left(\gamma e - A_\phi(x)^T w - A_I(x)^T w_I - A_E(x)^T w_E - A_M(x)^T w_M\right) \\ & + t\,\mathrm{lb}(-w_I) - t\,\mathrm{logdet}(-w_M) + \frac{t}{2}\left(\|w_E\|^2 + \|w_I\|^2 + \|w_M\|^2\right) \\ & + t\,(\log(t) - 1)\,(d + n_I + n_M)\,. \end{aligned} \tag{4.14}$$

Then, for $t > 0$,

$$G(x,t) = \min_{(\lambda,s,S)\in\mathbb{E}} p(\lambda, s, S; x, t)$$

and

$$D(x,t) = \min_{\substack{(w,w_E,w_I,w_M)\in\mathbb{F} \\ \gamma\in\mathbb{R}}} d(w, w_E, w_I, w_M, \gamma; x, t).$$

The next proposition states the first-order conditions for primal-dual solutions of $\texttt{BendersPrimal}(x,t)$ and $\texttt{BendersDual}(x,t)$. These conditions give useful relationships between the variables in a primal-dual solution. In particular, the first-order conditions are continuous as a function of the primal and dual variables.

Proposition 4.4. *Suppose $x \in \Omega^d$, $t \geq 0$, and there exists (λ, s, S) strictly feasible for* $\texttt{BendersPrimal}(x,t)$*. Then*

$$(\lambda, s, S, z, z_I, z_E, z_M, w, w_E, w_I, w_M, \gamma) \in \mathbb{E} \times \mathbb{F} \times \mathbb{F} \times \mathbb{R}$$

is a primal-dual solution pair to $\texttt{BendersPrimal}(x,t)$ *and* $\texttt{BendersDual}(x,t)$ *if and only if $\lambda \geq 0$, $s \leq 0$, $S \preceq 0$ and there exists $y \in \mathbb{R}^d$ with*

$$F_{(x,t)}(y, \lambda, s, S, z, z_E, z_I, z_M, w, w_E, w_I, w_M, \gamma) = 0 \tag{4.15}$$

where

$$F_{(x,t)}(y, \lambda, s, S, z, z_I, z_E, z_M, w, w_E, w_I, w_M, \gamma)$$

$$:= \begin{bmatrix} \lambda \circ y - te \\ y + {A_\phi}^T w + A_E^T w_E + A_I^T w_I + A_M^T w_M - \gamma e \\ w + \nabla\phi(z) \\ z_E + tw_E \\ z_I - s + tw_I \\ z_M - S + tw_M \\ s \circ w_I - te \\ Sw_M - tI \\ z - A_\phi\lambda \\ z_E - A_E\lambda \\ z_I - A_I\lambda \\ z_M - A_M\lambda \\ e^T\lambda - 1 \end{bmatrix}. \tag{4.16}$$

Proof. The stated conditions are the standard first order conditions for a convex problem, as stated in Theorem 2.23.

In more detail, the first derivatives of the Lagrangian

$$L : \mathbb{E} \times \mathbb{F} \times \mathbb{F} \times \mathbb{R} \to \overline{\mathbb{R}},$$

given by

$$
\begin{aligned}
L(\lambda, s, &S, z, z_I, z_E, z_M, w, w_E, w_I, w_M, \gamma) \\
&= \phi(z) + t\,\mathrm{lb}(\lambda) + t\,\mathrm{lb}(-s) - t\,\mathrm{logdet}(-S) \\
&+ \frac{1}{2t}\left\|z_E\right\|^2 + \frac{1}{2t}\left\|z_I - s\right\|^2 + \frac{1}{2t}\left\|z_M - S\right\|^2 \\
&+ \langle w,\, z - A_\phi(x)\lambda\rangle + \langle w_E,\, z_E - A_E(x)\lambda\rangle \\
&+ \langle w_I,\, z_I - A_I(x)\lambda\rangle + \langle w_M,\, z_M - A_M(x)\lambda\rangle \\
&+ \gamma(e^T\lambda - 1)
\end{aligned}
$$

are given below. The notation $\frac{\partial}{\partial \bullet}$ was used instead of $d_\bullet$ is used for readability due to the primal and dual variables with subscripts. (See Subsection 2.5.1.)

$$
\begin{aligned}
\frac{\partial L}{\partial \lambda} &= -t\lambda^{-1} - A_\phi(x)^T w - A_{\scriptscriptstyle E}(x)^T w_{\scriptscriptstyle E} - A_{\scriptscriptstyle I}(x)^T w_{\scriptscriptstyle I} - A_{\scriptscriptstyle M}(x)^T w_{\scriptscriptstyle M} + \gamma e \\
\frac{\partial L}{\partial s} &= -ts^{-1} + \frac{1}{t}(s - z_{\scriptscriptstyle I}) \\
\frac{\partial L}{\partial S} &= -tS^{-1} + \frac{1}{t}(S - z_{\scriptscriptstyle M}) \\
\frac{\partial L}{\partial z} &= \nabla\phi(z) + w \\
\frac{\partial L}{\partial z_{\scriptscriptstyle E}} &= \frac{1}{t} z_{\scriptscriptstyle E} + w_{\scriptscriptstyle E} \\
\frac{\partial L}{\partial z_{\scriptscriptstyle I}} &= \frac{1}{t}\left(z_{\scriptscriptstyle I} - s\right) + w_{\scriptscriptstyle I} \\
\frac{\partial L}{\partial z_{\scriptscriptstyle M}} &= \frac{1}{t}\left(z_{\scriptscriptstyle M} - S\right) + w_{\scriptscriptstyle M} \\
\frac{\partial L}{\partial w} &= z - A_\phi(x)\lambda \\
\frac{\partial L}{\partial w_{\scriptscriptstyle E}} &= z_{\scriptscriptstyle E} - A_{\scriptscriptstyle E}(x)\lambda \\
\frac{\partial L}{\partial w_{\scriptscriptstyle I}} &= z_{\scriptscriptstyle I} - A_{\scriptscriptstyle I}(x)\lambda \\
\frac{\partial L}{\partial w_{\scriptscriptstyle M}} &= z_{\scriptscriptstyle M} - A_{\scriptscriptstyle M}(x)\lambda \\
\frac{\partial L}{\partial \gamma} &= e^T\lambda - 1,
\end{aligned}
$$

where λ^{-1} and s^{-1} mean the vectors of component-wise inverses. The conditions in (4.16) are obtained by introducing a new variable $y = t\lambda^{-1}$ and rearranging

$$
-ts^{-1} + \frac{1}{t}(s - z_{\scriptscriptstyle I}) = 0 \quad \text{and} \quad -tS^{-1} + \frac{1}{t}(S - z_{\scriptscriptstyle M}) = 0
$$

into

$$
s \circ \left(\frac{1}{t}(s - z_{\scriptscriptstyle I})\right) = te \quad \text{and} \quad S\left(\frac{1}{t}(S - z_{\scriptscriptstyle M})\right) = tI,
$$

respectively. □

An immediate consequence of the first-order conditions in Theorem 4.4 is that to every dual solution there corresponds a unique primal solution, when $t > 0$.

Corollary 4.5. *If $t > 0$, then there is a unique primal solution λ associated with each dual solution $(w, w_E, w_I, w_M, \gamma)$.*

Proof. The result follows immediately from the optimality conditions

$$\lambda \circ y = te$$
$$y + {A_\phi}^T w + A_E^T w_E + A_I^T w_I + A_M^T w_M = \gamma e$$

□

The following sets and regularity conditions will be important for analyzing the behavior of solutions to BendersPrimal(x) and BendersDual(x). In particular, they are necessary for understanding non-singleton solution spaces as well as the behavior of solutions to BendersPrimal(x,t) and BendersDual(x,t) as $t \searrow 0$.

Definition 4.6. *1. For each $x \in \Omega^d$, $\lambda \in \Delta^{d-1}$, and $\beta \in \Omega$, define*

$$\text{Distinct}(x) := \{\, \beta \in \Omega \,|\, \beta = x_i \text{ for some } i \in \{1, \ldots, d\}\,\} \tag{4.17}$$
$$\text{Sol}_\lambda(x) := \{\, \lambda \,|\, \lambda \text{ optimal solution to } \texttt{BendersPrimal}(x)\,\} \tag{4.18}$$
$$\mathcal{I}(x,\beta) := \{\, i \,|\, x_i = \beta,\ 1 \le i \le d\,\} \tag{4.19}$$
$$\lambda_{\mathcal{I}(x,\beta)} := \sum_{i \in \mathcal{I}(x,\beta)} \lambda_i. \tag{4.20}$$

2. The point x is objective non-degenerate *if the vectors*

$$\left\{\, f(\beta) \,\middle|\, \beta \in \text{Distinct}(x),\ \exists \lambda \in \text{Sol}_\lambda(x) \text{ with } \lambda_{\mathcal{I}(x,\beta)} > 0 \,\right\}$$

are linearly independent.

3. *The pair* (λ, x) *is* constraint non-degenerate *if the set*

$$\left\{ (f_E(\beta), f_I(\beta), f_M(\beta)) \,\middle|\, \beta \in \mathrm{Distinct}(x),\ \lambda_{\mathcal{I}(x,\beta)} > 0 \right\} \subset \mathbb{R}^{n_E} \times \mathbb{R}^{n_I} \times \mathbb{S}^{n_M}$$

spans $\mathbb{R}^{n_E} \times \mathbb{R}^{n_I} \times \mathbb{S}^{n_M}$.

4. *The point* x *is* constraint non-degenerate *if* (λ, x) *is constraint non-degenerate for every* $\lambda \in \mathrm{Sol}_\lambda(x)$.

Typically the solutions to `BendersPrimal`(x) are not unique. For example, if $x = (\beta_1, \ldots, \beta_d) \in \Omega^d$ and $\beta_i = \beta_j$ for some $i \neq j$, then the corresponding weights λ_i and λ_j in `BendersPrimal`(x) are not uniquely determined. The theorem below shows that, subject to strict convexity of ϕ, this is the only way in which the primal solution can be non-unique. It also shows that, under various strict convexity and regularity assumptions, the solutions to `BendersDual`(x) and primal-dual solutions to `BendersPrimal`(x) and `BendersDual`(x) are unique.

Theorem 4.7. *The following are true.*

1. *If* ϕ *is strictly convex, then* $A_\phi(x)\lambda^1 = A_\phi(x)\lambda^2$ *for all* $\lambda^1, \lambda^2 \in \mathrm{Sol}_\lambda(x)$.

2. *If* ϕ *is strictly convex and* $x \in \Omega$ *is objective non-degenerate, then* $\lambda^1_{\mathcal{I}(x,\beta)} = \lambda^2_{\mathcal{I}(x,\beta)}$ *for all* $\lambda^1, \lambda^2 \in \mathrm{Sol}_\lambda(x)$.

3. *If* ϕ^* *is strictly convex, then the optimal solutions* w *and* γ *to* `BendersDual`(x) *are unique.*

4. *If* ϕ^* *is strictly convex,* $x \in \Omega$ *is constraint non-degenerate, and*

$$(\lambda^i, w^i, w^i_E, w^i_I, w^i_M, \gamma^i) \in \Delta^{d-1} \times \mathbb{F} \times \mathbb{R}, \quad i = 1, 2,$$

are two primal-dual solution pairs to `BendersPrimal`(x) *and* `BendersDual`(x) *with* $\lambda^1_{\mathcal{I}(x,\beta)} = \lambda^2_{\mathcal{I}(x,\beta)}$*, then* $w^1_E = w^2_E$*,* $w^1_I = w^2_I$*, and* $w^1_M = w^2_M$.

5. *If $t > 0$ and $x \in \Omega^d$, then the optimal solution to* `BendersPrimal`(x, t) *exists and is unique.*

6. *If ϕ^* is strictly convex, $t > 0$, and $x \in \Omega^d$, then the optimal solution to* `BendersDual`(x, t) *exists and is unique.*

Proof. (1) It is a standard fact for optimizing strictly convex functions over convex sets.

(2) This follows from (1) and that the expression

$$A_\phi(x)\lambda = \sum_{\substack{\beta \in \text{Distinct}(x) \\ \lambda_{\mathcal{I}(x,\beta)} > 0}} f_\phi(\beta)\lambda_{\mathcal{I}(x,\beta)}$$

is the same for all $\lambda \in \text{Sol}_\lambda(x)$. By assumption $\left\{ f(\beta) \middle| \beta \in \text{Distinct}(x) \text{ and } \exists \lambda \in \text{Sol}_\lambda(x) \text{ with } \lambda_{\mathcal{I}(x,\beta)} > 0 \right\}$ is a linearly independent set, which implies the $\lambda_{\mathcal{I}(x,\beta)}$'s are uniquely defined.

(3) If ϕ^* is strictly convex then the optimal solution w is unique. If $(\overline{w}, \overline{\gamma})$ are optimal solutions, then $\overline{\gamma} = D(x) - \phi^*(\overline{w})$, so $\overline{\gamma}$ is also unique.

(4) This follows from the KKT conditions

$$\begin{aligned} y \circ \lambda &= 0 \\ y + A_\phi(x)^T w + A_E(x)^T w_E + A_I(x)^T w_I + A_M(x)^T w_M - \gamma e &= 0. \end{aligned}$$

Let $\overline{\lambda} \in \text{Sol}_\lambda(x)$, and define $C = \left\{ i \middle| \overline{\lambda}_i > 0 \right\}$. Then $C \neq \emptyset$ because $\overline{\lambda} \in \Delta^{d-1}$, so there exists at least one index i with $\overline{\lambda}_i > 0$. Letting $A(:, C)$ denote restricting a linear operator $A \in \text{Lin}(\mathbb{R}^p, X)$ to the elements of $\mathbb{R}^p$ with indices in C, the second KKT condition above implies

$$(A_\phi(x)(:, C))^T w + (A_E(x)(:, C))^T w_E + (A_I(x)(:, C))^T w_I + (A_M(x)(:, C))^T w_M - \gamma e = 0.$$

Finally, the optimal solutions w and γ are unique by (3) and

$$\begin{bmatrix} (A_E(x)(:,C))^T \\ (A_I(x)(:,C))^T \\ (A_M(x)(:,C))^T \end{bmatrix}$$

is injective by the definition of constraint non-degeneracy. So the solution set of (w_E, w_I, w_M) that satisfy the KKT conditions is a singleton.

(5) The objective $p(\lambda, s, S; x, t)$ is strictly convex in (λ, s, S), so the optimal solution exists and is unique.

(6) The objective $d(w, w_E, w_I, w_M, \gamma; x, t)$ is strictly convex in (w, w_E, w_I, w_M), so their optimal solutions are unique. Then the first-order primal-dual optimality conditions

$$\begin{aligned} \lambda \circ y - te &= 0 \\ y + {A_\phi}^T w + A_E^T w_E + A_I^T w_I + A_M^T w_M - \gamma e &= 0 \end{aligned}$$

from Proposition 4.4 combined with uniqueness of the optimal solution λ from (5) imply the optimal solution set for γ is unique. □

Since the optimal solutions for `BendersPrimal`(x, t) and `BendersDual`(x, t) are unique for $t > 0$, there is a well-defined mapping from $(x, t) \in \Omega \times \mathbb{R}_{++}$ to the optimal solution $(\lambda(x, t), s(x, t), S(x, t))$ of `BendersPrimal`(x, t). Similarly, for $t > 0$ let $(w(x, t), w_E(x, t), w_I(x, t), w_M(x, t), \gamma(x, t))$ denote the unique optimal solution to `BendersDual`(x, t).

The next proposition shows that, for $t > 0$, solutions along the central path distribute weight equally to components of λ corresponding to the same support point β. However, this corollary does not imply the weights are equally distributed at $t = 0$.

Proposition 4.8. *If $x \in \Omega^d$ and $t > 0$, then*

$$\lambda_i(x,t) = \lambda_j(x,t)$$

whenever $x_i = x_j$.

Proof. The proof amounts to the simple observation that the KKT conditions (4.16) include the conditions

$$\lambda \circ y = te$$

and

$$y = \gamma e - \left({A_\phi}^T w + A_I^T w_I + A_E^T w_E + A_M^T w_M \right).$$

The linear operators A_ϕ, A_I, A_E, and A_M all take the form $Ax = \sum_j f(\beta_j) x_j$ for some continuous f. So components of y corresponding to columns of A_ϕ, A_E, A_I, and A_M with the same β will be equal. Then $\lambda \circ y = te$ implies this is also true for λ. □

4.4 Limiting Behavior of Solutions to BendersPrimal(x,t) and BendersDual(x,t)

The following section builds up an understanding of how the optimal values and optimal solutions to $\texttt{BendersPrimal}(x,t)$ and $\texttt{BendersDual}(x,t)$ behave as $(x,t) \to (\overline{x}, 0)$. In particular, it is shown that, subject to regularity conditions, the limit of solutions to $\texttt{BendersPrimal}(x,t)$ and $\texttt{BendersDual}(x,t)$ as $t \to 0$ are solutions to $\texttt{BendersPrimal}(x,0)$ and $\texttt{BendersDual}(x,0)$, respectively.

