

6-2-2020

Convex and Nonconvex Optimization Techniques for Multifacility Location and Clustering

Tuyen Dang Thanh Tran
Portland State University

Follow this and additional works at: https://pdxscholar.library.pdx.edu/open_access_etds



Part of the [Mathematics Commons](#)

Let us know how access to this document benefits you.

Recommended Citation

Tran, Tuyen Dang Thanh, "Convex and Nonconvex Optimization Techniques for Multifacility Location and Clustering" (2020). *Dissertations and Theses*. Paper 5482.

<https://doi.org/10.15760/etd.7356>

This Dissertation is brought to you for free and open access. It has been accepted for inclusion in Dissertations and Theses by an authorized administrator of PDXScholar. Please contact us if we can make this document more accessible: pdxscholar@pdx.edu.

Convex and Nonconvex Optimization Techniques
for Multifacility Location and Clustering

by
Tuyen Dang Thanh Tran

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

Doctor of Philosophy
in
Mathematical Sciences

Dissertation Committee:
Mau Nam Nguyen, Chair
Dacian Daescu
Bin Jiang
Erik Sanchez

Portland State University
2020

© 2020 Tuyen Dang Thanh Tran

Abstract

This thesis contains contributions in two main areas: calculus rules for generalized differentiation and optimization methods for solving nonsmooth nonconvex problems with applications to multifacility location and clustering. A variational geometric approach is used for developing calculus rules for subgradients and Fenchel conjugates of convex functions that are not necessarily differentiable in locally convex topological and Banach spaces. These calculus rules are useful for further applications to nonsmooth optimization from both theoretical and numerical aspects. Next, we consider optimization methods for solving nonsmooth optimization problems in which the objective functions are not necessarily convex. We particularly focus on the class of functions representable as differences of convex functions. This class of functions is broad enough to cover many problems in facility location and clustering, while the generalized differentiation tools from convex analysis can be applied. We develop algorithms for solving a number of multifacility location and clustering problems and computationally implement these algorithms via MATLAB. The methods used throughout this thesis involve DC programming, Nesterov's smoothing technique, and the DCA, a numerical algorithm for minimizing differences of convex functions to cope with the nonsmoothness and nonconvexity.

Dedication

*This thesis is dedicated to my parents.
For their endless love, patience, support and encouragement.*

Table of Contents

Abstract	i
Dedication	ii
List of Tables	v
List of Figures	vi
1 Introduction and Preliminaries	1
1.1 Introduction	1
1.2 Basic Tools of Convex Analysis and Optimization	2
1.2.1 Convexity of Sets	2
1.2.2 Convexity of Functions	4
1.2.3 Fenchel Conjugate and Subgradient of Convex Functions ..	9
1.2.4 Normal Cones to Convex Set and Euclidean Projections ...	10
1.2.5 DC Functions	11
1.2.6 Optimization Problem	17
2 Numerical Techniques and Algorithms	19
2.1 Nesterov Smoothing Technique	19
2.2 Overview of the DCA	23
2.3 A Penalty Method via Distance Functions	32
3 Calculus Rules for Fenchel conjugates and Subdifferentials	39
3.1 Introduction and Basic Definitions	39
3.2 Support Functions and Normal Cones to Set Intersections	41
3.3 Geometric Approach to Conjugate Calculus	52

	3.4 Geometric Approach to Convex Subdifferential Calculus	59
4	DC Programming for Constrained Multifacility Location and Clustering	64
	4.1 Overview	64
	4.2 Clustering with Constraints	65
	4.3 Set Clustering with Constraints	70
	4.4 Multifacility Location with Constraints	74
	4.5 Numerical Experiments	78
	4.5.1 Clustering with constraints	78
	4.5.2 Set Clustering with Constraints	79
	4.5.3 Multifacility Location with Constraints	80
5	DC Programming for Hierarchical Clustering	82
	5.1 Overview	82
	5.2 The Bilevel Hierarchical Clustering Problem	83
	5.2.1 The bilevel hierarchical clustering: Model I	83
	5.2.2 The bilevel hierarchical clustering: Model II	91
	5.2.3 Numerical Experiments	93
6	DC Programming for Multifacility Location via Mixed Integer Programming	97
	6.1 Overview	97
	6.2 Smooth Approximation by Continuous DC Problems	98
	6.3 Design and Implementation of the Solution Algorithm	104
7	Conclusion	111
	References	112

List of Tables

Table 5.1	Results for the 18 points artificial data set.	95
Table 5.2	Results for EIL76 data set.	95
Table 5.3	Results for PR1002 data set.	95
Table 6.1	Comparison between Algorithm 3 and k -means.	107
Table 6.2	Comparison between Algorithm 3 and k -means.	108
Table 6.3	Comparison between Algorithm 9 (combined with k -means) and standard k -means.	109
Table 6.4	Cost comparison between Algorithm 3 (combined with k -means) and k -means.	110

List of Figures

Figure 1.1	Convex set and nonconvex set.	2
Figure 1.2	Convex and nonconvex functions.	4
Figure 1.3	An example of the epigraph of a function.	5
Figure 2.1	Nesterov’s smoothing for $f(x) = x $	23
Figure 2.2	Convergence of the gradient method and DCA-1.	31
Figure 2.3	Graph and performance of Example 2.2.10.	32
Figure 4.1	A 2-center constrained clustering problem for dataset EIL76.	78
Figure 4.2	A 3-center set clustering problems with 50 most populous US cities.	79
Figure 4.3	A 4-center multifacility location with one ball constraint.	80
Figure 4.4	A 4-center constrained multifacility location problem with US cities dataset.	81
Figure 5.1	Plots of the three test data sets.	94
Figure 5.2	PR1002, The 1002 City Problem, with 6 Cluster Centers.	96
Figure 6.1	MFLP with 14 demand points and 2 centers.	107
Figure 6.2	MFLP with 200 demand points and 2 centers.	108
Figure 6.3	MFLP with 50 demand points and 3 centers.	110

Introduction and Preliminaries

1.1. Introduction

It has been well recognized that convex analysis is an important fundamental mathematical foundation for many applications in which convex optimization is the first to name. In recent years, convex analysis has become more and more important for applications to several new fields such as computational statistics, machine learning, and location science to cope with the state of the art problems arising in daily life. Although convex optimization techniques have been topics of extensive research for more than 50 years, solving large-scale optimization problems without the presence of convexity remains a challenge.

“The great watershed in optimization is not between linearity and nonlinearity, but convexity and nonconvexity”- R. T. Rockafellar.

In addition, many optimization techniques are based on the differentiability of data, while nondifferentiable structures also appear frequently and naturally in many optimization models. Motivated by applications to optimization problems of nondifferentiable nature, nonsmooth/variational analysis has been developed to study generalized differentiation properties of sets, functions, and set-valued mappings without making assumptions about the smoothness of the data.

This thesis firstly focuses on a geometric approach of variational analysis for the case of convex objects considered in locally convex topological spaces and in Banach space settings. Next, we consider optimization methods for solving nonsmooth optimization problems in which the objective functions are not necessarily convex, with applications to location problems and clustering.

1.2. Basic Tools of Convex Analysis and Optimization

Convex analysis was developed by Rockafellar [41] and Moreau [32] independently in the 1960s. Their development of the *subdifferential*, which is a useful concept in nonsmooth analysis and optimization theory, generalizes the idea of the derivative in classical calculus from differentiable functions to functions that are convex but not necessarily differentiable. This section provides the mathematical foundation for convex analysis that will be used throughout this thesis. More details of convex analysis can be found in [21, 25, 26, 41]. We begin with a systematic study of convex sets and functions with the discussion beginning with sets.

1.2.1. Convexity of Sets.

Throughout this subsection, we will consider topological vector spaces. Given $x, y \in X$, we define the *interval/line segment*

$$[x, y] := \{\lambda x + (1 - \lambda)y \mid \lambda \in [0, 1]\}.$$

If $x = y$ then this interval becomes a singleton $[x, y] = \{x\}$. Naturally, this leads us to the definition of convexity of sets.

Definition 1.2.1. A subset Ω of X is convex if for all $x, y \in \Omega$ and $\lambda \in (0, 1)$, we have

$$\lambda x + (1 - \lambda)y \in \Omega.$$

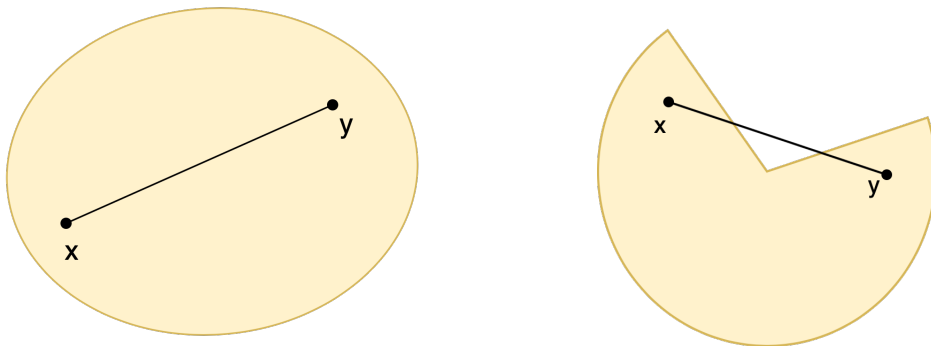


FIGURE 1.1. Convex set and nonconvex set.

Additionally, a mapping $B: X \rightarrow Y$ is *affine* if there exists a linear mapping $A: X \rightarrow Y$ and a vector $b \in Y$ such that

$$B(x) = Ax + b \text{ for all } x \in X.$$

It is not hard to check that the convexity of sets is preserved under operations using affine mappings. Thus, this promotes the following proposition for affine mappings and convex sets.

Proposition 1.2.2. *Let $B: X \rightarrow Y$ be an affine mapping between topological vector spaces X and Y . The following properties hold:*

- (a) *If Ω is a convex subset of X , then the direct image $B(\Omega)$ is a convex subset of Y .*
- (b) *If Θ is a convex subset of Y , then the inverse image $B^{-1}(\Theta)$ is a convex subset of X .*

Proceeding forward, for any subsets $\Omega_1, \Omega_2 \subset X$ and $\gamma \in \mathbb{R}$, we define

$$\gamma\Omega_1 := \{\gamma x \mid x \in \Omega_1\} \quad \text{and} \quad \Omega_1 + \Omega_2 := \{x_1 + x_2 \mid x_1 \in \Omega_1, x_2 \in \Omega_2\}.$$

The next corollary is a simple consequence of Proposition 1.2.2.

Corollary 1.2.3. *Let Ω_1 and Ω_2 be convex subsets of X , and $\gamma \in \mathbb{R}$. Then $\gamma\Omega_1$ and $\Omega_1 + \Omega_2$ are also convex.*

We next consider the convexity under the Cartesian product.

Proposition 1.2.4. *Let Ω_1 be a convex subset of X and let Ω_2 be a convex subset of Y . Then the Cartesian product $\Omega_1 \times \Omega_2$ is a convex subset of $X \times Y$.*

For any collection of convex sets $\{\Omega_i\}_{i \in I}$, their intersection $\bigcap_{i \in I} \Omega_i$ is also convex. This motivates us to define the *convex hull* of a set $\Omega \subset X$ by

$$\text{co}(\Omega) := \bigcap \left\{ C \mid C \text{ is convex and } \Omega \subset C \right\}.$$

Clearly, the convex hull of a set Ω is the smallest convex set containing Ω . The following useful observation is a direct consequence of the definition of the *convex hull*.

Proposition 1.2.5. For any nonempty subset Ω of X , its convex hull admits the representation

$$\text{co}(\Omega) = \left\{ \sum_{i=1}^m \lambda_i w_i \mid \sum_{i=1}^m \lambda_i = 1, \lambda_i \geq 0, w_i \in \Omega, m \in \mathbb{N} \right\}.$$

In all, we have given several definitions and theorems relating to convex sets. In the next subsection we will discuss the convexity in terms of functions.

1.2.2. Convexity of Functions.

This subsection provides the study of convex functions in real vector spaces. We will mainly focus on *extended-real-valued* functions $f: X \rightarrow \overline{\mathbb{R}} := (-\infty, \infty]$. From this definition it follows that $f(x)$ can be equal to ∞ for some elements $x \in X$.

Definition 1.2.6. Let X be a vector space and $f: \Omega \rightarrow \overline{\mathbb{R}}$ be a function defined on a convex set $\Omega \subset X$. We say that f is convex on Ω if for all $x, y \in \Omega$ and $\lambda \in (0, 1)$ the following inequality holds

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

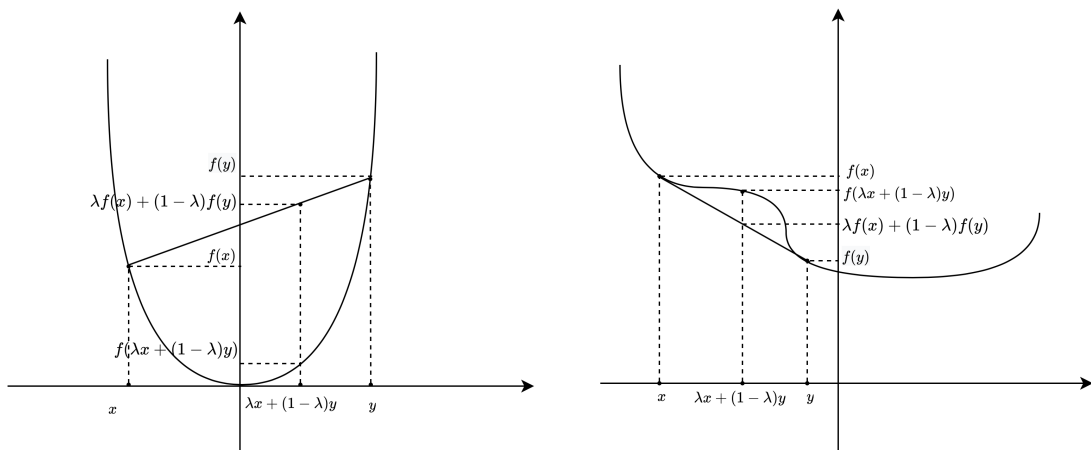


FIGURE 1.2. Convex and nonconvex functions.

We next introduce definitions of *domain* and *epigraph* associated with any extended-real-valued function f defined as

$$\text{dom } f := \{x \in X \mid f(x) < \infty\} \text{ and } \text{epi } f := \{(x, \alpha) \in X \times \mathbb{R} \mid f(x) \leq \alpha\},$$

respectively. Furthermore, we call a function f *proper* if $\text{dom } f \neq \emptyset$.

As a consequence, the next proposition shows the relation between the convexity of a function and its epigraphical set.

Proposition 1.2.7. *Let $f : X \rightarrow \overline{\mathbb{R}}$ be a function on a vector space X . Then, f is convex if and only if $\text{epi } f$ is a convex set in $X \times \mathbb{R}$.*

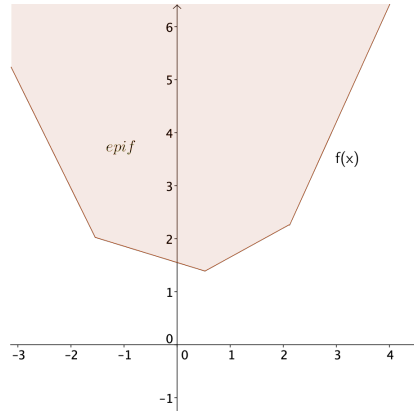


FIGURE 1.3. An example of the epigraph of a function.

We next consider several useful definitions and properties in finite dimensional case \mathbb{R}^d and Hilbert space H .

Definition 1.2.8. *Let H be a Hilbert space. A function $h : H \rightarrow \overline{\mathbb{R}}$ is called γ -convex ($\gamma \geq 0$) if there exists $\gamma \geq 0$ such that the function defined by $k(x) := h(x) - \frac{\gamma}{2}\|x\|^2$, $x \in H$, is convex. If there exists $\gamma > 0$ such that h is γ -convex, then h is called strongly convex.*

It is easy to see that if a function is strongly convex then it is also convex.

Definition 1.2.9. *A function $f : \mathbb{R}^d \rightarrow \mathbb{R}$ is called a $C^{1,1}$ function if it is differentiable on X and its gradient ∇f is Lipschitz continuous, i.e., if there exists $L > 0$ such that*

$$\|\nabla f(x) - \nabla f(y)\| \leq L\|x - y\|, \text{ for all } x, y \in H.$$

Note that any $C^{1,1}$ function is also C^1 (a continuously differentiable function).

Example 1.2.10. Let A be an $m \times n$ matrix and $b \in \mathbb{R}^m$. The function $f(x) := \|Ax - b\|^2$ is $C^{1,1}$ whose gradient is Lipschitz continuous with the Lipschitz constant determined by matrix norm.

Following the definition of a $C^{1,1}$ function, we shall provide relevant propositions.

Proposition 1.2.11. Let $f : \mathbb{R}^d \rightarrow \mathbb{R}$ be a $C^{1,1}$ function (not necessarily convex) with Lipschitz continuous gradient and Lipschitz constant L . Then,

$$f(y) \leq f(x) + \langle \nabla f(x), y - x \rangle + \frac{L}{2} \|y - x\|^2 \text{ for all } x, y \in \mathbb{R}^d.$$

Proof. For all $x, y \in \mathbb{R}^d$, define $\varphi(\lambda) := f(x + \lambda(y - x))$, $\lambda \in \mathbb{R}$. Then φ is differentiable on \mathbb{R} . Observe that

$$\varphi(0) = f(x), \quad \varphi(1) = f(y), \quad \varphi'(\lambda) = \langle \nabla f(x + \lambda(y - x)), y - x \rangle.$$

By the Fundamental Theorem of Calculus we have

$$\varphi(1) - \varphi(0) = \int_0^1 \varphi'(\lambda) d\lambda = \int_0^1 \langle \nabla f(x + \lambda(y - x)), y - x \rangle d\lambda.$$

Hence,

$$\begin{aligned} f(y) - f(x) &= \int_0^1 \langle \nabla f(x + \lambda(y - x)) - \nabla f(x), y - x \rangle d\lambda \\ &= \int_0^1 \langle \nabla f(x + \lambda(y - x)) - \nabla f(x), y - x \rangle d\lambda + \langle \nabla f(x), y - x \rangle. \end{aligned}$$

Rearranging the above equation by moving $f(x)$ to the right-hand side gives us

$$\begin{aligned} f(y) &= f(x) + \langle \nabla f(x), y - x \rangle + \int_0^1 \langle \nabla f(x + \lambda(y - x)) - \nabla f(x), y - x \rangle d\lambda \\ &\leq f(x) + \langle \nabla f(x), y - x \rangle + \int_0^1 \|\nabla f(x + \lambda(y - x)) - \nabla f(x)\| \cdot \|y - x\| d\lambda \\ &\leq f(x) + \langle \nabla f(x), y - x \rangle + \int_0^1 L\lambda \|y - x\|^2 d\lambda \\ &\leq f(x) + \langle \nabla f(x), y - x \rangle + \frac{L}{2} \|y - x\|^2 \end{aligned}$$

Which concludes the proof. \square

As mentioned earlier, any $C^{1,1}$ function is also C^1 . We provide the following relevant propositions associated with C^1 functions.

Proposition 1.2.12. *Suppose that $f : \mathbb{R}^d \rightarrow \mathbb{R}$ is a C^1 function. Then f is convex if and only if*

$$\langle \nabla f(y), x - y \rangle \leq f(x) - f(y) \text{ for all } x, y \in \mathbb{R}^d.$$

Proof. Fix any $x, y \in \mathbb{R}^d$ and $\lambda \in (0, 1)$. Then we have

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

which is equivalent to

$$f(y + \lambda(x - y)) \leq f(y) + \lambda[f(x) - f(y)].$$

Thus,

$$\frac{f(y + \lambda(x - y)) - f(y)}{\lambda} \leq f(x) - f(y).$$

Letting $\lambda \rightarrow 0^+$, gives us

$$\langle \nabla f(y), x - y \rangle \leq f(x) - f(y).$$

Conversely, for any $x, y \in \mathbb{R}^d$ and $\lambda \in (0, 1)$, define $z_\lambda := \lambda x + (1 - \lambda)y$. Then we will get that

$$\langle \nabla f(z_\lambda), x - z_\lambda \rangle \leq f(x) - f(z_\lambda) \quad \text{and} \quad \langle \nabla f(z_\lambda), y - z_\lambda \rangle \leq f(y) - f(z_\lambda).$$

It follows that

$$\begin{aligned} \lambda \langle \nabla f(z_\lambda), x - z_\lambda \rangle &\leq \lambda f(x) - \lambda f(z_\lambda), \\ (1 - \lambda) \langle \nabla f(z_\lambda), y - z_\lambda \rangle &\leq (1 - \lambda)f(y) - (1 - \lambda)f(z_\lambda). \end{aligned}$$

Adding these inequalities gives us

$$0 \leq \lambda f(x) + (1 - \lambda)f(y) - f(z_\lambda).$$

Therefore, $f(z_\lambda) \leq \lambda f(x) + (1 - \lambda)f(y)$, and thus f is convex. \square

Corollary 1.2.13. *Let $f : \mathbb{R}^d \rightarrow \mathbb{R}$ be a C^1 function. Then f is convex if and only if it is monotone in the sense that*

$$\langle \nabla f(x) - \nabla f(y), x - y \rangle \geq 0 \text{ for all } x, y \in \mathbb{R}^d.$$

Proof. Suppose that f is convex. Let us show that ∇f is monotone. For all $x, y \in \mathbb{R}^d$ we have

$$\langle \nabla f(x), y - x \rangle \leq f(y) - f(x),$$

$$\langle \nabla f(y), x - y \rangle \leq f(x) - f(y).$$

Adding these inequalities, we get

$$\langle \nabla f(x) - \nabla f(y), x - y \rangle \geq 0.$$

Thus, ∇f is monotone.

To prove the converse, fix any $x, y \in \mathbb{R}^d$ and let $z_\lambda := \lambda x + (1 - \lambda)y$. Using a similar technique as in the proof of Proposition 1.2.11, we have

$$\begin{aligned} f(x) - f(y) &= \int_0^1 \langle \nabla f(z_\lambda) - \nabla f(y) + \nabla f(y), x - y \rangle d\lambda \\ &= \int_0^1 \langle \nabla f(z_\lambda) - \nabla f(y), x - y \rangle d\lambda + \langle \nabla f(y), x - y \rangle. \end{aligned}$$

We next prove that $\langle \nabla f(z_\lambda) - \nabla f(y), x - y \rangle \geq 0$ for all $\lambda \in [0, 1]$. If $\lambda = 0$, then $z_\lambda = y$, and hence

$$\langle \nabla f(z_\lambda) - \nabla f(y), x - y \rangle = 0.$$

If $0 < \lambda \leq 1$, then $z_\lambda - y = \lambda(x - y)$ and $x - y = \frac{1}{\lambda}(z_\lambda - y)$. By the monotonicity of ∇f , we have

$$\langle \nabla f(z_\lambda) - \nabla f(y), x - y \rangle = \frac{1}{\lambda} \langle \nabla f(z_\lambda) - \nabla f(y), z_\lambda - y \rangle \geq 0.$$

Thus, $f(x) - f(y) \geq \langle \nabla f(y), x - y \rangle$ and Proposition 1.2.12 tells us that f is a convex function on \mathbb{R}^d . \square

We now conclude this subsection with a proposition, linking a C^1 function with the previous propositions and corollary.

Proposition 1.2.14. *Let $f : \mathbb{R}^d \rightarrow \mathbb{R}$ be a C^1 function. The following assertions are equivalent:*

- (a) f is strongly convex with the parameter γ .
- (b) For all $x, y \in \mathbb{R}^d$, we have the inequality

$$\langle \nabla f(y), x - y \rangle \leq f(x) - f(y) - \frac{\gamma}{2} \|x - y\|^2,$$

- (c) The gradient of the function f is strongly monotone in the sense that

$$\langle \nabla f(x) - \nabla f(y), x - y \rangle \geq \gamma \|x - y\|^2 \text{ for all } x, y \in \mathbb{R}^d.$$

1.2.3. Fenchel Conjugate and Subgradient of Convex Functions.

We now move our discussion forward from the convexity of sets and functions, to the Fenchel conjugate and subgradients of convex functions.

Definition 1.2.15. *Given a function $\varphi : \mathbb{R}^d \rightarrow \overline{\mathbb{R}}$ (not necessarily convex), the Fenchel conjugate of φ is defined by*

$$\varphi^*(v) := \sup \{ \langle v, x \rangle - \varphi(x) \mid x \in \mathbb{R}^d \}, \quad v \in \mathbb{R}^d.$$

Note that in the case where φ is a proper function, i.e.,

$$\text{dom}(\varphi) := \{x \in \mathbb{R}^d \mid \varphi(x) < \infty\} \neq \emptyset,$$

the Fenchel conjugate $\varphi^* : \mathbb{R}^d \rightarrow \overline{\mathbb{R}}$ is always convex and is an extended-real-valued convex function. Suppose further that φ is convex and lower semi-continuous, then the

Fenchel-Moreau Theorem states that $(\varphi^*)^* = \varphi$; see, e.g., [51]. Based on this theorem, we have the following relation between the subgradients of φ and its Fenchel conjugate:

$$(1.2.1) \quad x \in \partial\varphi^*(y) \iff y \in \partial\varphi(x).$$

With this relation in mind, the definition of the subgradient is as follows.

Definition 1.2.16. Let $\varphi: \mathbb{R}^d \rightarrow \overline{\mathbb{R}}$ be a convex function and let $x_0 \in \text{dom}(\varphi)$. The element $v \in \mathbb{R}^d$ is called a subgradient of φ at x_0 if

$$\langle v, x - x_0 \rangle \leq \varphi(x) - \varphi(x_0) \text{ for all } x \in \mathbb{R}^d.$$

The *subdifferential* in the sense of convex analysis of φ at x_0 , denoted by $\partial\varphi(x_0)$, consists of all subgradients of φ at this point. In the case when $x_0 \notin \text{dom}(\varphi)$, we set $\partial\varphi(x_0) = \emptyset$.

Additionally, this subdifferential concept possesses many calculus rules that are important for applications. In particular, for a finite number of convex functions $f_i: \mathbb{R}^d \rightarrow \overline{\mathbb{R}}$, $i = 1, \dots, m$, we have the following sum rule:

$$\partial(f_1 + \dots + f_m)(\bar{x}) = \partial f_1(\bar{x}) + \dots + \partial f_m(\bar{x}) \text{ for all } \bar{x} \in \mathbb{R}^d$$

provided that $\bigcap_{i=1}^m \text{ri}(\text{dom}(f_i)) \neq \emptyset$. Here $\text{ri}(\Omega)$ stands for the *relative interior* of Ω ; see, e.g., [25, Definition 1.68].

Moreover, if $f = \max_{i=1, \dots, m} f_i$, and f_i is continuous at \bar{x} for every $i = 1, \dots, m$, then for any $\bar{x} \in \mathbb{R}^d$ we have the following maximal rule:

$$(1.2.2) \quad \partial f(\bar{x}) = \text{co}\left(\bigcup_{i \in I(\bar{x})} \partial f_i(\bar{x})\right),$$

where $I(\bar{x}) = \{i \mid f_i(\bar{x}) = f(\bar{x})\}$ is called the active index set.

1.2.4. Normal Cones to Convex Set and Euclidean Projections.

In this subsection, we shall quickly mention normal cones and Euclidean projections since they have well-known applications to convexity.

Definition 1.2.17. Given a nonempty closed convex subset Ω of \mathbb{R}^d with $\bar{x} \in \Omega$, the normal cone to Ω at \bar{x} is defined by

$$N(\bar{x}; \Omega) = \{v \in \mathbb{R}^d \mid \langle v, x - \bar{x} \rangle \leq 0 \text{ for all } x \in \Omega\}.$$

If $\bar{x} \notin \Omega$, we set $N(\bar{x}, \Omega) = \emptyset$. It is well-known that an element $\bar{x} \in \mathbb{R}^d$ is an absolute minimizer of a convex function $f: \mathbb{R}^d \rightarrow \mathbb{R}$ on Ω if and only if \bar{x} is a local minimizer of f on Ω . Moreover, this happens if and only if the following optimality condition holds

$$0 \in \partial f(\bar{x}) + N(\bar{x}; \Omega).$$

The reader can find more details in [25, 41].

Further, given a nonempty set $\Omega \subseteq \mathbb{R}^n$ and a point $x \in \mathbb{R}^n$, the *Euclidean projection* from x onto Ω is the set, denoted by $P(x; \Omega)$, is defined by

$$P(x; \Omega) := \{w \in \Omega \mid \|x - w\| = d(x; \Omega)\},$$

where $d(x; \Omega)$ defines the *Euclidean distance* from x to Ω , i.e.,

$$d(x; \Omega) := \inf \{\|x - w\| \mid w \in \Omega\}.$$

We can show that if Ω is a nonempty closed set, then $P(x; \Omega)$ is nonempty, and it is a singleton if we assume in addition that Ω is convex. Moreover, we can show that if Ω is a convex set and $w \in P(x; \Omega)$, then $x - w \in N(w; \Omega)$; see, e.g., [25, 41] and the references therein.

1.2.5. DC Functions.

Convexity is a nice property of functions, but it is not preserved under simple algebraic operations such as scalar multiplication or minimum. This is a motivation to search for new optimization methods that are capable of handling broader classes of functions and sets where convexity is not assumed. One of the most successful approaches to go beyond convexity is to consider the class of functions representable as differences

of two convex functions. Functions of this type are called *DC functions*, where DC stands for *difference of convex*. It was first considered by Alexandrov (1940, 1950) and Landis (1951), and some time later by P. Hartman (1959) who recognized that the class of DC functions has many nice algebraic properties; see, e.g., [20]. In this section we introduce DC functions and a number of important properties of this class under operations which are commonly considered in nonconvex optimization problems. More details can be found in [22, 26, 48] in the proofs and other properties. We first begin with the definition of DC functions.