The following result shows that the solutions to $\texttt{BendersPrimal}(x,t)$ are bounded as $t \searrow 0$. This is straight-forward for λ since $\lambda \in \Delta^{d-1}$ for all $t \geq 0$, but slightly more complicated for the primal variables s and S.

Proposition 4.9. *Suppose $(x^\nu, t^\nu) \to (\overline{x}, \bar{t})$, with $x^\nu \in \Omega^d$ and $t^\nu > 0$. The corresponding sequence $(\lambda^\nu, s^\nu, S^\nu)$ of optimal solutions, i.e.*

$$(\lambda^\nu, s^\nu, S^\nu) \in \operatorname{argmin} p(\lambda, s, S; x^\nu, t^\nu),$$

is bounded. Furthermore, the sequence

$$(z^\nu, z^\nu_{E}, z^\nu_{I}, z^\nu_{M}) = (A_\phi(x^\nu)\lambda^\nu, A_{E}(x^\nu)\lambda^\nu, A_{I}(x^\nu)\lambda^\nu, A_{M}(x^\nu)\lambda^\nu),$$

is also bounded.

Proof. The sequence $\{\lambda^\nu\}$ is bounded because $\lambda^\nu \in \Delta^{d-1}$ for all ν. Since $\{x^\nu\}$ is convergent it is also a bounded sequence. Boundedness of $\{\lambda^\nu\}$ and $\{x^\nu\}$ combined with the continuity of the map $x \mapsto (A_\phi(x), A_{E}(x), A_{I}(x), A_{M}(x))$ immediately implies z, z_{E}, z_{I}, and z_{M} are also bounded. All that remains is showing the slack variables s and S are bounded. If $t^\nu = 0$ then $s^\nu = z^\nu_{I}$ and $S^\nu = z^\nu_{M}$. If $t > 0$, then the first-order equation for optimizing $p(\lambda, s, S; x, t)$ in s_j only is

$$-\frac{t}{s_j} + \frac{1}{2}\left(s_j - (z_{I})_j\right) = 0.$$

This can be re-arranged into a quadratic in s_j,

$$s_j^2 - (z_{I})_j s_j - 2t = 0,$$

so that the quadratic formula yields the following bound:

$$|s_j| \leq \frac{1}{2}\left(|(z_{I})_j| + \sqrt{(z_{I})_j^2 + 4t^2}\right).$$

This shows s is bounded because z_{I} and t are. Now consider the optimal S for given z_{M} and $t > 0$. Differentiating $p(\lambda, s, S; x, t)$ with respect to S gives

$$-tS^{-1} + \frac{1}{t}\left(S - z_{M}\right) = 0.$$

Since S is real symmetric there exist a real orthogonal matrix Q and real diagonal matrix Λ

with $S = Q\Lambda Q^T$. Substituting this into the above equation gives

$$-tQ\Lambda^{-1}Q^T + \frac{1}{t}\left(Q\Lambda Q^T - z_{\scriptscriptstyle M}\right) = 0.$$

Rearranging gives

$$-t\Lambda^{-1} + \frac{1}{t}\left(\Lambda - Q^T z_{\scriptscriptstyle M} Q\right) = 0.$$

This shows $Q^T z_{\scriptscriptstyle M} Q$ is diagonal and each eigenvalue Λ_{jj} of S satisfies

$$\Lambda_{jj}^2 - \Lambda_{jj}(Q^T z_{\scriptscriptstyle M} Q)_{jj} - t^2 = 0.$$

Solving for Λ_{jj} using the quadratic formula shows

$$|\Lambda_{jj}| \leq \frac{1}{2}\left(\left|(Q^T z_{\scriptscriptstyle M} Q)_{jj}\right| + \sqrt{(Q^T z_{\scriptscriptstyle M} Q)_{jj}^2 + 4t^2}\right).$$

Since Q is orthogonal, $z_{\scriptscriptstyle M}^\nu$ is bounded, and t^ν is bounded, this shows S^ν is bounded. □

In addition, the function $G(\cdot, t)$ is an epi-continuous mapping as a function of t. This fundamental property shows that a homotopy-based algorithm, where $G(x, t)$ is successively optimized for values of t with $t \searrow 0$, might yield a sequence of optimal solutions to `BendersPrimal`(x, t) which converge to an optimal solution to `BendersPrimal`(x). This is explored in the remainder of the chapter.

Theorem 4.10. *The mapping $t \mapsto G(\cdot, t)$ for $t \geq 0$ is epi-continuous (see Definition 2.37).*

Proof. We use the limit definition in [55, Exercise 7.40]. This requires we show for every $t^\nu \to \bar{t}$ and $\overline{x} \in \Omega^d$ that for all sequences $x^\nu \to \overline{x}$,

$$\liminf_{\nu \to \infty} G(x^\nu, t^\nu) \geq G(\overline{x}, \bar{t}),$$

and there exists a sequence $x^\nu \to \overline{x}$ such that

$$\limsup_{\nu\to\infty} G(x^\nu, t^\nu) \le G(\overline{x}, \overline{t}).$$

There are three distinct cases: $\overline{t} > 0$, $\overline{t} = 0$ and $\overline{x}$ feasible, and $\overline{t} = 0$ and $\overline{x}$ infeasible. We show each in turn. Throughout the proof $(\lambda^\nu, s^\nu, S^\nu)$ denotes the optimal solutions to `BendersPrimal`(x^ν, t^ν).

$\overline{t} > 0$: Suppose $x^\nu \to \overline{x}$. If $\liminf_{\nu\to\infty} G(x^\nu, t^\nu) = +\infty$, then there is nothing to prove. So assume $\liminf_{\nu\to\infty} G(x^\nu, t^\nu) < +\infty$. By Proposition 4.9, $(\lambda^\nu, s^\nu, S^\nu)$ is a bounded sequence. So, taking a subsequence of the subsequence attaining $\liminf_{\nu\to\infty} G(x^\nu, t^\nu)$ if necessary, we further assume $t^\nu > 0$ for all ν and $(\lambda^\nu, s^\nu, S^\nu) \to (\overline{\lambda}, \overline{s}, \overline{S})$. Lower-semicontinuity of $p(\lambda, s, S; x, t)$ then gives

$$\begin{aligned}\liminf_{\nu\to\infty} G(x^\nu, t^\nu) &= \liminf_{\nu\to\infty} p(\lambda^\nu, s^\nu, S^\nu; x^\nu, t^\nu)\\ &= p(\overline{\lambda}, \overline{s}, \overline{S}; \overline{x}, \overline{t})\\ &\ge G(\overline{x}, \overline{t}).\end{aligned}$$

Now consider the sequence $x^\nu = \overline{x}$ for all ν. If $(\overline{\lambda}, \overline{s}, \overline{S}) \in \operatorname{argmin} G(\overline{x}, \overline{t})$, then since $p(\lambda, s, S; x, t)$ is continuous in t for $t > 0$,

$$\begin{aligned}\limsup_{\nu\to\infty} G(\overline{x}, t^\nu) &\le \limsup_{\nu\to\infty} p(\overline{\lambda}, \overline{s}, \overline{S}; \overline{x}, t^\nu)\\ &= p(\overline{\lambda}, \overline{s}, \overline{S}; \overline{x}, \overline{t})\\ &= G(\overline{x}, \overline{t}).\end{aligned}$$

$\overline{t} = 0$ and `BendersPrimal`$(\overline{x}, 0)$ infeasible: It suffices to show for any $x^\nu \to \overline{x}$ that

$$\liminf_{\nu\to\infty} G(x^\nu, t^\nu) = +\infty.$$

We temporarily define

$$A_C(x) = \begin{bmatrix} A_E(x) \\ A_I(x) \\ A_M(x) \end{bmatrix}, \quad C^\nu = \begin{bmatrix} A_E(x^\nu) \\ A_I(x^\nu) \\ A_M(x^\nu) \end{bmatrix} \Delta^{d-1}, \quad \text{and} \quad \overline{C} = \begin{bmatrix} A_E(\overline{x}) \\ A_I(\overline{x}) \\ A_M(\overline{x}) \end{bmatrix} \Delta^{d-1}.$$

Continuity of A_C implies $A_C(x^\nu)e_j \to A_C(\overline{x})e_j$ for each standard basis vector e_j. Then [55, Proposition 4.30(b)] implies

$$C^\nu = \text{co}\{A_C(x^\nu)e_1, \ldots, A_C(x^\nu)e_d\} \to \text{co}\{A_C(\overline{x})e_1, \ldots, A_C(\overline{x})e_d\} = \overline{C}.$$

Let $\|\cdot\|$ temporarily be the norm on $\mathbb{R}^{n_E} \times \mathbb{R}^{n_I} \times \mathbb{S}^{n_M}$. If $V \subset \mathbb{R}^{n_E} \times \mathbb{R}^{n_I} \times \mathbb{S}^{n_M}$ and $u \in \mathbb{R}^{n_E} \times \mathbb{R}^{n_I} \times \mathbb{S}^{n_M}$, then define

$$d(u, V) := \inf_{v \in V} \|v - u\|.$$

The triangle inequality implies for any $\overline{c} \in \overline{C}$, $c^\nu \in C^\nu$, and $b \in \mathbb{K} = \{0\}^{n_E} \times \mathbb{R}^{n_I}_- \times \mathbb{S}^{n_M}_-$ that

$$\|\overline{c} - b\| \leq \|\overline{c} - c^\nu\| + \|c^\nu - b\|.$$

Taking the infimum of both sides over $\overline{c} \in \overline{C}$ yields

$$d(b, \overline{C}) \leq d(c^\nu, \overline{C}) + \|c^\nu - b\|.$$

Now, taking the infimum on both sides over $b \in \mathbb{K}$

$$\left[\inf_{\overline{c} \in \overline{C},\ b \in \mathbb{K}} \|\overline{c} - b\|\right] \leq d(c^\nu, \overline{C}) + d(c^\nu, \mathbb{K}).$$

The infeasibility of `BendersPrimal`$(\overline{x})$ implies $\inf_{\overline{c} \in \overline{C},\ b \in \mathbb{K}} \|\overline{c} - b\| \geq \delta > 0$ for some

$\delta > 0$, so that

$$\delta \leq d(c^\nu, \overline{C}) + d(c^\nu, \mathbb{K})$$

for ν sufficiently large.

Since $C^\nu \to \overline{C}$ with C^ν and $\overline{C}$ compact, [55, Theorem 4.35(b)] implies $\sup_{c^\nu \in C^\nu} d(c^\nu, \overline{C}) \leq \delta/2$ for all ν sufficiently large. So $d(c^\nu, \mathbb{K}) \geq \delta/2$ for all ν sufficiently large. Hence any feasible solution $(\overline{\lambda}^\nu, \overline{s}^\nu, \overline{S}^\nu)$ to $\texttt{BendersPrimal}(x^\nu, t^\nu)$ satisfies $d(A_C(x^\nu)\lambda^\nu, \mathbb{K}) \geq \delta/2$ for ν sufficiently large. Using convexity of ϕ, lb, and logdet, there exist subgradients g_i, $i = 1, \ldots, 4$ and a constant $\alpha \in \mathbb{R}$, such that

$$\begin{aligned} &p(\lambda^\nu, s^\nu, S^\nu; x^\nu, t^\nu) \\ &\geq \alpha + \langle g_1,\, A_\phi(x^\nu)\lambda^\nu \rangle + t^\nu \langle g_2,\, \lambda^\nu \rangle + t^\nu \langle g_3,\, -s^\nu \rangle - t^\nu \langle g_4,\, -S^\nu \rangle \\ &\quad + \frac{\|A_C(x^\nu)\lambda^\nu - (0, s^\nu, S^\nu)\|^2}{2t^\nu}. \end{aligned}$$

Proposition 4.9 implies the optimal solutions $(\lambda^\nu, s^\nu, S^\nu)$ to $\texttt{BendersPrimal}(x^\nu, t^\nu)$ are bounded, hence for the sequence of optimal solutions $(\lambda^\nu, s^\nu, S^\nu)$ there exists $\alpha' \in \mathbb{R}$ with

$$\min_{(\lambda^\nu, s^\nu, S^\nu) \in \mathbb{E}} p(\lambda^\nu, s^\nu, S^\nu; x^\nu, t^\nu) \geq \alpha' + \frac{\delta^2}{2t^\nu} \quad \forall \nu.$$

This shows $p(\lambda^\nu, s^\nu, S^\nu; x^\nu, t^\nu) \to \infty$ as $\nu \to \infty$.

$\overline{t} = 0$ **and** $\texttt{BendersPrimal}(\overline{x}, 0)$ **feasible:** Suppose $x^\nu \to \overline{x}$. As in the argument for $\overline{t} > 0$, we may assume $(\lambda^\nu, s^\nu, S^\nu)$ is a convergent sequence, with $\lambda^\nu \to \overline{\lambda}$. Then

$$\begin{aligned}
&\liminf_{\nu\to\infty} G(x^\nu, t^\nu) \\
&= \liminf_{\nu\to\infty} p(\lambda^\nu, s^\nu, S^\nu; x^\nu, t^\nu) \\
&\geq \liminf_{\nu\to\infty} \phi(A_\phi(x^\nu)\lambda^\nu) \\
&\qquad + \liminf_{\nu\to\infty} t^\nu \operatorname{lb}(\lambda^\nu) + t^\nu \operatorname{lb}(-s^\nu) - t \operatorname{logdet}(-S^\nu) \\
&\qquad + \liminf_{\nu\to\infty} \frac{1}{2t^\nu} \|A_E(x^\nu)\lambda^\nu\|^2 + \frac{1}{2t^\nu} \|A_I(x^\nu)\lambda^\nu - s^\nu\|^2 + \frac{1}{2t^\nu} \|A_M(x^\nu)\lambda - S^\nu\|^2 \\
&\geq \liminf_{\nu\to\infty} \phi(A_\phi(x^\nu)\lambda^\nu) \\
&\geq \phi(A_\phi(\overline{x})\overline{\lambda}) \\
&\geq G(\overline{x}, 0),
\end{aligned}$$

where the constraint relaxation terms are taken as 0 if $t^\nu = 0$. The second to last inequality follows because, by Proposition 4.9, the sequence $\{\lambda^\nu, s^\nu, S^\nu\}$ is bounded, hence

$$\liminf_{\nu\to\infty} t^\nu \operatorname{lb}(\lambda^\nu) + t^\nu \operatorname{lb}(-s^\nu) - t \operatorname{logdet}(-S^\nu) \geq 0$$

$$\liminf_{\nu\to\infty} \frac{1}{2t^\nu} \|A_E(x^\nu)\lambda^\nu\|^2 + \frac{1}{2t^\nu} \|A_I(x^\nu)\lambda^\nu - s^\nu\|^2 + \frac{1}{2t^\nu} \|A_M(x^\nu)\lambda - S^\nu\|^2 \geq 0.$$

This proves $\liminf_{\nu\to\infty} G(x^\nu, t^\nu) \geq G(\overline{x}, 0)$.

Next we must show there exists a sequence $x^\nu \to \infty$ such that $\limsup_{\nu\to\infty} G(x^\nu, t^\nu) \leq G(\overline{x}, 0)$. To do his we take $x^\nu = \overline{x}$ for all ν and bound $G(\overline{x}, t^\nu)$ above. Let $\overline{\lambda}$, $\overline{s}$, and $\overline{S}$

be optimal solutions to $G(\overline{x},0)$. Define

$$\begin{aligned}\lambda^\nu &= (1-(t^\nu)^2)\overline{\lambda} + (t^\nu)^2 e\\ s^\nu &= (1-(t^\nu)^2)\overline{s} - (t^\nu)^2 e\\ S^\nu &= (1-(t^\nu)^2)\overline{S} - (t^\nu)^2 I.\end{aligned}$$

Then

$$\begin{aligned}G(\overline{x},t^\nu) &\le p(\lambda^\nu, s^\nu, S^\nu; \overline{x}, t^\nu)\\ &= \phi\left(A_\phi(\overline{x})\lambda^\nu\right)\\ &\quad + t^\nu \operatorname{lb}(\lambda^\nu) + t^\nu \operatorname{lb}(-s^\nu)\\ &\quad - t^\nu \operatorname{logdet}(-S^\nu)\\ &\quad + \frac{t^\nu}{2}\left\|A_{\scriptscriptstyle E}(\overline{x})e\right\|^2 + \frac{t^\nu}{2}\left\|A_{\scriptscriptstyle I}(\overline{x})e - e\right\|^2\\ &\quad + \frac{t^\nu}{2}\left\|A_{\scriptscriptstyle M}(\overline{x})e - I\right\|^2,\end{aligned}$$

where we interpret $t^\nu \operatorname{lb}(\lambda^\nu) = 0$ if $t^\nu = 0$, and similarly for $t^\nu \operatorname{lb}(s^\nu)$ and $-t^\nu \operatorname{logdet}(S^\nu)$. Taking the limsup of the rightmost expression as $\nu \to \infty$, and using the continuity of ϕ, we obtain

$$\limsup_{\nu\to\infty} G(\overline{x},t^\nu) \le \phi(A_\phi(\overline{x})\overline{\lambda}) = G(\overline{x},0).$$

□

The following proposition shows that the existence of a strictly feasible point for `BendersPrimal`(x) ensures that the limit of solutions of `BendersPrimal`(x,t) as $t \searrow 0$ will be feasible for `BendersPrimal`(x).

Proposition 4.11. *Let $t^\nu > 0$ and $(\lambda^\nu, s^\nu, S^\nu)$ be the optimal solutions to* `BendersPrimal`(x^ν, t^ν). *Suppose $(x^\nu, t^\nu) \to (\overline{x}, 0)$ and there exists a strictly feasible point to* `BendersPrimal`$(\overline{x}, 0)$.

Then

$$\lim_{\nu\to\infty} A_E(x^\nu)\lambda^\nu = 0,$$
$$\lim_{\nu\to\infty} A_I(x^\nu)\lambda^\nu - s^\nu = 0,$$
$$\lim_{\nu\to\infty} A_M(x^\nu)\lambda^\nu - S^\nu = 0.$$

Proof. By Proposition 4.9 the sequence $\{(\lambda^\nu, s^\nu, S^\nu)\}$ is bounded. Furthermore, $\{\phi(A_\phi(x^\nu)\lambda^\nu)\}$ is bounded below because ϕ and A_ϕ are smooth, ϕ is proper and continuous on its domain, and $x \in \Omega^d$ and $\lambda \in \Delta^{d-1}$ are bounded. Therefore there exists $\alpha_1 \in \mathbb{R}$ with

$$\alpha_1 + \frac{1}{2t^\nu}\left\|A_E(x^\nu)\lambda^\nu\right\|^2 + \frac{1}{2t^\nu}\left\|A_I(x^\nu)\lambda^\nu - s^\nu\right\|^2 + \frac{1}{2t^\nu}\left\|A_M(x^\nu)\lambda^\nu - S^\nu\right\|^2 \le p(\lambda^\nu, s^\nu, S^\nu; x^\nu, t^\nu)$$

for all ν.

Let (λ^*, s^*, S^*) be a strictly feasible point for $\texttt{BendersPrimal}(\overline{x}, 0)$. Then for ν large enough there exists $\alpha_2 \in \mathbb{R}$ with

$$p(\lambda^*, s^*, S^*; x^\nu, t^\nu) \le \alpha_2 + \frac{1}{2t^\nu}\left\|A_E(x^\nu)\lambda^*\right\|^2 + \frac{1}{2t^\nu}\left\|A_I(x^\nu)\lambda^* - s^*\right\|^2 + \frac{1}{2t^\nu}\left\|A_M(x^\nu)\lambda^* - S^*\right\|^2.$$

Since $p(\lambda^\nu, s^\nu, S^\nu; x^\nu, t^\nu) \le p(\lambda^*, s^*, S^*; x^\nu, t^\nu)$ for each ν, the above inequalities give

$$\begin{aligned}&\alpha_1 + \frac{1}{2t^\nu}\left\|A_E(x^\nu)\lambda^\nu\right\|^2 + \frac{1}{2t^\nu}\left\|A_I(x^\nu)\lambda^\nu - s^\nu\right\|^2 + \frac{1}{2t^\nu}\left\|A_M(x^\nu)\lambda^\nu - S^\nu\right\|^2 \\ &\le \alpha_2 + \frac{1}{2t^\nu}\left\|A_E(x^\nu)\lambda^*\right\|^2 + \frac{1}{2t^\nu}\left\|A_I(x^\nu)\lambda^* - s^*\right\|^2 + \frac{1}{2t^\nu}\left\|A_M(x^\nu)\lambda^* - S^*\right\|^2.\end{aligned}$$

Multiplying both sides through by t^ν yields

$$\begin{aligned}&t^\nu\alpha_1 + \left\|A_E(x^\nu)\lambda^\nu\right\|^2 + \left\|A_I(x^\nu)\lambda^\nu - s^\nu\right\|^2 + \left\|A_M(x^\nu)\lambda^\nu - S^\nu\right\|^2 \\ &\le t^\nu\alpha_2 + \left\|A_E(x^\nu)\lambda^*\right\|^2 + \left\|A_I(x^\nu)\lambda^* - s^*\right\|^2 + \left\|A_M(x^\nu)\lambda^* - S^*\right\|^2.\end{aligned}$$

Taking the limit as $\nu \to \infty$ gives

$$\lim_{\nu\to\infty} \|A_E(x^\nu)\lambda^\nu\|^2 + \|A_I(x^\nu)\lambda^\nu - s^\nu\|^2 + \|A_M(x^\nu)\lambda^\nu - S^\nu\|^2 \leq 0,$$

completing the proof. □

The lemma below proves that if `BendersPrimal`$(\overline{x}, 0)$ has a strictly feasible solution and $(x^\nu, t^\nu) \to (\overline{x}, 0)$, then `BendersPrimal`$(x^\nu, t^\nu)$ has a strictly feasible solution for ν sufficiently large. Furthermore, in a sense made precise below, it's possible to uniformly bound such strict solutions away from the boundary of the feasible region.

Lemma 4.12. *Suppose $x^\nu \to \overline{x}$, $A_E(\overline{x})$ is surjective, and there exists $\overline{\lambda}$ which is strictly feasible for* `BendersPrimal`$(\overline{x}, 0)$, *i.e.*

$$\begin{aligned}
\overline{\lambda} &> 0,\\
e^T\overline{\lambda} &= 1,\\
A_\phi(\overline{x})\overline{\lambda} &\in \operatorname{int}(\operatorname{dom}(\phi)),\\
A_E(\overline{x})\overline{\lambda} &= 0,\\
A_I(\overline{x})\overline{\lambda} &< 0,\\
A_M(\overline{x})\overline{\lambda} &\prec 0.
\end{aligned}$$

Then for any $\epsilon > 0$ there exists a sequence $\{\lambda^\nu\}$ and a ν_0 such that each λ^ν is strictly feasible for `BendersPrimal`$(x^\nu, 0)$ *if $\nu \geq \nu_0$. Furthermore, ν_0 can also be chosen so that there exists $\delta > 0$ such that*

$$\begin{aligned}
\lambda^\nu &\geq \delta e,\\
A_I(x^\nu)\lambda^\nu &\leq -\delta e,\\
A_M(x^\nu)\lambda^\nu &\preceq -\delta I, \qquad (4.21)\\
\left\|\lambda^\nu - \overline{\lambda}\right\| &< \epsilon,\\
\left\|A_\phi(\overline{x})\overline{\lambda} - A_\phi(x^\nu)\lambda^\nu\right\| &< \epsilon
\end{aligned}$$

for all $\nu \geq \nu_0$.

Proof. Define

$$C(x) = \left\{ \lambda \in \mathbb{R}^d \middle| \, A_E(x)\lambda = 0 \right\}.$$

Then [55, Thm 4.32(b)] implies $\lim_{\nu\to\infty} C(x^\nu) = C(\overline{x})$. In particular this means there exists a sequence λ^ν such that $\lambda^\nu \to \overline{\lambda}$ with $A_E(x^\nu)\lambda^\nu = 0$ for all ν. So for large enough ν we have $\lambda^\nu > 0$. By scaling if necessary, assume without loss of generality that $e^T\lambda^\nu = 1$ for all ν. The remaining inequalities follow because $(x^\nu, \lambda^\nu) \to (\overline{x}, \overline{\lambda})$ and the mappings

$$(x, \lambda) \mapsto A_\phi(x)\lambda$$
$$(x, \lambda) \mapsto A_I(x)\lambda$$
$$(x, \lambda) \mapsto A_M(x)\lambda$$

are all continuous. □

Lemma 4.12 above is used in the next proposition to show any limit as $t \searrow 0$ of optimal solutions to `BendersPrimal`(x, t) is an optimal solution to `BendersPrimal`$(x, 0)$. This strengthens Proposition 4.11, which only showed the limit, if it exists, is feasible.

Proposition 4.13. *Suppose there exists a strictly feasible point to* `BendersPrimal`$(\overline{x}, 0)$*,* $(x^\nu, t^\nu) \to (\overline{x}, 0)$ *with* $t^\nu > 0$*, and* $A_E(\overline{x})$ *is surjective. Then any limit point of optimal solutions* $(\lambda^\nu, s^\nu, S^\nu)$ *to* `BendersPrimal`(x^ν, t^ν) *is an optimal solution to* `BendersPrimal`$(\overline{x}, 0)$*.*

Proof. To simplify notation, assume the entire sequence of primal optimal solutions $(\lambda^\nu, s^\nu, S^\nu)$ converges to $(\overline{\lambda}, \overline{s}, \overline{S})$. Proposition 4.11 implies $(\overline{\lambda}, \overline{s}, \overline{S})$ is feasible for `BendersPrimal`$(\overline{x}, 0)$.

Suppose $(\overline{\lambda}, \overline{s}, \overline{S})$ is not an optimal solution for `BendersPrimal`$(\overline{x}, 0)$. A point $D^* = (\lambda^*, s^*, S^*)$ which is strictly feasible for `BendersPrimal`$(\overline{x}, 0)$ and satisfies $\phi(A_\phi(\overline{x})\lambda^*) < \phi(A_\phi(\overline{x})\overline{\lambda})$ will be constructed. Let $\widetilde{D} = (\widetilde{\lambda}, \widetilde{s}, \widetilde{S})$ be an optimal solution for `BendersPrimal`$(\overline{x}, 0)$ and $\widehat{D} = (\widehat{\lambda}, \widehat{s}, \widehat{S})$ be strictly feasible for `BendersPrimal`$(\overline{x}, 0)$. If $\widetilde{D}$ is strictly feasible, then define $D^* := \widetilde{D}$. If $\widetilde{D}$ is not strictly feasible, then define $D^* := (1-\alpha)\widetilde{D} + \alpha\widehat{D}$, where $\alpha \in (0, 1)$

is chosen sufficiently small so that

$$\phi\left(A_\phi(\overline{x})((1-\alpha)\widetilde{\lambda}+\alpha\widehat{\lambda})\right) < \phi(A_\phi(\overline{x})\overline{\lambda}).$$

Such an α exists by continuity of ϕ.

A sequence of strictly feasible solutions to `BendersPrimal`(x^ν, t^ν) approximating D^* will be constructed. The objective function at these approximating points will eventually be lower than those of $(\lambda^\nu, s^\nu, S^\nu)$, contradicting the optimality of $(\lambda^\nu, s^\nu, S^\nu)$ for `BendersPrimal`(x^ν, t^ν) for every ν. This will complete the proof.

Define $\eta = \phi(A_\phi(\overline{x})\overline{\lambda}) - \phi(A_\phi(\overline{x})\lambda^*)$, so $\eta > 0$, and choose ϵ such that if $v \in \mathbb{R}^n$ and $\|v - A_\phi(\overline{x})\lambda^*\| < \epsilon$ then $|\phi(v) - \phi(A_\phi(\overline{x})\lambda^*)| < \eta/2$. Lemma 4.12 applied to ϵ and D^* yields a sequence of points $\big((\lambda^\sharp)^\nu, (s^\sharp)^\nu, (S^\sharp)^\nu\big)$ strictly feasible for `BendersPrimal`(x^ν, t^ν) with

$$\begin{aligned}
(s^\sharp)^\nu &= A_{I}(x^\nu)(\lambda^\sharp)^\nu \\
(S^\sharp)^\nu &= A_{M}(x^\nu)(\lambda^\sharp)^\nu \\
(\lambda^\sharp)^\nu &\geq \delta e \\
(s^\sharp)^\nu &\leq -\delta e \\
(S^\sharp)^\nu &\preceq -\delta e I \\
\big\|(\lambda^\sharp)^\nu - \lambda^*\big\| &< \epsilon \\
\big\|A_\phi(\overline{x})\lambda^* - A_\phi(x^\nu)(\lambda^\sharp)^\nu\big\| &< \epsilon
\end{aligned}$$

for some $\delta > 0$.

Since

$$p((\lambda^\sharp)^\nu, (s^\sharp)^\nu, (S^\sharp)^\nu; x^\nu, t^\nu) = \phi(A_\phi(x^\nu)(\lambda^\sharp)^\nu) + t^\nu \operatorname{lb}((\lambda^\sharp)^\nu) + t^\nu \operatorname{lb}((s^\sharp)^\nu) - t \operatorname{logdet}((S^\sharp)^\nu),$$

the later three terms go to 0 as $\nu \to \infty$ and $\phi(A_\phi(x^\nu)(\lambda^\sharp)^\nu) < \phi(A_\phi(\overline{x})\overline{\lambda})$. So, for ν sufficiently large,

$$p(x^\nu, t^\nu, (D^\sharp)^\nu) < p(x^\nu, t^\nu, D^\nu).$$

This contradicts optimality of $(\lambda^\nu, s^\nu, S^\nu)$ for `BendersPrimal`(x^ν, t^ν). Thus $(\overline{\lambda}, \overline{s}, \overline{S})$ is optimal for `BendersPrimal`$(\overline{x}, 0)$. □

In Proposition 4.13 the assumption $A_E(\overline{x})$ is surjective cannot be removed. The following provides a counter-example.

Example 4.14. Let $\Omega = [-5, 5]^2 \subset \mathbb{R}^2$ and

$$\begin{aligned} z &= (z_1, z_2) \in \Omega \\ \phi(z) &= z_1^2 \\ A_\phi(x^\nu) &= \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \\ A_E(x^\nu) &= \begin{bmatrix} \nu^{-1} & -\nu^{-1} \end{bmatrix} \\ t^\nu &= \nu^{-4} \end{aligned}$$

and there are no inequality or semidefinite constraints, then

$$\begin{aligned} p(\lambda, s, S; x^\nu, t^\nu) &= \lambda_1^2 + \frac{1}{2t^\nu} \left\| A_E(x^\nu)\lambda \right\|^2 \\ &= \lambda_1^2 + \frac{\nu^2}{2}(\lambda_1 - \lambda_2)^2. \end{aligned}$$

So the limit of optimal solutions λ^ν as $\nu \to \infty$ is $(0.5, 0.5)$ rather than the optimal solution $(0, 1)$.

To prove corresponding results for the dual variables we require a technical lemma.

Lemma 4.15. *For* $(\overline{x}, \overline{t}) \in \Omega \times \mathbb{R}_{++}$, *if* $(\overline{\lambda}, \overline{s}, \overline{S})$ *is the optimal solution to* `BendersPrimal`$(\overline{x}, \overline{t})$

and $(\overline{w}, \overline{w_E}, \overline{w_I}, \overline{w_M}, \overline{\gamma})$ *is the optimal solution to* $\texttt{BendersDual}(\overline{x}, \bar{t})$*, then for all* $\lambda \in \mathbb{R}^d$

$$\begin{aligned}\phi(A_\phi(\overline{x})\lambda) + \left\langle \lambda,\, -\overline{y} - A_E(\overline{x})^T\overline{w_E} - A_I(\overline{x})^T\overline{w_I} - A_M(\overline{x})^T\overline{w_M} + \overline{\gamma}e \right\rangle \\ \geq \phi(A_\phi(\overline{x})\overline{\lambda}) + \langle \overline{z},\, \overline{w} \rangle\,,\end{aligned}$$

where, as in Proposition 4.4, $\overline{z} = A_\phi(\overline{x})\overline{\lambda}$ *and* $\overline{\lambda} \circ \overline{y} = \bar{t}e$*.*

Proof. Let $f : \mathbb{R}^d \to \overline{\mathbb{R}}$ be $f(\lambda) := \phi(A_\phi(\overline{x})\lambda)$. Then f is convex and differentiable at $\overline{\lambda}$. Proposition 4.4 implies

$$\begin{aligned}\overline{w} &= -\nabla\phi(\overline{z}),\\ A_\phi(\overline{x})^T\overline{w} &= -\overline{y} - A_E(\overline{x})^T\overline{w_E} - A_I(\overline{x})^T\overline{w_I} - A_M(\overline{x})^T\overline{w_M} + \overline{\gamma}e,\end{aligned}$$

so,

$$\begin{aligned}\left\langle \nabla f(\overline{\lambda}),\, \overline{\lambda} \right\rangle &= \left\langle -A_\phi(\overline{x})^T\overline{w},\, \overline{\lambda} \right\rangle\\ &= \left\langle -\overline{w},\, A_\phi(\overline{x})\overline{\lambda} \right\rangle\\ &= -\left\langle \overline{w},\, \overline{z} \right\rangle\\ \left\langle \nabla f(\overline{\lambda}),\, \lambda \right\rangle &= \left\langle -A_\phi(\overline{x})^T\overline{w},\, \lambda \right\rangle\\ &= \left\langle \overline{y} + A_E(\overline{x})^T\overline{w_E} + A_I(\overline{x})^T\overline{w_I} + A_M(\overline{x})^T\overline{w_M} - \overline{\gamma}e,\, \lambda \right\rangle\end{aligned}$$

Substituting the above into the subdifferential inequality $f(\lambda) \geq f(\overline{\lambda}) + \left\langle \nabla f(\overline{\lambda}),\, \lambda - \overline{\lambda} \right\rangle$, which holds for for all $\lambda \in \mathbb{R}^d$, completes the proof.