Definition 1.2.18. *Let Ω be a nonempty convex subset of \mathbb{R}^d . We say that a function $f : \Omega \rightarrow \mathbb{R}$ is DC on Ω if there exists two convex function $g, h : \Omega \rightarrow \mathbb{R}$ such that $f = g - h$. In addition to this, f is a continuous DC function if there exist two continuous convex functions $g, h : \Omega \rightarrow \mathbb{R}$ such that $f = g - h$.*

Prompted by the definition of DC functions, we can show a connection between DC functions and Fréchet differentiable functions under the condition that the Fréchet gradient is Lipschitz continuous.

Proposition 1.2.19. *Let $f : \mathbb{R}^d \rightarrow \mathbb{R}$ be a Fréchet differentiable function and the Fréchet gradient be Lipschitz continuous with constant ℓ . Then f is a DC function.*

Proof. We can rewrite f as follows

$$f(x) = \frac{1}{2} \left[\left(\frac{\ell}{2} \|x\|^2 + f(x) \right) - \left(\frac{\ell}{2} \|x\|^2 - f(x) \right) \right], \quad x \in \mathbb{R}^d.$$

Next, we will show that both functions $g(x) := \frac{\ell}{2} \|x\|^2 + f(x)$ and $h(x) := \frac{\ell}{2} \|x\|^2 - f(x)$ are convex for all $x \in \mathbb{R}^d$. It suffices to prove that g is convex since the proof for h can be done similarly.

Taking any \bar{x} and $d \in \mathbb{R}^d$, we first define the function $\varphi := g(\bar{x} + td)$ for $t \in \mathbb{R}$ and prove that φ is convex. To see this, we differentiate φ on \mathbb{R} and get

$$\varphi'(t) = \ell \langle \bar{x} + td, d \rangle + \langle \nabla f(\bar{x} + td), d \rangle, \quad t \in \mathbb{R}.$$

Without loss of generality, consider $t_1 < t_2$. Then,

$$\begin{aligned}
\varphi'(t_2) - \varphi'(t_1) &= \ell(t_2 - t_1)\langle d, d \rangle + \langle \nabla f(\bar{x} + t_2d) - \nabla f(\bar{x} + t_1d), d \rangle \\
&\geq \ell(t_2 - t_1)\|d\|^2 - \|\langle \nabla f(\bar{x} + t_2d) - \nabla f(\bar{x} + t_1d) \rangle\| \|d\| \\
&\geq \ell(t_2 - t_1)\|d\|^2 - \ell(t_2 - t_1)\|d\|^2 = 0.
\end{aligned}$$

Trivially, this indicates that φ' is increasing and thus shows us that φ is convex which implies the convexity of g . \square

The next propositions show that the class of DC functions is closed under many operations usually considered in optimization such as taking the linear combination, maximum, or product of a finite number of DC functions.

Proposition 1.2.20. *Assume that f and $f_i : \mathbb{R}^d \rightarrow \mathbb{R}$ for $i = 1, \dots, k$ are DC functions. Then, the following functions are also DC:*

- (a) $\sum_{i=1}^k \gamma_i f_i(x)$, for any $\gamma_i \in \mathbb{R}$, $i = 1, \dots, k$,
- (b) $\max\{f_1(x), \dots, f_k(x)\}$, $\min\{f_1(x), \dots, f_k(x)\}$,
- (c) $|f|$, $f^+(x) := \max\{0, f(x)\}$, $f^-(x) := \min\{0, f(x)\}$.

Proof. (a) It is not hard to see that if f is a DC function then so is γf for any real number γ and thus (a) is a trivial consequence of properties of convex and concave functions.

(b) Let $f_i = g_i - h_i$, for $i = 1, \dots, k$, be DC decompositions of f_i . It is obvious that

$$f_i = g_i + \sum_{j=1, j \neq i}^k h_j - \sum_{j=1}^k h_j,$$

and

$$\max_{i=1, \dots, k} f_i = \max_{i=1, \dots, k} \left\{ g_i + \sum_{j=1, j \neq i}^k h_j \right\} - \sum_{j=1}^k h_j.$$

Since the maximum and the sum of a finite number of convex functions are convex, the functions $g := \max_{i=1, \dots, k} \left\{ g_i + \sum_{j=1, j \neq i}^k h_j \right\}$ and $h := \sum_{j=1}^k h_j$ are convex. Hence, $\max_{i=1, \dots, k} f_i$ is a DC function. The proof for $\min_{i=1, \dots, k} f_i$ is similar with the observation that

$$\min_{i=1,\dots,k} f_i = - \max_{i=1,\dots,k} (-f_i).$$

(c) Let $f = g - h$ with g, h convex. Then, we can rewrite f as

$$|f(x)| = 2 \max\{g, h\} - (g + h).$$

This is a DC decomposition of $|f|$. The assertion for f^+ and f^- follows from (b). \square

The next lemma is useful when it comes to the nonnegative DC decomposition.

Lemma 1.2.21. *If $f : \mathbb{R}^d \rightarrow \mathbb{R}$ is a DC function. Then there exist convex functions $g, h : X \rightarrow [0, \infty)$ such that $f = g - h$.*

Proof. Suppose that $f_1, f_2 : \mathbb{R}^d \rightarrow \mathbb{R}$ are convex functions in the DC decomposition such that $f = f_1 - f_2$. We can choose $x_1, x_2 \in \mathbb{R}^d$ and $b_1, b_2 \in \mathbb{R}$ such that

$$f_1(x) \geq \varphi_1(x) := \langle x_1, x \rangle + b_1$$

$$f_2(x) \geq \varphi_2(x) := \langle x_2, x \rangle + b_2.$$

for all $x \in \mathbb{R}^d$. Then $f = (f_1 - \varphi_1) - (f_2 - \varphi_2) + (\varphi_1 - \varphi_2)$. Let us define $\phi := \varphi_1 - \varphi_2$.

Then, it is an affine function with the representation

$$\phi = \max\{\phi, 0\} - \max\{-\phi, 0\}.$$

Let $\phi^+ := \max\{\phi, 0\}$ and $\phi^- := \max\{-\phi, 0\}$. It is easy to see that they are both nonnegative and convex. Hence, f can be rewritten as follows

$$f = (f_1 - \varphi_1 + \phi^+) - (f_2 - \varphi_2 + \phi^-).$$

Let $g := f_1 - \varphi_1 + \phi^+$ and $h := f_2 - \varphi_2 + \phi^-$. \square

Consequently, this lemma leads us to the next proposition.

Proposition 1.2.22. *Let $f : \mathbb{R}^d \rightarrow \mathbb{R}$ be a DC function. Then, f^2 is also a DC function.*

Proof. Based on Lemma 1.2.21, there exists two nonnegative convex functions $g, h : \mathbb{R}^d \rightarrow \mathbb{R}$ such that $f = g - h$. Then

$$f^2 = g^2 + h^2 - 2gh = 2g^2 + 2h^2 - (g + h)^2,$$

which is clearly a DC function because both $2g^2 + 2h^2$ and $(g + h)^2$ are convex. \square

More generally, we can see that the product of any two DC functions is also DC.

Proposition 1.2.23. *Let f_i be DC functions for $i = 1, \dots, m$. Then the product $\prod_{i=1}^m f_i$ is a DC function.*

Proof. It suffices to prove for the case where $m = 2$. That is, consider

$$f_1 f_2 = \frac{1}{2}((f_1 + f_2)^2 - f_1^2 - f_2^2).$$

It follows directly from Proposition 1.2.22 that $(f_1 + f_2)^2$, f_1^2 , and f_2^2 are DC functions.

Therefore, $f_1 f_2$ is a DC function. The result for $\prod_{i=1}^m f_i$ follows by induction. \square

Intuitively DC functions can be quite simple to picture in terms of the graphs, in the next example we provide a DC function in terms of a symmetric matrix.

Example 1.2.24. Let $A \in \mathbb{R}^{d \times d}$ be a symmetric matrix. Consider the quadratic form

$$f(x) := x^T A x \text{ for } x \in \mathbb{R}^d.$$

Then f is a DC function. Indeed, fix a real number δ such that $\delta \geq \rho(A)$ (here $\rho(A)$ is the spectral radius of A). Then

$$f(x) = \frac{1}{2}(x^T(\delta I + A)x - x^T(\delta I - A)x) \text{ for all } x \in \mathbb{R}^d.$$

Therefore, f is a DC function.

Another interesting result is the DC property of the distance function.

Proposition 1.2.25. *Let Ω be a nonempty closed subset of \mathbb{R}^d . Then the distance function $f(x) := d^2(x; \Omega)$ for $x \in \mathbb{R}^d$ is DC.*

Proof. Indeed, for any $x \in \mathbb{R}^d$ we have

$$\begin{aligned}
f(x) &= \inf_{u \in \Omega} \|x - u\|^2 \\
&= \|x\|^2 + \inf_{u \in \Omega} \{-\|x\|^2 + \|x - u\|^2\} \\
&= \|x\|^2 + \inf_{u \in \Omega} \{-2\langle x, u \rangle + \|u\|^2\} \\
&= \|x\|^2 - \sup_{u \in \Omega} \{2\langle u, x \rangle - \|u\|^2\}
\end{aligned}$$

Observe that the functions $g(x) := \|x\|^2$ and $h(x) := \sup_{u \in \Omega} (2\langle u, x \rangle - \|u\|^2)$ are convex for all $x \in \mathbb{R}^d$. Therefore, f is a DC function. \square

Lastly for this subsection, we shall define a locally DC function and briefly mention associated corollaries with locally DC functions.

Definition 1.2.26. We say that a function $f : \mathbb{R}^d \rightarrow \mathbb{R}$ is a locally DC function if for any $x_0 \in X$, there exists a ball $\mathbb{B}(x_0; \varepsilon)$ such that f is DC on $\mathbb{B}(x_0; \varepsilon)$. Equivalently, we can always find convex functions g, h such that

$$f(x) = g(x) - h(x) \text{ for all } x \in \mathbb{B}(x_0; \varepsilon).$$

Example 1.2.27. Every C^2 function $f : I \rightarrow \mathbb{R}$, where I is an open interval in \mathbb{R} is locally DC. Indeed, for any $x_0 \in I$, there exists $\delta > 0$ such that $J := [x_0 - \delta, x_0 + \delta] \subset I$. Then $S := \sup_{x \in J} |f''(x)| < \infty$. We can rewrite f as

$$f(x) = \frac{1}{2} [(\frac{\gamma}{2}x^2 + f(x)) - (\frac{\gamma}{2}x^2 - f(x))], \quad x \in J.$$

It follows that for $\gamma \geq S$,

$$g(x) := \frac{\gamma}{2}x^2 + f(x) \text{ and } h(x) := \frac{\gamma}{2}x^2 - f(x), \quad x \in J$$

are convex. Therefore, f is a locally DC function.

Theorem 1.2.28. Every locally DC function $f : \mathbb{R}^d \rightarrow \mathbb{R}$ is DC.

The proof of this theorem can be found in [22, 26].

Remark 1.2.29. *It can be shown that if $f : \Omega \rightarrow \mathbb{R}$ is a locally DC function, where $\Omega \subset \mathbb{R}^n$ is a nonempty convex set that is either open or closed, then f is a DC function. More about this can be found in [20].*

Theorem 1.2.28 leads to the following important corollaries.

Corollary 1.2.30. *Any C^2 function $f : \mathbb{R}^d \rightarrow \mathbb{R}$ is a DC function.*

Since polynomials have continuous derivatives of any order, the next result follows directly from Corollary 1.2.30.

Corollary 1.2.31. *Any polynomial with $x \in \mathbb{R}^d$ is a DC function on \mathbb{R}^d .*

We next consider the composition of two DC functions.

Proposition 1.2.32. *Let $f : \mathbb{R}^d \rightarrow I$ and let $g : I \rightarrow \mathbb{R}$ be DC functions, where I is an open interval in \mathbb{R} . Then $g \circ f$ is a DC function.*

Consequently, this proposition leads us to the following corollary.

Corollary 1.2.33. *If $f_1, f_2 : \mathbb{R}^d \rightarrow \mathbb{R}$ are DC functions, where $f_2(x) \neq 0$ for all $x \in \mathbb{R}^d$, then $\frac{f_1}{f_2}$ is a DC function.*

Proof. Without loss of generality we can assume that $f_2(x) > 0$ for all $x \in \mathbb{R}^d$. From Proposition 1.2.23 it suffices to show $\frac{1}{f_2}$ is DC. Since $f_2 : \mathbb{R}^d \rightarrow I := (0, \infty)$ is a DC function, and $g : I \rightarrow \mathbb{R}$ with $g(y) := \frac{1}{y}$ for $y \in I$ is also a DC function, Proposition 1.2.32 tells us that $\frac{1}{f_2}$ is a DC function. □

1.2.6. Optimization Problem.

After the discussion of the basic tools of convex analysis, this leads us to the optimization problems. Given a nonempty $\Omega \subset X$ and a function $f : \mathbb{R}^d \rightarrow \overline{\mathbb{R}}$, we consider the optimization problem

$$(1.2.3) \quad \text{minimize } f(x), \quad x \in \Omega.$$

We say that (1.2.3) is a *constrained optimization problem* if Ω is a proper subset of X . Otherwise, it is called an *unconstrained optimization problem*. In the case where both f and Ω are convex, (1.2.3) is a *convex optimization problem*.

The following propositions and corollary give us an effective tools to solving optimization problems when we have convexity.

Proposition 1.2.34. *Let $f : \mathbb{R}^d \rightarrow \overline{\mathbb{R}}$ be a continuously differentiable convex function (C^1 convex) and $\Omega \subset \mathbb{R}^d$ be a nonempty closed convex set. Then, x^* is the optimal solution to (1.2.3) if and only if*

$$\langle \nabla f(x^*), u - x^* \rangle \geq 0 \text{ for all } u \in \Omega.$$

Proof. Let x^* be the optimal solution. Then, $f(x) \geq f(x^*)$ for all $x \in \Omega$. For any $u \in \Omega$ and $\lambda \in (0, 1)$, since Ω is convex $x^* + \lambda(u - x^*) = \lambda u + (1 - \lambda)x^* \in \Omega$. This implies $f(x^* + \lambda(u - x^*)) \geq f(x^*)$, and hence

$$\frac{f(x^* + \lambda(u - x^*)) - f(x^*)}{\lambda} \geq 0.$$

Letting $\lambda \rightarrow 0^+$ gives us $\langle \nabla f(x^*), u - x^* \rangle \geq 0$.

Conversely, suppose $\langle \nabla f(x^*), u - x^* \rangle \geq 0$ for all $u \in \Omega$. The convexity of f gives us

$$\langle \nabla f(x^*), u - x^* \rangle \leq f(u) - f(x^*) \text{ for all } u \in \mathbb{R}^d.$$

Therefore,

$$0 \leq \langle \nabla f(x^*), u - x^* \rangle \leq f(u) - f(x^*)$$

for all $u \in \Omega$. It follows that x^* is the optimal solution to (1.2.3). \square

Corollary 1.2.35. *Let $f : \mathbb{R}^d \rightarrow \mathbb{R}$ be a C^1 concave function (i.e., f is convex) and Ω be a nonempty closed subset of \mathbb{R}^d . Then, x^* is the optimal solution of*

$$\text{maximize } f(x), \quad x \in \Omega$$

if and only if $\langle \nabla f(x^), u - x^* \rangle \leq 0$ for all $u \in \Omega$.*

Proposition 1.2.36. *If $\Omega \subset \mathbb{R}^d$ is a nonempty closed set and $f : \mathbb{R}^d \rightarrow \mathbb{R}$ is strongly convex, then (1.2.3) has a unique optimal solution.*

Numerical Techniques and Algorithms

2.1. Nesterov Smoothing Technique

In this chapter, we shall go over numerical techniques and algorithms. Let $f : \mathbb{R}^d \rightarrow \mathbb{R}$ be a convex function. Consider the unconstrained optimization problem

$$\text{minimize } f(x), \quad x \in \mathbb{R}^d,$$

where f (not necessarily differentiable) is given by

$$f(x) := \max\{\langle Ax, u \rangle - \phi(u) \mid u \in Q\}, \quad x \in \mathbb{R}^d.$$

Here, A is an $m \times d$ matrix, Q is a nonempty compact convex subset of \mathbb{R}^m , and ϕ is a continuous convex function on Q . We consider these class of functions for the Nesterov's smoothing technique. The reader can find more details in [37].

Let q be a continuous *strongly convex function* on \mathbb{R}^m with parameter $\sigma > 0$. The function q is called a *prox-function*. Proposition 1.2.36 tells us that q has a unique optimal solution on Q . Let us denote

$$\bar{x} := \operatorname{argmin}\{q(x) \mid x \in Q\}.$$

Without loss of generality, we can assume that $q(\bar{x}) = 0$. Since q is strongly convex, we also have

$$q(x) \geq \frac{\sigma}{2} \|x - \bar{x}\|^2, \quad \forall x \in Q.$$

For all $\mu > 0$, define

$$(2.1.4) \quad f_\mu(x) := \max\{\langle Ax, u \rangle - \phi(u) - \mu q(u) \mid u \in Q\}.$$

Then, μ is called the *smoothing parameter* and f_μ is the Nesterov smooth approximation of f .

Recall that for an $m \times n$ matrix $A = (a_{ij})$, if we define

$$\|A\| := \max\{\|Ax\| \mid \|x\| \leq 1\},$$

then

$$\|Ax\| \leq \|A\| \|x\| \text{ for all } x \in \mathbb{R}^d.$$

Theorem 2.1.1. *The function f_μ in (2.1.4) is well-defined, a C^1 function with Lipschitz gradient $L_\mu = \frac{1}{\mu\sigma}\|A\|^2$, and*

$$\nabla f_\mu(x) = A^T u_\mu(x).$$

Here, $u_\mu(x) \in Q$ is the element for which the maximum is attained in the definition of $f_\mu(x)$.

Proof. Let us consider the continuous function

$$\varphi(u, x) := \langle Ax, u \rangle - \phi(u) - \mu q(u).$$

Since Q is compact, $f_\mu(x) = \max_{u \in Q} \{\varphi(u, x)\}$ is clearly well-defined. For each $u \in Q$, $\varphi(u, \cdot)$ is a convex function which implies the convexity of $f_\mu(x)$; see, e.g., [25, Theorem 2.93]. In addition to this, for each $x \in \mathbb{R}^d$ the function $-\varphi(\cdot, x)$ is strongly convex on Q . Hence, there exists $u_\mu(x) \in M$ such that (2.1.4) attains the maximum and for each $x \in \mathbb{R}^d$ we have

$$S(x) = \{u \in M \mid \varphi(u, x) = f_\mu(x)\} = \{u_\mu(x)\}.$$

Notice that $\nabla_x \varphi(u, x) = A^T u$ and [25, Theorem 2.93] gives us

$$\partial f_\mu(x) = \text{co} \bigcup_{u \in S(x)} \partial_x \varphi(u, x) = \text{co} \bigcup_{u \in S(x)} \nabla_x \varphi(u, x) = \{A^T u_\mu(x)\}.$$

Since $\partial f_\mu(x)$ is a singleton, f_μ is differentiable on \mathbb{R}^d ; see, e.g., [25, Theorem 3.3].

We next prove that ∇f_μ is Lipschitz. By taking any $x, y \in \mathbb{R}^d$ and using Corollary 1.2.35 we get

$$\langle Ax - \nabla\phi(u_\mu(x)) - \mu\nabla q(u_\mu(x)), u_\mu(y) - u_\mu(x) \rangle \leq 0,$$

$$\langle Ay - \nabla\phi(u_\mu(y)) - \mu\nabla q(u_\mu(y)), u_\mu(x) - u_\mu(y) \rangle \leq 0.$$

Combining these inequalities gives us

$$\begin{aligned} \langle A(x - y), u_\mu(x) - u_\mu(y) \rangle &\geq \langle \nabla\phi(u_\mu(x)) - \nabla\phi(u_\mu(y)), u_\mu(x) - u_\mu(y) \rangle \\ &\quad + \mu \langle \nabla d(u_\mu(x)) - \nabla d(u_\mu(y)), u_\mu(x) - u_\mu(y) \rangle. \end{aligned}$$

Using the monotonicity of $\nabla\phi$ and the strong monotonicity of q from Corollary 1.2.13 and Proposition 1.2.14, respectively, we get

$$\langle A(x - y), u_\mu(x) - u_\mu(y) \rangle \geq \mu\sigma \|u_\mu(x) - u_\mu(y)\|^2.$$

Thus,

$$\begin{aligned} \|\nabla f_\mu(x) - \nabla f_\mu(y)\|^2 &= \|A^T u_\mu(x) - A^T u_\mu(y)\|^2 \\ &\leq \|A\|^2 \|u_\mu(x) - u_\mu(y)\|^2 \\ &\leq \frac{1}{\mu\sigma} \|A\|^2 \langle A(x - y), u_\mu(x) - u_\mu(y) \rangle \\ &= \frac{1}{\mu\sigma} \|A\|^2 \langle x - y, A^T u_\mu(x) - A^T u_\mu(y) \rangle \\ &\leq \frac{1}{\mu\sigma} \|A\|^2 \|x - y\| \|A^T u_\mu(x) - A^T u_\mu(y)\| \\ &= \frac{1}{\mu\sigma} \|A\|^2 \|x - y\| \|\nabla f_\mu(x) - \nabla f_\mu(y)\|. \end{aligned}$$

And so

$$\|\nabla f_\mu(x) - \nabla f_\mu(y)\| \leq \frac{1}{\mu\sigma} \|A\|^2 \|x - y\|.$$

That is, ∇f_μ is Lipschitz with the constant $L_\mu = \frac{1}{\mu\sigma} \|A\|^2$. □

The following proposition shows that f_μ provides a good smooth approximation for the nonsmooth function f .

Proposition 2.1.2. *Let $D := \max\{q(u) \mid u \in Q\}$. Then*

$$f_\mu(x) \leq f(x) \leq f_\mu(x) + \mu D \text{ for all } x \in \mathbb{R}^d.$$

Proposition 2.1.3. *Assume further that in the setting of Theorem 2.1.1, we have $q(u) := \frac{1}{2}\|u - u_0\|^2$, where $u_0 \in Q$, and $\phi(u) = \langle b, u \rangle$ for all $b \in \mathbb{R}^m$. Then*

$$f_\mu(x) = \frac{\|Ax - b\|}{2\mu} + \langle Ax - b, u_0 \rangle - \frac{\mu}{2} \left[d\left(u_0 + \frac{Ax - b}{\mu}; Q\right) \right].$$

Furthermore, f_μ is differentiable with $\nabla f_\mu(x) = A^T u_\mu(x)$, where $u_\mu(x)$ can be represented in terms of the Euclidean projection

$$u_\mu(x) := P\left(u_0 + \frac{Ax - b}{\mu}; Q\right),$$

and $\nabla f_\mu(x)$ is Lipschitz continuous with constant $L_\mu = \frac{1}{\mu}\|A\|^2$.

Example 2.1.4. Consider the function $f(x) = |x|$ which is not differentiable at 0 and can be represented as

$$f(x) = \max\{ux \mid u \in [-1, 1]\}, \quad x \in \mathbb{R}.$$

Using the prox-function $q(u) = \frac{1}{2}u^2$ gives us

$$f_\mu(x) = \max\left\{ux - \frac{\mu}{2}u^2 \mid u \in [-1, 1]\right\},$$

and

$$u_\mu(x) = \begin{cases} -1 & x \leq -\mu, \\ \frac{x}{\mu} & -\mu < x < \mu, \\ 1 & x \geq \mu. \end{cases}$$

Therefore,

$$f_\mu(x) = \begin{cases} -x - \frac{\mu}{2} & x \leq -\mu, \\ \frac{x^2}{2\mu} & -\mu < x < \mu, \\ x - \frac{\mu}{2} & x \geq \mu. \end{cases}$$

In this case the Lipschitz constant of ∇f_μ is $L_\mu = \frac{1}{\mu}$.

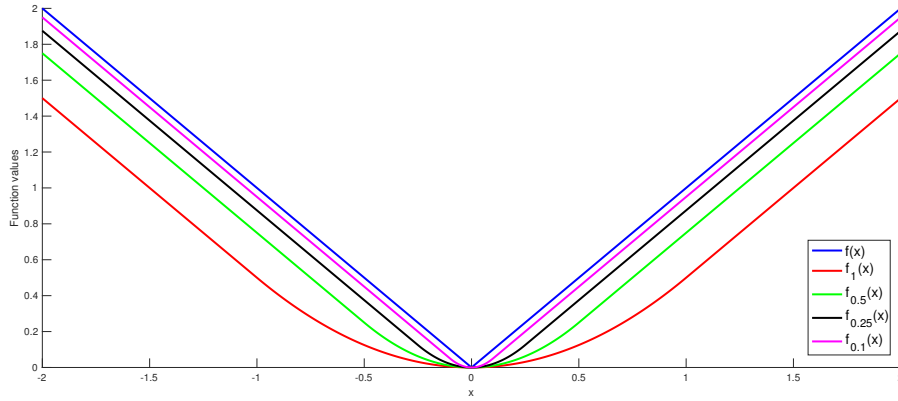


FIGURE 2.1. Nesterov's smoothing for $f(x) = |x|$.

Example 2.1.5. Given any $a \in \mathbb{R}^d$ and $\mu > 0$, a Nesterov smoothing approximation of $\varphi(x) := \|x - a\|$ has the representation

$$\varphi_\mu(x) = \frac{1}{2\mu} \|x - a\|^2 - \frac{\mu}{2} \left[d\left(\frac{x - a}{\mu}; \mathbb{B}\right) \right]^2.$$

Moreover, $\nabla \varphi_\mu(x) = P\left(\frac{x-a}{\mu}; \mathbb{B}\right)$ and

$$\varphi_\mu(x) \leq \varphi(x) \leq \varphi_\mu(x) + \frac{\mu}{2},$$

where \mathbb{B} is the closed unit ball of \mathbb{R}^d .

2.2. Overview of the DCA

Consider the difference of two convex functions $g - h$ on a Hilbert space H and assume that $g: H \rightarrow \overline{\mathbb{R}}$ is extended-real-valued while $h: H \rightarrow \mathbb{R}$ is real-valued on H . Then a

general problem of *DC optimization* is defined by

$$(2.2.5) \quad \text{minimize } f(x) := g(x) - h(x), \quad x \in H.$$

Note that problem (2.2.5) is written in the unconstrained format, but—due to the allowed infinite value for g —it actually contains the domain constraint $x \in \text{dom}(g) := \{u \in H \mid g(u) < \infty\}$. Furthermore, the explicit constraints of the type $x \in \Omega$ given by a nonempty convex set $\Omega \subset H$ can be incorporated into the format of (2.2.5) via the indicator function $\delta_\Omega(x)$ of Ω , which equals 0 for $x \in \Omega$ and ∞ otherwise.

The framework for DC programming and DC Algorithms (DCA) were introduced by Tao and An in the late 1990's. In this section, we consider the essential elements of DCA in finite dimensional case, \mathbb{R}^d . Firstly, we shall briefly overview two algorithms of the DCA type to solve DC problems (2.2.5) and the convergence of the DCA. Then we present numerical examples illustrating both algorithms. The reader can find more details and further developments in [45, 46].

Algorithm 1 DCA-1

INPUT: $x_0 \in \mathbb{R}^d, k = 0$.

Repeat

Find $y_k \in \partial h(x_k)$.

Find $x_{k+1} \in \partial g^*(y_k)$.

$k := k + 1$

Until convergence of $\{x_k\}$.

OUTPUT: x_{k+1}

Before presenting Algorithm 2, we shall do an analysis of the first algorithm. Since the convex function $h: \mathbb{R}^d \rightarrow \mathbb{R}$ in (2.2.5) is real-valued on the whole space \mathbb{R}^d , we always have $\partial h(x) \neq \emptyset$ for all $x \in \mathbb{R}^d$. At the same time, the other convex function $g: \mathbb{R}^d \rightarrow \overline{\mathbb{R}}$ in (2.2.5) is generally extended-real-valued, and so the subdifferential of its conjugate g^* may be empty. Let us present an efficient condition that excludes this possibility. Recall that a function $g: \mathbb{R}^d \rightarrow \overline{\mathbb{R}}$ is *coercive* if

$$\lim_{\|x\| \rightarrow \infty} \frac{g(x)}{\|x\|} = \infty.$$

Proposition 2.2.1. *Let $g: \mathbb{R}^d \rightarrow \overline{\mathbb{R}}$ be a proper lower semi-continuous convex function.*

Then

$$\partial g(\mathbb{R}^d) := \bigcup_{x \in \mathbb{R}^d} \partial g(x) = \text{dom } \partial g^* := \{y \in \mathbb{R}^d \mid \partial g^*(y) \neq \emptyset\}.$$

Proof. Let $x \in \mathbb{R}^d$ and $y \in \partial g(x)$. Then $x \in \partial g^*(y)$ which implies $\partial g^*(y) \neq \emptyset$, and so $y \in \text{dom } \partial g^*$. The opposite inclusion is obvious. \square

Proposition 2.2.2. *Let $g: \mathbb{R}^d \rightarrow \overline{\mathbb{R}}$ be a proper, lower semi-continuous, and convex function. If in addition g is coercive, then $\partial g^*(v) \neq \emptyset$ for all $v \in \mathbb{R}^d$.*

Proof. Since g is proper, the conjugate function g^* takes values in $[-\infty, \infty]$ being convex on \mathbb{R}^d . Taking into account that g is also lower semi-continuous and invoking the aforementioned biconjugate relationship, we find $w \in \mathbb{R}^d$ and $c \in \mathbb{R}$ such that

$$(2.2.6) \quad c + \langle w, x \rangle \leq g(x) \quad \text{for all } x \in \mathbb{R}^d.$$

The coercivity property of g ensures the existence of $\eta > 0$ for which

$$\|x\|(\|w\| + 1) \leq g(x) \quad \text{whenever } \|x\| \geq \eta.$$

Furthermore, it follows that

$$\sup \{ \langle v, x \rangle - g(x) \mid \|x\| \geq \eta \} \leq -\|x\| \quad \text{for any } v \in \mathbb{R}^d.$$

By using (2.2.6), we arrive at the estimates

$$\sup_{v \in \mathbb{R}^d} \{ \langle v, x \rangle - g(x) \mid \|x\| \leq \eta \} \leq \sup_{v \in \mathbb{R}^d} \{ \langle v, x \rangle - \langle w, x \rangle - c \mid \|x\| \leq \eta \} < \infty.$$

It follows that $g^*(v) < \infty$, and therefore $\text{dom}(g^*) = \mathbb{R}^d$. Since g^* is a convex function with finite values, it is continuous on \mathbb{R}^d and hence $\partial g^*(v) \neq \emptyset$ for all $v \in \mathbb{R}^d$. \square

The following proposition gives us a two-sided relationship between the Fenchel conjugates and subgradients of convex functions.

Proposition 2.2.3. *Let $\varphi: \mathbb{R}^d \rightarrow \overline{\mathbb{R}}$ be a proper, lower semi-continuous, and convex function. Then $v \in \partial\varphi^*(y)$ if and only if*

$$(2.2.7) \quad v \in \operatorname{argmin} \{ \varphi(x) - \langle y, x \rangle \mid x \in \mathbb{R}^d \}.$$

Furthermore, we have that $w \in \partial\varphi(x)$ if and only if

$$(2.2.8) \quad w \in \operatorname{argmin} \{ \varphi^*(y) - \langle x, y \rangle \mid y \in \mathbb{R}^d \}.$$

Proof. To verify the first assertion, suppose that (2.2.7) is satisfied. Then we get that $0 \in \partial\psi(v)$, where $\psi(x) := \varphi(x) - \langle y, x \rangle$ as $x \in \mathbb{R}^d$. This implies the following

$$0 \in \partial\varphi(v) - y,$$

and hence $y \in \partial\varphi(v)$, which is equivalent to $v \in \partial\varphi^*(y)$ due to the biconjugate relationship $\varphi^{**} = \varphi$ valued under the assumptions made.

To prove the opposite direction, assuming $v \in \partial\varphi^*(y)$ gives us by the proof above that $0 \in \partial\psi(v)$, which clearly yields (2.2.7) and thus justifies the first assertion.

To verify the second assertion, suppose that (2.2.8) holds and then get $0 \in \partial\psi(w)$, where $\psi(y) := \varphi^*(y) - \langle x, y \rangle$ as $y \in \mathbb{R}^d$. This clearly implies that

$$0 \in \partial\varphi^*(w) - x,$$

and hence $x \in \partial\varphi^*(w)$, which is equivalent to $w \in \partial\varphi(x)$ due to the biconjugate relationship. The proof of the opposite implication in (2.2.8) is similar to the one given above. □

Proposition 2.2.4. *Let $h: \mathbb{R}^d \rightarrow \overline{\mathbb{R}}$ be γ -convex with $\bar{x} \in \operatorname{dom} h$. Then $v \in \partial h(\bar{x})$ if and only if*

$$\langle v, x - \bar{x} \rangle + \frac{\gamma}{2} \|x - \bar{x}\|^2 \leq h(x) - h(\bar{x}).$$

Proof. Since h is γ -convex, we can find a convex function $k: \mathbb{R}^d \rightarrow \overline{\mathbb{R}}$ such that $k(x) = h(x) - \frac{\gamma}{2} \|x\|^2$. For any $v \in \partial h(\bar{x})$, we have $v \in \partial\phi(\bar{x})$, where $\phi(x) = k(x) + \frac{\gamma}{2} \|x\|^2$

for $x \in \mathbb{R}^d$. The subdifferential sum rule gives us

$$v \in \partial k(\bar{x}) + \gamma \bar{x} \text{ or } v - \gamma \bar{x} \in \partial k(\bar{x}).$$

Thus, using the definition of the subdifferential of k we obtain

$$\langle v - \gamma \bar{x}, x - \bar{x} \rangle \leq k(x) - k(\bar{x}) \text{ for all } x \in \mathbb{R}^d.$$

That is,

$$\begin{aligned} \langle v, x - \bar{x} \rangle &\leq \gamma \langle \bar{x}, x \rangle - \gamma \langle \bar{x}, \bar{x} \rangle + h(x) - \frac{\gamma}{2} \|x\|^2 - (h(\bar{x}) - \frac{\gamma}{2} \|\bar{x}\|^2) \\ &\leq h(x) - h(\bar{x}) - \frac{\gamma}{2} (\|x\|^2 - 2\langle x, \bar{x} \rangle + \|\bar{x}\|^2) \\ &= h(x) - h(\bar{x}) - \frac{\gamma}{2} \|x - \bar{x}\|^2. \end{aligned}$$

This completes the proof. \square

Proposition 2.2.5. *Consider the function f defined in (2.2.5) and consider the sequence $\{x_k\}$ generated by Algorithm 1. Suppose that g is γ_1 -convex and h is γ_2 -convex. Then*

$$(2.2.9) \quad f(x_k) - f(x_{k+1}) \geq \frac{\gamma_1 + \gamma_2}{2} \|x_{k+1} - x_k\|^2 \text{ for all } k \in \mathbb{N}.$$

Proof. Starting with $y_k \in \partial h(x_k)$, Proposition 2.2.4 gives us

$$\langle y_k, x - x_k \rangle + \frac{\gamma_2}{2} \|x - x_k\|^2 \leq h(x) - h(x_k) \text{ for all } x \in \mathbb{R}^d.$$

In particular, we can let $x = x_{k+1}$

$$\langle y_k, x_{k+1} - x_k \rangle + \frac{\gamma_2}{2} \|x_{k+1} - x_k\|^2 \leq h(x_{k+1}) - h(x_k).$$

Moreover, $x_{k+1} \in \partial g^*(y_k)$ is equivalent to $y_k \in \partial g(x_{k+1})$. Similarly, we have

$$\langle y_k, x_k - x_{k+1} \rangle + \frac{\gamma_1}{2} \|x_k - x_{k+1}\|^2 \leq g(x_k) - g(x_{k+1}).$$

Adding these inequalities gives us the inequality in (2.2.9). \square

Lemma 2.2.6. *Suppose that $h : \mathbb{R}^d \rightarrow \mathbb{R}$ is a convex function. If $w_k \in \partial h(x_k)$ and $\{x_k\}$ is a bounded sequence, then $\{w_k\}$ is also bounded.*

Proof. Let us first fix $\bar{x} \in \mathbb{R}^d$. Then there exists $\ell > 0$ and $\delta > 0$ such that

$$|h(x) - h(y)| \leq \ell \|x - y\|$$

whenever $x, y \in \mathbb{B}(\bar{x}; \delta)$ by the fact that h is locally Lipschitz continuous around \bar{x} . So it follows that whenever $w \in \partial h(u)$ for $u \in \mathbb{B}(\bar{x}; \frac{\delta}{2})$, we have $\|w\| \leq \ell$. Indeed,

$$\langle w, x - u \rangle \leq h(x) - h(u) \text{ for all } x \in \mathbb{R}^d.$$

Now, whenever $\|x - u\| \leq \gamma$, for a chosen sufficiently small $\gamma > 0$ such that $\mathbb{B}(u; \gamma) \subset \mathbb{B}(\bar{x}; \delta)$, we have that

$$\langle w, x - u \rangle \leq h(x) - h(u) \leq \ell \|x - u\|$$

Hence, $\|w\| \leq \ell$.

By contradiction, assume that $\{w_k\}$ is not bounded. Then trivially we can assume that $\|w_k\| \rightarrow \infty$. Now, since $\{x_k\}$ is bounded, it has a subsequence $\{x_{k_p}\}$ that converges to $x_0 \in \mathbb{R}^d$. Letting $\ell > 0$ be the Lipschitz constant of f around x_0 , it follows that

$$\|w_{k_p}\| \leq \ell \text{ for sufficiently large } p,$$

a contradiction. Hence, $\{w_k\}$ is bounded. □

We next recall a *stationary point* of a DC function and present a theorem, which can be derived from [45, 46], and summarize some convergence results of the DCA. Deeper studies of the convergence of this algorithm and its generalizations involving the Kurdyka-Lojasiewicz (KL) inequality are given in [3, 4].

Definition 2.2.7. *An element $\bar{x} \in \mathbb{R}^d$ is called a stationary point of the function f defined by (2.2.5) if $\partial g(\bar{x}) \cap \partial h(\bar{x}) \neq \emptyset$.*

Theorem 2.2.8. *Let f be a DC function taken from (2.2.5), and let $\{x_k\}$ be an iterative sequence generated by Algorithm 1. The following assertions hold:*

(a) The sequence $\{f(x_k)\}$ is always monotone decreasing.

(b) Suppose that f is bounded from below, g is lower semi-continuous and γ_1 -convex, and that h is γ_2 -convex with $\gamma_1 + \gamma_2 > 0$. If $\{x_k\}$ is bounded, then the limit of any convergent subsequence of $\{x_k\}$ is a stationary point of f .

Proof. (a) Using (2.2.9) we can see that $\{f(x_k)\}$ is a decreasing sequence.

(b) Since f is bounded from below and by part (a), it follows that $\{f(x_k)\}$ will converge to a real number. Then, $f(x_k) - f(x_{k+1}) \rightarrow 0$ as $k \rightarrow \infty$ and hence $\|x_{k+1} - x_k\| \rightarrow 0$ by (2.2.9). Suppose that $x_{k_\ell} \rightarrow x^*$ as $\ell \rightarrow \infty$. By definition, we have that

$$y_k \in \partial g(x_{k+1}) \text{ for all } k \in \mathbb{N}.$$

It follows from Lemma 2.2.6 that $\{y_k\}$ is also a bounded sequence since $\{x_k\}$ is bounded. Furthermore, by extracting a further subsequence, we can assume without loss of generality that $y_{k_\ell} \rightarrow y^*$ as $\ell \rightarrow \infty$. Then we have that

$$y^* \in \partial h(x^*),$$

since $y_{k_\ell} \in \partial h(x_{k_\ell})$ for all $\ell \in \mathbb{N}$. Indeed, by the definition of $\partial h(x^*)$ we get

$$(2.2.10) \quad \langle y_{k_\ell}, x - x_{k_\ell} \rangle \leq h(x) - h(x_{k_\ell}) \text{ for all } x \in \mathbb{R}^d, \ell \in \mathbb{N}.$$

Replacing x by x^* , we have

$$\langle y_{k_\ell}, x^* - x_{k_\ell} \rangle \leq h(x^*) - h(x_{k_\ell}).$$

Rearranging the above equation gives us that $h(x_{k_\ell}) \leq \langle y_{k_\ell}, x_{k_\ell} - x^* \rangle + h(x^*)$, and hence $\limsup h(x_{k_\ell}) \leq h(x^*)$. Then, the lower semicontinuity of h implies $h(x_{k_\ell}) \rightarrow h(x^*)$. Furthermore, by letting $\ell \rightarrow \infty$ in (2.2.10) we get $y^* \in \partial h(x^*)$. Lastly, we have that $x_k \rightarrow x^*$ since $\|x_{k+1} - x_k\| \rightarrow 0$ and $x_{k_\ell} \rightarrow x^*$ and from the relation $y_{k_\ell} \in \partial g(x_{k_\ell+1})$, we have $y^* \in \partial g(x^*)$ by a similar argument. Therefore, x^* is a stationary point of f . Which completes the proof. \square

In many practical applications of Algorithm 1, for a given DC decomposition of f , it is possible to find subgradient vectors from $\partial h(x_k)$ based on available formulas and calculus rules of convex analysis. However, it may not be possible to explicitly calculate an element of $\partial g^*(y_k)$. Such a situation requires either constructing a more suitable DC decomposition of f , or finding $x_{k+1} \in \partial g^*(y_k)$ by using the description of Proposition 2.2.3. This leads us to the following modified version of our second algorithm.

Algorithm 2 DCA-2

INPUT: $x_0 \in \text{dom } g$, $k = 0$

Repeat

 Find $y_k \in \partial h(x_k)$.

 Find x_{k+1} by solving the problem

$$\text{minimize } \varphi_k(x) := g(x) - \langle y_k, x \rangle, \quad x \in \mathbb{R}^d.$$

$k := k + 1$

Until convergence of $\{x_k\}$.

OUTPUT: x_{k+1}

Let us now present two examples illustrating the performances of Algorithms 1 and 2.

The first example concerns with a polynomial function of one variable.

Example 2.2.9. Consider the function $f: \mathbb{R} \rightarrow \mathbb{R}$ given by

$$f(x) := x^4 - 2x^2 + 2x - 3 \quad \text{for } x \in \mathbb{R}.$$

This function admits the DC representation $f = g - h$ with $g(x) := x^4$ and $h(x) := 2x^2 - 2x + 3$. To minimize f , apply first the *gradient method* with constant stepsize. Clearly the derivative of f is $f'(x) = 4x^3 - 4x + 2$ and picking any starting point $x_0 \in \mathbb{R}$, we get the sequence of iterates

$$x_{k+1} = x_k - t(4x_k^3 - 4x_k + 2) \quad \text{for } k = 0, 1, \dots$$

constructed by the gradient method with stepsize $t > 0$. The usage of the DC *Algorithm 1* (DCA-1) gives us $y_k = \nabla h(x_k) = 4x_k - 2$ and then $g^*(x) = 3(x/4)^{4/3}$ with $\nabla g^*(x) = (x/4)^{1/3}$. Thus the iterates of DCA-1 are as follows

$$x_{k+1} = \nabla g^*(y_k) = \left(\frac{y_k}{4}\right)^{1/3} = \left(\frac{4x_k - 2}{4}\right)^{1/3} = \left(\frac{2x_k - 1}{2}\right)^{1/3}, \quad k = 0, 1, \dots$$

Figure 2.2 provides the visualization and comparison between the DCA-1 and the gra-

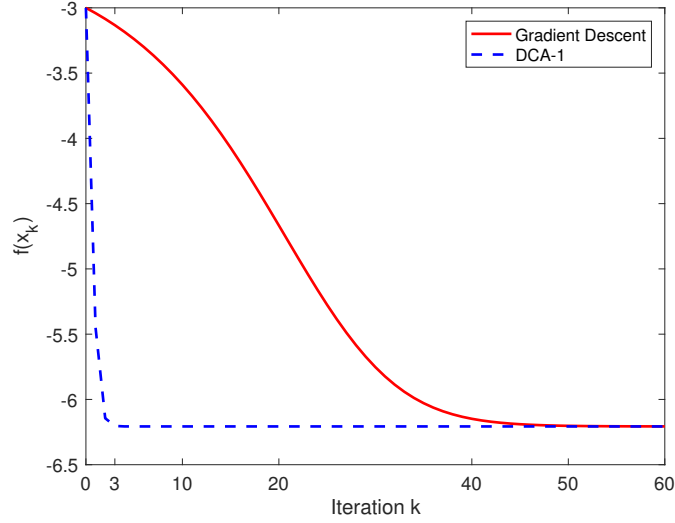


FIGURE 2.2. Convergence of the gradient method and DCA-1.

dent method. It shows that for $x_0 = 0$ and $t = 0.01$ the DCA-1 exhibits much faster convergence.

The next two-dimensional example illustrates the performance of the DCA-2.

Example 2.2.10. Consider the nonsmooth optimization problem defined by

$$\text{minimize } f(x_1, x_2) := x_1^4 + x_2^2 - 2x_1^2 - |x_2| \text{ over } x = (x_1, x_2) \in \mathbb{R}^2.$$

The graph of the function f is depicted in Figure 2.3a. Observe that this function has four global minimizers, which are $(1, 0.5)$, $(1, -0.5)$, $(-1, 0.5)$, and $(-1, -0.5)$. It is easy to see that f admits a DC representation $f = g - h$ with $g(x_1, x_2) := x_1^4 + x_2^2$ and $h(x_1, x_2) := 2x_1^2 + |x_2|$. We get the gradient $\nabla g(x) = [4x_1^3, 2x_2]^T$ and the Hessian

$$\nabla^2 g(x) = \begin{bmatrix} 12x_1^2 & 0 \\ 0 & 2 \end{bmatrix},$$

while an explicit formula to calculate $\partial g^*(y_k)$ is not available. Let us apply the DCA-2 to solve this problem. The subdifferential of h is calculated by

$$\partial h(x) = [4x_1, \text{sign}(x_2)]^T \text{ for any } x = (x_1, x_2) \in \mathbb{R}^2.$$

Having y_k , we proceed with solving the subproblem

$$(2.2.11) \quad \text{minimize } \varphi_k(x) := g(x) - \langle y_k, x \rangle \quad \text{over } x \in \mathbb{R}^2$$

by the classical Newton method with $\nabla^2 \varphi_k(x) = \nabla^2 g(x)$ and observe that the DCA-2 shows its superiority in convergence with different choices of initial points. Figure 2.3b presents the results of computation by using the DCA-2 with the starting point $x_0 = (-2, 2)$ and employing the Newton method with $\varepsilon = 10^{-8}$ to solve subproblem (2.2.11).

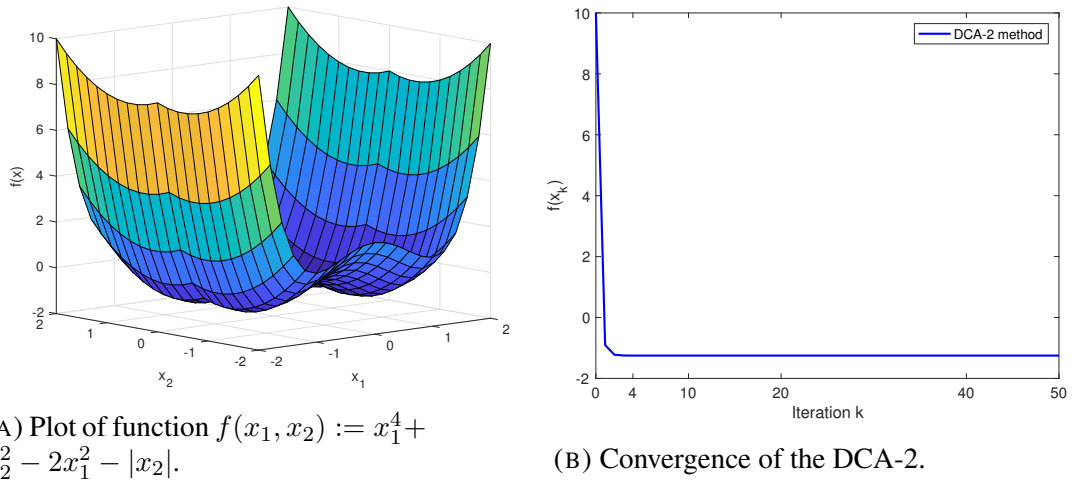


FIGURE 2.3. Graph and performance of Example 2.2.10.

In summary, both algorithms work well for the case when we do not have convexity in the objective functions. In particular, Algorithm 2 is very useful when it's not easy to explicitly compute an element of $\partial g^*(y_k)$.

2.3. A Penalty Method via Distance Functions

In this section, we study a penalty method using distance functions for solving constrained optimization problems and apply them specifically to DC programming. This method is based on the *quadratic penalty method*; see, e.g, [15, 38]. This section presents the work we have done in [34]. The reader can find more details in there. Let $f: \mathbb{R}^d \rightarrow \mathbb{R}$ be a function and let Ω_i for $i = 1, \dots, q$ be nonempty closed subsets of

\mathbb{R}^d with $\bigcap_{i=1}^q \Omega_i \neq \emptyset$. Consider the optimization problem

$$(2.3.12) \quad \begin{aligned} & \text{minimize} && f(x) \\ & \text{subject to} && x \in \bigcap_{i=1}^q \Omega_i. \end{aligned}$$

Let us first study the relation between this problem and the unconstrained one given by

$$(2.3.13) \quad \text{minimize } f_\lambda(x) = f(x) + \frac{\lambda}{2} \sum_{i=1}^q [d(x; \Omega_i)]^2, \quad x \in \mathbb{R}^d.$$

The theorem below provides a relation between optimal solutions of the constrained optimization problem (2.3.12) and the unconstrained optimization problem (2.3.13) obtained by a penalty method based on distance functions. The proof follows [38, Theorem 17.1].

Theorem 2.3.1. *Consider (2.3.12) in which $f: \mathbb{R}^d \rightarrow \mathbb{R}$ is a lower semi-continuous function. Suppose that (2.3.12) has an optimal solution. If $\lim_{n \rightarrow \infty} \lambda_n = \infty$ and $x_n \in \mathbb{R}^d$ is an absolute minimizer of the function f_{λ_n} defined in (2.3.13) for all $n \in \mathbb{N}$, then every subsequential limit of $\{x_n\}$ is a solution of (2.3.12).*

Proof. Let $\bar{x} \in \mathbb{R}^d$ be an optimal solution of (2.3.12). That is, $\bar{x} \in \Omega_i$ for $i = 1, \dots, q$ and

$$f(\bar{x}) \leq f(x) \text{ whenever } x \in \Omega_i \text{ for all } i = 1, \dots, q.$$

Since $x_n \in \mathbb{R}^d$ is an absolute minimizer of the function f_{λ_n} ,

$$f_{\lambda_n}(x_n) \leq f_{\lambda_n}(\bar{x}).$$

This implies, with the observation that $d(\bar{x}; \Omega_i) = 0$ for $i = 1, \dots, q$, that

$$(2.3.14) \quad f(x_n) + \frac{\lambda_n}{2} \sum_{i=1}^q [d(x_n; \Omega_i)]^2 \leq f(\bar{x}).$$

Then, $\sum_{i=1}^q [d(x_n; \Omega_i)]^2 \leq \frac{2}{\lambda_n} (f(\bar{x}) - f(x_n))$ for all $n \in \mathbb{N}$.

Let $x^* \in \mathbb{R}^d$ be a subsequential limit of $\{x_n\}$. Without loss of generality, we can assume that $\lim_{n \rightarrow \infty} x_n = x^*$. By the continuity of the distance function and the lower

semicontinuity of f ,

$$\sum_{i=1}^q [d(x^*; \Omega_i)]^2 = \lim_{n \rightarrow \infty} \sum_{i=1}^q [d(x_n; \Omega_i)]^2 \leq \liminf_{n \rightarrow \infty} \frac{2}{\lambda_n} \left(f(\bar{x}) - f(x_n) \right) \leq 0.$$

It follows that $d(x^*; \Omega_i) = 0$, and so $x^* \in \Omega_i$ for $i = 1, \dots, q$. In addition, by (2.3.14) and the lower semicontinuity of f we have

$$f(x^*) \leq \liminf_{n \rightarrow \infty} f(x_n) \leq \liminf_{n \rightarrow \infty} \left(f(x_n) + \frac{\lambda_n}{2} \sum_{i=1}^q [d(x_n; \Omega_i)]^2 \right) \leq f(\bar{x}).$$

Therefore, x^* is an optimal solution of (2.3.12). \square

Now we discuss a direct consequence of Theorem 2.3.1 that will be used in the Chapter 4. Let $F: \mathbb{R}^{k \times d} \rightarrow \mathbb{R}$ be a function and let Ω_i^ℓ for $\ell = 1, \dots, k$ and $i = 1, \dots, q$ be nonempty closed subsets of \mathbb{R}^d . Consider the problem

$$(2.3.15) \quad \begin{aligned} & \text{minimize} && F(x^1, \dots, x^k) \\ & \text{subject to} && x^\ell \in \bigcap_{i=1}^q \Omega_i^\ell, x^\ell \in \mathbb{R}^d \text{ for } \ell = 1, \dots, k. \end{aligned}$$

We clarify the relation between this problem and the unconstrained problem given by

$$(2.3.16) \quad \text{minimize} \quad F_\lambda(x^1, \dots, x^k) = F(x^1, \dots, x^k) + \frac{\lambda}{2} \sum_{\ell=1}^k \sum_{i=1}^q [d(x^\ell; \Omega_i^\ell)]^2$$

for $x^\ell \in \mathbb{R}^d$ and $\ell = 1, \dots, k$. In what follows, we identify $\mathbf{X} = (x^1, \dots, x^k) \in \mathbb{R}^{k \times d}$ with the matrix $\mathbf{X} \in \mathbb{R}^{k \times d}$, whose ℓ^{th} row is x^ℓ for $\ell = 1, \dots, k$.

Corollary 2.3.2. *Consider (2.3.15) in which $F: \mathbb{R}^{k \times d} \rightarrow \mathbb{R}$ is a lower semi-continuous function. Suppose that (2.3.15) has an optimal solution. If $\lim_{n \rightarrow \infty} \lambda_n = \infty$ and $X_n = (x_n^1, \dots, x_n^k) \in \mathbb{R}^{k \times d}$ is an absolute minimizer of the function F_{λ_n} , then every subsequential limit of $\{X_n\}$ is a solution of (2.3.15).*

Proof. Let $\mathbf{X} = (x^1, \dots, x^k) \in \mathbb{R}^{k \times d}$ and let $\Omega_i = \Omega_i^1 \times \dots \times \Omega_i^k \subset \mathbb{R}^{k \times d}$ for $i = 1, \dots, q$.

Note that

$$\bigcap_{i=1}^q \Omega_i = \bigcap_{i=1}^q \prod_{\ell=1}^k \Omega_i^\ell = \prod_{\ell=1}^k \bigcap_{i=1}^q \Omega_i^\ell.$$

It follows that $x^\ell \in \bigcap_{i=1}^q \Omega_i^\ell$ for $\ell = 1, \dots, k$ if and only if $\mathbf{X} \in \bigcap_{i=1}^q \Omega_i$, and thus (2.3.15) reduces to the following optimization problem

$$\begin{aligned} & \text{minimize} && F(\mathbf{X}) \\ & \text{subject to} && \mathbf{X} \in \bigcap_{i=1}^q \Omega_i. \end{aligned}$$

Based on (4.1.26), we can rewrite the objective function F_λ in (2.3.16) as follows

$$F_\lambda(\mathbf{X}) = F(\mathbf{X}) + \frac{\lambda}{2} \sum_{i=1}^q [d(\mathbf{X}; \Omega_i)]^2.$$

The conclusion now follows directly from Theorem 2.3.1. \square

Let us continue with a known result on DC decompositions of squared distance functions. The proof of the following result can be found in [36, Proposition 5.1].

Proposition 2.3.3. *Let Ω be a nonempty closed set in \mathbb{R}^d (not necessarily convex).*

Define the function

$$\varphi_\Omega(x) = \sup \{ \langle 2x, w \rangle - \|w\|^2 \mid w \in \Omega \} = 2 \sup \{ \langle x, w \rangle - \frac{1}{2} \|w\|^2 \mid w \in \Omega \}.$$

Then we have the following conclusions:

(a) *The function φ_Ω is always convex. If we assume in addition that Ω is convex, then φ_Ω is differentiable with $\nabla \varphi_\Omega(x) = 2P(x; \Omega)$.*

(b) *The function $f(x) = [d(x; \Omega)]^2$ is a DC function with $f(x) = \|x\|^2 - \varphi_\Omega(x)$ for all $x \in \mathbb{R}^d$.*

We now consider (2.3.1) in which $f(x) = g(x) - h(x)$ is a DC function where $g, h: \mathbb{R}^d \rightarrow \mathbb{R}$ are convex functions. We also assume additionally that all constraint sets are convex and satisfy $\bigcap_{i=1}^q \text{ri}(\Omega_i) \neq \emptyset$. By [28, Theorem 5.3], this condition ensures that

$$(2.3.17) \quad N(\bar{x}; \bigcap_{i=1}^q \Omega_i) = \sum_{i=1}^q N(\bar{x}; \Omega_i) \text{ for every } \bar{x} \in \bigcap_{i=1}^q \Omega_i.$$

Recall from [45] that an element $\bar{x} \in \mathbb{R}^d$ is a *critical point* of a DC function f with DC decomposition $f = g - h$ if $\partial g(\bar{x}) \cap \partial h(\bar{x}) \neq \emptyset$. Observe that (2.3.1) can be written as

an unconstrained optimization problem using the indicator function as follows

$$\text{minimize } \left(g(x) + \delta(x, \bigcap_{i=1}^q \Omega_i) \right) - h(x), x \in \mathbb{R}^d.$$

Define $v(x) = g(x) + \delta(x, \bigcap_{i=1}^q \Omega_i)$ for $x \in \mathbb{R}^d$. By (2.3.17),

$$\partial v(\bar{x}) = \partial g(\bar{x}) + N(\bar{x}; \bigcap_{i=1}^q \Omega_i) = \partial g(\bar{x}) + \sum_{i=1}^q N(\bar{x}; \Omega_i) \text{ for all } \bar{x} \in \mathbb{R}^d.$$

Thus, we call an element $\bar{x} \in \mathbb{R}^d$ a critical point of (2.3.1) if

$$(2.3.18) \quad \left(\partial g(\bar{x}) + \sum_{i=1}^q N(\bar{x}; \Omega_i) \right) \cap \partial h(\bar{x}) \neq \emptyset.$$

The objective function of (2.3.13) now becomes

$$f_\lambda(x) = g(x) + \frac{\lambda}{2} \sum_{i=1}^q d(x; \Omega_i)^2 - h(x).$$

Using Proposition 3.2, we have

$$(2.3.19) \quad f_\lambda(x) = \left(g(x) + \frac{\lambda q}{2} \|x\|^2 \right) - \left(h(x) + \frac{\lambda}{2} \sum_{i=1}^q \varphi_{\Omega_i}(x) \right) = \tilde{g}_\lambda(x) - \tilde{h}_\lambda(x),$$

where \tilde{g}_λ and \tilde{h}_λ are functions defined on \mathbb{R}^d by

$$\tilde{g}_\lambda(x) = g(x) + \frac{\lambda q}{2} \|x\|^2 \text{ and } \tilde{h}_\lambda(x) = h(x) + \frac{\lambda}{2} \sum_{i=1}^q \varphi_{\Omega_i}(x), x \in \mathbb{R}^d.$$

Proposition 2.3.4. *Suppose that $\lim_{n \rightarrow \infty} \lambda_n = \infty$ and x_n is a critical point of the DC function $f_{\lambda_n} = \tilde{g}_{\lambda_n} - \tilde{h}_{\lambda_n}$ given in (2.3.19). Then every subsequential limit of the sequence $\{x_n\}$ is a critical point of (2.3.1).*

Proof. Since x_n is a critical point of f_{λ_n} and by Proposition 2.3.3, there exist $v_n \in \partial g(x_n)$ and $w_n \in \partial h(x_n)$ such that

$$(2.3.20) \quad v_n + \lambda_n q x_n = w_n + \lambda_n \sum_{i=1}^q P(x_n; \Omega_i).$$

Let \bar{x} be a subsequential limit of $\{x_n\}$. Without loss of generality, we can assume that $\{x_n\}$ converges to \bar{x} . Since any finite convex function is locally Lipschitz continuous, we can assume that both g and h are locally Lipschitz continuous around \bar{x} with Lipschitz constant $L > 0$. Then

$$(2.3.21) \quad \|v_n\| \leq L \text{ and } \|w_n\| \leq L \text{ for sufficiently large } n.$$

By (2.3.20), (2.3.21) and the assumption that $\lambda_n \rightarrow \infty$ as $n \rightarrow \infty$,

$$\sum_{i=1}^q (x_n - P(x_n; \Omega_i)) = \frac{1}{\lambda_n} (w_n - v_n) \rightarrow 0 \text{ as } n \rightarrow \infty.$$

Letting $n \rightarrow \infty$ yields $\sum_{i=1}^q (\bar{x} - P(\bar{x}; \Omega_i)) = 0$, due to the continuity of projection operators onto convex sets. Note also that $\lambda_n \sum_{i=1}^q (x_n - P(x_n; \Omega_i)) \in \sum_{i=1}^q N(x_n; \Omega_i)$.