□

The following proposition is the dual version of Proposition 4.9. It shows that as $t \searrow 0$ the optimal solutions of $\texttt{BendersDual}(x, t)$ are bounded.

Proposition 4.16. *Suppose* $(x^\nu, t^\nu) \to (\overline{x}, 0)$ *and* $\texttt{BendersPrimal}(\overline{x}, 0)$ *is strictly feasible. Define the corresponding sequence* $\{(\lambda^\nu, s^\nu, S^\nu)\}$ *and* $\{(w^\nu, w_E^\nu, w_I^\nu, w_M^\nu, \gamma^\nu)\}$ *of primal-dual* *solutions to* `BendersPrimal`(x^ν, t^ν) *and* `BendersDual`(x^ν, t^ν). *The sequence* $\{(w^\nu, \gamma^\nu)\}$ *is bounded. If* $A_E(\overline{x})$ *is surjective, then the sequence* $\{(w_E^\nu, w_I^\nu, w_M^\nu, y^\nu)\}$ *is also bounded, where, as in* (4.16), $y^\nu \circ \lambda^\nu - t^\nu e = 0$ *and for each* ν.

Proof. We prove this in several pieces.

w **is bounded:** Suppose w^ν is not bounded. Taking a subsequence, if necessary, we may assume $|w^\nu| \to \infty$. By Proposition 4.9 we may also assume, taking a further subsequence if necessary, $z^\nu \to \overline{z}$ for some $\overline{z}$ where $\overline{z}$ is part of an optimal solution to $p(\overline{x}, 0)$. So

$$\lim_{\nu\to\infty} w^\nu = \lim_{\nu\to\infty} -\nabla\phi(z^\nu) = -\nabla\phi(\overline{z})$$

where the first equality follows from the first-order optimality conditions and the second because ϕ is continuously differentiable. This contradiction shows that w^ν must be bounded.

γ **is bounded:** Recall the first-order optimality condition

$$\gamma e = y + {A_\phi}^T w + A_I^T w_I + A_E^T w_E + A_M^T w_M$$

from Proposition 4.4. Taking the inner-product of that condition with λ and applying

Proposition 4.4 again gives

$$\begin{aligned}\gamma &= t\,\langle y,\,\lambda\rangle + \langle w,\,z\rangle + \langle z_I,\,w_I\rangle + \langle z_E,\,w_E\rangle + \langle z_M,\,w_M\rangle \\ &= t\,\langle y,\,\lambda\rangle + \langle w,\,z\rangle + \langle z_E,\,w_E\rangle \\ &\qquad + \langle z_I - s,\,w_I\rangle + \langle s,\,w_I\rangle \\ &\qquad + \langle z_M - S,\,w_M\rangle + \langle S,\,w_M\rangle \\ &= td + \langle w,\,z\rangle - \frac{1}{t}\,\|z_E\|^2 \\ &\qquad - \frac{1}{t}\,\|z_I - s\|^2 + tn_I \\ &\qquad - \frac{1}{t}\,\|z_M - S\|^2 + tn_M\end{aligned}$$

The final expression is bounded by Proposition 4.9.

y, w_I, and w_M are bounded: This follows by applying Lemma 4.15, which implies for any λ that

$$\begin{aligned}&\phi(A_\phi(x^\nu)\lambda) + \left\langle \lambda,\, -y^\nu - A_E(x^\nu)^T w_E^\nu - A_I(x^\nu)^T w_I^\nu - A_M(x^\nu)^T w_M^\nu + \gamma^\nu e\right\rangle \\ &\qquad \geq \phi(A_\phi(x^\nu)\lambda^\nu) + \langle z^\nu,\, w^\nu\rangle\,.\end{aligned}$$

Let $\{\widehat{\lambda}^\nu\}$ and $\delta > 0$ be obtained from Lemma 4.12 so that (4.21) is satisfied. Inserting $\widehat{\lambda}^\nu$ into the above inequality yields

$$\begin{aligned}&\phi(A_\phi(x^\nu)\widehat{\lambda}^\nu) + \left\langle \widehat{\lambda}^\nu,\, -y^\nu - A_E(x^\nu)^T w_E^\nu - A_I(x^\nu)^T w_I^\nu - A_M(x^\nu)^T w_M^\nu + \gamma^\nu e\right\rangle \\ &\qquad = \phi(A_\phi(x^\nu)\widehat{\lambda}^\nu) + \left\langle \widehat{\lambda}^\nu,\, -y^\nu - A_I(x^\nu)^T w_I^\nu - A_M(x^\nu)^T w_M^\nu\right\rangle + \gamma^\nu \\ &\qquad \geq \phi(A_\phi(x^\nu)\lambda^\nu) + \langle z^\nu,\, w^\nu\rangle\,.\end{aligned}$$

If $(y^\nu, w_I^\nu, w_M^\nu)$ is not bounded then, taking a subsequence if necessary to ensure

$\|(y^\nu, w_I^\nu, w_M^\nu)\| \to \infty$, divide both sides above by $\|(y^\nu, w_I^\nu, w_M^\nu)\|$. Since $\widehat{\lambda}^\nu$ is bounded, as are λ^ν, z^ν, γ^ν and w^ν, as $\nu \to \infty$ we may take a convergence subsequence if necessary so that

$$\begin{aligned} &\left\langle \widehat{\lambda},\, -\widehat{y} - A_I(\overline{x})^T \widehat{w_I} - A_M(\overline{x})^T \widehat{w_M} \right\rangle \\ &= \left\langle \widehat{\lambda},\, -\widehat{y} \right\rangle + \left\langle -A_I(\overline{x})\widehat{\lambda},\, \widehat{w_I} \right\rangle + \left\langle -A_M(\overline{x})\widehat{\lambda},\, \widehat{w_M} \right\rangle \\ &\geq 0, \end{aligned} \tag{4.22}$$

where

$$\lim_{\nu\to\infty} \frac{(y^\nu, w_I^\nu, w_M^\nu)}{\|(y^\nu, w_I^\nu, w_M^\nu)\|} = (\widehat{y}, \widehat{w_I}, \widehat{w_M}) \quad \text{and} \quad \lim_{\nu\to\infty} \widehat{\lambda}^\nu = \widehat{\lambda}.$$

By construction $\widehat{\lambda}$ and $-A_I(\overline{x})\widehat{\lambda}$ are entry-wise positive, while $-A_M(\overline{x})\widehat{\lambda}$ is positive definite. By the definition of `BendersDual`(x, t), $y^\nu \geq 0$, $w_I^\nu \leq 0$, and $w_M^\nu \preceq 0$. Therefore each term in(4.22) is non-positive, and hence each term must be zero. This is a contradiction because $(\widehat{y}, \widehat{w_I}, \widehat{w_M})$ has unit norm while $\widehat{\lambda}$ and $-A_I(\overline{x})\widehat{\lambda}$ have all positive entries and $-A_M(\overline{x})\widehat{\lambda}$ is positive definite. This contradiction shows that $(y^\nu, w_I^\nu, w_M^\nu)$ must be bounded as $\nu \to \infty$.

w_E **is bounded:** This follows because

$$A_E^T(x^\nu) w_E^\nu = \gamma^\nu e - y^\nu - A_\phi{}^T(x^\nu) w^\nu - A_I^T(x^\nu) w_I^\nu - A_M^T(x^\nu) w_M^\nu$$

and $A_E(\overline{x})$ has full rank. In particular, because singular values depend continuously on a matrix, this implies that there exists $\nu_0 > 0$ and $C > 0$ such that for all $\nu \geq \nu_0$

$$\|w_E^\nu\| \leq C \left\| \gamma^\nu e - A_\phi{}^T(x^\nu) w^\nu - A_I^T(x^\nu) w_I^\nu - A_M^T(x^\nu) w_M^\nu \right\|.$$

But γ^ν, w^ν, w_I^ν, and w_M^ν are bounded and $A_\phi(x^\nu) \to A_\phi(\overline{x})$, $A_I(x^\nu) \to A_I(\overline{x})$, and $A_M(x^\nu) \to A_M(\overline{x})$. So the above inequality implies w_E^ν is also bounded.

□

The theorem below builds on all of the previous propositions to give results on the behavior of optimal solutions to `BendersPrimal`(x,t) and `BendersDual`(x,t) as $t \searrow 0$. In addition, it shows the optimal value function $G(x,t)$ is differentiable for $t > 0$ and, subject to regularity conditions, continuous at $t = 0$.

Theorem 4.17. *Assume that ϕ is twice differentiable strictly convex on its domain $\operatorname{dom}\phi$, which we assume to be open. Moreover, assume there exists a set $\mathcal{O}$ with $\Omega \subset \mathcal{O}$ on which f_ϕ, f_E, f_I, and f_M are continuously differentiable. Also, assume $A_\phi(\mathcal{O}) \operatorname{int} \Delta^{d-1} \subset \operatorname{dom}\phi$, that $(x^\nu, t^\nu) \to (\overline{x}, 0)$ with $t^\nu > 0$ for each ν, and that* `BendersPrimal`$(\overline{x}, 0)$ *is strictly feasible.*

1. *For $(x,t) \in \Omega^d \times \mathbb{R}_{++}$ the primal-dual solutions*

$$(\lambda, s, S, z, z_I, z_E, z_M, w, w_E, w_I, w_M, \gamma) \in \mathbb{E} \times \mathbb{F} \times \mathbb{F} \times \mathbb{R}$$

   and $G(x,t)$ are continuously differentiable.

2. *Let $\{(\lambda^\nu, s^\nu, S^\nu), (w^\nu, w_E^\nu, w_I^\nu, w_M^\nu, \gamma^\nu)\}$ be the sequence of primal-dual solutions to* `BendersPrimal`(x^ν, t^ν) *and* `BendersDual`(x^ν, t^ν)*. If $A_E(\overline{x})$ is surjective, then there is a convergent subsequence of this sequence of primal-dual solutions. Any limit point of primal-dual solution pairs to* `BendersPrimal`(x^ν, t^ν) *is a primal-dual solution pair for* `BendersPrimal`$(\overline{x}, 0)$*.*

3. *Let $\{(\lambda^\nu, s^\nu, S^\nu), (w^\nu, w_E^\nu, w_I^\nu, w_M^\nu, \gamma^\nu)\}$ be the sequence of primal-dual solutions to* `BendersPrimal`(x^ν, t^ν) *and* `BendersDual`(x^ν, t^ν) *and assume $A_E(\overline{x})$ is surjective. Then $\{z^\nu\}$, $\{z_E^\nu\}$, $\{w^\nu\}$, and $\{\gamma^\nu\}$ are convergent sequences. If $\overline{x}$ is objective non-degenerate, then $\{\lambda^\nu_{\mathcal{I}(\overline{x},\beta)}\}$ is convergent for each $\beta \in \operatorname{Distinct}(\overline{x})$. If $\overline{x}$ is objective and constraint non-degenerate, then $\{y^\nu\}$, $\{w_E^\nu\}$, $\{w_I^\nu\}$, and $\{w_M^\nu\}$ are also convergent.*

4. *If $A_E(\overline{x})$ is surjective and $\overline{x}$ is constraint non-degenerate, then $\lim_{\nu\to\infty} G(x^\nu, t^\nu) = G(\overline{x}, 0)$.*

Proof. (1) This is a consequence of Theorem 2.42, an instance of the implicit function theorem. In the notation of that theorem, take $Q = (\operatorname{int}\Delta^{d-1} \times \mathbb{R}^{n_I}_{--} \times \mathbb{S}^{n_M}_{--}) \times (\mathcal{O}^d \times \mathbb{R}_{++})$ and f to be $p(\lambda, s, S; x, t)$. Then p is obviously smooth for $t > 0$, and the transversality condition on Q is trivially true. The final requirement to apply the theorem is (λ, s) is a strong minimizer of $p(x, t, \lambda, s)$. This follows because $\nabla^2_{(\lambda,s,S)} p$ is positive definite. This can be seen by writing $p(\lambda, s, S; x, t)$ as

$$p(\lambda, s, S; x, t) = t\,\mathrm{lb}(\lambda) + t\,\mathrm{lb}(-s) - t\,\mathrm{logdet}(-S) + \widetilde{p}(\lambda, s, S; x, t), \tag{4.23}$$

where $\widetilde{p}$ is convex and the hessian of the first 3 terms is positive definite on its domain. So we may apply Theorem 2.42. This gives a local continuously differentiable function $\Phi(x, t) = (\lambda(x, t), s(x, t), S(x, t))$. Moreover, since the minimizers $\lambda(x, t)$, $s(x, t)$, and $S(x, t)$ are unique for each $(x, t) \in \mathcal{O}^d$ these local mappings can be used to construct a global continuously differentiable extension of Φ to all of $\mathcal{O}^d \times \mathbb{R}_{++}$ by continuation. Thus $\lambda(x, t)$, $s(x, t)$ and $S(x, t)$ are continuously differentiable on $\Omega^d \times \mathbb{R}_{++}$. Finally, The optimality conditions in Proposition 4.4 demonstrate the remaining primal and dual variables are also continuously differentiable. It follows immediately that $G(x, t)$ is continuously differentiable on $\Omega^d \times \mathbb{R}_{++}$.

(2) Boundedness follows from Propositions 4.9 and 4.16. Continuity of the first order optimality conditions, (4.16) in Proposition 4.4, prove any limit of solutions is an optimal solution of the limit.

(3) Recall that any bounded sequence with a single limit point must converge to that limit point. By (2), limit points of the optimal solutions to $\texttt{BendersPrimal}(x^\nu, t^\nu)$ or $\texttt{BendersDual}(x^\nu, t^\nu)$ are optimal solutions to $\texttt{BendersPrimal}(\overline{x}, 0)$ or $\texttt{BendersDual}(\overline{x}, 0)$. So convergence follows for a component of the primal-dual solutions when the corresponding component of the optimal solution to $\texttt{BendersPrimal}(\overline{x}, 0)$ or $\texttt{BendersDual}(\overline{x}, 0)$ is unique. Let $\{(\overline{\lambda}, \overline{s}, \overline{S}, \overline{z}, \overline{z_E}, \overline{z_I}, \overline{z_M}), (\overline{w}, \overline{w_E}, \overline{w_I}, \overline{w_M}, \overline{\gamma})\}$ be a primal-dual optimal solution to $\texttt{BendersDual}(\overline{x}, 0)$. By definition, $\overline{z_E}$ is zero, and hence unique. By Theorem 4.7(1), the

optimal solution $\overline{z}$ is unique. By [55, Theorem 11.13], ϕ^* is strictly convex. Hence, by Theorem 4.7(3), the optimal solutions $\overline{w}$ and $\overline{\gamma}$ are unique as well. If $\overline{x}$ is also objective non-degenerate, then Theorem 4.7(2) shows $\{\lambda^{\nu}_{\mathcal{I}(\overline{x},\beta)}\}$ is unique. Finally, if $\overline{x}$ is constraint non-generate, then Theorem 4.7(4) gives that the optimal solutions $\overline{w_E}$, $\overline{w_I}$, and $\overline{w_M}$ are also unique. The first-order optimality conditions in Proposition 4.4 show that if $\overline{w}$, $\overline{\gamma}$, $\overline{w_E}$, $\overline{w_I}$, and $\overline{w_M}$ are unique, then $\overline{y}$ is also unique.