This implies

$$w_n - v_n \in \sum_{i=1}^q N(x_n; \Omega_i) = N(x_n; \bigcap_{i=1}^q \Omega_i).$$

By (2.3.21), we can assume without loss of generality that $v_n \rightarrow \bar{v}$ and $w_n \rightarrow \bar{w}$ as $n \rightarrow \infty$. Then by passing to a limit, we have

$$\bar{w} - \bar{v} \in N(\bar{x}; \bigcap_{i=1}^q \Omega_i) = \sum_{i=1}^q N(\bar{x}; \Omega_i).$$

Observe also that $\bar{v} \in \partial g(\bar{x})$ and $\bar{w} \in \partial h(\bar{x})$. Therefore, (2.3.18) is satisfied and thus \bar{x} is a critical point of (2.3.1). \square

We continue by considering (2.3.15) in which

$$F(x^1, \dots, x^k) = G(x^1, \dots, x^k) - H(x^1, \dots, x^k)$$

is a DC function, where $G, H: \mathbb{R}^{k \times d} \rightarrow \mathbb{R}$ are convex functions. From the proof of Corollary 2.3.2, we can rewrite (2.3.15) as

$$\begin{aligned} & \text{minimize} && F(\mathbf{X}) = G(\mathbf{X}) - H(\mathbf{X}) \\ & \text{subject to} && \mathbf{X} \in \bigcap_{i=1}^q \Omega_i. \end{aligned}$$

Recall that a point $\bar{\mathbf{X}} = (\bar{x}^1, \dots, \bar{x}^k)$ is called a critical point of this problem if

$$\left(\partial G(\bar{\mathbf{X}}) + \sum_{i=1}^q N(\bar{\mathbf{X}}; \Omega_i) \right) \cap \partial H(\bar{\mathbf{X}}) \neq \emptyset,$$

where $N(\bar{\mathbf{X}}; \Omega_i) = N(\bar{x}^1; \Omega_i^1) \times \dots \times N(\bar{x}^k; \Omega_i^k)$. For $\Omega \subset \mathbb{R}^{k \times d}$, based on Frobenious norm, define

$$\varphi_{\Omega}(\mathbf{X}) = \|\mathbf{X}\|_F^2 - d(\mathbf{X}; \Omega)^2 = 2 \sup \left\{ \langle \mathbf{X}, \mathbf{Y} \rangle - \frac{\|\mathbf{Y}\|_F^2}{2} \mid \mathbf{Y} \in \Omega \right\}, \mathbf{X} \in \mathbb{R}^{k \times d}.$$

Observe that if $\Omega_i = \Omega_i^1 \times \Omega_i^2 \dots \times \Omega_i^k$ and $\mathbf{X} = (x^1, \dots, x^k) \in \mathbb{R}^{k \times d}$, then

$$d(\mathbf{X}; \Omega_i)^2 = \sum_{\ell=1}^k d(x^\ell; \Omega_i^\ell)^2 = \sum_{\ell=1}^k \left(\|x^\ell\|^2 - \varphi_{\Omega_i^\ell}(x^\ell) \right) = \|\mathbf{X}\|_F^2 - \sum_{\ell=1}^k \varphi_{\Omega_i^\ell}(x^\ell).$$

Therefore, $\varphi_{\Omega_i}(\mathbf{X}) = \sum_{\ell=1}^k \varphi_{\Omega_i^\ell}(x^\ell)$.

In this new notation, the function F_λ in (2.3.16) can be rewritten as

$$F_\lambda(\mathbf{X}) = \left(G(\mathbf{X}) + \frac{\lambda q}{2} \|\mathbf{X}\|_F^2 \right) - \left(H(\mathbf{X}) + \sum_{i=1}^q \varphi_{\Omega_i}(\mathbf{X}) \right) = G_1(\mathbf{X}) - H_1(\mathbf{X}),$$

where

$$G_1(\mathbf{X}) = G(\mathbf{X}) + \frac{\lambda q}{2} \|\mathbf{X}\|_F^2 \text{ and } H_1(\mathbf{X}) = H(\mathbf{X}) + \sum_{i=1}^q \varphi_{\Omega_i}(\mathbf{X}), \mathbf{X} \in \mathbb{R}^{k \times d}.$$

We also recall that $\bar{\mathbf{X}} \in \mathbb{R}^{k \times d}$ is a critical point of (2.3.16) if

$$\partial G_1(\bar{\mathbf{X}}) \cap \partial H_1(\bar{\mathbf{X}}) \neq \emptyset.$$

The proof of the following result is similar to that of Proposition 2.3.4.

Proposition 2.3.5. *Suppose that $\lim_{n \rightarrow \infty} \lambda_n = \infty$ and $\mathbf{X}_n = (x_n^1, \dots, x_n^k) \in \mathbb{R}^{k \times d}$ is a critical point of the function F_{λ_n} . Then every subsequential limit of $\{\mathbf{X}_n\}$ is a critical point of (2.3.15).*

Calculus Rules for Fenchel conjugates and Subdifferentials

3.1. Introduction and Basic Definitions

Although most applications of convex analysis typically take place in the finite dimensional space \mathbb{R}^d with the objective function defined in an $\mathbb{R}^d \rightarrow \mathbb{R}$ setting, the development of generalized calculus in a more general setting can lead to greater insight and a more cohesive theory. In this chapter, we will provide basic definitions in the more abstract setting of locally convex topological vector (LCTV) spaces and discuss the development of major calculus rules for Fenchel conjugates of convex functions on these spaces and Banach spaces.

Throughout this chapter, unless otherwise stated, that all the spaces under consideration are separated LCTV. Given $\Omega \subset X$, we write

$$\mathbb{R}^+(\Omega) := \{tv \in X \mid t \in \mathbb{R}_+, v \in \Omega\},$$

where \mathbb{R}_+ signifies the collection of positive numbers, and use the symbol $\overline{\Omega}$ for the topological closure of the set Ω . We also recall the symbols \mathbb{B} and \mathbb{B}^* stand for the closed unit balls in the normed space X and its topological dual X^* , respectively. In what follows we deal with *extended-real-valued* functions $f: X \rightarrow \overline{\mathbb{R}}$ with the assumption that they are *proper*. A class of these functions plays a highly important role in many aspects of variational analysis, optimization, and their applications; see, e.g., [24, 43, 47] and the references therein. This chapter presents the work we have done in [29]. The reader can find more details in there. We first recall several definitions from Chapter 1 but in the LCTV setting.

Given a convex set $\Omega \subset X$ with $\bar{x} \in \Omega$, the *normal cone* to Ω at \bar{x} is

$$(3.1.1) \quad N(\bar{x}; \Omega) := \{x^* \in X^* \mid \langle x^*, x - \bar{x} \rangle \leq 0 \text{ for all } x \in \Omega\}.$$

Given an extended-real-valued function $f: X \rightarrow \overline{\mathbb{R}}$ with its epigraph

$$\text{epi } f := \{(x, \mu) \in X \times \mathbb{R} \mid \mu \geq f(x)\},$$

we know that the standard notions of convexity and lower semicontinuity of f can be equivalently described via the convexity and closedness of its epigraph, respectively.

An element $x^* \in X^*$ is a *subgradient* of f at $\bar{x} \in \text{dom } f$ if

$$\langle x^*, x - \bar{x} \rangle \leq f(x) - f(\bar{x}) \text{ for all } x \in X,$$

while the collection $\partial f(\bar{x})$ of all such x^* is called the *subdifferential* of f at \bar{x} . More than that the subdifferential $\partial f(\bar{x})$ can be represented geometrically as

$$\partial f(\bar{x}) = \{x^* \in X^* \mid (x^*, -1) \in N((\bar{x}, f(\bar{x})); \text{epi } f)\}.$$

We also have $N(\bar{x}; \Omega) = \partial \delta(\bar{x}; \Omega)$, where $\delta(x; \Omega) = \delta_\Omega(x)$ is the indicator function of the set Ω equal to 0 if $x \in \Omega$ and to ∞ if $x \notin \Omega$. The *support function* $\sigma_\Omega: X^* \rightarrow \overline{\mathbb{R}}$ of a nonempty subset $\Omega \subset X$ is defined by

$$\sigma_\Omega(x^*) := \sup \{\langle x^*, x \rangle \mid x \in \Omega\}, \quad x^* \in X^*.$$

Note that σ_Ω is always convex on X^* regardless of the convexity of Ω .

The *Fenchel conjugate* $f^*: X^* \rightarrow [-\infty, \infty]$ of $f: X \rightarrow \overline{\mathbb{R}}$ is defined by

$$f^*(x^*) := \sup \{\langle x^*, x \rangle - f(x) \mid x \in X\}.$$

Note that if f is proper, then $f^*: X^* \rightarrow \overline{\mathbb{R}}$ is convex and lower semi-continuous on X^* regardless of the convexity and lower semicontinuity of f . Recall also the *second*

Fenchel conjugate (or *biconjugate*) of $f: X \rightarrow \overline{\mathbb{R}}$ is given by

$$f^{**}(x) := (f^*)^*(x) = \sup \{ \langle x^*, x \rangle - f^*(x^*) \mid x^* \in X^* \}, \quad x \in X.$$

It is well-known that $f^{**} = f$ on X provided that f is proper, lower semi-continuous, and convex.

Another important operation on functions is the infimal convolution. Given two functions $f, g: X \rightarrow \overline{\mathbb{R}}$, their *infimal convolution* $f \oplus g$ is defined by

$$(3.1.2) \quad \begin{aligned} (f \oplus g)(x) &:= \inf \{ f(x_1) + g(x_2) \mid x_1 + x_2 = x \} \\ &= \inf \{ f(u) + g(x - u) \mid u \in X \}, \quad x \in X. \end{aligned}$$

Assuming that $(f \oplus g)(x) > -\infty$ for all $x \in X$, we clearly have the convexity of the infimal convolution $f \oplus g$ provided that both functions f, g are convex.

3.2. Support Functions and Normal Cones to Set Intersections

In this section, we study support functions to set intersections largely used in what follows. The key observation of this section is the extended version of the convex extremal principle applied to deriving a refined support and normal cone intersection rules in arbitrary LCTV spaces.

We first formulate the enhanced notion of set extremality, which is the most appropriate for convex sets in LCTV spaces, and then establish the *convex extremal principle* in this setting. The main applications of the convex extremal principle are given here to deriving refined support and normal cone intersection rules.

We start with some auxiliary properties of support functions needed in what follows. Recall that the *support function* $\sigma_\Omega: X^* \rightarrow \overline{\mathbb{R}}$ of a nonempty subset $\Omega \subset X$ is defined by

$$(3.2.3) \quad \sigma_\Omega(x^*) := \sup \{ \langle x^*, x \rangle \mid x \in \Omega \}, \quad x^* \in X^*.$$

Note that σ_Ω is always convex on X^* regardless of the convexity of Ω and that we have the conjugacy relationship $\sigma_\Omega(x^*) = \delta_\Omega^*(x^*)$ as $x^* \in X^*$ with the indicator function of Ω .

Let us first present the following known result widely used below. We provide a direct proof of it for the reader's convenience.

Proposition 3.2.1. *Let the sets Ω_1 and Ω_2 be nonempty, closed, and convex in X with $\Omega_1 \cap \Omega_2 \neq \emptyset$. Then we have the representation*

$$(3.2.4) \quad (\sigma_{\Omega_1} \oplus \sigma_{\Omega_2})^*(x) = \delta_{\Omega_1 \cap \Omega_2}(x) \quad \text{for all } x \in X.$$

Proof. Consider the infimal convolution (3.1.2) of the support functions of Ω_i given by

$$f(x^*) := (\sigma_{\Omega_1} \oplus \sigma_{\Omega_2})(x^*), \quad x^* \in X^*.$$

It follows from [41, Corollary 2.4.7] and a well-known subdifferential result for support functions on LCTV spaces that

$$\partial f(0) = \partial \sigma_{\Omega_1}(0) \cap \partial \sigma_{\Omega_2}(0) = \Omega_1 \cap \Omega_2.$$

It is easy to see that $f(x^*)$ is convex and positively homogeneous with $f(0) = 0$ under the assumptions made, and so by [51, Theorem 2.4.11(i)] we get

$$f^*(x) = \delta_{\partial f(0)}(x) \quad \text{whenever } x \in X,$$

which readily verifies the infimal convolution representation in (3.2.4). □

The proofs of the next two lemmas follow the arguments from [8] with some simplifications in the case of the support functions under consideration. They are essentially based on the imposed Banach space and closedness assumptions.

Lemma 3.2.2. *Let Ω_1 and Ω_2 be nonempty closed subsets of a Banach space X . Suppose that $\mathbb{R}^+(\Omega_1 - \Omega_2) = X$. Then for any numbers $\alpha, \beta \geq 0$ the restricted level*

set

$$K = K_{\alpha, \beta} := \{(x_1^*, x_2^*) \in X^* \times X^* \mid \sigma_{\Omega_1}(x_1^*) + \sigma_{\Omega_2}(x_2^*) \leq \alpha, \|x_1^* + x_2^*\| \leq \beta\}$$

is compact in the weak* topology of the dual product space $X^* \times X^*$.

Proof. The closedness of the set K in the weak* topology of $X^* \times X^*$ is obvious. By the classical Alaoglu-Bourbaki theorem it remains to show that this set is norm-bounded in $X^* \times X^*$. Having in mind the uniform boundedness principle, we need to verify that the set of linear continuous functionals from K is bounded pointwise. To proceed, take any $(x_1, x_2) \in X \times X$ and find by the imposed assumption $\mathbb{R}^+(\Omega_1 - \Omega_2) = X$ elements $\lambda > 0$, $w_1 \in \Omega_1$, and $w_2 \in \Omega_2$ for which $x_1 - x_2 = \lambda(w_1 - w_2)$. Then we have

$$\begin{aligned} \langle x_1^*, x_1 \rangle + \langle x_2^*, x_2 \rangle &= \lambda \langle x_1^*, w_1 \rangle + \lambda \langle x_2^*, w_2 \rangle + \langle x_1^* + x_2^*, x_2 - \lambda w_2 \rangle \\ &\leq \lambda(\sigma_{\Omega_1}(x_1^*) + \sigma_{\Omega_2}(x_2^*)) + \|x_1^* + x_2^*\| \cdot \|x_2 - \lambda w_2\| \\ &\leq \lambda\alpha + \beta\|x_2 - \lambda w_2\|. \end{aligned}$$

Since the latter also holds for $(-x_1, -x_2)$, we arrive at the conclusion of the lemma. \square

Lemma 3.2.3. *In the setting of Lemma 3.2.2 we have that the infimal convolution $(\sigma_{\Omega_1} \oplus \sigma_{\Omega_2}): X^* \rightarrow \overline{\mathbb{R}}$ is lower semi-continuous with respect to the weak* topology of X^* .*

Proof. It suffices to prove that for any $\gamma \in \mathbb{R}$ the set

$$C := \{x^* \in X^* \mid (\sigma_{\Omega_1} \oplus \sigma_{\Omega_2})(x^*) \leq \gamma\}$$

is weak* closed in X^* . Consider the parametric set family

$$C_\varepsilon := \{x^* \in X^* \mid x^* = x_1^* + x_2^*, \sigma_{\Omega_1}(x_1^*) + \sigma_{\Omega_2}(x_2^*) \leq \gamma + \varepsilon\}, \quad \varepsilon > 0,$$

with $C = \bigcap_{\varepsilon > 0} C_\varepsilon$ and show that each set C_ε as $\varepsilon > 0$ is weak* closed in X^* . Using the seminal Banach-Dieudonné-Krein-Šmulian theorem (see, e.g., [16, Theorem V.5.7]), we only need to check that the intersection $C_\varepsilon \cap r\mathbb{B}^*$ is weak* closed in X^* for all

$r > 0$. To this end, define the mapping $T: X^* \times X^* \rightarrow X^*$ by

$$T(x_1^*, x_2^*) = x_1^* + x_2^*,$$

which is continuous in the weak* topology of $X^* \times X^*$. It is easy to observe that

$$C_\varepsilon \cap r\mathbb{B}^* = T(K_{\gamma+\varepsilon, r}),$$

where $K_{\gamma+\varepsilon, r}$ is defined in Lemma 3.2.2. It follows from Lemma 3.2.2 that $C_\varepsilon \cap r\mathbb{B}^*$ is weak* compact and hence weak* closed, which verifies that $\sigma_{\Omega_1} \oplus \sigma_{\Omega_2}$ is weak* lower semi-continuous on X^* . \square

To proceed further, let us first formulate an appropriate LCTV version of set extremality defined recently in [27] in normed spaces. It seems to be more efficient to deal with convex sets in LCTV spaces in comparison with its local extremality counterpart developed in [24].

Definition 3.2.4. *We say that two nonempty sets $\Omega_1, \Omega_2 \subset X$ form an extremal system in the space X if for any neighborhood V of the origin there exists a vector $a \in X$ such that*

$$(3.2.5) \quad a \in V \quad \text{and} \quad (\Omega_1 + a) \cap \Omega_2 = \emptyset.$$

Note that (3.2.5) does not require that the sets Ω_1, Ω_2 have the nonempty intersection in contrast to [24]. As discussed in [27] similarly to [24, Section 2.1] for the local counterpart, the extremality notion (3.2.5) covers various concepts of optimal solutions to problems of scalar and multiobjective optimization, equilibrium systems, etc. with numerous applications.

The following result is an extension of [27, Theorem 2.2] for the case of normed spaces.

Theorem 3.2.5. *Let the sets $\Omega_1, \Omega_2 \subset X$ be nonempty and convex. Then we have the assertions:*

(a) The sets Ω_1 and Ω_2 form an extremal system in X if and only if $0 \notin \text{int}(\Omega_1 - \Omega_2)$. Furthermore, the extremality of Ω_1, Ω_2 implies that $(\text{int } \Omega_1) \cap \Omega_2 = \emptyset$ and $(\text{int } \Omega_2) \cap \Omega_1 = \emptyset$.

(b) If Ω_1 and Ω_2 form an extremal system in X and the set difference $\Omega_1 - \Omega_2$ is solid, i.e.,

$$(3.2.6) \quad \text{int}(\Omega_1 - \Omega_2) \neq \emptyset,$$

then the set Ω_1 and Ω_2 are separated, i.e.,

$$(3.2.7) \quad \sup_{x \in \Omega_1} \langle x^*, x \rangle \leq \inf_{x \in \Omega_2} \langle x^*, x \rangle \text{ for some } x^* \in X^*, x^* \neq 0.$$

(c) The separation property (3.2.7) always yields the set extremality (3.2.5) even if (3.2.6) fails.

Proof. (a) Let the sets Ω_1 and Ω_2 form an extremal system in X . Arguing by contradiction, suppose that $0 \in \text{int}(\Omega_1 - \Omega_2)$. Then there is a balanced neighborhood V of $0 \in X$ with

$$V \subset \Omega_1 - \Omega_2.$$

For any $a \in V$ we have $-a \in V \subset \Omega_1 - \Omega_2$ and arrive at the contradiction $(\Omega_1 + a) \cap \Omega_2 \neq \emptyset$.

Now suppose that $0 \notin \text{int}(\Omega_1 - \Omega_2)$. Then for any neighborhood V of the origin we get

$$V \cap [X \setminus (\Omega_1 - \Omega_2)] \neq \emptyset.$$

Assume without loss of generality that V is balanced, which yields $(-V) \cap [X \setminus (\Omega_1 - \Omega_2)] \neq \emptyset$. Choosing $a \in V$ gives us the inclusion

$$-a \in [X \setminus (\Omega_1 - \Omega_2)],$$

and hence $(\Omega_1 + a) \cap \Omega_2 = \emptyset$. This verifies that the sets Ω_1 and Ω_2 form an extremal system.

For the second part of (a), suppose for contradiction that $(\text{int } \Omega_1) \cap \Omega_2 \neq \emptyset$. Then there exists a vector $x \in \text{int } \Omega_1$ with $x \in \Omega_2$. We can always choose a balanced neighborhood V of the origin such that $x + V \subset \Omega_1$. For any $a \in V$ it shows that $-a \in V$ and $x - a \in \Omega_1$. Hence $(a + \Omega_1) \cap \Omega_2 \neq \emptyset$, which is a contradiction.

(b) If Ω_1 and Ω_2 form an extremal system, then it follows from (i) that $0 \notin \text{int}(\Omega_1 - \Omega_2)$. In addition to this, the assumption of (b) on the solidness of the set difference $\Omega_1 - \Omega_2$ allows us to use the convex separation theorem, which yields (3.2.7).

(c) Suppose that (3.2.7) holds, which gives us a vector $c \in X$ with $\langle x^*, c \rangle > 0$. Fix any neighborhood V of the origin. Since V is always absorbing, we can select a natural number k sufficiently large such that $a := -c/k \in V$. Let us show that (3.2.5) is satisfied with this vector a . Indeed, the negation of this means that then there exists $\hat{x} \in \Omega_2$ such that $\hat{x} - a \in \Omega_1$. By the separation property from (3.2.7) we have

$$\langle x^*, \hat{x} - a \rangle \leq \sup_{x \in \Omega_1} \langle x^*, x \rangle \leq \inf_{x \in \Omega_2} \langle x^*, x \rangle \leq \langle x^*, \hat{x} \rangle,$$

On the other hand, the construction of a tells us that

$$\langle x^*, \hat{x} \rangle - \langle x^*, a \rangle = \langle x^*, \hat{x} \rangle + k \langle x^*, c \rangle \leq \langle x^*, \hat{x} \rangle \text{ for all } k \in \mathbb{N}.$$

This shows that $\langle x^*, c \rangle \leq 0$, which contradicts the above choice of the vector c . \square

It is worth mentioning that [27, Theorem 2.2] offers a version of Theorem 3.2.5(b) in Banach spaces with the replacement of (3.2.6) by the *sequential normal compactness* (SNC) property imposed on one of the sets Ω_1, Ω_2 . The SNC property is effectively characterized in [24, Theorem 1.21] for convex sets. Furthermore, in [27, Theorem 2.5] the reader can find an *approximate* version of the convex extremal principle without imposing either (3.2.5) or the SNC property. The latter approximate version yields, in particular, the celebrated Bishop-Phelps theorem establishing the density of support points on the boundary of any closed convex subset of a Banach space; see [39, Theorem 3.18].

The next theorem presents major geometric results on representing the support function for convex set intersections via the infimal convolution of the support functions for each component. We derive these representations under three generally independent groups of assumptions, which constitute the corresponding assertions of the theorem. The first assertion is established in the general LCTV setting without any closedness requirements on the sets in question under the *difference interiority condition*

$$(3.2.8) \quad 0 \in \text{int}(\Omega_1 - \Omega_2)$$

and the boundedness of one of the sets by using the convex extremal principle from Theorem 3.2.5. The second assertion, which is proved similarly in any LCTV space, does not impose any set boundedness while requires a more restrictive qualification condition; see more discussions after the proof of the theorem. The third assertion holds for closed sets in Banach spaces with replacing (3.2.8) by the weaker *Attouch-Brézis qualification condition*

$$(3.2.9) \quad \mathbb{R}^+(\Omega_1 - \Omega_2) \text{ is a closed subspace of } X$$

by using the above lemmas and Proposition 3.2.1.

Recall that a subset Ω of an LCTV space X is *bounded* if for any neighborhood V of the origin there exists $\gamma > 0$ such that $\Omega \subset \alpha V$ whenever $|\alpha| > \gamma$.

Theorem 3.2.6. *Let the sets $\Omega_1, \Omega_2 \subset X$ be nonempty and convex. Suppose that one of the following conditions is satisfied:*

- (a) *The difference interiority condition (3.2.8) is satisfied and the set Ω_2 is bounded.*
- (b) *Either $(\text{int } \Omega_2) \cap \Omega_1 \neq \emptyset$, or $(\text{int } \Omega_1) \cap \Omega_2 \neq \emptyset$.*
- (c) *The space X is Banach, both sets Ω_1 and Ω_2 are closed, and the Attouch-Brézis qualification condition (3.2.9) holds.*

Then the support function (3.2.3) to the intersection $\Omega_1 \cap \Omega_2$ is represented as

$$(3.2.10) \quad (\sigma_{\Omega_1 \cap \Omega_2})(x^*) = (\sigma_{\Omega_1} \oplus \sigma_{\Omega_2})(x^*) \text{ for all } x^* \in X^*.$$

Furthermore, for any $x^* \in \text{dom}(\sigma_{\Omega_1 \cap \Omega_2})$ there are $x_1^*, x_2^* \in X^*$ such that $x^* = x_1^* + x_2^*$ and

$$(3.2.11) \quad (\sigma_{\Omega_1 \cap \Omega_2})(x^*) = \sigma_{\Omega_1}(x_1^*) + \sigma_{\Omega_2}(x_2^*).$$

Proof. First we justify both formula (3.2.10) and representation (3.2.11) in case (a). To verify the inequality “ \leq ” in (3.2.10), fix $x^* \in X^*$ and pick x_1^*, x_2^* with $x^* = x_1^* + x_2^*$. Then

$$\langle x^*, x \rangle = \langle x_1^*, x \rangle + \langle x_2^*, x \rangle \leq \sigma_{\Omega_1}(x_1^*) + \sigma_{\Omega_2}(x_2^*) \quad \text{for any } x \in \Omega_1 \cap \Omega_2.$$

Taking the infimum on the right-hand side above with respect to all such x_1^*, x_2^* yields

$$\langle x^*, x \rangle \leq (\sigma_{\Omega_1} \oplus \sigma_{\Omega_2})(x^*) \quad \text{whenever } x \in \Omega_1 \cap \Omega_2,$$

which clearly implies the inequality “ \leq ” in (3.2.10) without imposing the assumptions in (a).

Next we prove the inequality “ \geq ” in (3.2.10) under the assumptions in (a). To proceed, fix any $x^* \in \text{dom}(\sigma_{\Omega_1 \cap \Omega_2})$, denote $\alpha := \sigma_{\Omega_1 \cap \Omega_2}(x^*) \in \mathbb{R}$ for which we have

$$\langle x^*, x \rangle \leq \alpha \quad \text{whenever } x \in \Omega_1 \cap \Omega_2,$$

and define the nonempty convex subsets of $X \times \mathbb{R}$ by

$$(3.2.12) \quad \Theta_1 := \Omega_1 \times [0, \infty) \quad \text{and} \quad \Theta_2 := \{(x, \lambda) \in X \times \mathbb{R} \mid x \in \Omega_2, \lambda \leq \langle x^*, x \rangle - \alpha\}.$$

It is easy to see from the constructions of Θ_1 and Θ_2 that

$$(\Theta_1 + (0, \gamma)) \cap \Theta_2 = \emptyset \quad \text{for any } \gamma > 0,$$

and thus these sets form an *extremal system* in $X \times \mathbb{R}$. We deduce from Theorem 3.2.5(a) that $0 \notin \text{int}(\Theta_1 - \Theta_2)$. To apply Theorem 3.2.5(b) to these sets, we need to verify that the set difference $\Theta_1 - \Theta_2$ is solid. The property $0 \in \text{int}(\Omega_1 - \Omega_2)$ allows us to find a

neighborhood U of the origin such that $U \subset \Omega_1 - \Omega_2$. By the continuity of the function $\varphi(x) := \langle -x^*, x \rangle + \alpha$ for $x \in X$, there exists a neighborhood W of the origin such that φ is bounded above on W . Since Ω_2 is bounded, we can find $t > 0$ such that $\Omega_2 \subset tW$. Note that φ is also bounded above on tW , so we can find $\bar{\lambda} \in \mathbb{R}$ such that

$$\bar{\lambda} \geq \sup_{x \in tW} \langle -x^*, x \rangle + \alpha \geq \sup_{x \in \Omega_2} \langle -x^*, x \rangle + \alpha.$$

Let us check that $U \times (\bar{\lambda}, \infty) \subset \Theta_1 - \Theta_2$, and so (3.2.6) holds. Indeed, for $(x, \lambda) \in U \times (\bar{\lambda}, \infty)$ we have $x \in U \subset \Omega_1 - \Omega_2$ and $\lambda > \bar{\lambda}$. Hence $x = w_1 - w_2$ with some $w_1 \in \Omega_1, w_2 \in \Omega_2$ and

$$(x, \lambda) = (w_1, \lambda - \bar{\lambda}) - (w_2, -\bar{\lambda}).$$

It follows from $\lambda - \bar{\lambda} > 0$ that $(w_1, \lambda - \bar{\lambda}) \in \Theta_1$. We deduce from (3.2.12) and the choice of $\bar{\lambda}$ that $(w_2, -\bar{\lambda}) \in \Theta_2$, and thus $\text{int}(\Theta_1 - \Theta_2) \neq \emptyset$. Theorem 3.2.5(b) tells us now that there exists a pair $(0, 0) \neq (y^*, \beta) \in X^* \times \mathbb{R}$ for which we get

$$(3.2.13) \quad \langle y^*, x \rangle + \beta \lambda_1 \leq \langle y^*, y \rangle + \beta \lambda_2 \quad \text{whenever } (x, \lambda_1) \in \Theta_1, (y, \lambda_2) \in \Theta_2.$$

It follows from the structure of Θ_1 that $\beta \leq 0$. By (3.2.13), assuming that $\beta = 0$ gives us

$$\langle y^*, x \rangle \leq \langle y^*, y \rangle \quad \text{for all } x \in \Omega_1, y \in \Omega_2.$$

This yields $y^* = 0$ due to $0 \in \text{int}(\Omega_1 - \Omega_2)$, a contradiction. Take now the pairs $(x, 0) \in \Theta_1$ and $(y, \langle x^*, y \rangle - \alpha) \in \Theta_2$ in (3.2.13) and get

$$\langle y^*, x \rangle \leq \langle y^*, y \rangle + \beta(\langle x^*, y \rangle - \alpha) \quad \text{with } \beta < 0,$$

which brings us to the estimate

$$\alpha \geq \langle y^*/\beta + x^*, y \rangle + \langle -y^*/\beta, x \rangle \quad \text{for all } x \in \Omega_1, y \in \Omega_2.$$

By putting $x_1^* := y^*/\beta + x^*$ and $x_2^* := -y^*/\beta$ we arrive at the inequality “ \geq ” in (3.2.10) and representation (3.2.11). This justifies the conclusions of the theorem in case (a).

To verify the results under (b), it clearly suffices to examine only the first case therein. Considering the sets Θ_1, Θ_2 from (3.2.12), we see that

$$\text{int } \Theta_2 = \{(x, \lambda) \in X \times \mathbb{R} \mid x \in \text{int } \Omega_2, \lambda < \langle x^*, x \rangle - \alpha\} \neq \emptyset,$$

and so $\text{int}(\Theta_1 - \Theta_2) \neq \emptyset$. Furthermore, it follows from the assumption $(\text{int } \Omega_2) \cap \Omega_1 \neq \emptyset$ that $0 \in \text{int}(\Omega_1 - \Omega_2)$, which allows us to proceed as in the above proof in case (a).

Consider finally the Banach space case (c) of the theorem and first verify its conclusions when $\mathbb{R}^+(\Omega_1 - \Omega_2) = X$. Taking the Fenchel conjugate of both sides of formula (3.2.4) from Proposition 3.2.1, which holds even without the assumptions of (b), and then using Lemma 3.2.3 gives us the equalities

$$\delta_{\Omega_1 \cap \Omega_2}^*(x^*) = \sigma_{\Omega_1 \cap \Omega_2}(x^*) = (\sigma_{\Omega_1} \oplus \sigma_{\Omega_2})^{**}(x^*) = (\sigma_{\Omega_1} \oplus \sigma_{\Omega_2})(x^*)$$

for all $x^* \in X^*$, which justify the representation in (3.2.10) when the assumption $\mathbb{R}^+(\Omega_1 - \Omega_2) = X$ is satisfied. Treating further the general case of (3.2.9) in (c), denote $L := \mathbb{R}^+(\Omega_1 - \Omega_2)$ the closed subspace of X in question. Since $\Omega_1 \cap \Omega_2 \neq \emptyset$ by (3.2.9), we can always translate the situation to $0 \in \Omega_1 \cap \Omega_2$, and hence suppose that $\Omega_1, \Omega_2 \subset L$. This reduces the general case under (3.2.9) to the one $\mathbb{R}^+(\Omega_1 - \Omega_2)$ treated above; so (3.2.10) is justified.

Representation (3.2.11) under (c) for $x^* \in \text{dom}(\sigma_{\Omega_1} \cap \sigma_{\Omega_2})$ follows from the weak* compactness of the set $K_{\alpha, \beta}$ in Lemma 3.2.2 with $\alpha := (\sigma_{\Omega_1} \oplus \sigma_{\Omega_2})(x^*) + \varepsilon$, where $\varepsilon > 0$ is arbitrary and where $\beta := \|x^*\|$. This completes the proof of the theorem. \square

The results of type (3.2.11) in Theorem 3.2.6 (i.e., ensuring that the infimal convolution is *exact*) go back to Moreau [31] in Hilbert spaces. More recently, in [17], Ernst and Théra established necessary and sufficient conditions for this property in finite-dimensional spaces.

If X is Banach and both sets Ω_i are closed with $\text{int}(\Omega_1 - \Omega_2) \neq \emptyset$, then (3.2.8) reduces to the *core qualification condition* $0 \in \text{core}(\Omega_1 - \Omega_2)$ developed by Rockafellar [42], where

$$\text{core } \Omega := \{x \in \Omega \mid \forall v \in X \exists \gamma > 0 \text{ such that } x + tv \in \Omega \text{ whenever } |t| < \gamma\}.$$

This follows from the well-known facts that $\text{int } \Omega = \text{core } \Omega$ for closed convex subsets of Banach spaces (see, e.g., [13, Theorem 4.1.8]) and that $\text{int } \overline{\Omega} = \text{int } \Omega$ provided that $\text{int } \Omega \neq \emptyset$. Note that the Attouch-Brézis condition (3.2.9) from [8] essentially supersedes the equivalent core and difference interior qualification conditions for closed sets in Banach spaces, while the latter works in general LCTV spaces. As shown in [27], the conventional interiority conditions in (b) of Theorem 3.2.6 *strictly* imply the conditions in (a), including the set boundedness, provided that the space X is *normed*. We cannot conclude this in general, unless the LCTV space X in question has a basis of *bounded* neighborhoods of the origin.

The next result provides various qualification conditions in LCTV and Banach space settings under which the important normal cone intersection formula holds. This result under the qualification condition (3.2.14) has been derived in [27] directly from the convex extremal principle in normed spaces. Below we present a unified derivation of the normal cone intersection rule from the corresponding conditions of Theorem 3.2.6.

Theorem 3.2.7. *Let the sets $\Omega_1, \Omega_2 \subset X$ be convex with $\bar{x} \in \Omega_1 \cap \Omega_2$. Suppose that one of the following conditions (a)–(c) is satisfied:*

(a) *There exists a bounded convex neighborhood V of \bar{x} such that*

$$(3.2.14) \quad 0 \in \text{int}(\Omega_1 - (\Omega_2 \cap V)).$$

(b) *Either $(\text{int } \Omega_2) \cap \Omega_1 \neq \emptyset$, or $(\text{int } \Omega_1) \cap \Omega_2 \neq \emptyset$.*

(c) *The space X is Banach, both sets Ω_1 and Ω_2 are closed, and the Attouch-Brézis qualification condition (3.2.9) holds.*

Then we have the normal cone intersection rule

$$(3.2.15) \quad N(\bar{x}; \Omega_1 \cap \Omega_2) = N(\bar{x}; \Omega_1) + N(\bar{x}; \Omega_2).$$

Proof. First we verify (3.2.15) under (a). Denote $A := \Omega_1$, $B := \Omega_2 \cap V$ and observe by (3.2.14) that $0 \in \text{int}(A - B)$ and B is bounded. It follows from the normal cone definition (3.1.1) that $x^* \in N(\bar{x}; \Omega)$ for $\bar{x} \in \Omega$ if and only if $\sigma_\Omega(x^*) = \langle x^*, \bar{x} \rangle$. Then pick $x^* \in N(\bar{x}; A \cap B)$ and get $\langle x^*, \bar{x} \rangle = \sigma_{A \cap B}(x^*)$. By Theorem 3.2.6(a) there are $x_1^*, x_2^* \in X^*$ with $x^* = x_1^* + x_2^*$ and

$$\langle x_1^*, \bar{x} \rangle + \langle x_2^*, \bar{x} \rangle = \langle x^*, \bar{x} \rangle = \sigma_{A \cap B}(x^*) = \sigma_A(x_1^*) + \sigma_B(x_2^*).$$

This implies that $\langle x_1^*, \bar{x} \rangle = \sigma_A(x_1^*)$ and $\langle x_2^*, \bar{x} \rangle = \sigma_B(x_2^*)$. Thus we have $x_1^* \in N(\bar{x}; A)$ and $x_2^* \in N(\bar{x}; B)$, which show that $N(\bar{x}; A \cap B) \subset N(\bar{x}; A) + N(\bar{x}; B)$. Observe that $N(\bar{x}; A \cap B) = N(\bar{x}; \Omega_1 \cap \Omega_2)$, $N(\bar{x}; A) = N(\bar{x}; \Omega_1)$, and $N(\bar{x}; B) = N(\bar{x}; \Omega_2)$; hence we arrive at the inclusion “ \subset ” in (3.2.15). The opposite inclusion therein is obvious. Similar arguments allow us to deduce normal cone intersection rule (3.2.15) under assumptions in (b) and (c) from the corresponding assertions of Theorem 3.2.6. \square Observe that (3.2.14) reduces to the difference interiority condition (3.2.8) provided that one of the sets Ω_i , say Ω_2 , is bounded. Note that there are other qualification conditions for the validity of (3.2.15) in various space settings studied in connection with the so-called “strong conic hull intersection property” (strong CHIP); see, e.g., [9, 17, 14, 23, 30].

3.3. Geometric Approach to Conjugate Calculus

In this section we develop a geometric approach, based on the set intersection rules obtained above, to easily derive some basic calculus results for Fenchel conjugates of extended-real-valued convex functions in LCTV and Banach spaces settings. Known proofs of such functional results are more involved and employ analytic arguments; see, e.g., [12, 44, 51].

First we present a simple lemma relating the conjugate and epigraphical support functions.

Lemma 3.3.1. *For any $f: X \rightarrow \overline{\mathbb{R}}$ we have the conjugate function representation*

$$f^*(x^*) = \sigma_{\text{epi } f}(x^*, -1) \text{ whenever } x^* \in X^*.$$

Proof. It follows directly from the definitions that

$$\begin{aligned} f^*(x^*) &= \sup \{ \langle x^*, x \rangle - f(x) \mid x \in \text{dom } f \} = \sup \{ \langle x^*, x \rangle - \lambda \mid (x, \lambda) \in \text{epi } f \} \\ &= \sigma_{\text{epi } f}(x^*, -1), \end{aligned}$$

which therefore verifies the claimed relationship. □

This lemma allows us to derive the conjugate sum rules and other conjugate calculus results from Theorem 3.2.6. In what follows we concentrate for simplicity on applying this theorem under the assumptions imposed in (b) and (c) therein.

Theorem 3.3.2. *For convex functions $f, g: X \rightarrow \overline{\mathbb{R}}$, we assume the followings:*

- (a) *either one of the functions f, g is continuous at some point $\bar{x} \in \text{dom } f \cap \text{dom } g$,*
- (b) *or X is a Banach space, f and g are lower semi-continuous, and $\mathbb{R}^+(\text{dom } f - \text{dom } g)$ is a closed subspace of X .*

Then we have the conjugate to the sum $f + g$ is represented by

$$(3.3.16) \quad (f + g)^*(x^*) = (f^* \oplus g^*)(x^*) \text{ for all } x^* \in X^*.$$

Moreover, in both cases the infimum in $(f^ \oplus g^*)(x^*)$ is attained, i.e., for any $x^* \in \text{dom } (f + g)^*$, there are $x_1^*, x_2^* \in X^*$ such that $x^* = x_1^* + x_2^*$ and*

$$(f + g)^*(x^*) = f^*(x_1^*) + g^*(x_2^*).$$

Proof. Fix $x^* \in X^*$ and easily observe that $(f + g)^*(x^*) \leq (f^* \oplus g^*)(x^*)$. Indeed, for any $x_1^*, x_2^* \in X^*$ with $x_1^* + x_2^* = x^*$ we have

$$\begin{aligned}
f^*(x_1^*) + g^*(x_2^*) &= \sup \{ \langle x_1^*, x \rangle - f(x) \mid x \in X \} + \sup \{ \langle x_2^*, x \rangle - g(x) \mid x \in X \} \\
&\geq \sup \{ \langle x_1^*, x \rangle - f(x) + \langle x_2^*, x \rangle - g(x) \mid x \in X \} \\
&= \sup \{ \langle x^*, x \rangle - (f + g)(x) \mid x \in X \} = (f + g)^*(x^*).
\end{aligned}$$

Taking the infimum with respect to all such x_1^*, x_2^* justifies the claimed inequality.

Let us verify the opposite inequality in (3.3.16) for $x^* \in \text{dom } (f + g)^*$ in each case (a) and (b).

(a) Define the convex subsets Ω_1, Ω_2 of $X \times \mathbb{R} \times \mathbb{R}$ by

$$\begin{aligned}
\Omega_1 &:= \{ (x, \lambda_1, \lambda_2) \in X \times \mathbb{R} \times \mathbb{R} \mid \lambda_1 \geq f(x) \}, \\
\Omega_2 &:= \{ (x, \lambda_1, \lambda_2) \in X \times \mathbb{R} \times \mathbb{R} \mid \lambda_2 \geq g(x) \}
\end{aligned}$$

and show that $(\text{int } \Omega_1) \cap \Omega_2 \neq \emptyset$ if f is continuous at some $\bar{x} \in \text{dom } f \cap \text{dom } g$. Indeed,

$$\text{int } \Omega_1 = \{ (x, \lambda_1, \lambda_2) \in X \times \mathbb{R} \times \mathbb{R} \mid x \in \text{int}(\text{dom } f), \lambda_1 > f(x) \}.$$

It follows from the continuity of f at \bar{x} that $\bar{x} \in \text{int}(\text{dom } f)$. Letting $\bar{\lambda}_1 := f(\bar{x}) + 1$ and $\bar{\lambda}_2 := g(\bar{x}) + 1$ yields $(\bar{x}, \bar{\lambda}_1, \bar{\lambda}_2) \in \text{int } \Omega_1$ and $(\bar{x}, \bar{\lambda}_1, \bar{\lambda}_2) \in \Omega_2$, and hence $(\text{int } \Omega_1) \cap \Omega_2 \neq \emptyset$. Similarly to Lemma 3.3.1 we have the representation

$$(f + g)^*(x^*) = \sigma_{\Omega_1 \cap \Omega_2}(x^*, -1, -1).$$

Then Theorem 3.2.6(b) gives us triples $(x_1^*, -\alpha_1, -\alpha_2), (x_2^*, -\beta_1, -\beta_2) \in X^* \times \mathbb{R} \times \mathbb{R}$ such that $(x^*, -1, -1) = (x_1^*, -\alpha_1, -\alpha_2) + (x_2^*, -\beta_1, -\beta_2)$ and

$$(f + g)^*(x^*) = \sigma_{\Omega_1 \cap \Omega_2}(x^*, -1, -1) = \sigma_{\Omega_1}(x_1^*, -\alpha_1, -\alpha_2) + \sigma_{\Omega_2}(x_2^*, -\beta_1, -\beta_2).$$

If $\alpha_2 \neq 0$, then $\sigma_{\Omega_1}(x_1^*, -\alpha_1, -\alpha_2) = \infty$, which is not possible since $x^* \in \text{dom } (f + g)^*$, and so $\alpha_2 = 0$. Similarly we get $\beta_1 = 0$. Employing Lemma 3.3.1 again and taking

into account the structures of Ω_1 and Ω_2 tell us that

$$\begin{aligned}(f + g)^*(x^*) &= \sigma_{\Omega_1 \cap \Omega_2}(x^*, -1, -1) = \sigma_{\Omega_1}(x_1^*, -1, 0) + \sigma_{\Omega_2}(x_2^*, 0, -1) \\ &= \sigma_{\text{epi } f}(x_1^*, -1) + \sigma_{\text{epi } g}(x_2^*, -1) = f^*(x_1^*) + g^*(x_2^*) \geq (f^* \oplus g^*)(x^*),\end{aligned}$$

which verifies (3.3.16) and also the last statement of the theorem.

(b) To prove the theorem under the conditions in (b), observe first that the sets Ω_1, Ω_2 above are closed by the lower semi-continuous assumption on f, g and then check that

$$(3.3.17) \quad \mathbb{R}^+(\Omega_1 - \Omega_2) = \mathbb{R}^+(\text{dom } f - \text{dom } g) \times \mathbb{R} \times \mathbb{R}.$$

Indeed, consider $u \in \mathbb{R}^+(\Omega_1 - \Omega_2)$ and find $t > 0, v \in (\Omega_1 - \Omega_2)$ with $u = tv$; therefore $v = (x_1, \lambda_1, \lambda_2) - (x_2, \gamma_1, \gamma_2)$ with $(x_1, \lambda_1, \lambda_2) \in \Omega_1$ and $(x_2, \gamma_1, \gamma_2) \in \Omega_2$. Note that $x_1 \in \text{dom } f$ and $x_2 \in \text{dom } g$ due to $f(x_1) \leq \lambda_1 < \infty$ and $g(x_2) \leq \gamma_2 < \infty$.

Hence we get

$$tv = t(x_1 - x_2, \lambda_1 - \gamma_1, \lambda_2 - \gamma_2) \in \mathbb{R}^+(\text{dom } f - \text{dom } g) \times \mathbb{R} \times \mathbb{R}.$$

To verify the opposite inclusion in (3.3.17), fix $x \in \mathbb{R}^+(\text{dom } f - \text{dom } g) \times \mathbb{R} \times \mathbb{R}$ and find $t > 0, x_1 \in \text{dom } f, x_2 \in \text{dom } g, \gamma_1, \gamma_2 \in \mathbb{R}$, and λ_1, λ_2 with

$$\begin{aligned}x &= (t(x_1 - x_2), \gamma_1, \gamma_2) = t(x_1 - x_2, \lambda_1, \lambda_2) \\ &= t[(x_1, f(x_1), \lambda_2 + g(x_2)) - (x_2, -\lambda_1 + f(x_1), g(x_2))],\end{aligned}$$

which readily yields $x \in t(\Omega_1 - \Omega_2) \subset \mathbb{R}^+(\Omega_1 - \Omega_2)$. Applying now Theorem 3.2.6(c), we arrive at both conclusions of the theorem under (b) and thus complete the proof. \square

The next result derives geometrically the chain rules for Fenchel conjugates.

Theorem 3.3.3. *Let $A: X \rightarrow Y$ be a linear continuous mapping, and let $g: Y \rightarrow \overline{\mathbb{R}}$ be a convex function. Assume that:*

(a) *either g is finite and continuous at some point of AX ,*

(b) or X and Y are Banach spaces, g is lower semi-continuous, and $\mathbb{R}^+(\text{dom } g - AX)$ is a closed subspace of Y .

Then the conjugate of the composition $g \circ A$ is represented by

$$(3.3.18) \quad (g \circ A)^*(x^*) = \inf \{g^*(y^*) \mid y^* \in (A^*)^{-1}(x^*)\}.$$

Moreover, in both cases we have that for any $x^* \in \text{dom } (g \circ A)^*$ there is $y^* \in (A^*)^{-1}(x^*)$ with

$$(g \circ A)^*(x^*) = g^*(y^*).$$

Proof. Fix any $y^* \in (A^*)^{-1}(x^*)$ and deduce the following relationships from the definitions:

$$\begin{aligned} g^*(y^*) &= \sup \{ \langle y^*, y \rangle - g(y) \mid y \in Y \} \\ &\geq \sup \{ \langle y^*, A(x) \rangle - g(A(x)) \mid x \in X \} \\ &= \sup \{ \langle A^* y^*, x \rangle - (g \circ A)(x) \mid x \in X \} \\ &= \sup \{ \langle x^*, x \rangle - (g \circ A)(x) \mid x \in X \} = (g \circ A)^*(x^*). \end{aligned}$$

This readily implies the inequality “ \leq ” in (3.3.18). To verify the opposite inequality therein, take $x^* \in \text{dom } (g \circ A)^*$ and construct the following convex sets:

$$\Omega_1 := \text{gph } A \times \mathbb{R} \subset X \times Y \times \mathbb{R} \quad \text{and} \quad \Omega_2 := X \times \text{epi } g \subset X \times Y \times \mathbb{R}.$$

It follows directly from the above constructions that

$$(g \circ A)^*(x^*) = \sigma_{\Omega_1 \cap \Omega_2}(x^*, 0, -1) < \infty.$$

Let us further proceed with proof in each case (a) and (b).

(a) Under the assumption in (a) we clearly have

$$\text{int } \Omega_2 = \{ (x, y, \lambda) \in X \times Y \times \mathbb{R} \mid x \in X, y \in \text{int}(\text{dom } g), \lambda > g(y) \}$$

and easily observe that $\Omega_1 \cap (\text{int } \Omega_2) \neq \emptyset$ for the sets Ω_1, Ω_2 above. Then Theorem 3.2.6 (b) tells us that there are triples $(x_1^*, y_1^*, \alpha_1), (x_2^*, y_2^*, \alpha_2) \in X^* \times Y^* \times \mathbb{R}$ satisfying

$$(x^*, 0, -1) = (x_1^*, y_1^*, \alpha_1) + (x_2^*, y_2^*, \alpha_2),$$

and $\sigma_{\Omega_1 \cap \Omega_2}(x^*, 0, -1) = \sigma_{\Omega_1}(x_1^*, y_1^*, \alpha_1) + \sigma_{\Omega_2}(x_2^*, y_2^*, \alpha_2)$.

It follows from the structures of Ω_1, Ω_2 that $\alpha_1 = 0$ and $x_2^* = 0$. This yields

$$\sigma_{\Omega_1 \cap \Omega_2}(x^*, 0, -1) = \sigma_{\Omega_1}(x^*, -y_2^*, 0) + \sigma_{\Omega_2}(0, y_2^*, -1)$$

for some $y_2^* \in Y^*$. Thus we arrive at the representations

$$\begin{aligned} \sigma_{\Omega_1 \cap \Omega_2}(x^*, 0, -1) &= \sup \{ \langle x^*, x \rangle - \langle y_2^*, A(x) \rangle \mid x \in X \} + \sigma_{\text{epi } g}(y_2^*, -1) \\ &= \sup \{ \langle x^* - A^* y_2^*, x \rangle \mid x \in X \} + g^*(y_2^*), \end{aligned}$$

which allow us to conclude that $x^* = A^* y_2^*$ and therefore

$$\sigma_{\Omega_1 \cap \Omega_2}(x^*, 0, -1) = g^*(y_2^*) \geq \inf \{ g^*(y^*) \mid y^* \in (A^*)^{-1}(x^*) \}.$$

This justifies both statements of the theorem in case (a).

(b) It is easy to check the equality

$$\mathbb{R}^+(\Omega_1 - \Omega_2) = X \times \mathbb{R}^+(AX - \text{dom } g) \times \mathbb{R}.$$

Then we apply Theorem 3.2.6(c) and get both claimed statement in this case as well. \square

Finally in this section, we present a simple geometric proof of the Fenchel conjugate representation for the *maximum* of two convex functions $f, g: X \rightarrow \overline{\mathbb{R}}$ defined by

$$(3.3.19) \quad (f \vee g)(x) := \max \{ f(x), g(x) \}, \quad x \in X.$$

In the next theorem we use the convention that $0f := \delta_{\text{dom } f}$ and similarly for g ; see, e.g., [12].

Theorem 3.3.4. *Given convex functions $f, g: X \rightarrow \overline{\mathbb{R}}$, suppose that either the assumptions in (a) or the assumptions in (b) of Theorem 3.3.2 are satisfied. Then we have the representation*

$$(3.3.20) \quad (f \vee g)^*(x^*) = \inf_{\lambda \in [0,1]} [\lambda f + (1 - \lambda)g]^*(x^*).$$

If furthermore $(f \vee g)^(x^*) \in \mathbb{R}$, then the minimum in (3.3.20) is achieved.*

Proof. Let us first check that the inequality “ \leq ” always holds in (3.3.20) whenever $\lambda \in [0, 1]$. Indeed, it follows directly from the definitions that

$$\begin{aligned} [\lambda f + (1 - \lambda)g]^*(x^*) &= \sup_{x \in X} [\langle x^*, x \rangle - \lambda f(x) - (1 - \lambda)g(x)] \\ &\geq \sup_{x \in X} [\langle x^*, x \rangle - \lambda(f \vee g)(x) - (1 - \lambda)(f \vee g)(x)] \\ &= \sup_{x \in X} [\langle x^*, x \rangle - (f \vee g)(x)] = (f \vee g)^*(x^*), \quad x^* \in X^*. \end{aligned}$$

To verify the opposite inequality, observe that $\text{epi}(f \vee g) = \text{epi} f \cap \text{epi} g$, and hence we deduce from Lemma 3.3.1 the relationships

$$(f \vee g)^*(x^*) = \sigma_{\text{epi}(f \vee g)}(x^*, -1) = \sigma_{\Omega_1 \cap \Omega_2}(x^*, -1) \quad \text{with } \Omega_1 := \text{epi} f, \Omega_2 = \text{epi} g.$$

It follows from Theorem 3.2.6 under the assumptions in either (a) or (b) that

$$(f \vee g)^*(x^*) = [\sigma_{\Omega_1} \oplus \sigma_{\Omega_2}](x^*, -1), \quad x^* \in X^*.$$

Observe that the continuity of f at $\bar{x} \in \text{dom} f \cap \text{dom} g$ yields $(\text{int } \Omega_1) \cap \Omega_2 \neq \emptyset$. In the case where $\mathbb{R}^+(\text{dom} f - \text{dom} g)$ is a closed subspace we have that the set $\mathbb{R}^+(\Omega_1 - \Omega_2) = \mathbb{R}^+(\text{dom} f - \text{dom} g) \times \mathbb{R}$ is a closed subspace as well.

Suppose now that $(f \vee g)^*(x^*) \in \mathbb{R}$, i.e., $x^* \in \text{dom}(f \vee g)^*$. Then Theorem 3.2.6 gives us pairs $(x_1^*, -\lambda_1), (x_2^*, -\lambda_2) \in X^* \times \mathbb{R}$ such that $x^* = x_1^* + x_2^*$, $1 = \lambda_1 + \lambda_2$, and

$$(f \vee g)^*(x^*) = \sigma_{\text{epi}(f \vee g)}(x^*, -1) = \sigma_{\Omega_1 \cap \Omega_2}(x^*, -1) = \sigma_{\Omega_1}(x_1^*, -\lambda_1) + \sigma_{\Omega_2}(x_2^*, -\lambda_2).$$

Note that if either $\lambda_1 < 0$ or $\lambda_2 < 0$, then $\sigma_{\Omega_1}(x_1^*, -\lambda_1) = \infty$ or $\sigma_{\Omega_2}(x_1^*, -\lambda_1) = \infty$, which is a contradiction. Thus $\lambda_1, \lambda_2 \geq 0$. In the case where $\lambda_1, \lambda_2 > 0$ it follows that

$$\begin{aligned} (f \vee g)^*(x^*) &= \lambda_1 \sigma_{\Omega_1}\left(\frac{x_1^*}{\lambda_1}, -1\right) + \lambda_2 \sigma_{\Omega_2}\left(\frac{x_2^*}{\lambda_2}, -1\right) = \lambda_1 f^*\left(\frac{x_1^*}{\lambda_1}\right) + \lambda_2 g^*\left(\frac{x_2^*}{\lambda_2}\right) \\ &= (\lambda_1 f)^*(x_1^*) + (\lambda_2 g)^*(x_2^*) \quad \text{with } x^* \in \text{dom}(f \vee g)^*. \end{aligned}$$

Furthermore, we obviously have the estimate

$$(\lambda_1 f)^*(x_1^*) + (\lambda_2 g)^*(x_2^*) \geq [\lambda_1 f + (1 - \lambda_1)g]^*(x^*) \geq \inf_{\lambda \in [0,1]} [\lambda f + (1 - \lambda)g]^*(x^*).$$

Plugging in $\lambda_1 = 0$ gives us $\lambda_2 = 1$, and hence

$$\begin{aligned} \sigma_{\Omega_1}(x_1^*, -\lambda_1) + \sigma_{\Omega_2}(x_2^*, -\lambda_2) &= \sigma_{\text{dom } f}(x_1^*) + g^*(x_2^*) \\ &= \delta_{\text{dom } f}^*(x_1^*) + g^*(x_2^*) \geq [\delta_{\text{dom } f} + g]^*(x^*) \\ &\geq \inf_{\lambda \in [0,1]} [\lambda f + (1 - \lambda)g]^*(x^*). \end{aligned}$$

Since the latter inequality also holds if $(f \vee g)^*(x^*) = \infty$, this completes the proof. \square

3.4. Geometric Approach to Convex Subdifferential Calculus

This section presents a geometric design of major rules of convex subdifferential calculus (including some new results) that is based on applications of the normal cone intersection rule taken from Theorem 3.2.7. Having in mind a clear illustration of the developed geometric approach, we mainly confine ourselves to the applications of the easily formulated qualification conditions in (b) and (c) therein. Let us start with the fundamental sum rule in both LCTV and Banach space settings.

Theorem 3.4.1. *Given convex functions $f, g: X \rightarrow \overline{\mathbb{R}}$, we have the subdifferential sum rule*

$$(3.4.21) \quad \partial(f + g)(\bar{x}) = \partial f(\bar{x}) + \partial g(\bar{x}) \quad \text{for all } \bar{x} \in \text{dom } f \cap \text{dom } g$$

provided the validity of the following:

- (a) either f is finite and continuous at some point in $\text{dom } f \cap \text{dom } g$,
- (b) or X is a Banach space, f and g are lower semi-continuous, and $\mathbb{R}^+(\text{dom } f - \text{dom } g)$ is a subspace of X .