(4) Applying the first-order optimality conditions in Proposition 4.4 to the definition of $G(x,t)$ gives

$$\begin{aligned} G(x,t) &= \phi(z(x,t)) + t\,\mathrm{lb}(\lambda(x,t)) + t\,\mathrm{lb}(s(x,t)) - t\,\mathrm{logdet}(S(x,t)) \\ &\quad + \frac{t}{2}\left\|w_E(x,t)\right\|^2 + \frac{t}{2}\left\|w_I(x,t)\right\|^2 + \frac{t}{2}\left\|w_M(x,t)\right\|^2 . \end{aligned} \tag{4.24}$$

The first term tends to $\phi(z(\overline{x},0))$ as $\nu \to \infty$. Boundedness of $\{w_E^{\nu}\}$, $\{w_I^{\nu}\}$, and $\{w_M^{\nu}\}$ as $\nu \to \infty$ implies the last three terms tend to zero. To complete the proof we need only show the logarithmic barrier terms vanish as $\nu \to \infty$. For all $(x,t) \in \Omega^d \times \mathbb{R}_{++}$, the equation $\lambda(x,t)_i = t/y(x,t)_i$ holds. So

$$t\log(\lambda(x,t)_i) + t\log(y(x,t)_i) = t\log(t).$$

Taking the limit as $\nu \to \infty$ gives

$$\lim_{\nu\to\infty} t^{\nu}\log(\lambda_i^{\nu}) + t^{\nu}\log(y_i^{\nu}) = 0.$$

By Proposition 4.9 and 4.16, both $\{\lambda^{\nu}\}$ and $\{y^{\nu}\}$ are bounded, so the above implies $t\,\mathrm{lb}(\lambda^{\nu}) \to 0$. A similar argument can be applied to $\{s^{\nu}\}$ and $\{S^{\nu}\}$ using, respectively, that $s(x,t)_i \circ w_I(x,t)_i = t$ and $S(x,t)w_M(x,t) = tI$.

□

4.5 Variational Analysis

The previous section showed that the optimal value function $G(x,t)$ is continuously differentiable as a function of x for $t > 0$. This section considers the variational properties of $G(x,t)$ as a function of x for $t > 0$ and $t = 0$. Finally, it is shown that the limit of first-order stationary points for $G(x,t)$ in x as $t \searrow 0$ is a first-order Clarke stationary point for $G(x,0)$.

If $\mathcal{E}$ is a Euclidian space and $g : \Omega \to \mathcal{E}$, matrix-valued functions $A_g : \Omega^d \to \mathcal{E}$ of the form $A_g(x)e_i = g(\beta_i)$ for $i = 1, \ldots, d$ have particularly simple first and second derivatives. In particular, this applies to the linear operator-valued functions A_ϕ, $A_{\scriptscriptstyle E}$, $A_{\scriptscriptstyle I}$, and $A_{\scriptscriptstyle M}$. For future reference, these derivatives are derived below. This proposition uses the notation for differentials introduced in Section 2.5.1

Proposition 4.18. *Let $\mathcal{E}$ be a Euclidian space and $g : \Omega \to \mathcal{E}$ smooth on an open set containing Ω. Define $A_g : \Omega^d \to \mathrm{Lin}(\mathbb{R}^d, \mathcal{E})$ as in* (4.5)*. Then the first and second derivatives of A_g at $x = (\beta_1, \ldots, \beta_d) \in \Omega^d$ are given by*

$$d_x A_g(x)[\Delta x] = \begin{bmatrix} d_\beta g(\beta_1)[\Delta\beta_1] & \cdots & d_\beta g(\beta_d)[\Delta\beta_d] \end{bmatrix}$$

and

$$d^2_{xx} A_g(x)[\Delta x, \Delta x] = \begin{bmatrix} d^2_{\beta\beta} g(\beta_1)[\Delta\beta_1, \Delta\beta_1] & \cdots & d^2_{\beta\beta} g(\beta_d)[\Delta\beta_d, \Delta\beta_d] \end{bmatrix}.$$

Proof. The given derivatives are straight-forward from the form of A_g. □

The proposition below gives the first and second derivatives of the optimal value function $G(x,t)$. These derivatives involve both the chain rule and, if written formally, tensors. This proposition uses the notation for differentials introduced in Section 2.5.1. The differential is used instead of the more traditional gradient to simplify the notation for the chain rule.

Proposition 4.19. *For $t > 0$ the first and second derivatives of $G(x,t)$ are given by*

$$\begin{aligned} d_x G(x,t)[\Delta x] &= d\phi(A_\phi(x)\lambda(x,t)) d_x A_\phi(x)[\Delta x]\lambda(x,t) \\ &+ \frac{1}{t}\left(A_E(x)\lambda(x,t)\right)^T d_x A_E(x)[\Delta x]\lambda(x,t) \\ &+ \frac{1}{t}(A_I(x)\lambda(x,t) - s(x,t))^T d_x A_I(x)[\Delta x]\lambda(x,t) \\ &+ \frac{1}{t}\left\langle A_M(x)\lambda(x,t) - S(x,t)),\ d_x A_M(x)[\Delta x]\lambda(x,t)\right\rangle \end{aligned} \tag{4.25}$$

$$\begin{aligned} &= -w(x,t)^T d_x A_\phi(x)[\Delta x]\lambda(x,t) \\ &- w_E(x,t)^T d_x A_E(x)[\Delta x]\lambda(x,t) \\ &- w_I(x,t)^T d_x A_I(x)[\Delta x]\lambda(x,t) \\ &- \left\langle w_M(x,t),\ d_x A_M(x)[\Delta x]\lambda(x,t)\right\rangle . \end{aligned} \tag{4.26}$$

and

$$\begin{aligned} d_{xx}G(x,t) &= d_{xx}p(\lambda(x,t), s(x,t), S(x,t); x, t) \\ &+ d_{x,(\lambda,s,S)}p(\lambda(x,t), s(x,t), S(x,t); x, t) d_x(\lambda(x,t), s(x,t), S(x,t)) \end{aligned} \tag{4.27}$$

where

$$\begin{aligned} d_{xx}p[\Delta x, \Delta x] &= (d_x A_\phi(x)[\Delta x]\lambda)^T d^2\phi(A_\phi(x)\lambda) d_x A_\phi(x)[\Delta x]\lambda \\ &+ d\phi(A_\phi(x)\lambda) d_{xx} A_\phi(x)[\Delta x, \Delta x]\lambda \\ &+ \frac{1}{t}\left\|d_x A_E(x)[\Delta x]\lambda\right\|^2 + \frac{1}{t}(A_E(x)\lambda)^T d_{xx} A_E(x)[\Delta x, \Delta x]\lambda \\ &+ \frac{1}{t}\left\|d_x A_I(x)[\Delta x]\lambda\right\|^2 + \frac{1}{t}(A_I(x)\lambda)^T d_{xx} A_I(x)[\Delta x, \Delta x]\lambda \\ &+ \frac{1}{t}\left\|d_x A_M(x)[\Delta x]\lambda\right\|^2 + \frac{1}{t}\left\langle A_M(x)\lambda,\ d_{xx} A_M(x)[\Delta x, \Delta x]\lambda\right\rangle \\ &- \frac{1}{t}s^T d_{xx} A_I(x)[\Delta x, \Delta x]\lambda - \frac{1}{t}\left\langle S,\ d_{xx} A_M(x)[\Delta x, \Delta x]\lambda\right\rangle \end{aligned} \tag{4.28}$$

$$\begin{aligned} d_{\lambda x}p[\Delta\lambda, \Delta x] &= (A_\phi(x)\Delta\lambda)^T d^2\phi(A_\phi(x)\lambda) d_x A_\phi(x)[\Delta x]\lambda \\ &+ d\phi(A_\phi(x)\lambda) d_x A_\phi(x)[\Delta x]\Delta\lambda \\ &+ \frac{1}{t}\left[(A_E(x)\lambda)^T d_x A_E(x)[\Delta x]\Delta\lambda + (d_x A_E(x)[\Delta x]\lambda)^T A_E(x)\Delta\lambda\right] \\ &+ \frac{1}{t}\left[(A_I(x)\lambda)^T d_x A_I(x)[\Delta x]\Delta\lambda + (d_x A_I(x)[\Delta x]\lambda)^T A_I(x)\Delta\lambda\right] \\ &+ \frac{1}{t}\left[\langle A_M(x)\lambda,\ d_x A_M(x)[\Delta x]\Delta\lambda\rangle + \langle d_x A_M(x)[\Delta x]\lambda,\ A_M(x)\Delta\lambda\rangle\right] \\ &- \frac{1}{t} s^T d_x A_I(x)[\Delta x]\Delta\lambda - \frac{1}{t}\langle S,\ d_x A_M(x)[\Delta x]\Delta\lambda\rangle \end{aligned} \tag{4.29}$$

$$d_{sx}p[\Delta s, \Delta x] = -\frac{1}{t}\Delta s^T d_x A_I(x)[\Delta x]\lambda \tag{4.30}$$

$$d_{Sx}p[\Delta S, \Delta x] = -\frac{1}{t}\langle \Delta S,\ d_x A_M(x)[\Delta x]\lambda\rangle \tag{4.31}$$

and $d_x(\lambda(x,t), s(x,t), S(x,t))$ *is given by solving*

$$\begin{bmatrix} d_{\lambda\lambda}p & d_{\lambda s}p & d_{\lambda S}p & -e \\ d_{s\lambda}p & d_{ss}p & d_{sS}p & 0 \\ d_{S\lambda}p & d_{Ss}p & d_{SS}p & 0 \\ e^T & 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} d_x\lambda(x,t) \\ d_x s(x,t) \\ d_x S(x,t) \\ d_x\gamma(x,t) \end{bmatrix} = \begin{bmatrix} -d_{\lambda x}p \\ -d_{sx}p \\ -d_{Sx}p \\ 0 \end{bmatrix}, \tag{4.32}$$

where

$$\begin{aligned} d_{\lambda\lambda}p[\Delta\lambda, \Delta\lambda] &= (A_\phi(x)\Delta\lambda)^T d^2\phi(A_\phi(x)\lambda)(A_\phi(x)\Delta\lambda) \\ &+ t\Delta\lambda\,\mathrm{diag}(\lambda)^{-2}\Delta\lambda \\ &+ \frac{1}{t}\|A_E(x)\Delta\lambda\|^2 + \frac{1}{t}\|A_I(x)\Delta\lambda\| + \frac{1}{t}\|A_M(x)\Delta\lambda\| \end{aligned} \tag{4.33}$$

$$d_{ss}p[\Delta s, \Delta s] = t\Delta s^T \operatorname{diag}(s)^{-2}\Delta s + \frac{1}{t}\left\|\Delta s\right\|^2 \tag{4.34}$$

$$d_{SS}p[\Delta S, \Delta S] = t\left\langle S^{-1}(\Delta S),\, S^{-1}(\Delta S)\right\rangle + \frac{1}{t}\left\|\Delta S\right\|^2 \tag{4.35}$$

$$d_{\lambda s}p[\Delta\lambda, \Delta s] = \frac{1}{t}\Delta s^T A_{\scriptscriptstyle I}(x)\Delta\lambda \tag{4.36}$$

$$d_{\lambda S}p[\Delta\lambda, \Delta S] = \frac{1}{t}\left\langle \Delta S,\, A_{\scriptscriptstyle M}(x)\Delta\lambda\right\rangle \tag{4.37}$$

$$d_{sS}p[\Delta s, \Delta S] = 0 \tag{4.38}$$

Proof. Before deriving (4.25)-(4.27), several derivatives of $p(\lambda, s, S; x, t)$ are necessary. The following derivatives are straight-forward, though tedious, to verify:

$$\begin{aligned} d_xp[\Delta x] = {} & d\phi(A_\phi(x)\lambda)d_xA_\phi(x)[\Delta x]\lambda \\ & + \frac{1}{t}\left(A_{\scriptscriptstyle E}(x)\lambda\right)^T d_xA_{\scriptscriptstyle E}(x)[\Delta x]\lambda \\ & + \frac{1}{t}\left(A_{\scriptscriptstyle I}(x)\lambda - s\right)^T d_xA_{\scriptscriptstyle I}(x)[\Delta x]\lambda \\ & + \frac{1}{t}\left\langle A_{\scriptscriptstyle M}(x)\lambda - S,\, d_xA_{\scriptscriptstyle M}(x)[\Delta x]\lambda\right\rangle, \end{aligned} \tag{4.39}$$

$$\begin{aligned} d_\lambda p[\Delta\lambda] = {} & d\phi(A_\phi(x)\lambda)A_\phi(x)\Delta\lambda - t(\lambda^{-1})^T\Delta\lambda \\ & + \frac{1}{t}(A_{\scriptscriptstyle E}(x)\lambda)^T A_{\scriptscriptstyle E}(x)\Delta\lambda \\ & + \frac{1}{t}(A_{\scriptscriptstyle I}(x)\lambda - s)^T A_{\scriptscriptstyle I}(x)\Delta\lambda \\ & + \frac{1}{t}\left\langle A_{\scriptscriptstyle M}(x)\lambda - S,\, A_{\scriptscriptstyle M}(x)\Delta\lambda\right\rangle, \end{aligned} \tag{4.40}$$

$$
\begin{aligned}
d_s p[\Delta s] &= -t(s^{-1})^T \Delta s \\
&\quad - \frac{1}{t}(A_I(x)\lambda - s)^T \Delta s,
\end{aligned}
\tag{4.41}
$$

$$
\begin{aligned}
d_S p[\Delta S] &= -t \left\langle S^{-1},\, \Delta S \right\rangle \\
&\quad - \frac{1}{t} \left\langle A_M(x)\lambda - S,\, \Delta S \right\rangle .
\end{aligned}
\tag{4.42}
$$

Similarly, the derivatives (4.28)-(4.31) and (4.33)-(4.38) can be computed directly.

Theorem 2.43(3) applied to $p(\lambda, s, S; x, t)$ yields (4.25) and (4.27). In the theorem's notation, z is (λ, s, S) and y is x. Furthermore, in the constraint $Az = b$, take $A(\lambda, s, S) = e^T\lambda$ and $b = 1$. In particular, Theorem 2.43(3) gives

$$
d_x G(x,t)[\Delta x] = d_x p(\lambda(x,t), s(x,t), S(x,t); x, t)[\Delta x],
\tag{4.43}
$$

(4.27), and (4.32). Equations (4.39) and (4.43) yield (4.25). The alternative expression (4.26) for $d_x G(x,t)[\Delta x]$ is derived from the first-order conditions in Proposition 4.4.

□

The optimal value function $G(x,t)$ is not differentiable when $t = 0$. To compute the subdifferential will require the subdifferential of the objective of `BendersPrimal`$(x, 0)$ as a function of λ. This is derived in the following proposition.

Proposition 4.20. *If*

$$
p(\lambda; x, 0) = \phi(A_\phi(x)\lambda) + \delta_{\{0\}^{n_E}}(A_E(x)\lambda) + \delta_{\mathbb{R}_-^{n_I}}(A_I(x)\lambda) + \delta_{\mathbb{S}_-^{n_M}}(A_M(x)\lambda) + \delta_{\Delta^{d-1}}(\lambda)
$$

and

1. $A_\phi(x)\lambda \in \operatorname{int}\operatorname{dom}\phi$

2. $p(\lambda; x, 0)$ *is finite at* (λ, x)

3. The pair (λ, x) *is constraint non-degenerate,*

then $p(\lambda; x, 0)$ *is amenable at* (λ, x)*. In particular, this implies* $p(\lambda; x, 0)$ *is subdifferentially regular at* (λ, x) *with*

$$\begin{bmatrix} A_\phi(x)^T v_1 + A_E(x)^T v_2 + A_I(x)^T v_3 + A_M(x)^T v_4 + v_5 \\ d_x(A_\phi(x)\lambda)^T v_1 + d_x(A_E(x)\lambda)^T v_2 + d_x(A_I(x)\lambda)^T v_3 + d_x(A_M(x)\lambda)^T v_4 \end{bmatrix} \in \partial h(\lambda, x)$$

and

$$\begin{bmatrix} A_E(x)^T v_2 + A_I(x)^T v_3 + A_M(x)^T v_4 + v_5 \\ d_x(A_E(x)\lambda)^T v_2 + d_x(A_I(x)\lambda)^T v_3 + \langle d_x(A_M(x)\lambda),\, v_4\rangle \end{bmatrix} \in \partial^\infty h(\lambda, x)$$

for all

$$\begin{aligned} v_1 &\in \partial\phi(A_\phi(x)\lambda) \\ v_2 &\in N_{\{0\}^{n_E}}(A_E(x)\lambda) \\ v_3 &\in N_{\mathbb{R}_-^{n_I}}(A_I(x)\lambda) \\ v_4 &\in N_{\mathbb{S}_-^{n_M}}(A_M(x)\lambda) \\ v_5 &\in N_{\Delta^{d-1}}(\lambda). \end{aligned}$$

Finally, if $(0, y) \in \partial^\infty h(\lambda, x)$*, then* $y = 0$*.*

Proof. We apply [55, Definition 10.23] to show $p(\lambda; x, 0)$ is amenable as a function of (λ, x). The function $p(\lambda; x, 0)$ can be written as $g \circ F$ with $g : \mathbb{F} \times \mathbb{R}^d \to \overline{\mathbb{R}}$ proper lsc convex and

$F : \Delta^{d-1} \times \Omega^d \to \mathbb{F} \times \mathbb{R}^d$ smooth by setting

$$g(a_1, a_2, a_3, a_4, a_5) = \phi(a_1) + \delta_{\{0\}^{n_E}}(a_2) + \delta_{\mathbb{R}_-^{n_I}}(a_3) + \delta_{\mathbb{S}_-^{n_M}}(a_4) + \delta_{\Delta^{d-1}}(a_5), \tag{4.44}$$

$$F(\lambda, x) = \begin{bmatrix} A_\phi(x)\lambda \\ A_E(x)\lambda \\ A_I(x)\lambda \\ A_M(x)\lambda \\ \lambda \end{bmatrix}. \tag{4.45}$$

All that remains is to show if

$$\begin{aligned} v &= (v_1, v_2, v_3, v_4, v_5) \\ &\in N_{\operatorname{dom}\phi}(A_\phi(x)\lambda) \times N_{\{0\}^{n_E}}(A_E(x)\lambda) \times N_{\mathbb{R}_-^{n_I}}(A_I(x)\lambda) \times N_{\mathbb{S}_-^{n_M}}(A_M(x)\lambda) \times N_{\Delta^{d-1}}(\lambda) \end{aligned}$$

and $dF(x, \lambda)^T v = 0$, then $v = 0$.

We immediately have $v_1 = 0$ since $A_\phi(x)\lambda \in \operatorname{int}\operatorname{dom}\phi$. Since

$$N_{\Delta^{d-1}}(\lambda) = \mathbb{R}e + N_{\mathbb{R}_+^d}(\lambda),$$

there exists $\alpha \in \mathbb{R}$ and $v_5' \in N_{\mathbb{R}_-^d}(\lambda)$ such that $v_5 = \alpha e + v_5'$. If $dF(x, \lambda)^T v = 0$ then $d_\lambda F(x, \lambda)^T v = 0$. So

$$\begin{aligned} 0 &= \lambda^T d_\lambda F(x, \lambda)^T v \\ &= \lambda^T A_\phi(x)^T v_2 + \lambda^T A_I(x)^T v_3 + \lambda^T A_M(x)^T v_4 + \alpha\lambda^T e + \lambda^T v_5' \\ &= \alpha, \end{aligned}$$

using that $a^T b = 0$ if $a \in N_K(b)$ for $K = \{0\}^{n_E}$, $\mathbb{R}_-^{n_I}$, $\mathbb{S}_-^{n_M}$ [12, Exercise 5.11]. Since $\lambda \in \Delta^{d-1}$, there exist indices $B \subset \{1, \ldots, d\}$ such that $\lambda_i > 0$ for all $i \in B$. Then $(v_5')_i = 0$

for all $i \in B$. So

$$[d_\lambda F(x,\lambda)^T v]_i = [A_\phi(x)^T v_2 + A_I(x)^T v_3 + A_M(x)^T v_4]_i \quad \text{for all } i \in B. \tag{4.46}$$

Then non-degeneracy on constraints implies $v_2,\ v_3,\ v_4 = 0$, so $v = 0$. Thus $p(\lambda; x, t)$ is amenable.

Amenable functions are subdifferentially regular by [55, Exercise 10.25(a)]. So the chain rule, [55, Theorem 10.6], applies to amenable functions, which yields the formulas for the subdifferential and horizon subdifferential given. The previous paragraph proves the only element in $\partial^\infty h(\lambda, x)$ whose first component vanishes is $(0,0)$. □

We can combine the previous proposition with the perturbation theory of optimal value functions to compute the subdifferential of $G(x) = G(x, 0)$.

Proposition 4.21. *For any $\overline{x} \in \operatorname{dom} G$,*

$$\partial G(\overline{x}) \subset \mathcal{A}(\overline{x})$$

where

$$\begin{aligned}\mathcal{A}(\overline{x}) := \Big\{ -d_x(A_\phi(x)\overline{\lambda})^T \overline{w} - d_x(A_E(x)\overline{\lambda})^T \overline{w_E} - d_x(A_I(x)\overline{\lambda})^T \overline{w_I} - \big\langle d_x(A_M(x)\overline{\lambda}),\, \overline{w_M} \big\rangle \\ \Big|\ \overline{\lambda} \in \operatorname{argmin}_{\lambda \in \Delta^{d-1}} h(\lambda, \overline{x}),\ (\overline{w}, \overline{w_E}, \overline{w_I}, \overline{w_M}) \textit{ dual optimal} \Big\}.\end{aligned} \tag{4.47}$$

If $\overline{x}$ is constraint non-degenerate, then $\partial^\infty G(\overline{x}) = \{0\}$ and $\partial G(\overline{x}) \neq \emptyset$.

If $\overline{x}$ is both objective non-degenerate and constraint non-degenerate, then for each $\beta \in$ $\mathrm{Distinct}(\overline{x})$ and $v^1, v^2 \in \mathcal{A}(\overline{x})$,

$$\sum_{i \in \mathcal{I}(\overline{x},\beta)} v_i^1 = \sum_{i \in \mathcal{I}(\overline{x},\beta)} v_i^2.$$

Also, if $v \in \mathcal{A}(\overline{x})$, then

$$\left[\frac{1}{|\mathcal{I}(x,x_j)|}\left(\sum_{i\in\mathcal{I}(x,x_j)} v_i\right)\right]_{j=1,\dots,d} \in \mathcal{A}(\overline{x}),$$

and this also is true when $\mathcal{A}(\overline{x})$ is replaced with $\widehat{\partial}G(\overline{x})$ or $\overline{\partial}G(\overline{x})$.

Proof. The derivative formula an immediate consequence of [55, Theorem 10.13] combined with the KKT conditions in Proposition 4.4. The assumptions of [55, Theorem 10.13] are satisfied because $\lambda \in \Delta^{d-1}$, and hence is bounded for all $x \in \operatorname{dom} G$.

Similarly, [55, Theorem 10.13] also implies

$$\partial^\infty G(\overline{x}) \subset \left\{ y \middle| \, (0,y) \in \partial^\infty h(\overline{\lambda},\overline{x}), \quad \overline{\lambda} \in \operatorname{argmin}_{\lambda\in\Delta^{d-1}} h(\lambda,\overline{x}) \right\}.$$

If $\overline{x}$ is constraint non-degenerate, then Proposition 4.20 implies $\partial^\infty h(\overline{\lambda},\overline{x}) = \{(0,0)\}$, so $\partial^\infty G(\overline{x}) = \{0\}$. Then [55, Corollary 8.10] implies $\partial G(\overline{x}) \neq \emptyset$.

Note that if $v \in \mathcal{A}(\overline{x})$ and $\beta \in \text{Distinct}(\overline{x})$, then

$$\sum_{i\in\mathcal{I}(\overline{x},\beta)} v_j = -\lambda_{\mathcal{I}(\overline{x},\beta)} \left(d_\beta f_\phi(\beta)^T\overline{w} + d_\beta f_E(\beta)^T\overline{w_E} + d_\beta f_I(\beta)^T\overline{wi} + \langle d_\beta f_M(\beta), \, \overline{w_M}\rangle\right)$$

If $\overline{x}$ is both objective non-degenerate and constraint non-degenerate, then Theorem 4.7 implies $\lambda_{\mathcal{I}(\overline{x},\beta)}$ and the dual solutions $(\overline{w},\overline{w_E},\overline{w_I},\overline{w_M})$ are unique, hence the above sum is the same for all $v \in \mathcal{A}(\overline{x})$.

Finally, note that elements of $\mathcal{A}(\overline{x})$ are invariant under permutations of indices $i \in \lambda_{\mathcal{I}(\overline{x},\beta)}$ with $\beta \in \text{Distinct}(\overline{x})$. If $v \in \mathcal{A}(\overline{x})$, then

$$\left[\frac{1}{|\mathcal{I}(x,x_j)|}\left(\sum_{i\in\mathcal{I}(x,x_j)} v_i\right)\right]_{j=1,\dots,d} \in \mathcal{A}(\overline{x}),$$

is simply the average of v over all such permutations. It is also an element of $\mathcal{A}(\overline{x})$ by convexity. This same argument applies if $v \in \widehat{\partial}G(\overline{x})$ or $v \in \overline{\partial}G(\overline{x})$. □

The next theorem shows that a limit of stationary points in x for $G(x,t)$ as $t \searrow 0$ is Clarke stationary for $G(x,0)$. This shows that any limit point $\overline{x}$ of a numerical method which finds a local optimum x^ν to $G(x,t)$ for a sequence of $t^\nu \searrow 0$ will give a Clarke stationary point for `BendersPrimal`$(\overline{x})$. Note that if $G(x,t)$ was convex over $x \in \Omega$, then Theorem 4.10 and Attouch's Theorem [55, Theorem 12.35] would give the desired result. However, Attouch's Theorem does not apply as $G(x,t)$ is not convex in x.

Theorem 4.22. *Suppose* $(x^\nu, t^\nu) \to (\overline{x}, 0)$ *and*

$$0 \in \nabla_x G(x^\nu, t^\nu) + N_{\Omega^d}(x^\nu) \quad \forall \nu.$$

Then

1. *With* $\mathcal{A}(x)$ *as in Proposition 4.21,*

$$0 \in \mathcal{A}(\overline{x}) + N_{\Omega^d}(\overline{x}).$$

2. *If, in addition,* $\overline{x}$ *is objective non-degenerate and constraint non-degenerate, then*

$$0 \in \overline{\partial}G(\overline{x}) + N_{\Omega^d}(\overline{x}).$$

3. *If, in addition to the previous assumptions,* $\widehat{\partial}G(\overline{x}) \neq \emptyset$, *then*

$$0 \in \widehat{\partial}G(\overline{x}) + N_{\Omega^d}(\overline{x}).$$

Proof. We prove the claims in the order (1), (3), (2), since the proof of (3) is a simpler version of the proof of (2).

(1) By Proposition 4.19,

$$\begin{aligned} d_x G(x,t)\Delta x = &-w(x,t)^T d_x A_\phi(x)[\Delta x]\lambda(x,t) \\ &- w_E(x,t)^T d_x A_E(x)[\Delta x]\lambda(x,t) \\ &- w_I(x,t)^T d_x A_I(x)[\Delta x]\lambda(x,t) \\ &- \langle w_M(x,t),\, d_x A_M(x)[\Delta x]\lambda(x,t)\rangle\,. \end{aligned}$$

Theorem 4.17(2) implies there exists a subsequence $J \subset \mathbb{N}$ such that

$$\begin{aligned} \lim_{\nu\to\infty,\ \nu\in J} d_x G(x^\nu,t^\nu)\Delta x = &-w^T d_x A_\phi(\overline{x})[\Delta x]\lambda \\ &- w_E^T d_x A_E(\overline{x})[\Delta x]\lambda \\ &- w_I^T d_x A_I(\overline{x})[\Delta x]\lambda \\ &- \langle w_M,\, d_x A_M(\overline{x})[\Delta x]\lambda\rangle \end{aligned}$$

where λ is a solution to `BendersPrimal`$(\overline{x})$ and (w, w_E, w_I, w_M) are dual solutions to `BendersDual`$(\overline{x})$. That expression is the same as for elements of $\mathcal{A}(\overline{x})$ in Proposition 4.21, so

$$\lim_{\nu\to\infty,\ \nu\in J} \nabla_x G(x^\nu,t^\nu) \in \mathcal{A}(\overline{x}).$$

The normal cone $N_{\Omega^d}(x)$ is outer semi-continuous as a function of x, so $-\nabla_x G(x^\nu,t^\nu) \in N_{\Omega^d}(x^\nu)$ for every ν implies

$$-\lim_{\nu\to\infty,\ \nu\in J} \nabla_x G(x^\nu,t^\nu) \in N_{\Omega^d}(\overline{x}).$$

So

$$0 \in \mathcal{A}(\overline{x}) + N_{\Omega^d}(\overline{x}).$$

(3) The general strategy will be to combine the components of x^ν with the same limit, since then the corresponding sum of weights $\lambda_{\mathcal{I}(x^\nu,\beta)}$ will have a meaningful limit. Then

Proposition 4.21 can be used to unpack this limit into components in Ω^d.

To make the above more concrete, note for every $\overline{\beta} \in \text{Distinct}(\overline{x})$,

$$
\begin{aligned}
& w(x,t)^T \left(\sum_{j \in \mathcal{I}(\overline{x},\overline{\beta})} d_x(A_\phi(x))_{:,j}[\Delta x] \lambda(x,t)_j \right) \\
=& w(x,t)^T \left(\sum_{j \in \mathcal{I}(\overline{x},\overline{\beta})} d_\beta f(\beta_j)[\Delta \beta_j] \lambda(x,t)_j \right) \\
=& w(x,t)^T d_\beta f(\overline{\beta}) \lambda_{\mathcal{I}(\overline{x},\overline{\beta})}(x,t) + w(x,t)^T \left[\sum_{j \in \mathcal{I}(\overline{x},\overline{\beta})} \left(d_\beta f(\beta_j)[\Delta \beta_j] - d_\beta f(\overline{\beta})[\Delta \beta_j] \right) \lambda(x,t)_j \right]
\end{aligned}
$$

The second term in the final expression evaluated at (x^ν, t^ν) goes to zero as $\nu \to \infty$ because $\{\lambda(x^\nu, t^\nu)\}$ and $\{w(x^\nu, t^\nu)\}$ are bounded while $d_\beta f(\beta_j^\nu) \to d_\beta f(\overline{\beta})$. Then continuity of $w(x,t)$ and $\lambda_{\mathcal{I}(\overline{x},\overline{\beta})}(x,t)$ at $(\overline{x}, 0)$ gives

$$
w(x^\nu, t^\nu)^T \left(\sum_{j \in \mathcal{I}(\overline{x},\overline{\beta})} d_x(A_\phi(x^\nu))_{:,j}[\Delta x] \lambda_j(x^\nu, t^\nu) \right) \to w(\overline{x}, 0)^T d_\beta f(\overline{\beta}) \lambda_{\mathcal{I}(\overline{x},\overline{\beta})}(\overline{x}, 0)
$$

as $\nu \to \infty$.

The same reasoning can be applied when A_ϕ and w are replaced by A_E and w_E or A_I and w_I or A_M and w_M. Doing so gives that for each $\overline{\beta} \in \text{Distinct}(\overline{x})$ the sequence $\{\sum_{j \in \mathcal{I}(\overline{x},\overline{\beta})} [\nabla_x G(x^\nu, t^\nu)]_j\}$ has a limit, namely

$$
\begin{aligned}
\sum_{j \in \mathcal{I}(\overline{x},\overline{\beta})} (\nabla_x G(x^\nu, t^\nu))_j \to & \lambda_{\mathcal{I}(\overline{x},\overline{\beta})}(\overline{x}, 0) d_\beta f_\phi(\beta_s^\infty)^T w(\overline{x}, 0) \\
& + \lambda_{\mathcal{I}(\overline{x},\overline{\beta})}(\overline{x}, 0) d_\beta f_E(\beta_s^\infty)^T w_E(\overline{x}, 0) \\
& + \lambda_{\mathcal{I}(\overline{x},\overline{\beta})}(\overline{x}, 0) d_\beta f_I(\beta_s^\infty)^T w_I(\overline{x}, 0) \\
& + \lambda_{\mathcal{I}(\overline{x},\overline{\beta})}(\overline{x}, 0) d_\beta f_M(\beta_s^\infty)^T w_M(\overline{x}, 0).
\end{aligned}
$$

If $v \in \widehat{\partial} G(\overline{x}, 0)$, then comparing the above expression with the expression in Proposition

4.21 shows

$$\lim_{\nu\to\infty} \sum_{j\in\mathcal{I}(\overline{x},\overline{\beta})} (\nabla_x G(x^\nu, t^\nu))_j = \sum_{j\in\mathcal{I}(\overline{x},\overline{\beta})} v_j$$

and, furthermore,

$$\left[\frac{1}{|\mathcal{I}(\overline{x},\overline{x}_j)|}\left(\sum_{i\in\mathcal{I}(\overline{x},\overline{x}_j)} v_i\right)\right]_{j=1,\dots,d} \in \widehat{\partial} G(\overline{x}, 0).$$

On the other hand, $-\nabla_x G(x^\nu, t^\nu) \in N_{\Omega^d}(x^\nu)$ for each each ν. So the limit from the previous paragraph implies

$$-\lim_{\nu\to\infty} \sum_{j\in\mathcal{I}(\overline{x},\overline{\beta})} (\nabla_x G(x^\nu, t^\nu))_j \in N_\Omega(\overline{\beta}),$$

where the inclusion follows because the normal cone operator is osc. In particular, this shows

$$-\frac{1}{|\mathcal{I}(x, x_j)|}\left(\sum_{i\in\mathcal{I}(x,x_j)} v_i\right) \in N_\Omega(\overline{x}_j)$$

for each $j \in \{1,\dots,d\}$, so $0 \in \widehat{\partial} G(\overline{x}) + N_{\Omega^d}(\overline{x})$.