Proof. Let $\bar{x} \in \text{dom } f \cap \text{dom } g$ be fixed for the entire proof. Since the inclusion “ \supset ” in (3.4.21) can be easily checked by the definition, we now concentrate on proving the opposite inclusion therein. Pick any $x^* \in \partial(f + g)(\bar{x})$ and show how the geometric results of Theorem 3.2.7 can be used in verifying $x^* \in \partial f(\bar{x}) + \partial g(\bar{x})$. Having

$$\langle x^*, x - \bar{x} \rangle \leq (f + g)(x) - (f + g)(\bar{x}) \text{ for all } x \in X,$$

define the following convex subsets of $X \times \mathbb{R} \times \mathbb{R}$

$$\begin{aligned} \Omega_1 &:= \{(x, \lambda_1, \lambda_2) \in X \times \mathbb{R} \times \mathbb{R} \mid \lambda_1 \geq f(x)\}, \\ \Omega_2 &:= \{(x, \lambda_1, \lambda_2) \in X \times \mathbb{R} \times \mathbb{R} \mid \lambda_2 \geq g(x)\}. \end{aligned}$$

It follows from the definition that $(x^*, -1, -1) \in N((\bar{x}, f(\bar{x}), g(\bar{x})); \Omega_1 \cap \Omega_2)$. The application of Theorem 3.2.7 under an appropriate qualification condition yields

$$(3.4.22) \quad (x^*, -1, -1) \in N((\bar{x}, f(\bar{x}), g(\bar{x})); \Omega_1) + N((\bar{x}, f(\bar{x}), g(\bar{x})); \Omega_2),$$

which tells us therefore that

$$(x^*, -1, -1) = (x_1^*, -\lambda_1, -\lambda_2) + (x_2^*, -\gamma_1, -\gamma_2)$$

with

$$(x_1^*, -\lambda_1, -\lambda_2) \in N((\bar{x}, f(\bar{x}), g(\bar{x})); \Omega_1) \text{ and } (x_2^*, -\gamma_1, -\gamma_2) \in N((\bar{x}, f(\bar{x}), g(\bar{x})); \Omega_2).$$

By the construction of Ω_1 and Ω_2 we have $\lambda_2 = \gamma_1 = 0$ and hence find dual elements $(x_1^*, -1) \in N((\bar{x}, f(\bar{x})); \text{epi } f)$ and $(x_2^*, -1) \in N((\bar{x}, g(\bar{x})); \text{epi } g)$ satisfying the relationships

$$x_1^* \in \partial f(\bar{x}), \quad x_2^* \in \partial g(\bar{x}), \quad \text{and} \quad x^* = x_1^* + x_2^*.$$

This shows that $x^* \in \partial f(\bar{x}) + \partial g(\bar{x})$, and thus (3.4.21) holds, provided that the corresponding conditions of Theorem 3.2.7 are satisfied under the assumptions imposed in the theorem.

To this end, we easily observe that (a) yields $(\text{int } \Omega_1) \cap \Omega_2 \neq \emptyset$ and that (b) ensures the closedness of the subspace

$$\mathbb{R}^+(\Omega_1 - \Omega_2) = \mathbb{R}^+(\text{dom } f - \text{dom } g) \times \mathbb{R} \times \mathbb{R}$$

for the sets Ω_1, Ω_2 defined above. □

Next we employ the geometric approach to obtain a chain rule for convex subgradients under different qualification conditions.

Theorem 3.4.2. *Let $A: X \rightarrow Y$ be a continuous linear mapping, and let $g: Y \rightarrow \overline{\mathbb{R}}$ be convex. Suppose that:*

- (a) *either g is finite and continuous at some point of AX .*
- (b) *or X is Banach, g is lower semi-continuous, and $\mathbb{R}^+(AX - \text{dom } g)$ is a closed subspace of X .*

Then we have the subdifferential chain rule

(3.4.23)

$$\partial(g \circ A)(\bar{x}) = A^* \partial g(A\bar{x}) := \{A^* y^* \mid y^* \in \partial g(A\bar{x})\} \quad \text{for all } \bar{x} \in \text{dom}(g \circ A).$$

Proof. First we verify the inclusion “ \subset ” in (3.4.23) under the validity of (a) and (b). Pick any $y^* \in \partial(g \circ A)(\bar{x})$ and form the subsets of $X \times Y \times \mathbb{R}$ by

$$\Omega_1 := (\text{gph } A) \times \mathbb{R} \quad \text{and} \quad \Omega_2 := X \times (\text{epi } g).$$

If (a) holds, then it follows from the proof of Theorem 3.3.3(a) that $\Omega_1 \cap (\text{int } \Omega_2) \neq \emptyset$. Using the definitions of subgradients and normals of convex analysis, we easily conclude that $(y^*, 0, -1) \in N((\bar{x}, \bar{y}, \bar{z}); \Omega_1 \cap \Omega_2)$ with $\bar{z} := g(\bar{y})$ and $\bar{y} := A\bar{x}$. The

intersection rule from Theorem 3.2.7(b) tells us that

$$(y^*, 0, -1) \in N((\bar{x}, \bar{y}, \bar{z}); \Omega_1) + N((\bar{x}, \bar{y}, \bar{z}); \Omega_2),$$

which ensures the existence of

$$(y^*, -w^*) \in N((\bar{x}, \bar{y}); \text{gph } A) \text{ and } (w^*, -1) \in N((\bar{y}, \bar{z}); \text{epi } g)$$

satisfying $(y^*, 0, -1) = (y^*, -w^*, 0) + (0, w^*, -1)$. We can directly check that

$$N((\bar{x}, \bar{y}); \text{gph } A) = \{(x^*, y^*) \in X^* \times Y^* \mid x^* = -A^*y^*\}$$

and therefore deduce from the above that

$$y^* = A^*w^* \text{ and } w^* \in \partial g(\bar{y}).$$

This gives us $y^* \in A^*\partial g(\bar{y})$ and hence justifies the claimed inclusion “ \subset ” in (3.4.23).

To verify the opposite inclusion in (3.4.23), fix $y^* \in \partial g(A\bar{x})$ and deduce from the definitions of convex subgradients and adjoint linear operators that

$$\langle y^*, Ax - A\bar{x} \rangle \leq g(Ax) - g(A\bar{x}) \iff \langle A^*y^*, x - \bar{x} \rangle \leq (g \circ A)(x) - (g \circ A)(\bar{x})$$

for every $x \in X$. This implies that $A^*\partial g(\bar{y}) \subset \partial(g \circ A)(\bar{x})$, which completes the proof of the theorem under the assumptions in (a)

If now the assumptions in (b) are satisfied, then we can easily see that

$$\mathbb{R}^+(\Omega_1 - \Omega_2) = X \times \mathbb{R}^+(AX - \text{dom } g) \times \mathbb{R}.$$

It remains to employ Theorem 3.2.7(c) in this setting by using the arguments above. \square

The final result of this section employs the geometric approach to derive formulas for subdifferentiation of maximum of convex functions. For simplicity we consider the case of two functions; the maximum of finitely many ones can be deduced by induction.

Theorem 3.4.3. *Let $\bar{x} \in \text{dom } f \cap \text{dom } g$, where the functions $f, g: X \rightarrow \overline{\mathbb{R}}$ are convex. Then the following assertions hold:*

(a) *If $f(\bar{x}) > g(\bar{x})$ and g is upper semi-continuous at \bar{x} , then*

$$\partial(f \vee g)(\bar{x}) = \partial f(\bar{x}).$$

(b) *If $f(\bar{x}) < g(\bar{x})$ and f is upper semi-continuous at \bar{x} , then*

$$\partial(f \vee g)(\bar{x}) = \partial g(\bar{x}).$$

(c) *If $f(\bar{x}) = g(\bar{x})$, then*

$$(3.4.24) \quad \partial(f \vee g)(\bar{x}) = \text{co}[\partial f(\bar{x}) \cup \partial g(\bar{x})].$$

provided that either the assumptions in (a) or the assumptions in (b) of Theorem 3.3.2 are satisfied. Here the symbol ‘co’ stands for the convex hull of a set.

Proof. We obviously have $\text{epi}(f \vee g) = \text{epi } f \cap \text{epi } g$, and thus it follows from the subgradient definition that $x^* \in \partial(f \vee g)(\bar{x})$ if and only if

$$(3.4.25) \quad (x^*, -1) \in N((\bar{x}, (f \vee g)(\bar{x})); \text{epi}(f \vee g)) = N((\bar{x}, (f \vee g)(\bar{x})); \text{epi } f \cap \text{epi } g).$$

Denote $\bar{\lambda} := (f \vee g)(\bar{x})$. Under the assumption given in (a), it is easy to check that $(\bar{x}, \bar{\lambda}) \in \text{int}(\text{epi } g)$ due to the imposed upper semicontinuity assumption on g . This allows us to exclude the latter function from the intersection in (3.4.25) and arrive at the conclusion. The proof of assertion (b) is the same.

It remains to verify formula (3.4.24) under the assumptions given in (c). To furnish this, we only need to apply Theorem 3.2.7 and proceed similarly to the proof of Theorem 3.3.4. □

DC Programming for Constrained Multifacility Location and Clustering

4.1. Overview

In the current time of “big data”, clustering is a very important problem that helps classify data in many fields such as machine learning, pattern recognition, image analysis, data compression, and computer graphics. Given a finite number of data points with a measurement distance, a centroid-based clustering problem seeks a finite number of cluster centers with each data point assigned to the nearest cluster center in a way that a certain measurement distance is minimized.

It is well-known that the k -mean algorithm is one of the simplest clustering algorithms, providing an easy way to classify a given data set through a certain number of clusters. However, it possesses certain drawbacks: the k -mean algorithm depends heavily on the initial choice of cluster centers; there is no guarantee that the k -mean algorithm converges to a global optimal solution; the number of clusters k is an input parameter: an inappropriate choice of k may yield poor results; the results depend heavily on the measurement distance; the algorithm may not be applicable for handling constraints imposed on the cluster centers.

This chapter focuses on solving a number of clustering and multifacility location problems with constraints. We use a penalty method with squared Euclidean distance functions to convert constrained problems to unconstrained problems. Then appropriate DC decompositions and the DCA are used to minimize the penalized objective functions. In the case where the measurement distance is defined by the Euclidean norm instead of the squared Euclidean norm, we use Nesterov’s smoothing techniques for *reducing the nonsmoothness* of the model and for providing a DC decomposition that is favorable for applying the DCA. Our method opens up the possibility of using distance function

penalty methods for other problems of DC programming. This chapter presents the work we have done in [34]. The reader can find more details in there.

For convenience, we define the *data matrix* $\mathbf{A} \in \mathbb{R}^{m \times d}$ as the matrix whose i^{th} row is $a^i \in \mathbb{R}^d$ for $i = 1, \dots, m$. Similarly, we define the *variable matrix* $\mathbf{X} \in \mathbb{R}^{k \times d}$ as the matrix whose ℓ^{th} row is $x^\ell \in \mathbb{R}^d$ for $\ell = 1, \dots, k$. We equip the linear space $\mathbb{R}^{k \times d}$ with the inner product $\langle \mathbf{X}, \mathbf{Y} \rangle = \text{trace}(\mathbf{X}^T \mathbf{Y})$. Recall that the *Frobenius norm* on $\mathbb{R}^{k \times d}$ is defined by

$$\|\mathbf{X}\|_F = \langle \mathbf{X}, \mathbf{X} \rangle^{1/2} = \left(\sum_{\ell=1}^k \langle x^\ell, x^\ell \rangle \right)^{1/2} = \left(\sum_{\ell=1}^k \|x^\ell\|^2 \right)^{1/2}.$$

Observe that the square of the Frobenius norm is differentiable with

$$\nabla \|\mathbf{X}\|_F^2 = 2\mathbf{X} \text{ for } \mathbf{X} \in \mathbb{R}^{k \times d}.$$

Let $\Omega^\ell \subset \mathbb{R}^d$ for $\ell = 1, \dots, k$ be nonempty closed convex sets and let $\Omega = \Omega^1 \times \Omega^2 \times \dots \times \Omega^k$. For $\mathbf{X} \in \mathbb{R}^{k \times d}$, the projection from \mathbf{X} to Ω is the matrix \mathbf{Y} whose ℓ^{th} row is $y^\ell = P(x^\ell; \Omega^\ell)$. We thus have

$$(4.1.26) \quad [d(\mathbf{X}; \Omega)]^2 = \|\mathbf{X} - \mathbf{Y}\|_F^2 = \sum_{\ell=1}^k \|x^\ell - y^\ell\|^2 = \sum_{\ell=1}^k d(x^\ell; \Omega^\ell)^2.$$

4.2. Clustering with Constraints

In this section, we study problems of *clustering with constraints* in which the measurement distance is defined by the squared Euclidean norm. We seek k centers $x^1, \dots, x^k \in \mathbb{R}^d$ of m data nodes $a^1, \dots, a^m \in \mathbb{R}^d$ and impose the restriction that each $x^\ell \in \bigcap_{i=1}^q \Omega_i^\ell$ for some nonempty closed convex set $\Omega_i^\ell \subset \mathbb{R}^d$ with $\ell = 1, \dots, k$ and $i = 1, \dots, q$. Here, without loss of generality, we assume that the numbers of constraints for each center is equal to each other. The problem we are concerned with is given by

$$(4.2.27) \quad \begin{aligned} &\text{minimize} && \psi(x^1, \dots, x^k) = \sum_{i=1}^m \min_{\ell=1, \dots, k} \|x^\ell - a^i\|^2 \\ &\text{subject to} && x^\ell \in \bigcap_{j=1}^q \Omega_j^\ell \text{ for } \ell = 1, \dots, k. \end{aligned}$$

This problem can be converted to an unconstrained minimization problem

$$(4.2.28) \quad \text{minimize } f(x^1, \dots, x^k) = \frac{1}{2} \sum_{i=1}^m \min_{\ell=1, \dots, k} \|x^\ell - a^i\|^2 + \frac{\tau}{2} \sum_{\ell=1}^k \sum_{i=1}^q [d(x^\ell; \Omega_i^\ell)]^2,$$

$x^1, \dots, x^k \in \mathbb{R}^d$ and $\tau > 0$ is a penalty parameter.

Recall from Proposition 2.3.3 that for any nonempty closed convex set Ω in \mathbb{R}^d ,

$$[d(x; \Omega)]^2 = \|x\|^2 - \varphi_\Omega(x),$$

where $\varphi_\Omega(x) = 2 \sup \{ \langle x, w \rangle - \frac{1}{2} \|w\|^2 \mid w \in \Omega \}$ is a differentiable function with $\nabla \varphi_\Omega(x) = 2P(x; \Omega)$. Let us use the *minimum-sum principle* for k real numbers α_ℓ for $\ell = 1, \dots, k$:

$$\min_{\ell=1, \dots, k} \alpha_\ell = \sum_{\ell=1}^k \alpha_\ell - \max_{r=1, \dots, k} \sum_{\ell=1, \ell \neq r}^k \alpha_\ell$$

to obtain a DC decomposition of f as follows

$$\begin{aligned} f(x^1, \dots, x^k) &= \left(\frac{1}{2} \sum_{i=1}^m \sum_{\ell=1}^k \|x^\ell - a^i\|^2 + \frac{\tau q}{2} \sum_{\ell=1}^k \|x^\ell\|^2 \right) \\ &\quad - \left(\frac{1}{2} \sum_{i=1}^m \max_{r=1, \dots, k} \sum_{\ell=1, \ell \neq r}^k (\|x^\ell - a^i\|)^2 + \frac{\tau}{2} \sum_{\ell=1}^k \sum_{i=1}^q \varphi_{\Omega_i^\ell}(x^\ell) \right). \end{aligned}$$

We see that $f = g - h$ by defining

$$g_1(x^1, \dots, x^k) = \frac{1}{2} \sum_{i=1}^m \sum_{\ell=1}^k \|x^\ell - a^i\|^2, \quad g_2(x^1, \dots, x^k) = \frac{\tau q}{2} \sum_{\ell=1}^k \|x^\ell\|^2,$$

$$h_1(x^1, \dots, x^k) = \frac{1}{2} \sum_{i=1}^m \max_{r=1, \dots, k} \sum_{\ell=1, \ell \neq r}^k \|x^\ell - a^i\|^2, \quad h_2(x^1, \dots, x^k) = \frac{\tau}{2} \sum_{\ell=1}^k \sum_{i=1}^q \varphi_{\Omega_i^\ell}(x^\ell),$$

and setting $g = g_1 + g_2$ and $h = h_1 + h_2$.

As discussed in the overview, we may collect x^j into the variable matrix \mathbf{X} and denote $\Omega_i = \Omega_i^1 \times \Omega_i^2 \times \dots \times \Omega_i^k \in \mathbb{R}^{k \times d}$ for $i = 1, \dots, q$. Then (4.2.27) becomes

$$\begin{aligned} & \text{minimize} && \psi(\mathbf{X}) \\ & \text{subject to} && \mathbf{X} \in \bigcap_{i=1}^q \Omega_i. \end{aligned}$$

We also collect a^i into the data matrix \mathbf{A} , and upon doing so we may express g in terms of the Frobenius norm, namely,

$$\begin{aligned} g_1(\mathbf{X}) &= \frac{1}{2} \sum_{i=1}^m \sum_{\ell=1}^k \|x^\ell - a^i\|^2 = \frac{1}{2} \sum_{i=1}^m \sum_{\ell=1}^k \left(\|x^\ell\|^2 - 2\langle x^\ell, a^i \rangle + \|a^i\|^2 \right) \\ &= \frac{m}{2} \sum_{\ell=1}^k \|x^\ell\|^2 - \sum_{i=1}^m \sum_{\ell=1}^k \langle x^\ell, a^i \rangle + \frac{k}{2} \sum_{i=1}^m \|a^i\|^2 \\ &= \frac{m}{2} \|\mathbf{X}\|_F^2 - \langle \mathbf{X}, \mathbf{EA} \rangle + \frac{k}{2} \|\mathbf{A}\|_F^2, \end{aligned}$$

where $\mathbf{E} \in \mathbb{R}^{k \times m}$ is the matrix with all entries equal to 1. In this form, it is easy to obtain the gradient of g_1 as

$$\nabla g_1(\mathbf{X}) = m\mathbf{X} - \mathbf{EA}.$$

Similarly, g_2 can be equivalently written as

$$g_2(\mathbf{X}) = \frac{\tau q}{2} \sum_{\ell=1}^k \|x^\ell\|^2 = \frac{\tau q}{2} \|\mathbf{X}\|_F^2.$$

Hence, g_2 is differentiable and its gradient is given by $\nabla g_2(\mathbf{X}) = \tau q \mathbf{X}$. Therefore,

$$\nabla g(\mathbf{X}) = \nabla g_1(\mathbf{X}) + \nabla g_2(\mathbf{X}) = (m + \tau q)\mathbf{X} - \mathbf{EA}.$$

Based on the relation (1.2.1), finding $\mathbf{X} \in \partial g^*(\mathbf{Y})$ is equivalent to solving the equation

$$\mathbf{Y} = (m + \tau q)\mathbf{X} - \mathbf{EA}.$$

It follows that

$$\mathbf{X} = \frac{\mathbf{Y} + \mathbf{EA}}{m + \tau q} \in \partial g^*(\mathbf{Y}).$$

Our goal, then, is to find $\mathbf{Y}_p \in \partial h(\mathbf{X}_p)$ from which we will obtain \mathbf{X}_{p+1} and thereby compute the first N terms of the sequence $\{\mathbf{X}_p\}$ via Algorithm 1. Toward the end we will find subgradients of the convex function h .

For each $i = 1, \dots, m$, let $r(i) \in \{1, \dots, k\}$ be an index for which

$$\sum_{\ell=1, \ell \neq r(i)}^k \|x^\ell - a^i\|^2 = \max_{r=1, \dots, k} \sum_{\ell=1, \ell \neq r}^k \|x^\ell - a^i\|^2,$$

in which case we see that a subgradient $\mathbf{W} \in \partial h_1(\mathbf{X})$ is given by

(4.2.29)

$$\mathbf{W} = \sum_{i=1}^m \left(\mathbf{X} - \mathbf{A}_i - e_{r(i)}(x^{r(i)} - a^i) \right) = m\mathbf{X} - \mathbf{EA} - \sum_{i=1}^m e_{r(i)}(x^{r(i)} - a^i),$$

where $\mathbf{A}_i \in \mathbb{R}^{k \times d}$ is the matrix whose all rows are a^i and e_r is the $k \times 1$ column vector with a one in the r^{th} position and zeros elsewhere.

Now for $h_2(\mathbf{X}) = \frac{\tau}{2} \sum_{\ell=1}^k \sum_{i=1}^q \varphi_{\Omega_i^\ell}(x^\ell)$, we have

$$\frac{\partial h_2}{\partial x^j}(\mathbf{X}) = \frac{\tau}{2} \sum_{\ell=1}^k \sum_{i=1}^q \frac{\partial}{\partial x^j} \varphi_{\Omega_i^\ell}(x^\ell) = \tau \sum_{i=1}^q P(x^j; \Omega_i^j)$$

with $j = 1, \dots, k$. Then $\mathbf{U} = \frac{1}{\tau} \nabla h_2(\mathbf{X})$ is the $k \times d$ matrix whose rows are $u^j = \sum_{i=1}^q P(x^j; \Omega_i^j)$.

The form of the DCA instructs us to find $\mathbf{Y}_p \in \partial h(\mathbf{X}_p)$ at the p th iteration, so we set $\mathbf{Y}_p = \mathbf{W} + \tau \mathbf{U}$. Combining the above results gives $\mathbf{X}_{p+1} = (\mathbf{W} + \tau \mathbf{U} + \mathbf{EA}) / (m + \tau q)$.

Substituting (4.2.29) for \mathbf{W} , we obtain the recursive relation

$$\mathbf{X}_{p+1} = \frac{1}{m + \tau q} \left(m\mathbf{X}_p + \tau \mathbf{U} - \sum_{i=1}^m e_{r(i)}(x_p^{r(i)} - a^i) \right),$$

where x_p^ℓ denotes the ℓ^{th} row of \mathbf{X}_p . The following algorithm summarizes the DCA-based procedure we just derived.

Algorithm 3 : DC program for (4.2.28)

INPUT: $\mathbf{A}, \mathbf{X}_0, \{\Omega_j^\ell\}_{j=1,\dots,q}^{\ell=1,\dots,k}, N, \tau$
for $p = 1, \dots, N$ **do**
 for $i = 1, \dots, m$ **do**
 Find $r(i)$ s.t. $\|x_{p-1}^{r(i)} - a^i\|^2 = \min\{\|x_{p-1}^\ell - a^i\|^2 \mid \ell = 1, \dots, k\}$
 Set $\mathbf{W}_i := e_{r(i)}(x_{p-1}^{r(i)} - a^i)$
 end for
 for $\ell = 1, \dots, k$ **do**
 Find $u^\ell := \sum_{j=1}^q P(x_{p-1}^\ell; \Omega_j^\ell)$
 end for
 Set $\mathbf{X}_p := \frac{1}{m+\tau q} \left(m\mathbf{X}_{p-1} + \tau\mathbf{U} - \sum_{i=1}^m \mathbf{W}_i \right)$
end for
OUTPUT: \mathbf{X}_N

Inspecting (4.2.28), we see that for small τ our problem begins to resemble the associated unconstrained problem. For solving the clustering (4.2.27), we may gradually increase the value of the penalty parameter $\tau > 0$ by periodically multiplying by some $\sigma > 1$ and terminate whenever $\tau > \tau_f$. This may be accomplished by Algorithm 4. Notice that for the initial choice of τ , the maximum number of overall iterations of Algorithm 4 is $N \lceil \log_\sigma(\tau_f/\tau) \rceil$, where $\lceil \cdot \rceil$ denotes the ceiling function.

Algorithm 4 : Penalty DC program for (4.2.27)

INPUT: $\mathbf{A}, \mathbf{X}_0, \{\Omega_j^\ell\}_{j=1,\dots,q}^{\ell=1,\dots,k}, N, \tau, \sigma, \tau_f$
while $\tau < \tau_f$ **do**
 Find \mathbf{X}_N by executing **Algorithm 3** with $\mathbf{A}, \mathbf{X}_0, \{\Omega_j^\ell\}_{j=1,\dots,q}^{\ell=1,\dots,k}, N, \tau$
 Reassign $\mathbf{X}_0 := \mathbf{X}_N$
 Reassign $\tau := \sigma\tau$
end while
OUTPUT: \mathbf{X}_N

Remark 4.2.1. Assume that in problem (4.2.27) the constraint $Q_\ell = \cap_{j=1}^q \Omega_j^\ell$ for each center x^ℓ for $\ell = 1, \dots, k$ is simple enough. Then we can use the *projected k-means algorithm* which is similar to the *k-means algorithm* as follows. After assigning each data point to its nearest centroid, we update each centroid by

$$x^\ell \leftarrow P \left(\frac{\sum_{i \in \mathbf{A}(x^\ell)} a^i}{|\mathbf{A}(x^\ell)|}; Q_\ell \right),$$

where $\mathbf{A}(x^\ell) = \{i \in \{1, \dots, m\} \mid \|a^i - x^\ell\| = \min_{j=1, \dots, k} \|a^i - x^j\|\}$ is the cluster associated with x^ℓ . In this case, to get a good starting point \mathbf{X}_0 for Algorithm 3, we perform the projected k -means algorithm in several initial steps. This procedure has shown its efficiency in clustering without constraint; see, e.g., [1].

4.3. Set Clustering with Constraints

In this section, we turn our attention to a model of *set clustering* with constraints, i.e., for given m subsets $\Lambda_1, \dots, \Lambda_m \subset \mathbb{R}^d$, we seek k cluster centers $x^\ell \in \bigcap_{j=1}^q \Omega_j^\ell$ for $\ell = 1, \dots, k$, where each Ω_j^ℓ is a subset of \mathbb{R}^d . The measurement distance is defined by the squared distance functions to the sets involved. The optimization modeling of the problem to be solved is given by

$$(4.3.30) \quad \begin{aligned} & \text{minimize} && \psi(x^1, \dots, x^k) = \sum_{i=1}^m \min_{\ell=1, \dots, k} [d(x^\ell; \Lambda_i)]^2 \\ & \text{subject to} && x^\ell \in \bigcap_{j=1}^q \Omega_j^\ell \text{ for } \ell = 1, \dots, k. \end{aligned}$$

Throughout this section, we assume that Λ_i for $i = 1, \dots, m$ and Ω_j^ℓ for $j = 1, \dots, q$ and $\ell = 1, \dots, k$ are nonempty, closed and convex.

Using the penalty method based on distance functions with a parameter $\tau > 0$, we consider the constrained set clustering model:

$$(4.3.31) \quad \text{minimize } f(x^1, \dots, x^k) = \frac{1}{2} \sum_{i=1}^m \min_{\ell=1, \dots, k} [d(x^\ell; \Lambda_i)]^2 + \frac{\tau}{2} \sum_{\ell=1}^k \sum_{j=1}^q [d(x^\ell; \Omega_j^\ell)]^2,$$

$x^1, \dots, x^k \in \mathbb{R}^d$. Using the *minimum-sum principle*, we will now find a DC decomposition of $f = g - h$ as follows. For each $i = 1, \dots, m$, we have

$$\min_{\ell=1, \dots, k} [d(x^\ell; \Lambda_i)]^2 = \sum_{\ell=1}^k [d(x^\ell; \Lambda_i)]^2 - \max_{r=1, \dots, k} \sum_{\ell=1, \ell \neq r}^k [d(x^\ell; \Lambda_i)]^2$$

Thus,

$$\begin{aligned} \min_{\ell=1,\dots,k} [d(x^\ell; \Lambda_i)]^2 &= \left(\|x^\ell\|^2 - \varphi_{\Lambda_i}(x^\ell) \right) - \max_{r=1,\dots,k} \sum_{\ell=1, \ell \neq r}^k [d(x^\ell; \Lambda_i)]^2 \\ &= \|\mathbf{X}\|_F^2 - \left(\sum_{\ell=1}^k \varphi_{\Lambda_i}(x^\ell) + \max_{r=1,\dots,k} \sum_{\ell=1, \ell \neq r}^k [d(x^\ell; \Lambda_i)]^2 \right). \end{aligned}$$

Furthermore, we have

$$\sum_{\ell=1}^k \sum_{j=1}^q [d(x^\ell; \Omega_j^\ell)]^2 = \sum_{\ell=1}^k \sum_{j=1}^q \left(\|x^\ell\|^2 - \varphi_{\Omega_j^\ell}(x^\ell) \right) = q \|\mathbf{X}\|_F^2 - \sum_{\ell=1}^k \sum_{j=1}^q \varphi_{\Omega_j^\ell}(x^\ell).$$

Let

$$\begin{aligned} g_1(\mathbf{X}) &= \frac{m}{2} \|\mathbf{X}\|_F^2, \quad g_2(\mathbf{X}) = \frac{\tau q}{2} \|\mathbf{X}\|_F^2, \\ h_1(\mathbf{X}) &= \sum_{i=1}^m \left(\frac{1}{2} \sum_{\ell=1}^k \varphi_{\Lambda_i}(x^\ell) + \frac{1}{2} \max_{r=1,\dots,k} \sum_{\ell=1, \ell \neq r}^k [d(x^\ell; \Lambda_i)]^2 \right), \end{aligned}$$

and $h_2(\mathbf{X}) = \frac{\tau}{2} \sum_{\ell=1}^k \sum_{j=1}^q \varphi_{\Omega_j^\ell}(x^\ell)$. In this case we have the DC decomposition $f = g - h$, where $g = g_1 + g_2$ and $h = h_1 + h_2$ are convex.

Using the relation (1.2.1), we can easily see that $\mathbf{X} = \frac{1}{m+\tau q} \mathbf{Y} \in \partial g^*(\mathbf{Y})$. To apply the DCA from Algorithm 1, we also need to find $\mathbf{Y} \in \partial h(\mathbf{X})$ as $\mathbf{Y} = \mathbf{V} + \mathbf{U}$, where $\mathbf{V} \in \partial h_1(\mathbf{X})$ and $\mathbf{U} \in \partial h_2(\mathbf{X})$.