(2) This can be proven using nearly the exact same proof as for (3), with the following modification. By assumption $\overline{x}$ is constraint non-degenerate, so Proposition 4.47 implies $\partial^\infty G(\overline{x}) = \{0\}$. Then [55, Theorem 8.49] implies $\overline{\partial} G(\overline{x}, 0) = \mathrm{co}\,(\partial f(\overline{x}))$. So if $v \in \overline{\partial} G(\overline{x}, 0)$, then by Caratheodory's Theorem there exist $dp+1$ vectors $v^\ell \in \partial G(\overline{x}, 0)$, where $\ell \in \{1,\dots,dp+1\}$, and $\alpha \in \mathbb{R}^{dp+1}$ with $\alpha \geq 0$ and $\sum_i \alpha_i = 1$, such that

$$v = \sum_{\ell=1}^{dp+1} \alpha_\ell v^\ell.$$

As in (3),

$$\sum_{j\in\mathcal{I}(\overline{x},\overline{\beta})} (\nabla_x G(x^\nu, t^\nu))_j \to \sum_{j\in\mathcal{I}(\overline{x},\overline{\beta})} v_j^\ell,$$

so

$$\begin{aligned}\sum_{j\in\mathcal{I}(\overline{x},\overline{\beta})} v_j &= \sum_{j\in\mathcal{I}(\overline{x},\overline{\beta})}\left(\sum_{\ell=1}^{dp+1}\alpha_\ell v_j^\ell\right)\\ &= \sum_{\ell=1}^{dp+1}\alpha_\ell\left(\sum_{j\in\mathcal{I}(\overline{x},\overline{\beta})} v_j^\ell\right)\\ &= \sum_{\ell=1}^{dp+1}\alpha_\ell\left(\lim_{\nu\to\infty}\sum_{j\in\mathcal{I}(\overline{x},\overline{\beta})}\left(\nabla_x G(x^\nu,t^\nu)\right)_j\right)\\ &= \lim_{\nu\to\infty}\sum_{j\in\mathcal{I}(\overline{x},\overline{\beta})}\left(\nabla_x G(x^\nu,t^\nu)\right)_j .\end{aligned}$$

The above argument shows $\lim_{\nu\to\infty}\sum_{j\in\mathcal{I}(\overline{x},\overline{\beta})}\left(\nabla_x G(x^\nu,t^\nu)\right)_j = \sum_{j\in\mathcal{I}(\overline{x},\overline{\beta})} v_j$. Otherwise the argument for (2) is the same as for (3).

□

4.6 Conclusion

This chapter presented a homotopy method using Benders' Decomposition for obtaining optimal solutions to `BendersPrimal`, a specialized case of `F-Primal`. The convergence properties of iterates of this algorithm were investigated along with the optimality properties of any limit points of the iterates. Under regularity conditions it was shown the primal-dual solutions of the relaxed problems are bounded as the homotopy parameter decreases to 0, any primal limit point is a solution, and any limit of first-order critical points is itself a Clarke first-order critical point.

There are several extensions of the material that can be investigated in future work. In practice, only approximate critical points of $G(x,t)$ will be found. For such cases, Theorem 4.22 should be extended to use approximate critical points. Additionally, Theorem 4.22 examined the limit of critical points of $G(x,t)$, but not those of the related problem $D(x,t)$. Extending the analysis of the limiting behavior of critical points of $D(x,t)$ as $t \searrow 0$ is another open question.

Chapter 5

NUMERICAL EXPERIMENTS

5.1 *Introduction*

This chapter presents numerical results of the algorithm presented in Chapter 4 applied to the non-parametric maximum likelihood (`NPMLE`) and optimal experimental design (`OptD`) problems. The experiments were run on Ubuntu 14.04 with a Core i5-2500K 3.3 GHz processor. All code was written using Python 3.4 and the packages numpy 1.9.1 [59], scipy 0.15.0 [33], cvxopt 1.1.7 [1], and scikit-learn 0.15.2 [49]. Numpy, scipy, and scikit-learn used the Intel Matrix Kernel Library (MKL) 11.1 for matrix operations.

5.2 *Algorithms*

5.2.1 *Algorithm for* `BendersPrimal`(x)

The results in Chapter 4 suggest the following general algorithm for finding an approximate optimal solution to `BendersPrimal`(x).

Require: $x \in \Omega^d$, $t > 0$, $\alpha \in (0, 1)$

repeat

$x \leftarrow \operatorname{argmin}_{x \in \Omega^d} G(x, t)$

$t \leftarrow \alpha t$

until Converged

The test for convergence is left unspecified. It could be, for example, measuring a sufficiently small change in optimal solution $\operatorname{argmin}_x G(x, t)$ or decrease in the the optimal value of t. Unfortunately, the theory in Chapter 4 does not give guidance on good choices for an initial t or when t is sufficiently small. Determining a good initial value of $x \in \Omega^d$ is likewise problem dependent.

The optimization $\min_{x \in \Omega^d} G(x,t)$ can be done with standard black-box methods. In the examples below, box-constrained Newton and L-BFGS-B methods were used [46, 64]. The L-BFGS-B implementation is from [64] and is called from the scipy.optimize package. The box-constrained Newton algorithm modifies the hessian matrix by adding adding a sufficiently large non-negative number to the diagonal to ensure the resulting matrix is positive-definite with condition number less than 10^3. This number is determined by computing the minimum and maximum eigenvalues of the hessian using scipy. If the pure Newton step is feasible, then an Armijo line search is performed in that direction. Otherwise, the direction of descent is found by solving a bound-constrained quadratic program using cvxopt, and the Armijo line search is used on that direction. The algorithm is halted when any of

$$\frac{-d^T \nabla_x G(x,t)}{2\max(1, G(x^{\text{prev}},t))}, \ \|\nabla_x G(x,t)\|_\infty, \ \left|\frac{G(x^{\text{curr}},t) - G(x^{\text{prev}},t)}{\max(1, G(x^{\text{prev}},t))}\right|, \quad \text{or} \ \frac{\|x^{\text{curr}} - x^{\text{prev}}\|_\infty}{\max(\|x^{\text{prev}}\|_\infty, 1)}$$

are less than 10^{-8}, where d is the descent direction.

5.2.2 EM Algorithm

The EM algorithm [19] is a framework for creating likelihood maximizing algorithms for statistical models with missing data. The EM algorithm alternates between *expectation* steps, which take the expected valued of the log-likelihood given estimated parameters, and *maximization* steps, which maximize the expected log-likelihood. The expectation step takes a particularly simple form for mixture models with a fixed number of components. Using the notation of Section 3.3.1, if there are D observations y_i, K components, and indicator variables z_{ik} for whether the , then the log-likelihood for the observations y_i given the unobserved component indicators z_{ik} is

$$\begin{aligned}\ell(\lambda, x, z) &= \log\left(\prod_{i=1}^{D}\prod_{k=1}^{K}\left(\lambda_k p(y_i|z_{ik}=1,\beta_k)\right)^{z_{ik}}\right)\\ &= \sum_{i=1}^{D}\sum_{k=1}^{K} z_{ik}\left(\log\lambda_k + \log\left(p(y_i|z_{ik}=1,\beta_k)\right)\right).\end{aligned}$$

The expected log-likelihood for previously estimated parameters λ^{prev} and x^{prev} is

$$E\left[\ell(\lambda, x, z)|y, \lambda^{\text{prev}}, x^{\text{prev}}\right] = \sum_{i=1}^{N}\sum_{k=1}^{K} E[z_{ik}|y_i, \lambda^{\text{prev}}, x^{\text{prev}}]\left(\log\lambda_k + \log\left(p(y_i|z_{ik}=1,\beta_k)\right)\right),$$

where the expected value $E[z_{ik}|y_i, \lambda, \beta]$ is given by

$$E[z_{ik}|y_i, \lambda, x] = \frac{\lambda_k p(y_i|z_{ik}=1,\beta_k)}{\sum_{j=1}^{K}\lambda_j p(y_i|z_{ij}=1,\ \beta_j)}.$$

With no further constraints on the mixture model, the optimization of $E\left[\ell(\lambda, x, z)|y, \lambda^{\text{prev}}, x^{\text{prev}}\right]$ can be done over λ and x separately. To simplify notation, define $\rho_{ik} := E[z_{ik}|y_i, \lambda, x]$. A straight-forward computation shows the optimal weights λ for $E\left[\ell(\lambda, x, z)|y, \lambda^{\text{prev}}, x^{\text{prev}}\right]$ such that $\lambda \in \Delta^{K-1}$ are

$$\lambda_j = \frac{\sum_{i=1}^{N}\rho_{ij}}{\sum_{i=1}^{N}\sum_{k=1}^{K}\rho_{ik}}.$$

In several common models, such as Poisson and Gaussian mixture models, there is a closed form solution for $\max_{x\in\Omega^K}\sum_{i=1}^{N}\sum_{k=1}^{K}\rho_{ik}\log\left(p(y_i|z_{ik}=1,\beta_k)\right)$. These closed forms are given in later sections. Note that these closed form solutions do not exist in the presence of additional constraints on the mixture model.

5.2.3 Multiplicative Algorithm

The multiplicative algorithm [41] finds weights for given matrices in optimal design problems. Specifically, given matrices $A_1, \ldots, A_d \in \mathbb{S}_+^n$ and a differentiable convex loss function Φ :

$\mathbb{S}_+^n \to \mathbb{R}$, it solves

$$\inf_{\lambda \in \Delta^{d-1}} \Phi\left(\sum_{i=1}^{d} \lambda_i A_i\right)$$

via multiplicative updates

$$\lambda_i^{\text{next}} \propto -\operatorname{tr}\left(d\Phi\left(\sum_{j=1}^{d} \lambda_j^{\text{prev}} A_j\right) A_i\right) \lambda_i^{\text{prev}}$$

to give the next iterate λ^{next} from the previous iterate λ^{prev} That is, each component λ_i^{prev} is multiplied by $-\operatorname{tr}\left(d\Phi\left(\sum_{j=1}^{d} \lambda_j^{\text{prev}} A_j\right) A_i\right)$, and the resulting vector is re-normalized to sum to one, i.e.

$$\lambda_i^{\text{next}} = \frac{\operatorname{tr}\left(d\Phi\left(\sum_{j=1}^{d} \lambda_j^{\text{prev}} A_j\right) A_i\right)}{\sum_{k=1}^{d} \operatorname{tr}\left(d\Phi\left(\sum_{j=1}^{d} \lambda_j^{\text{prev}} A_j\right) A_k\right) \lambda_k^{\text{prev}}} \lambda_i^{\text{prev}}.$$

5.3 Non-Parametric Maximum Likelihood

5.3.1 Mixture of Poisson Distributions

The Benders decomposition algorithm was applied to a mixture of Poisson distributions for the non-parametric maximum likelihood problem `NPMLE`. The Poisson distribution $\text{Poiss}(\beta)$, where $\beta \in \mathbb{R}_{++}$, is a distribution on the non-negative integers with probability mass function

$$p(y|\beta) = \frac{\beta^y}{y!} e^{-\beta}.$$

Using this notation, the maximization step in the EM algorithm over $x = (\beta_1, \dots, \beta_K) \in \Omega^K$ is

$$\beta_k = \frac{\sum_{i=1}^{N} \rho_{ik} y_i}{\sum_{i=1}^{N} \rho_{ik}}, \quad k = 1, \dots, K.$$

This has the simple interpretation of β_k being updated as the weighted mean of the data points y_i according to the weights $\rho_{ik} = E[z_{ik}|y_i, \lambda, x]$. The weights ρ_{ik} can be interpreted as the probability observation i belongs to component k based on the previous iteration's components.

A sample of size 200 was drawn from the mixture distribution

$$0.1\text{Poiss}(2) + 0.15\text{Poiss}(10) + 0.05\text{Poiss}(12) + 0.2\text{Poiss}(15) + 0.25\text{Poiss}(37) + 0.05\text{Poiss}(41).$$

Three different algorithms were used to fit the mixture likelihood: Benders decomposition using L-BFGS-B to optimize $G(x,t)$ for fixed t, Benders decomposition using a bound-constrained Newton method to optimize $G(x,t)$ for fixed t, and the EM algorithm to optimize the likelihood for a fixed number of mixture components. For the Benders decomposition, $\Omega = [0.5, 50]$ and penalty parameters t in the Benders algorithms were $t = 1$, 10^{-1}, 10^{-2}, 10^{-3}, 10^{-4}, 10^{-5}, 5×10^{-6} with associated stopping tolerances tol $= 10^{-9}$, 10^{-9}, 10^{-9}, 10^{-10}, 10^{-10}, 10^{-10}. The penalty values were truncated at 5×10^{-6} because the optimization frequently failed at $t = 10^{-6}$. The Benders algorithm were initialized with a mesh grid of points between 0.5 and 50 that were spaced 0.5, 1.5, and 4 apart. The EM algorithm with N components was initialized to be the middle N points of an N+2 evenly spaced mesh between the smallest and largest observed values. No convergence tests were run on the EM algorithm, and it was stopped after 400 iterations.

As the results in the table show, for the unconstrained Poisson mixture measure problem the Benders decomposition gives results with similar negative log-likelihood and marginal means and variances. The computational cost is, depending on the optimization method used and the number of support points, between 2× and 10× as expensive computationally.

5.3.2 Mixture of Normal Distributions

The Benders decomposition algorithm and the EM algorithm were also applied to a mixture of Gaussian distributions for the non-parametric maximum likelihood problem `NPMLE`. The Gaussian distribution $n(m, \Sigma)$ is a distribution on $\mathbb{R}^n$ with probability density function

$$p(y|m, \Sigma) = \frac{1}{\det(2\pi\Sigma)} \exp\left(-\frac{1}{2}(y-m)^T\Sigma^{-1}(y-m)\right).$$

Method	cpu time (s)	objective value	mean	variance
Benders-LBFGSB-99	3.82	3.8500	7e-5	0.0067
Benders-LBFGSB-33	1.75	3.8521	8e-5	0.0066
Benders-LBFGSB-13	0.79	3.8500	3e-5	0.0069
Benders-Newton-99	1.35	3.8500	4e-7	0.0070
Benders-Newton-33	0.55	3.8500	3e-6	0.0069
Benders-Newton-13	0.38	3.8500	6e-7	0.0070
EM-4	0.16	3.8501	<1e-15	0.0061
EM-5	0.19	3.8501	<1e-15	0.0061
EM-6	0.22	3.8501	<1e-15	0.0061
EM-7	0.26	3.8500	<1e-15	0.0070
EM-8	0.29	3.8500	<1e-15	0.0071

Table 5.1: Algorithm Comparison for Poisson Mixture Model. Benders-LFBGSB-N is the Benders decomposition algorithm on an N-point mesh using L-BFGS-B to optimize $G(x,t)$. Benders-Newton-N is the same as Benders-LBFGSB-N, except a bound constraint Newton metod is used to optimize $G(x,t)$. EM-N fits the sample using the EM algorithm with N components. The column "mean" gives the relative error between the mean of the fitted distribution and the sample, and likewise for "variance".

Similar to the Poisson distribution, the maximization step in the EM algorithm over $x = (\beta_1, \dots, \beta_K) \in \Omega^K$ where $\beta_k = (m_k, \Sigma_k)$ gives

$$m_k = \frac{\sum_{i=1}^N \rho_{ik} y_i}{\sum_{i=1}^N \rho_{ik}}, \quad k = 1, \dots, K$$

$$\Sigma_k = \frac{1}{\sum_{i=1}^N \rho_{ik}} \sum_{i=1}^N \rho_{ik}(y_i - m_k)(y_i - m_k)^T, \quad k = 1, \dots, K.$$

Similar to the EM algorithm for mixtures of Poisson distributions, the above formulas have simple interpretations as weighted sample means and sample variances where the weights are the probability that y_i was generated by component k according to the means and variances in the previous iteration.

A sample of size 500 was drawn from the mixture distribution

$$0.7n\left(\begin{bmatrix}1\\2\end{bmatrix},\begin{bmatrix}1&0\\0&1\end{bmatrix}\right)+0.2n\left(\begin{bmatrix}1\\0\end{bmatrix},\begin{bmatrix}1&0.5\\0.5&1\end{bmatrix}\right)+0.1n\left(\begin{bmatrix}0\\3\end{bmatrix},\begin{bmatrix}0.5&-0.25\\-0.25&0.5\end{bmatrix}\right).$$

Three different algorithms were used to fit the mixture likelihood: Benders decomposition using L-BFGS-B to optimize $G(x,t)$ for fixed t, the EM algorithm to optimize the likelihood for a fixed number of mixture components, and a hybrid algorithm that optimizes $G(x,t)$ for fixed t using L-BFGS-B and that is initialized for the first (i.e., largest) t using the output of the EM algorithm. For Benders decomposition, $\Omega \subset \mathbb{R}^5$ uses the parameterization $(m_1, m_2, \ell_1, \ell_2, \ell_3)$ for the normal distribution

$$n\left(\begin{bmatrix}m_1\\m_2\end{bmatrix},\begin{bmatrix}\ell_1&0\\\ell_3&\ell_2\end{bmatrix}\begin{bmatrix}\ell_1&\ell_3\\0&\ell_2\end{bmatrix}\right).$$

That is, the means are parameterization directly while the covariance matrices are parameterized using their Cholesky factorization. Doing so ensures points $\beta \in \Omega$ yield valid parameters for a normal distribution. In particular, the covariance matrices are always positive definite. A lower bound of -10 was placed on the means and off-diagonal entries of the Cholesky factors, and a lower bound of 0.1 was placed on the diagonal entries of the Cholesky factors. An upper bound of 10 was placed on each component of β. The non-hybrid Benders decomposition initialized means by uniformly randomly sampling from the box whose lower bounds were the entry-wise minima over all the data, and similarly whose upper bounds were the entry-wise maxima over all the data. The non-hybrid Benders decomposition covariance matrices were initialized as diagonal matrices with diagonals equal to the entry-wise sample variance of the data. The hybrid Benders decomposition was initialized by using the EM algorithm with parameters as described below, except with a maximum of 100 iterations rather than 400 iterations. The penalty parameters t used for the non-hybrid Benders were $t = 10^{-1}, 10^{-2}, 10^{-3}, 10^{-4}, 10^{-5}$ with associated stopping tolerances of tol $= 10^{-4}, 10^{-5}, 10^{-5}, 10^{-5}, 10^{-5}$. The hybrid Benders algorithm used $t = 10^{-3}, 10^{-4}, 10^{-5}$ with stopping tolerances tol $= 10^{-5}, 10^{-5}, 10^{-5}$. The EM algorithm used was sklearn.mixture.GMM from scikit-learn and was run for up to 400 iterations with unstructured (i.e., 'full') covariance matrices. The default convergence test for the EM algorithm was used, namely the EM algorithm was terminated when the change in the objective in subsequent iterations was less than 10^{-3}.

Compared to the Poisson mixture model problem, the Benders decomposition for the Gaussian mixture model problem takes substantially longer compared to the EM algorithm. There are several potential reasons. One is that evaluation of $G(x, t)$ is more complex due to transformations between the parameterization, a flat vector of means and variances, and extracting the means and covariance matrices to form $A_\phi(x)$. Another is the usage of a Quasi-Newton method instead of a Newton method. The number of function evaluations of $G(x, t)$ becomes much larger as $t \searrow 0$ at the same time the cost of each such evaluation becomes more expensive. For example, $G(x, t)$ typically takes 5 primal-dual iterations in cvxopt for larger t, but closer to 25 primal-dual iterations for smaller t.

5.4 Optimal Experimental Design

The Benders decomposition algorithm described in Chapter 4 was applied to an instance of the optimal design problem `OptD`. The specific instance was determining a D-optimal design for a model as in Example 3.4, with

$$\begin{aligned} \Omega &= [0, 3] \subset \mathbb{R} \\ q(\beta) &= (e^{-\beta}, \beta e^{-\beta}, e^{-2\beta}, \beta e^{-2\beta}) \\ w(\beta) &= 1. \end{aligned}$$

This example is taken from [41].

The Benders decomposition algorithm is compared against the multiplicative algorithm for D-Optimality. The Benders decomposition implementation used L-BFGS-B to optimize

Method	cpu time (s)	objective value	objective value stderr
Benders-LBFGSB-2	0.70	1.5181	9e-3
Benders-LBFGSB-3	1.15	1.5104	3e-3
Benders-LBFGSB-4	1.17	1.5050	2e-3
Benders-LBFGSB-5	1.45	1.5013	2e-3
Benders-LBFGSB-6	1.64	1.4959	4e-3
Benders-LBFGSB-7	2.44	1.4905	4e-3
Benders-LBFGSB-8	2.45	1.4845	4e-3
Benders-EM-LBFGSB-2	0.25	1.5146	1e-6
Benders-EM-LBFGSB-3	0.34	1.5123	1e-16
Benders-EM-LBFGSB-4	0.37	1.5062	1e-17
Benders-EM-LBFGSB-5	0.83	1.4980	1e-6
Benders-EM-LBFGSB-6	1.01	1.4951	3e-4
Benders-EM-LBFGSB-7	1.27	1.4944	1e-4
Benders-EM-LBFGSB-8	2.31	1.4821	4e-3
EM-2	0.05	1.5147	1e-6
EM-3	0.07	1.5124	1e-16
EM-4	0.10	1.5062	6e-7
EM-5	0.32	1.4989	1e-6
EM-6	0.37	1.4968	1e-5
EM-7	0.60	1.4946	3e-4
EM-8	0.64	1.4894	1e-3

Table 5.2: Algorithm Comparison for Gaussian Mixture Model. Benders-LFBGSB-N is the Benders decomposition algorithm on N randomly initialized points $\beta \in \Omega$ using L-BFGS-B to optimize $G(x, t)$. Benders-EM-LBFGSB-N is the Benders decomposition algorithm initialized with the output of the EM algorithm as described in the text. EM-N fits the sample using the EM algorithm with N components. The column "objective valued stderr" gives the standard error of the objective value when the algorithm, including initialization, was run multiple times on the same data.

Method	cpu time (s)	objective value
Multiplicative-50	0.09	21.232
Multiplicative-100	0.09	20.871
Multiplicative-500	0.16	20.582
Multiplicative-1000	0.19	20.546
Benders-LBFGS-10	0.44	20.509

Table 5.3: Algorithm Comparison for Optimal Design. Benders-LBFGS-10 is the Benders decomposition algorithm optimized using L-BFGS-B with 16 support points. Multiplicative-N is the multiplicative algorithm with N support points.

$G(x,t)$ for fixed t. The penalties $t = 10,\ 5,\ 1,\ 10^{-1},\ 10^{-2},\ 10^{-3},\ 10^{-4}$ were used with the default stopping tolerance. The multiplicative algorithm requires fixed support points. The support points used are $\{q(s_i),\ 1 \leq i \neq n\}$ where $s_i = \frac{3i}{n}$.

As the table shows, the Benders decomposition algorithm took longer than the multiplicative algorithm but terminated with a lower objective value. Furthermore, the Benders decomposition algorithm terminated with only 4 distinct support points.

5.5 Conclusion

This chapter has presented numerical experiments on problems without convex cone constraints comparing the Benders decomposition algorithm to popular alternative algorithms on similar problems. These experiments suggest that, in general, the Benders decomposition algorithm found lower objective values, but was slower than the alternative algorithms. The Poisson mixture model experiment suggests using a Quasi-Newton solve to optimize $G(x,t)$ for fixed t will take long and reach a similar objective value compared to using a Newton method.

There are several open questions regarding the numerical performance of the Benders decomposition algorithm presented. The performance of the Benders decomposition algorithm on problems with moments constraints needs quantified. For the `NPMLE` problem, the effect of sample size on the algorithms compared, distances between the components, and selection procedures for the number of components are all areas of further study. For the `OptD` problem, numerical comparisons on problems with larger design matrices and different design criteria require further investigation.

BIBLIOGRAPHY

[1] Martin S Andersen, Joachim Dahl, and Lieven Vandenberghe. Cvxopt: A python package for convex optimization, version 1.1.7. `http://www.cvxopt.org`, 2013.

[2] Y. Baek. *An interior point approach to constrained nonparametric mixture models*. PhD thesis, University of Washington, 2007.

[3] A. Barvinok. *A Course In Convexity*, volume 54. American Mathematical Society, 2002.

[4] Aharon Ben-Tal, Dick Den Hertog, Anja De Waegenaere, Bertrand Melenberg, and Gijs Rennen. Robust solutions of optimization problems affected by uncertain probabilities. *Management Science*, 59(2):341–357, 2013.

[5] Aharon Ben-Tal, Laurent El Ghaoui, and Arkadi Nemirovski. *Robust Optimization*. Princeton University Press, 2009.

[6] A.L. Berger, V.J.D. Pietra, and S.A.D. Pietra. A maximum entropy approach to natural language processing. *Computational linguistics*, 22(1):39–71, 1996.

[7] Dimitri P. Bertsekas. *Convex Optimization Theory*. Athena Scientific, 2009.

[8] John R Birge and François V Louveaux. *Introduction to stochastic programming*. Springer, 2011.

[9] Christopher M Bishop. *Pattern Recognition and Machine Learning*. Springer New York, 2006.

[10] D. Böhning. A vertex-exchange-method in d-optimal design theory. *Metrika*, 33(1):337–347, 1986.

[11] J.F. Bonnans and A. Shapiro. *Perturbation Analysis of Optimization Problems*. Springer Verlag, 2000.

[12] Jonathan M Borwein and Adrian S Lewis. *Convex analysis and nonlinear optimization: theory and examples*, volume 3. Springer, 2010.

[13] S. Boyd and L. Vandenberghe. *Convex Optimization*. Cambridge University Press, 2004.

[14] J.V. Burke. Second order necessary and sufficient conditions for convex composite NDO. *Mathematical Programming*, 38(3):287–302, 1987.

[15] J.V. Burke and R.A. Poliquin. Optimality conditions for non-finite valued convex composite functions. *Mathematical Programming*, 57(1):103–120, 1992.

[16] David A Cohn, Zoubin Ghahramani, and Michael I Jordan. Active learning with statistical models. *arXiv preprint cs/9603104*, 1996.

[17] I Csisz et al. Information-type measures of difference of probability distributions and indirect observations. *Studia Sci. Math. Hungar.*, 2:299–318, 1967.

[18] Erick Delage and Yinyu Ye. Distributionally robust optimization under moment uncertainty with application to data-driven problems. *Operations Research*, 2010.

[19] A.P. Dempster, N.M. Laird, and D.B. Rubin. Maximum likelihood from incomplete data via the em algorithm. *Journal of the Royal Statistical Society. Series B (Methodological)*, pages 1–38, 1977.

[20] P. Diggle, P. Heagerty, K.Y. Liang, and S. Zeger. *Analysis of longitudinal data*, volume 25. Oxford University Press, 2002.

[21] Richard O Duda, Peter E Hart, and David G Stork. *Pattern classification*. John Wiley & Sons, 2001.

[22] V.V. Fedorov. *Theory of Optimal Experiments*. Academic Press, 1972.

[23] G.B. Folland. *Real Analysis: Modern Techniques and Their Applications*, volume 361. Wiley New York, 1999.

[24] D.A. Follmann and D. Lambert. Identifiability of finite mixtures of logistic regression models. *Journal of Statistical Planning and Inference*, 27(3):375–381, 1991.

[25] J. Friedman, T. Hastie, and R. Tibshirani. *The Elements of Statistical Learning: Data Mining, Inference, and Prediction*. New York, NY: Springer-Verlag New York, 2009.

[26] V. Ganapathi, D. Vickrey, J. Duchi, and D. Koller. Constrained approximate maximum entropy learning of markov random fields. *arXiv preprint arXiv:1206.3257*, 2012.

[27] A. Gelman, J. Hill, et al. *Data Analysis Using Regression and Multilevel/Hierarchical Models*, volume 1. Cambridge University Press, 2007.

[28] A.M. Geoffrion. Generalized benders decomposition. *Journal of optimization theory and applications*, 10(4):237–260, 1972.

[29] Joel Goh and Melvyn Sim. Distributionally robust optimization and its tractable approximations. *Operations Research*, 58(4-Part-1):902–917, 2010.

[30] B. Grün and F. Leisch. Identifiability of finite mixtures of multinomial logit models with varying and fixed effects. *Journal of classification*, 25(2):225–247, 2008.

[31] C. Hennig. Identifiablity of models for clusterwise linear regression. *Journal of Classification*, 17(2):273–296, 2000.

[32] Roger A Horn and Charles R Johnson. *Matrix analysis*. Cambridge University Press, 1990.

[33] E. Jones, T. Oliphant, P. Peterson, et al. Scipy: Open source scientific tools for python. `http://www.scipy.org`, 2001–. [Online;, accessed 2015-02-09].

[34] A. Karr. Extreme points of certain sets of probability measures, with applications. *Mathematics of Operations Research*, 8(1):74–85, 1983.

[35] A. Lewis. Convex analysis on the hermitian matrices. *SIAM Journal on Optimization*, 6(1):164–177, 1996.

[36] A. Lewis. Active sets, nonsmoothness, and sensitivity. *SIAM Journal on Optimization*, 13(3):702–725, 2002.

[37] B.G. Lindsay. The geometry of mixture likelihoods: a general theory. *The Annals of Statistics*, 11(1):86–94, 1983.

[38] B.G. Lindsay. Mixture models: Theory, geometry and applications. In *NSF-CBMS Regional Conference Series in Probability and Statistics*, 1995.

[39] B.G. Lindsay and K. Roeder. Uniqueness of estimation and identifiability in mixture models. *Canadian Journal of Statistics*, 21(2):139–147, 1993.

[40] David Love and Güzin Bayraksan. Phi-divergence constrained ambiguous stochastic programs. 2013.

[41] Zhaosong Lu and Ting Kei Pong. Computing optimal experimental designs via interior point method. *SIAM Journal on Matrix Analysis and Applications*, 34(4):1556–1580, 2013.

[42] L.S. Magder and S.L. Zeger. A smooth nonparametric estimate of a mixing distribution using mixtures of gaussians. *Journal of the American Statistical Association*, pages 1141–1151, 1996.

[43] A. Mallet. A maximum likelihood estimation method for random coefficient regression models. *Biometrika*, 73(3):645–656, 1986.

[44] P. McCullagh and J.A. Nelder. *Generalized Linear Models*. Chapman & Hall/CRC, 1989.

[45] K. Murphy. *Machine Learning: A Probabilistic Perspective*. MIT press, 2012.

[46] J. Nocedal and S. Wright. *Numerical Optimization (2nd edition)*. Springer, 2006.

[47] Panos Y Papalambros and Douglass J Wilde. *Principles of optimal design: modeling and computation*. Cambridge University Press, 2000.

[48] Dávid Papp. Optimal designs for rational function regression. *Journal of the American Statistical Association*, 107(497):400–411, 2012.

[49] F. Pedregosa, G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blondel, P. Prettenhofer, R. Weiss, V. Dubourg, J. Vanderplas, A. Passos, D. Cournapeau, M. Brucher, M. Perrot, and E. Duchesnay. Scikit-learn: Machine learning in Python. *Journal of Machine Learning Research*, 12:2825–2830, 2011.

[50] I. Popescu. Robust mean-covariance solutions for stochastic optimization. *Operations Research*, 55(1):98–112, 2007.

[51] András Prékopa. Probabilistic programming. *Handbooks in operations research and management science*, 10:267–351, 2003.

[52] Friedrich Pukelsheim. *Optimal Design of Experiments*, volume 50. SIAM, 1993.

[53] R.T. Rockafellar. *Conjugate Duality and Optimization*, volume 16. SIAM, 1974.

[54] R.T. Rockafellar. *Convex Analysis*, volume 28. Princeton University Press, 1997.

[55] R.T. Rockafellar and J.B.Wets Roger. *Variational Analysis*, volume 317. Springer, 2004.

[56] Herbert Scarf, KJ Arrow, and S Karlin. A min-max solution of an inventory problem. *Studies in the mathematical theory of inventory and production*, 10:201–209, 1958.

[57] A. Shapiro and A. Kleywegt. Minimax analysis of stochastic problems. *Optimization Methods and Software*, 17(3):523–542, 2002.

[58] Aad W Van der Vaart. *Asymptotic Statistics*, volume 3. Cambridge University Press, 2000.

[59] Stefan Van Der Walt, S Chris Colbert, and Gael Varoquaux. The numpy array: a structure for efficient numerical computation. *Computing in Science & Engineering*, 13(2):22–30, 2011.

[60] Zizhuo Wang, Peter Glynn, and Yinyu Ye. Likelihood robust optimization for data-driven problems. *arXiv preprint arXiv:1307.6279*, 2013.

[61] Gerhard Winkler. Extreme points of moment sets. *Mathematics of Operations Research*, 13(4):581–587, 1988.

[62] Huan Xu, Constantine Caramanis, and Shie Mannor. A distributional interpretation of robust optimization. In *Communication, Control, and Computing (Allerton), 2010 48th Annual Allerton Conference on*, pages 552–556. IEEE, 2010.

[63] S.J. Yakowitz and J.D. Spragins. On the identifiability of finite mixtures. *The Annals of Mathematical Statistics*, 39(1):209–214, 1968.

[64] Ciyou Zhu, Richard H Byrd, Peihuang Lu, and Jorge Nocedal. Algorithm 778: L-bfgs-b: Fortran subroutines for large-scale bound-constrained optimization. *ACM Transactions on Mathematical Software (TOMS)*, 23(4):550–560, 1997.

VITA

Chris Jordan-Squire is a mathematics graduate student, which continually confuses everyone that talks to him.