Now, we focus on finding $\mathbf{V} \in \partial h_1(\mathbf{X})$. Define

$$D_i(\mathbf{X}) = \frac{1}{2} \sum_{\ell=1}^k \varphi_{\Lambda_i}(x^\ell),$$

and

$$F_i(\mathbf{X}) = \frac{1}{2} \max_{r=1,\dots,k} \sum_{\ell=1, \ell \neq r}^k [d(x^\ell; \Lambda_i)]^2, \quad i = 1, \dots, m.$$

Then $h_1(\mathbf{X}) = \sum_{i=1}^m [D_i(\mathbf{X}) + F_i(\mathbf{X})]$. Based on Proposition 2.3.3, we see that $\nabla D_i(\mathbf{X})$ is the $k \times d$ matrix given by

$$\nabla D_i(\mathbf{X}) = \begin{bmatrix} P(x^1; \Lambda_i) \\ \vdots \\ P(x^k; \Lambda_i) \end{bmatrix}.$$

For each $i = 1, \dots, m$, choose an index $r(i)$ such that

$$\max_{r=1, \dots, k} \sum_{\ell=1, \ell \neq r}^k [d(x^\ell; \Lambda_i)]^2 = \sum_{\ell=1, \ell \neq r(i)}^k [d(x^\ell; \Lambda_i)]^2.$$

Now, for $j = 1, \dots, k$, define

$$v_i^j = \begin{cases} x^j - P(x^j; \Lambda_i) & \text{if } j \neq r(i), \\ 0 & \text{if } j = r(i). \end{cases}$$

By (1.2.2) and the fact that $\nabla[d(x; \Lambda)]^2 = 2(x - P(x; \Lambda))$ for a nonempty closed convex set Λ , the matrix V_i whose j^{th} row is v_i^j defines a subgradient of F_i at \mathbf{X} . It follows that such a subgradient \mathbf{V} is

$$\mathbf{V} = m\mathbf{X} - \sum_{i=1}^m e_{r(i)} \left(x^{r(i)} - P(x^{r(i)}; \Lambda_i) \right).$$

As computed in the previous section, $\nabla h_2(\mathbf{X}) = \tau\mathbf{U}$, where \mathbf{U} is the $k \times d$ matrix whose ℓ^{th} row is $\sum_{j=1}^k P(x^\ell; \Omega_j^\ell)$ for $\ell = 1, \dots, k$. Consequently, the $k \times d$ matrix

$$\begin{aligned} \mathbf{Y} &= m\mathbf{X} - \sum_{i=1}^m e_{r(i)} \left(x^{r(i)} - P(x^{r(i)}; \Lambda_i) \right) + \tau\mathbf{U} \\ &= m\mathbf{X} + \tau\mathbf{U} - \sum_{i=1}^m e_{r(i)} \left(x^{r(i)} - P(x^{r(i)}; \Lambda_i) \right) \end{aligned}$$

belongs to $\partial h(\mathbf{X})$.

Now, for $p \in \mathbb{N}$ such that \mathbf{X}_{p-1} is given, one has

$$\mathbf{Y}_{p-1} = m\mathbf{X}_{p-1} + \tau\mathbf{U}_p - \sum_{i=1}^m e_{r(i)} \left(x_{p-1}^{r(i)} - P(x_{p-1}^{r(i)}; \Lambda_i) \right) \in \partial H(\mathbf{X}_{p-1}),$$

where x_p^ℓ is the ℓ^{th} row of \mathbf{X}_p and \mathbf{U}_p is the $k \times d$ matrix whose ℓ^{th} row is $\sum_{j=1}^k P(x_p^\ell; \Omega_j^\ell)$ for $\ell = 1, \dots, k$. It follows that \mathbf{X}_p from the DCA in Algorithm 1 can be determined by

$$\mathbf{X}_p = \frac{1}{\tau q + m} \left(m\mathbf{X}_{p-1} + \tau\mathbf{U}_p - \sum_{i=1}^m e_{r(i)}(x_{p-1}^{r(i)} - P(x_{p-1}^{r(i)}; \Lambda_i)) \right).$$

We are ready to give an algorithm for (4.3.31).

Algorithm 5 : DC program for (4.3.31)

INPUT: $\mathbf{X}_0, \Lambda_i, \{\Omega_j^\ell\}_{j=1, \dots, q}^{\ell=1, \dots, k}, N, \tau$
for $p = 1, \dots, N$ **do**
 for $i = 1, \dots, m$ **do**
 for $\ell = 1, \dots, k$ **do**
 Set $w_i^\ell := P(x_{p-1}^\ell; \Lambda_i)$
 end for
 Find $r(i)$ s.t. $\|x_{p-1}^{r(i)} - a^i\|^2 = \min\{\|x_{p-1}^\ell - a^i\|^2 \mid \ell = 1, \dots, k\}$
 end for
 for $\ell = 1, \dots, k$ **do**
 Find $u^\ell := \sum_{j=1}^q P(x_{p-1}^\ell; \Omega_j^\ell)$
 end for
 $\mathbf{X}_p := \frac{1}{\tau q + m} \left(m\mathbf{X}_{p-1} + \tau\mathbf{U}_p - \sum_{i=1}^m e_{r(i)}(x_{p-1}^{r(i)} - w_i^{r(i)}) \right)$
end for
OUTPUT: \mathbf{X}_N

We now adapt Algorithm 5 to solve our set clustering problem. Just as in the previous section, we gradually increase the value of the penalty parameter $\tau > 0$ by periodically multiplying it by some $\sigma > 1$ and stopping when $\tau > \tau_f > 0$. This may be accomplished by Algorithm 6. We again see that for an initial choice of $\tau = \tau_0$, the maximum number of overall iterations of Algorithm 6 is $N \lceil \log_\sigma(\tau_f/\tau_0) \rceil$.

Algorithm 6 : Penalty DC program for (4.3.30)

INPUT: $\mathbf{X}_0, \{\Lambda_i\}_{i=1}^m, \{\Omega_j^\ell\}_{j=1, \dots, q}^{\ell=1, \dots, k}, N, \tau, \tau_f, \sigma$
while $\tau < \tau_f$ **do**
 Find \mathbf{X}_N by executing **Algorithm 5** with $\mathbf{X}_0, \{\Lambda_i\}_{i=1}^m, \{\Omega_j^\ell\}_{j=1, \dots, q}^{\ell=1, \dots, k}, \tau, N$
 Reassign $\mathbf{X}_0 := \mathbf{X}_N$
 Reassign $\tau := \sigma\tau$
end while
OUTPUT: \mathbf{X}_N

4.4. Multifacility Location with Constraints

Given a set of m points (nodes) a^1, a^2, \dots, a^m in \mathbb{R}^d , our goal is find k centers x^ℓ for $\ell = 1, \dots, k$, which must be in constraint sets $\bigcap_{i=1}^q \Omega_i^\ell$ for $\ell = 1, \dots, k$, such that the *transportation cost* to the nodes is minimized. The same setting in Section 4.2 gives us the constrained minimization problem

$$(4.4.32) \quad \begin{aligned} & \text{minimize} && \psi(\mathbf{X}) \\ & \text{subject to} && \mathbf{X} \in \bigcap_{i=1}^q \Omega_i, \end{aligned}$$

where the total cost now is given by

$$\psi(\mathbf{X}) = \psi(x^1, \dots, x^k) = \sum_{i=1}^m \min_{\ell=1, \dots, k} \|x^\ell - a^i\|.$$

This problem can be converted to an unconstrained minimization problem

$$(4.4.33) \quad \text{minimize} \quad f_\tau(x^1, \dots, x^k) = \sum_{i=1}^m \min_{\ell=1, \dots, k} \|x^\ell - a^i\| + \frac{\tau}{2} \sum_{\ell=1}^k \sum_{i=1}^q [d(x^\ell; \Omega_i^\ell)]^2,$$

where $x^1, \dots, x^k \in \mathbb{R}^d$ and $\tau > 0$ is a parameter.

We apply Nesterov's smoothing techniques from [36] to approximate the objective function f_τ by a new DC function which is favorable for applying the DCA.

$$\begin{aligned} f_{\tau, \mu}(x^1, \dots, x^k) &= \left(\frac{\mu}{2} \sum_{i=1}^m \sum_{\ell=1}^k \left\| \frac{x^\ell - a^i}{\mu} \right\|^2 + \frac{\tau q}{2} \sum_{\ell=1}^k \|x^\ell\|^2 \right) \\ &\quad - \left(\frac{\mu}{2} \sum_{i=1}^m \sum_{\ell=1}^k \left[d \left(\frac{x^\ell - a^i}{\mu}; \mathbb{B} \right) \right]^2 \right) \\ &\quad + \sum_{i=1}^m \max_{r=1, \dots, k} \sum_{\ell=1, \ell \neq r}^k \|x^\ell - a^i\| + \frac{\tau}{2} \sum_{\ell=1}^k \sum_{i=1}^q \varphi_{\Omega_i^\ell}(x^\ell). \end{aligned}$$

In what follows, we use f instead of $f_{\tau, \mu}$ for the simplicity of notations. The original clustering problem now can be solved using a DC programming

$$\text{minimize} \quad f(x^1, \dots, x^k) = g(x^1, \dots, x^k) - h(x^1, \dots, x^k), \quad x^1, \dots, x^k \in \mathbb{R}^d.$$

In this formulation, g and h are convex functions on $(\mathbb{R}^d)^k$ defined by

$$g(x^1, \dots, x^k) = g_1(x^1, \dots, x^k) + g_2(x^1, \dots, x^k),$$

$$h(x^1, \dots, x^k) = h_1(x^1, \dots, x^k) + h_2(x^1, \dots, x^k) + h_3(x^1, \dots, x^k),$$

with their respective components defined as

$$g_1 = \frac{\mu}{2} \sum_{i=1}^m \sum_{\ell=1}^k \left\| \frac{x^\ell - a^i}{\mu} \right\|^2, \quad g_2 = \frac{\tau q}{2} \sum_{\ell=1}^k \|x^\ell\|^2,$$

$$h_1 = \frac{\mu}{2} \sum_{i=1}^m \sum_{\ell=1}^k \left[d \left(\frac{x^\ell - a^i}{\mu}; \mathbb{B} \right) \right]^2, \quad h_2 = \sum_{i=1}^m \max_{r=1, \dots, k} \sum_{\ell=1, \ell \neq r}^k \|x^\ell - a^i\|,$$

$$h_3 = \frac{\tau}{2} \sum_{\ell=1}^k \sum_{i=1}^q \varphi_{\Omega_i^\ell}(x^\ell).$$

The function g_1 can be equivalently written as

$$\begin{aligned} g_1(\mathbf{X}) &= \frac{1}{2\mu} \sum_{i=1}^m \sum_{\ell=1}^k \|x^\ell - a^i\|^2 \\ &= \frac{1}{2\mu} \sum_{i=1}^m \sum_{\ell=1}^k \left(\|x^\ell\|^2 - 2\langle x^\ell, a^i \rangle + \|a^i\|^2 \right) \\ &= \frac{1}{2\mu} \left(m \|\mathbf{X}\|_F^2 - 2\langle \mathbf{X}, \mathbf{EA} \rangle + k \|\mathbf{A}\|_F^2 \right). \end{aligned}$$

Note that g_1 is differentiable and its gradient is given by

$$\nabla g_1(\mathbf{X}) = \frac{1}{\mu} [m\mathbf{X} - \mathbf{EA}].$$

The function g_2 is the same as before so its gradient is given by

$$\nabla g_2(\mathbf{X}) = \tau q \mathbf{X}.$$

Since $g(\mathbf{X}) = g_1(\mathbf{X}) + g_2(\mathbf{X})$, its gradient can be computed by

$$\nabla g(\mathbf{X}) = \nabla g_1(\mathbf{X}) + \nabla g_2(\mathbf{X}) = \frac{1}{\mu} (m\mathbf{X} - \mathbf{EA}) + \tau q \mathbf{X} = \left(\frac{m}{\mu} + \tau q \right) \mathbf{X} - \frac{1}{\mu} \mathbf{S},$$

where $\mathbf{S} = \mathbf{E}\mathbf{A}$. The latter can equivalently be written as

$$\mathbf{Y} = \left(\frac{m}{\mu} + \tau q\right)\mathbf{X} - \frac{1}{\mu}\mathbf{S}.$$

Our goal now is to compute $\nabla g^*(\mathbf{Y})$, which can be accomplished by the relation (1.2.1).

Then with some algebraic manipulations, we can show that

$$\nabla g^*(\mathbf{Y}) = \mathbf{X} = \frac{\mu\mathbf{Y} + \mathbf{S}}{m + \mu\tau q}.$$

Next, we will demonstrate in more details the techniques we used in finding a subgradient for the convex function h . Recall that h is defined by

$$h(\mathbf{X}) = \sum_{i=1}^3 h_i(\mathbf{X}).$$

We will start with the function h_1 given by

$$h_1(\mathbf{X}) = \frac{\mu}{2} \sum_{i=1}^m \sum_{\ell=1}^k \left[d\left(\frac{x^\ell - a^i}{\mu}; \mathbb{B}\right) \right]^2.$$

Similar to the situation in [36], we get

$$\frac{\partial h_1}{\partial x^\ell}(\mathbf{X}) = \sum_{i=1}^m \left(\frac{x^\ell - a^i}{\mu} - P\left(\frac{x^\ell - a^i}{\mu}; \mathbb{B}\right) \right).$$

Thus, for $\ell = 1, 2, \dots, k$, $\nabla h_1(\mathbf{X}) = \mathbf{Z}$ is the $k \times d$ matrix whose ℓ^{th} row is $\frac{\partial h_1}{\partial x^\ell}(\mathbf{X})$.

Let us compute a subgradient of h_2 as in [36]

$$h_2(\mathbf{X}) = \sum_{i=1}^m \max_{\ell=1, \dots, k} \sum_{j=1, j \neq \ell}^k \|x^j - a^i\| = \sum_{i=1}^m \gamma_i(\mathbf{X}),$$

where $\gamma_i(\mathbf{X}) = \max_{\ell=1, \dots, k} \sum_{j=1, j \neq \ell}^k \|x^j - a^i\|$. For each $i = 1, \dots, m$, define

$$\gamma_{i\ell}(\mathbf{X}) = \sum_{j=1, j \neq \ell}^k \|x^j - a^i\|, \ell = 1, \dots, k.$$

Then $\gamma_i(\mathbf{X}) = \max_{\ell=1, \dots, k} \gamma_{i\ell}(\mathbf{X})$.

Based on the subdifferential formula for maximum functions, for each $i = 1, \dots, m$, we find $\mathbf{W}_i \in \partial\gamma_i(\mathbf{X})$. Then define $\mathbf{W} = \sum_{i=1}^m \mathbf{W}_i$ to get a subgradient of the function h_2 at \mathbf{X} by the subdifferential sum rule. To accomplish this goal, we first choose an index $\ell^* = 1, \dots, k$ such that $\gamma_i(\mathbf{X}) = \gamma_{i\ell^*}(\mathbf{X}) = \sum_{j=1, j \neq \ell^*}^k \|x^j - a^i\|$. Using the familiar subdifferential formula of the Euclidean norm function, the j^{th} row w_i^j for $j \neq \ell^*$ of the matrix \mathbf{W}_i is determined as follows

$$w_i^j = \begin{cases} \frac{x^j - a^i}{\|x^j - a^i\|} & \text{if } x^j \neq a^i, \\ 0 & \text{if } x^j = a^i. \end{cases}$$

The ℓ^{th} row of the matrix \mathbf{W}_i is $w_i^{\ell^*} = 0$.

The procedure for computing $\partial h_3(\mathbf{X})$ is the same in Section 4.2. Let \mathbf{U} be the matrix whose rows are $\sum_{i=1}^q P(x^\ell; \Omega_i^\ell)$, for $\ell = 1, \dots, k$, then $\nabla h_3(\mathbf{X}) = \tau \mathbf{U}$.

At this point, we are ready to give a new DCA-based algorithm for our problem.

Algorithm 7 : DC program for (4.4.33)

INPUT: $\mathbf{A}, \mathbf{X}_0, \{\Omega_j^\ell\}_{j=1, \dots, q}^{\ell=1, \dots, k}, \tau, \mu, N \in \mathbb{N}$.
for $p = 1, \dots, N$ **do**
 Find $\mathbf{Y}_p := \mathbf{Z}_p + \mathbf{W}_p + \tau \mathbf{U}_p$ where
 $\mathbf{Z}_p := \nabla h_1(\mathbf{X}_p)$
 $\mathbf{W}_p \in \partial h_2(\mathbf{X}_p)$
 $\mathbf{U}_p := \nabla h_3(\mathbf{X}_p)$
 Find $\mathbf{X}_{p+1} := \frac{\mu(\mathbf{Z}_p + \mathbf{W}_p + \tau \mathbf{U}_p) + \mathbf{S}}{m + \mu\tau q}$
end for
OUTPUT: \mathbf{X}_N

We also present below an adapted version of Algorithm 7 for solving (4.4.32). We may improve Algorithm 7 by gradually increasing and decreasing the value of the penalty parameter τ and the smoothing parameter μ respectively. This can be done by periodically multiplying them by some $\sigma > 1$, $0 < \delta < 1$ and stopping when $\tau > \tau_f$, $\mu < \mu_f$.

Algorithm 8 : Penalty DC program for (4.4.32)

INPUT: $\mathbf{A}, \mathbf{X}_0, \{\Omega_j^\ell\}_{j=1, \dots, q}^{\ell=1, \dots, k}, N, \tau, \sigma, \tau_f, \mu, \delta, \mu_f$
while $\tau < \tau_f$ and $\mu > \mu_f$ **do**
 Find \mathbf{X}_N by executing **Algorithm 7** with $\mathbf{A}, \mathbf{X}_0, \{\Omega_j^\ell\}_{j=1, \dots, q}^{\ell=1, \dots, k}, N, \tau, \mu$
 Reassign $\mathbf{X}_0 := \mathbf{X}_N$
 Reassign $\tau := \sigma\tau$
 Reassign $\mu := \delta\mu$
end while
OUTPUT: \mathbf{X}_N

4.5. Numerical Experiments

4.5.1. Clustering with constraints.

Example 4.5.1. We now consider the dataset EIL76 taken from the Traveling Salesman Problem Library [40]. We impose the following constraints on the solution:

- (a) The first center is a common point of a box whose vertices are $(40, 40)$; $(40, 60)$; $(20, 60)$; $(20, 40)$ and a ball of radius $r = 7$ centered at $(20, 60)$.
- (b) The second center is in the intersection of two balls of the same radius $r = 7$, centered at $(35, 20)$ and $(45, 22)$, respectively.

Choosing $\tau = 1, \sigma = 10, \tau_f = 10^8$, Algorithm 4 yields an approximate solution:

$$\mathbf{X} = \begin{pmatrix} 26.69959 & 57.97125 \\ 41.06910 & 23.48799 \end{pmatrix}, \text{ with the cost } \psi(\mathbf{X}) = 33576.25387; \text{ see Figure 4.1.}$$

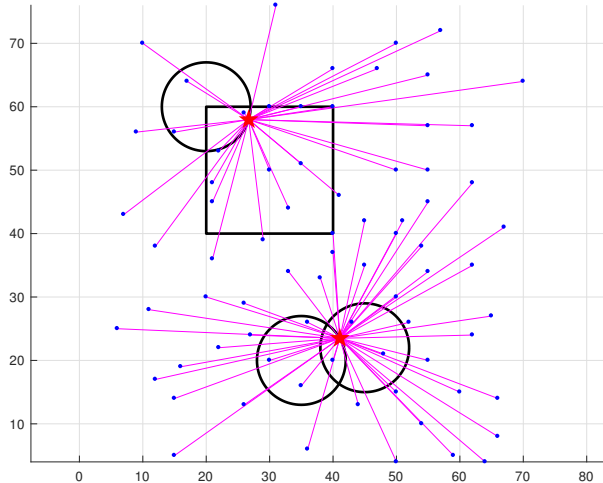


FIGURE 4.1. A 2-center constrained clustering problem for dataset EIL76.

4.5.2. Set Clustering with Constraints.

Example 4.5.2. We now use Algorithm 5 to solve a set clustering problem with constraints. We consider the latitude and longitude of the 50 most populous US cities taken from 2014 United States Census Bureau data ¹, and approximate each city by a ball with radius $0.1\sqrt{\frac{A}{\pi}}$ where A is the city's reported area in square miles.

We use Algorithm 5 for solving 3-center problem generated by this 50-set dataset with requirement that each center must belong to the intersection of two balls. The centers of these constrained balls are the columns of the matrix below

$$\begin{pmatrix} -80 & -80 & -92 & -90 & -115 & -110 \\ 34 & 38 & 37 & 40 & 45 & 40 \end{pmatrix}$$

with corresponding radii given by $(2 \ 3 \ 4 \ 3 \ 4 \ 4)$. The result is plotted in Figure 4.2 using a plate Carrée projection ². Each city is approximated by a ball proportional to its area.

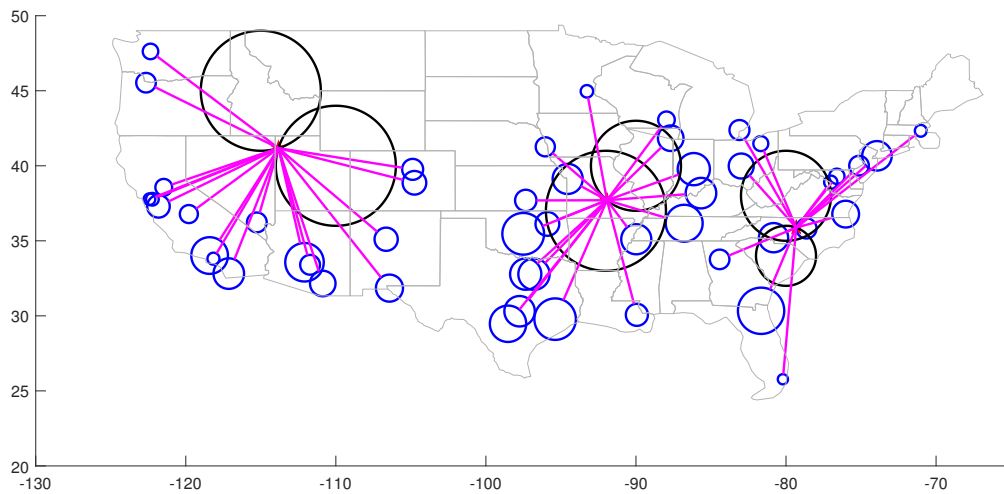


FIGURE 4.2. A 3-center set clustering problems with 50 most populous US cities.

¹https://en.wikipedia.org/wiki/List_of_United_States_cities_by_population

²<https://www.mathworks.com/help/map/pcarree.html>

We again choose $\tau = 1$, $\sigma = 10$, $\tau_f = 10^8$, Algorithm 6 yields an approximate optimal value $\psi(\mathbf{X}) = 2271.09657$ at an approximate solution given by

$$\mathbf{X} = \begin{pmatrix} -79.32172 & 35.88148 \\ -91.93134 & 37.70436 \\ -113.82289 & 41.17711 \end{pmatrix}.$$

4.5.3. Multifacility Location with Constraints.

Example 4.5.3. We now test Algorithm 8 on a data set \mathbf{A} containing random points in 4 balls of radius $r = 0.3$ centered at $(2, 2)$, $(4, 2)$, $(4, 4)$ and $(2, 4)$. Let $k = 4$ and the constraint be the ball with the same radius, centered at $(3, 3)$. We use the “*kmeans*” (a MATLAB built in function) to partition the nodes into 4 clusters first, and then we selected the 4 cluster centroid locations as starting centers. We choose $\tau = 1$, $\sigma = 10$, $\tau_f = 10^8$, $\mu = 1$, $\delta = 0.75$, $\mu_f = 10^{-6}$. Typical centers are the intersections of the constraint ball boundary and the line connecting centers of each ball to the center of the constraint one. A visualization is shown in Figure 4.3.

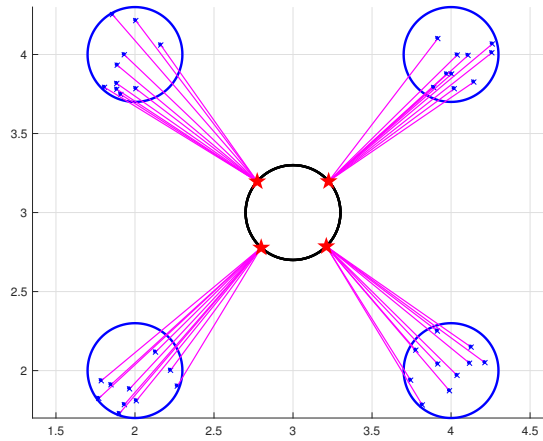


FIGURE 4.3. A 4-center multifacility location with one ball constraint.

Example 4.5.4. Next we consider the latitude and longitude data of the $m = 988$ most-populated cities in the contiguous 48 United States [49]. We impose the following constraints on the solution:

- (a) One center is to lie west of -115° longitude and within 4° latitude/longitude of Caldwell, Idaho.
- (b) One center is to lie within the state of Colorado and within 6° latitude/longitude of Oklahoma City, Oklahoma.
- (c) One center is to lie within 2° latitude/longitude of Skokie, Illinois and the triangle with vertices at Cleveland, Ohio; Atlanta, Georgia; and Des Moines, Iowa.
- (d) One center is to lie within 4° latitude/longitude of New York, NY and Washington, DC.

Employing Algorithms 7 and 8 with $\tau = 1$, $\sigma = 100$, $\tau_f = 10^8$, $\mu = 1$, $\delta = 0.85$, $\mu_f = 10^{-6}$, we terminate when $\|\mathbf{X}_{p+1} - \mathbf{X}_p\|_F < 10^{-6}$ and find final centers at

$$\mathbf{X} = \begin{pmatrix} -118.03185 & 39.89550 \\ -102.04996 & 36.99996 \\ -87.93854 & 40.90443 \\ -76.63980 & 38.67968 \end{pmatrix}$$

with an objective value $\psi(\mathbf{X}) = 42586.65060$; see Figure 4.4.

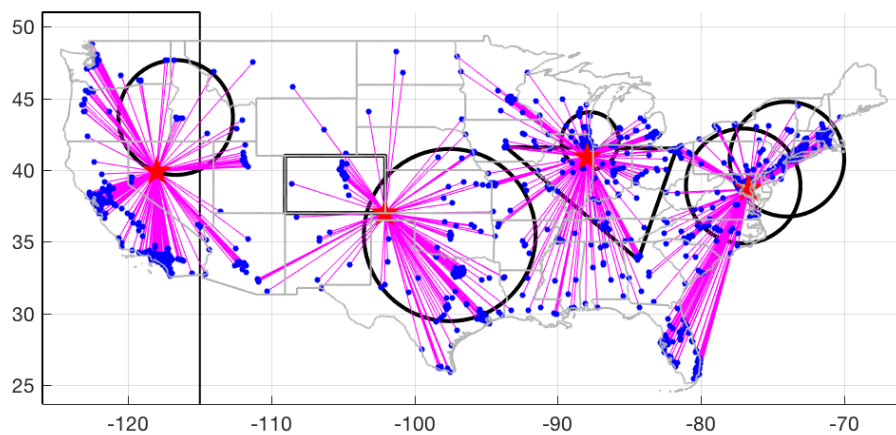


FIGURE 4.4. A 4-center constrained multifacility location problem with US cities dataset.

DC Programming for Hierarchical Clustering

5.1. Overview

A hierarchical clustering of a set of points/nodes can be described as a tree, in which the leaves are nodes to be clustered. Multilevel hierarchical clustering has a long history; see, e.g., [18, 19, 33], and has enormous important applications in data mining and statistics. Since many kinds of data encountered in practical applications have nested structures, they are required to use multilevel hierarchical clustering which involves grouping a data set into a hierarchy of clusters. Bilevel hierarchical clustering is a model of this class and is commonly used in designing optimal multicast networks. While mathematical programming is widely used for nonhierarchical clustering problems, there exist a few optimization models and techniques for multilevel hierarchical clustering ones. Except the work in [2, 11], we have not found other approaches using mathematical programming model for multilevel hierarchical clustering. In this chapter, we apply the mathematical optimization approach to the bilevel hierarchical clustering problem. In fact, using mathematical optimization in clustering is a very promising approach to overcome many disadvantages of the *k-mean algorithm* commonly used in clustering; see, e.g., [1, 2, 11, 36] and the references therein. In particular, the DCA was successfully applied in [2] to a bilevel hierarchical clustering problem in which the distance measurement is defined by the squared Euclidean distance. Although the DCA in [2] provides an effective way to solve the bilevel hierarchical clustering in high dimensions, it has not been used to solve the original model defined by the Euclidean distance measurement proposed in [11] which is *not suitable for the resulting DCA* according to the authors of [2]. By applying Nesterov's smoothing technique and the DCA, we

are able to solve the original model proposed in [11] in high dimensions. This chapter presents the work we have done in [35]. The reader can find more details in there.

5.2. The Bilevel Hierarchical Clustering Problem

Given a set of m points (nodes) a^1, a^2, \dots, a^m in \mathbb{R}^d , our goal is to decompose this set into k clusters. In each cluster, we would like to find a point x^i among the nodes and assign it as the center for this cluster with all points in the cluster connected to this center. Then we will find a *total center* x^* among the given points a^1, a^2, \dots, a^m , and all centers are connected to this total center. The goal is to minimize the *total transportation cost* in this tree computed by the sum of the distances from the total center to each center and from each center to the nodes in each cluster. This is a discrete optimization problem which can be shown to be *NP-hard*. We will solve this problem based on continuous optimization techniques.

5.2.1. The bilevel hierarchical clustering: Model I. The difficulty in solving this hierarchical clustering problem lies in the fact that the centers and total center have to be among the nodes. We first relax this condition with the use of artificial centers x^1, x^2, \dots, x^k that could be anywhere in \mathbb{R}^d . Then we define the total center as the centroid of x^1, x^2, \dots, x^k given by

$$x^* := \frac{1}{k}(x^1 + x^2 + \dots + x^k).$$

The total cost of the tree is given by

$$\varphi(x^1, \dots, x^k) := \sum_{i=1}^m \min_{\ell=1, \dots, k} \|x^\ell - a^i\| + \sum_{\ell=1}^k \|x^\ell - x^*\|.$$

Note that here each a^i is assigned to its closest center. However, what we expect are the real centers, which can be approximated by trying to minimize the difference between the artificial centers and the real centers. To achieve this goal, define the function

$$\phi(x^1, \dots, x^k) := \sum_{\ell=1}^k \min_{i=1, \dots, m} \|x^\ell - a^i\|, x^1, \dots, x^k \in \mathbb{R}^d.$$

Observe that $\phi(x^1, \dots, x^k) = 0$ if and only if for every $\ell = 1, \dots, k$, there exists $i \in \{1, \dots, m\}$ such that $x^\ell = a^i$, which means that x^ℓ is a real node. Therefore, we consider the constrained minimization problem

$$\begin{aligned} & \text{minimize} && \varphi(x^1, \dots, x^k) \\ & \text{subject to} && \phi(x^1, \dots, x^k) = 0. \end{aligned}$$

This problem can be converted to an unconstrained minimization problem

$$(5.2.34) \quad \text{minimize } f_\lambda(x^1, \dots, x^k) := \varphi(x^1, \dots, x^k) + \lambda\phi(x^1, \dots, x^k),$$

where $x^1, \dots, x^k \in \mathbb{R}^d$ and $\lambda > 0$ is a penalty parameter. Similar to the situation with the clustering problem, this new problem is nonsmooth and nonconvex, which can be solved by smoothing techniques and the DCA. A particular case of this model in two dimensions was considered in [11] where the problem was solved using the derivative-free discrete gradient method established in [10], but this method is not suitable for large-scale settings in high dimensions. The DCA was successfully applied in [2] to solve these models in which the similarity measure is defined by the squared Euclidean distance. Although the DCA in [2] provides an effective way to solve the bilevel hierarchical clustering in high dimensions, it has not been used to solve the original model defined by the Euclidean distance measurement proposed in [11] which is not suitable for the resulting DCA according to the authors of [2]. Nevertheless, we will show in what follows that the DCA is applicable to this model when combined with Nesterov's smoothing technique.

Note that the functions φ and ϕ in (5.2.34) belong to the class of DC functions with the following DC decompositions

$$\begin{aligned}\varphi(x^1, \dots, x^k) &= \sum_{i=1}^m \left[\sum_{\ell=1}^k \|x^\ell - a^i\| - \max_{r=1, \dots, k} \sum_{\ell=1, \ell \neq r}^k \|x^\ell - a^i\| \right] + \sum_{\ell=1}^k \|x^\ell - x^*\| \\ &= \left[\sum_{i=1}^m \sum_{\ell=1}^k \|x^\ell - a^i\| + \sum_{\ell=1}^k \|x^\ell - x^*\| \right] - \sum_{i=1}^m \max_{r=1, \dots, k} \sum_{\ell=1, \ell \neq r}^k \|x^\ell - a^i\|, \\ \phi(x^1, \dots, x^k) &= \sum_{\ell=1}^k \sum_{i=1}^m \|x^\ell - a^i\| - \sum_{\ell=1}^k \max_{s=1, \dots, m} \sum_{i=1, i \neq s}^m \|x^\ell - a^i\|.\end{aligned}$$

It follows that the objective function f_λ in (5.2.34) has the DC decomposition

$$\begin{aligned}f_\lambda(x^1, \dots, x^k) &= \left[(1 + \lambda) \sum_{\ell=1}^k \sum_{i=1}^m \|x^\ell - a^i\| + \sum_{\ell=1}^k \|x^\ell - x^*\| \right] \\ &\quad - \left[\sum_{i=1}^m \max_{r=1, \dots, k} \sum_{\ell=1, \ell \neq r}^k \|x^\ell - a^i\| + \lambda \sum_{\ell=1}^k \max_{s=1, \dots, m} \sum_{i=1, i \neq s}^m \|x^\ell - a^i\| \right].\end{aligned}$$

This DC decomposition is not suitable for applying the DCA because there is no closed form for a subgradient of the function g^* involved.

In the next step, we apply Nesterov's smoothing technique to approximate the objective function f_λ by a new DC function favorable for applying the DCA.

$$\begin{aligned}f_{\lambda\mu}(x^1, \dots, x^k) &:= \frac{(1 + \lambda)\mu}{2} \sum_{i=1}^m \sum_{\ell=1}^k \left\| \frac{x^\ell - a^i}{\mu} \right\|^2 + \frac{\mu}{2} \sum_{\ell=1}^k \left\| \frac{x^\ell - x^*}{\mu} \right\|^2 \\ &\quad - \frac{(1 + \lambda)\mu}{2} \sum_{i=1}^m \sum_{\ell=1}^k \left[d \left(\frac{x^\ell - a^i}{\mu}; \mathbb{B} \right) \right]^2 - \frac{\mu}{2} \sum_{\ell=1}^k \left[d \left(\frac{x^\ell - x^*}{\mu}; \mathbb{B} \right) \right]^2 \\ &\quad - \sum_{i=1}^m \max_{r=1, \dots, k} \sum_{\ell=1, \ell \neq r}^k \|x^\ell - a^i\| - \lambda \sum_{\ell=1}^k \max_{s=1, \dots, m} \sum_{i=1, i \neq s}^m \|x^\ell - a^i\|.\end{aligned}$$

The original bilevel hierarchical clustering problem now can be solved using a DC program

$$\text{minimize } f_{\lambda\mu}(x^1, \dots, x^k) = g_{\lambda\mu}(x^1, \dots, x^k) - h_{\lambda\mu}(x^1, \dots, x^k), \quad x^1, \dots, x^k \in \mathbb{R}^n.$$

In this formulation, $g_{\lambda\mu}$ and $h_{\lambda\mu}$ are convex functions on $(\mathbb{R}^n)^k$ defined by

$$g_{\lambda\mu}(x^1, \dots, x^k) := g_{\lambda\mu}^1(x^1, \dots, x^k) + g_{\lambda\mu}^2(x^1, \dots, x^k),$$

and

$$\begin{aligned} h_{\lambda\mu}(x^1, \dots, x^k) &:= h_{\lambda\mu}^1(x^1, \dots, x^k) + h_{\lambda\mu}^2(x^1, \dots, x^k) \\ &\quad + h_{\lambda\mu}^3(x^1, \dots, x^k) + h_{\lambda\mu}^4(x^1, \dots, x^k). \end{aligned}$$

Their respective components defined as

$$g_{\lambda\mu}^1(x^1, \dots, x^k) := \frac{1+\lambda}{2\mu} \sum_{i=1}^m \sum_{\ell=1}^k \|x^\ell - a^i\|^2, \quad g_{\lambda\mu}^2(x^1, \dots, x^k) := \frac{1}{2\mu} \sum_{\ell=1}^k \|x^\ell - x^*\|^2,$$

$$h_{\lambda\mu}^1(x^1, \dots, x^k) := \frac{(1+\lambda)\mu}{2} \sum_{i=1}^m \sum_{\ell=1}^k \left[d\left(\frac{x^\ell - a^i}{\mu}; \mathbb{B}\right) \right]^2,$$

$$h_{\lambda\mu}^2(x^1, \dots, x^k) := \frac{\mu}{2} \sum_{\ell=1}^k \left[d\left(\frac{x^\ell - x^*}{\mu}; \mathbb{B}\right) \right]^2,$$

$$h_{\lambda\mu}^3(x^1, \dots, x^k) := \sum_{i=1}^m \max_{r=1, \dots, k} \sum_{\ell=1, \ell \neq r}^k \|x^\ell - a^i\|,$$

$$h_{\lambda\mu}^4(x^1, \dots, x^k) := \lambda \sum_{\ell=1}^k \max_{s=1, \dots, m} \sum_{i=1, i \neq s}^m \|x^\ell - a^i\|.$$

To facilitate the gradient and subgradient calculations for the DCA, we introduce a *data matrix* \mathbf{A} and a *variable matrix* \mathbf{X} . The data \mathbf{A} is formed by putting each a^i , $i = 1, \dots, m$, in the i^{th} row, i.e.,

$$\mathbf{A} = \begin{pmatrix} a_{11} & a_{12} & a_{13} & \dots & a_{1n} \\ a_{21} & a_{22} & a_{23} & \dots & a_{2n} \\ \vdots & \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & a_{m3} & \dots & a_{mn} \end{pmatrix}.$$

Similarly, if x^1, \dots, x^k are the k cluster centers, then the variable \mathbf{X} is formed by putting each x^ℓ , $\ell = 1, \dots, k$, in the ℓ^{th} row, i.e.,

$$\mathbf{X} = \begin{pmatrix} x_{11} & x_{12} & x_{13} & \dots & x_{1n} \\ x_{21} & x_{22} & x_{23} & \dots & x_{2n} \\ \vdots & \vdots & \vdots & & \vdots \\ x_{k1} & x_{k2} & x_{k3} & \dots & x_{kn} \end{pmatrix}.$$

Then the variable matrix \mathbf{X} of the optimization problem belongs to $\mathbb{R}^{k \times n}$, the linear space of $k \times n$ real matrices equipped with the inner product $\langle \mathbf{X}, \mathbf{Y} \rangle := \text{trace}(\mathbf{X}^T \mathbf{Y})$. We represent the average of the k cluster centers by x^* , i.e., $x^* := \frac{1}{k} \sum_{j=1}^k x^j$.

Gradient and Subgradient Calculations for the DCA

Let us start by computing the gradient of

$$g_{\lambda\mu}(\mathbf{X}) = g_{\lambda\mu}^1(\mathbf{X}) + g_{\lambda\mu}^2(\mathbf{X}).$$

Using the Frobenius norm, the function $g_{\lambda\mu}^1$ can equivalently be written as

$$\begin{aligned} g_{\lambda\mu}^1(\mathbf{X}) &= \frac{1 + \lambda}{2\mu} \sum_{i=1}^m \sum_{\ell=1}^k \|x^\ell - a^i\|^2 \\ &= \frac{1 + \lambda}{2\mu} \sum_{i=1}^m \sum_{\ell=1}^k [\|x^\ell\|^2 - 2\langle x^\ell, a^i \rangle + \|a^i\|^2] \\ &= \frac{1 + \lambda}{2\mu} \left[m \|\mathbf{X}\|_F^2 - 2\langle \mathbf{X}, \mathbf{E}_{km} \mathbf{A} \rangle + k \|\mathbf{A}\|_F^2 \right], \end{aligned}$$

where \mathbf{E}_{km} is a $k \times m$ matrix whose entries are all ones. Hence, one can see that $g_{\lambda\mu}^1$ is differentiable and its gradient is given by

$$\nabla g_{\lambda\mu}^1(\mathbf{X}) = \frac{1 + \lambda}{\mu} [m\mathbf{X} - \mathbf{E}_{km} \mathbf{A}].$$

Similarly, $g_{\lambda\mu}^2$ can equivalently be written as

$$\begin{aligned}
g_{\lambda\mu}^2(\mathbf{X}) &= \frac{1}{2\mu} \sum_{\ell=1}^k \|x^\ell - x^*\|^2 \\
&= \frac{1}{2\mu} \sum_{\ell=1}^k [\|x^\ell\|^2 - 2\langle x^\ell, x^* \rangle + \|x^*\|^2] \\
&= \frac{1}{2\mu} \left[\|\mathbf{X}\|_F^2 - \frac{2}{k} \langle \mathbf{X}, \mathbf{E}_{kk} \mathbf{X} \rangle + \frac{1}{k} \langle \mathbf{X}, \mathbf{E}_{kk} \mathbf{X} \rangle \right] \\
&= \frac{1}{2\mu} \left[\|\mathbf{X}\|_F^2 - \frac{1}{k} \langle \mathbf{X}, \mathbf{E}_{kk} \mathbf{X} \rangle \right],
\end{aligned}$$

where \mathbf{E}_{kk} is a $k \times k$ matrix whose entries are all ones. Hence, $g_{\lambda\mu}^2$ is differentiable and its gradient is given by

$$\nabla g_{\lambda\mu}^2(\mathbf{X}) = \frac{1}{\mu} \left[\mathbf{X} - \frac{1}{k} \mathbf{E}_{kk} \mathbf{X} \right].$$

Since $g_{\lambda\mu}(\mathbf{X}) = g_{\lambda\mu}^1(\mathbf{X}) + g_{\lambda\mu}^2(\mathbf{X})$, its gradient can be computed by

$$\begin{aligned}
\nabla g_{\lambda\mu}(\mathbf{X}) &= \nabla g_{\lambda\mu}^1(\mathbf{X}) + \nabla g_{\lambda\mu}^2(\mathbf{X}) \\
&= \frac{1+\lambda}{\mu} [m\mathbf{X} - \mathbf{E}_{km}\mathbf{A}] + \frac{1}{\mu} \left[\mathbf{X} - \frac{1}{k} \mathbf{E}_{kk} \mathbf{X} \right] \\
&= \frac{1}{\mu} \left[(1+\lambda)m\mathbf{X} - (1+\lambda)\mathbf{E}_{km}\mathbf{A} + \mathbf{X} - \frac{1}{k} \mathbf{E}_{kk} \mathbf{X} \right] \\
&= \frac{1}{\mu} \left[\left[(1+\lambda)m + 1 \right] \mathbf{I}_{kk} - \frac{1}{k} \mathbf{E}_{kk} \right] \mathbf{X} - (1+\lambda)\mathbf{E}_{km}\mathbf{A}.
\end{aligned}$$

Therefore,

$$\nabla g_{\lambda\mu}(\mathbf{X}) = \frac{1}{\mu} \left[\left(\left[(1+\lambda)m + 1 \right] \mathbf{I}_{kk} - \mathbf{J} \right) \mathbf{X} - (1+\lambda)\mathbf{S} \right],$$

where

$$\mathbf{J} = \frac{1}{k} \mathbf{E}_{kk} \quad \text{and} \quad \mathbf{S} = \mathbf{E}_{km} \mathbf{A}.$$

Our goal now is to compute $\nabla g^*(Y)$, which can be accomplished by the relation

$$\mathbf{X} = \nabla g^*(\mathbf{Y}) \text{ if and only if } \mathbf{Y} = \nabla g(\mathbf{X}).$$

The latter can equivalently be written as

$$[[1 + (1 + \lambda)m] \mathbf{I}_{kk} - \mathbf{J}] \mathbf{X} = [(1 + \lambda)\mathbf{S} + \mu\mathbf{Y}].$$

Then with some algebraic manipulation we can show that

$$\nabla g^*(\mathbf{Y}) = \mathbf{X} = \left[\frac{1}{1 + (1 + \lambda)m} \mathbf{I}_{kk} + \frac{1}{[1 + (1 + \lambda)m](1 + \lambda)m} \mathbf{J} \right] [(1 + \lambda)\mathbf{S} + \mu\mathbf{Y}].$$

Next, we will demonstrate in more detail the techniques we used in finding a subgradient for the convex function $h_{\lambda\mu}$. Recall that $h_{\lambda\mu}$ is defined by

$$h_{\lambda\mu}(\mathbf{X}) = \sum_{i=1}^4 h_{\lambda\mu}^i(\mathbf{X}).$$

We will start with the function $h_{\lambda\mu}^1$ given by

$$h_{\lambda\mu}^1(\mathbf{X}) = \frac{(1 + \lambda)\mu}{2} \sum_{i=1}^m \sum_{\ell=1}^k \left[d \left(\frac{x^\ell - a^i}{\mu}; \mathbb{B} \right) \right]^2.$$

From its representation, one can see that $h_{\lambda\mu}^1$ is differentiable, and hence its subgradient coincides with its gradient, that can be computed by the partial derivatives with respect to x^1, \dots, x^k , i.e.,

$$\frac{\partial h_{\lambda\mu}^1}{\partial x^\ell}(\mathbf{X}) = (1 + \lambda) \sum_{i=1}^m \left[\frac{x^\ell - a^i}{\mu} - P \left(\frac{x^\ell - a^i}{\mu}; \mathbb{B} \right) \right].$$

Thus, for $\ell = 1, 2, \dots, k$, $\nabla h_{\lambda\mu}^1(\mathbf{X})$ is the $k \times n$ matrix \mathbf{U} whose ℓ^{th} row is $\frac{\partial h_{\lambda\mu}^1}{\partial x^\ell}(\mathbf{X})$.

Similarly, one can see that the function $h_{\lambda\mu}^2$ given by

$$h_{\lambda\mu}^2(\mathbf{X}) = \frac{\mu}{2} \sum_{\ell=1}^k \left[d \left(\frac{x^\ell - x^*}{\mu}; \mathbb{B} \right) \right]^2$$

is differentiable with its partial derivatives computed by

$$\frac{\partial h_\mu^2}{\partial x^\ell}(\mathbf{X}) = \left[\frac{\mathbf{x}^\ell - x^*}{\mu} - P\left(\frac{x^\ell - x^*}{\mu}; \mathbb{B}\right) \right] - \frac{1}{k} \sum_{j=1}^k \left[\frac{x^j - x^*}{\mu} - P\left(\frac{x^j - x^*}{\mu}; \mathbb{B}\right) \right].$$

Hence, for $\ell = 1, 2, \dots, k$, $\nabla h_\mu^2(\mathbf{X})$ is the $k \times n$ matrix \mathbf{V} whose ℓ^{th} row is $\frac{\partial h_\mu^2}{\partial x^\ell}(\mathbf{X})$.

Unlike $h_{\lambda\mu}^1$ and $h_{\lambda\mu}^2$, the convex functions $h_{\lambda\mu}^3$ and $h_{\lambda\mu}^4$ are not differentiable, but both can be written as a finite sum of the maximum of a finite number of convex functions.

Let us compute a subgradient of $h_{\lambda\mu}^3$ as an example. We have

$$h_{\lambda\mu}^3(\mathbf{X}) = \sum_{i=1}^m \max_{r=1, \dots, k} \sum_{\ell=1, \ell \neq r}^k \|x^\ell - a^i\| = \sum_{i=1}^m \gamma_i(\mathbf{X}),$$

where, for $i = 1, \dots, m$,

$$\gamma_i(\mathbf{X}) := \max \left\{ \gamma_{ir}(\mathbf{X}) = \sum_{\ell=1, \ell \neq r}^k \|x^\ell - a^i\|, \quad r = 1, \dots, k \right\}.$$

Then, for each $i = 1, \dots, m$, we find $\mathbf{W}_i \in \partial \gamma_i(\mathbf{X})$ according to the subdifferential rule for the maximum of convex functions. Then define $\mathbf{W} := \sum_{i=1}^m \mathbf{W}_i$ to get a subgradient of the function $h_{\lambda\mu}^3$ at \mathbf{X} by the subdifferential sum rule. To accomplish this goal, we first choose an index r^* from the index set $\{1, \dots, k\}$ such that

$$\gamma_i(\mathbf{X}) = \gamma_{ir^*}(\mathbf{X}) = \sum_{\ell=1, \ell \neq r^*}^k \|x^\ell - a^i\|.$$

Using the familiar subdifferential formula of the Euclidean norm function, the ℓ^{th} row w_i^ℓ for $\ell \neq r^*$ of the matrix \mathbf{W}_i is determined as follows

$$w_i^\ell := \begin{cases} \frac{x^\ell - a^i}{\|x^\ell - a^i\|_2} & \text{if } x^\ell \neq a^i, \\ u \in \mathbb{B} & \text{if } x^\ell = a^i. \end{cases}$$

The $r^{*\text{th}}$ row of the matrix \mathbf{W}_i is $w_i^{r^*} := 0$.

The procedure for calculating a subgradient of the function $h_{\lambda\mu}^4$ given by

$$h_{\lambda\mu}^4(x^1, \dots, x^k) = \lambda \sum_{\ell=1}^k \max_{s=1, \dots, m} \sum_{i=1, i \neq s}^m \|x^\ell - a^i\|,$$

is very similar to what we just demonstrated for $h_{\lambda\mu}^3$.

At this point, we can give a new DCA based algorithm for Model I using Algorithm 1. The reader can find more details in [35].

5.2.2. The bilevel hierarchical clustering: Model II. In this section, we introduce the second model to solve the bilevel hierarchical clustering problem. In this model, we use an additional variable x^{k+1} to denote the total center. At first we allow the total center x^{k+1} to be a free point in \mathbb{R}^d , the same as the k cluster centers. Then the total cost of the tree is given by

$$\varphi(x^1, \dots, x^{k+1}) := \sum_{i=1}^m \min_{\ell=1, \dots, k} \|x^\ell - a^i\| + \sum_{\ell=1}^k \|x^\ell - x^{k+1}\|, x^1, \dots, x^{k+1} \in \mathbb{R}^d.$$

To force the $k+1$ centers to be chosen from the given nodes (or to make them as close to the nodes as possible), we set the constraint

$$\phi(x^1, \dots, x^{k+1}) := \sum_{\ell=1}^{k+1} \min_{i=1, \dots, m} \|x^\ell - a^i\| = 0.$$

Our goal is to solve the optimization problem

$$\begin{aligned} & \text{minimize} && \varphi(x^1, \dots, x^{k+1}) \\ & \text{subject to} && \phi(x^1, \dots, x^{k+1}), x^1, \dots, x^{k+1} \in \mathbb{R}^d. \end{aligned}$$

Similar to the first model, this problem formulation can be converted to an unconstrained minimization problem involving a penalty parameter $\lambda > 0$:

$$(5.2.35) \quad \text{minimize } f_\lambda(x^1, \dots, x^{k+1}) := \varphi(x^1, \dots, x^{k+1}) + \lambda\phi(x^1, \dots, x^{k+1}),$$

where $x^1, \dots, x^{k+1} \in \mathbb{R}^d$.

Next, we apply Nesterov's smoothing technique to get an approximation of the objective function f given in (5.2.35) which involves two parameter $\lambda > 0$ and $\mu > 0$:

$$\begin{aligned}
f_{\lambda\mu}(\mathbf{X}) &:= \frac{1+\lambda}{2\mu} \sum_{i=1}^m \sum_{\ell=1}^{k+1} \|x^\ell - a^i\|^2 + \frac{1}{2\mu} \sum_{\ell=1}^k \|x^\ell - x^{k+1}\|^2 - \frac{1}{2\mu} \sum_{i=1}^m \|x^{k+1} - a^i\|^2 \\
&- \frac{\lambda\mu}{2} \sum_{i=1}^m \left[d\left(\frac{x^{k+1} - a^i}{\mu}; \mathbb{B}\right) \right]^2 - \frac{(1+\lambda)\mu}{2} \sum_{i=1}^m \sum_{\ell=1}^k \left[d\left(\frac{x^\ell - a^i}{\mu}; \mathbb{B}\right) \right]^2 \\
&- \frac{\mu}{2} \sum_{\ell=1}^k \left[d\left(\frac{x^\ell - x^{k+1}}{\mu}; \mathbb{B}\right) \right]^2 - \sum_{i=1}^m \max_{r=1, \dots, k} \sum_{\ell=1, \ell \neq r}^k \|x^\ell - a^i\| \\
&- \lambda \sum_{\ell=1}^{k+1} \max_{s=1, \dots, m} \sum_{i=1, i \neq s}^m \|x^\ell - a^i\|.
\end{aligned}$$

As we will show in what follows, it is convenient to apply the DCA to minimize the function $f_{\lambda\mu}$. This function can be represented as the differences of two convex functions defined on $\mathbb{R}^{(k+1) \times n}$ using a variable X whose i^{th} row is x^i for $i = 1, \dots, k+1$

$$f_{\lambda\mu}(\mathbf{X}) = g_{\lambda\mu}(\mathbf{X}) - h_{\lambda\mu}(\mathbf{X}), \quad \mathbf{X} \in \mathbb{R}^{(k+1) \times n}.$$

In this formulation, $g_{\lambda\mu}$ and $h_{\lambda\mu}$ are convex functions defined on $\mathbb{R}^{(k+1) \times n}$ by

$$g_{\lambda\mu}(\mathbf{X}) := g_{\lambda\mu}^1(\mathbf{X}) + g_{\lambda\mu}^2(\mathbf{X})$$

and

$$h_{\lambda\mu}(\mathbf{X}) := h_{\lambda\mu}^1(\mathbf{X}) + h_{\lambda\mu}^2(\mathbf{X}) + h_{\lambda\mu}^3(\mathbf{X}) + h_{\lambda\mu}^4(\mathbf{X}) + h_{\lambda\mu}^5(\mathbf{X}) + h_{\lambda\mu}^6(\mathbf{X}),$$

with their respective components given by

$$\begin{aligned}
g_{\lambda\mu}^1(\mathbf{X}) &:= \frac{1+\lambda}{2\mu} \sum_{i=1}^m \sum_{\ell=1}^{k+1} \|x^\ell - a^i\|^2, \quad g_{\lambda\mu}^2(\mathbf{X}) := \frac{1}{2\mu} \sum_{\ell=1}^k \|x^\ell - x^{k+1}\|^2, \\
h_{\lambda\mu}^1(\mathbf{X}) &:= \frac{1}{2\mu} \sum_{i=1}^m \|x^{k+1} - a^i\|^2, \quad h_{\lambda\mu}^2(\mathbf{X}) := \frac{\lambda\mu}{2} \sum_{i=1}^m \left[d\left(\frac{x^{k+1} - a^i}{\mu}; \mathbb{B}\right) \right]^2,
\end{aligned}$$

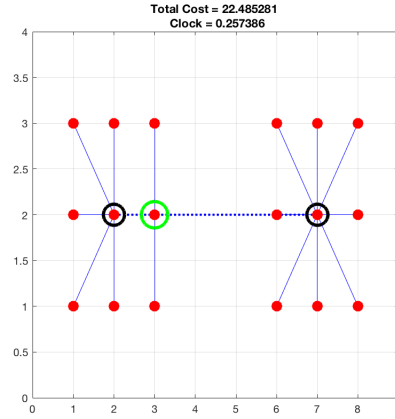
$$\begin{aligned}
h_{\lambda\mu}^3(\mathbf{X}) &:= \frac{(1+\lambda)\mu}{2} \sum_{i=1}^m \sum_{\ell=1}^k \left[d \left(\frac{x^\ell - a^i}{\mu}; \mathbb{B} \right) \right]^2, \\
h_{\lambda\mu}^4(\mathbf{X}) &:= \frac{\mu}{2} \sum_{\ell=1}^k \left[d \left(\frac{x^\ell - x^{k+1}}{\mu}; \mathbb{B} \right) \right]^2, \\
h_{\lambda\mu}^5(\mathbf{X}) &:= \sum_{i=1}^m \max_{r=1, \dots, k} \sum_{\ell=1, \ell \neq r}^k \|x^\ell - a^i\|, \quad h_{\lambda\mu}^6(\mathbf{X}) := \lambda \sum_{\ell=1}^{k+1} \max_{s=1, \dots, m} \sum_{i=1, i \neq s}^m \|x^\ell - a^i\|.
\end{aligned}$$

A similar calculation for gradient and subgradient can be done in the same way of section 5.2.1. More details can be found in [35].

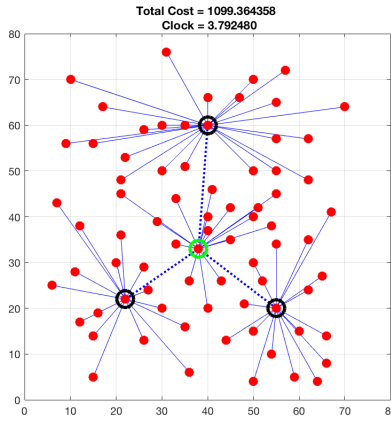
5.2.3. Numerical Experiments. We use MATLAB to code our algorithms and perform numerical experiments on a MacBook Pro with 2.2 GHz Intel Core i7 Processor. For our numerical experiments, we use three data sets: one artificial data set with 18 data points in \mathbb{R}^2 (see Figure 5.1a), the EIL76 and the PR1002 from [40] (see Figure 5.1b and Figure 5.1c), respectively.

The two MATLAB codes used to implement our two algorithms have two major parts: an *outer* loop for updating the penalty and the smoothing parameters and an *inner* loop for updating the cluster centers. At the beginning of each outer loop execution, the penalty and smoothing parameter are updated respectively as $\lambda_{i+1} = \sigma_1 \lambda_i$ and $\mu_{i+1} = \sigma_2 \mu_i$, $i \geq 0$. After we set values $\mu_0 > 0$, $\lambda_0 > 0$, $\sigma_1 > 1$, and $\sigma_2 \in (0, 1)$, we used $\mu < 10^{-6}$ as stopping criteria for the outer loop.

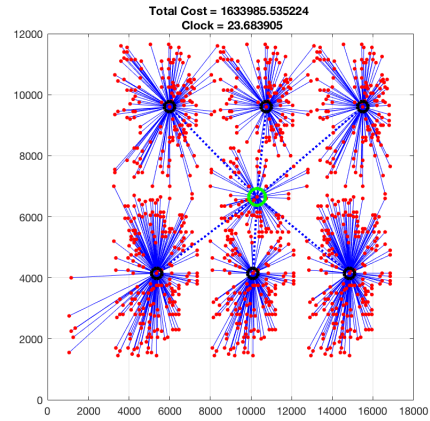
By trial and error we find out that the values chosen for λ_0 and μ_0 in large part determine the performance of the two algorithms for each data set. Intuitively, we see that very large values of λ_0 will over-penalize the distance between an artificial center and its nearest data node and may prevent the algorithm from clustering properly. We therefore use $\lambda_0 \leq 1$ so that the algorithm has a chance to cluster the data before the penalty parameter takes effect. Similarly, we choose $\mu_0 \geq 1$.



(A) 18 Data Points, 2 Centers



(B) 76 Data Points, 3 Centers



(C) 1002 Data Points, 6 Centers

FIGURE 5.1. Plots of the three test data sets.

The starting center \mathbf{X}_0 is selected at a certain radius, $\gamma \text{rad}(\mathbf{A})$, from the median point, $\text{median}(\mathbf{A})$, of the entire data set, i.e.,

$$\mathbf{X}_0 = \text{median}(\mathbf{A}) + \gamma \text{rad}(\mathbf{A}) \mathbf{U},$$

where $\text{rad}(\mathbf{A}) := \max\{\|a^i - \text{median}(\mathbf{A})\| \mid a^i \in \mathbf{A}\}$, γ is a randomly chosen real number from the standard uniform distribution on the open interval $(0,1)$, \mathbf{U} is a $k \times n$ matrix whose k rows are randomly generated unit vectors in \mathbb{R}^n , and the sum is in the sense of adding a vector to each row of a matrix.

As showed in Tables 5.1, 5.2, and 5.3, both algorithms identify the optimal solutions with reasonable amount of time for both DS18 and EIL76 with two and three cluster

centers, respectively. To get a good starting point which yields a better estimate of the optimal value for bigger data sets such as PR1002, we use a method called *radial search* described as follows. Given initial radius $r_0 > 0$ and $m \in \mathbb{N}$, set $\gamma = ir_0$ for $i = 1, \dots, m$. Then we test the algorithm with different starting points given by $\mathbf{X}_0(i) = \text{median}(\mathbf{A}) + ir_0(\text{rad}(\mathbf{A})\mathbf{U})$ for $i = 1, \dots, m$. Figure 5.2 shows the result of the method applied to PR1002 with six cluster centers, where the y -axis represents the optimal value returned by ALG1 with different starting points $\mathbf{X}_0(i)$, as represented on the x -axis.

$$\mu_0 = 5.70, \lambda_0 = 0.001, \sigma_1 = 7500, \sigma_2 = 0.5$$

	Cost1	Cost2	Time1	Time2	Iter1	Iter2	k	m	n
ADS18	22.4853	22.4853	0.07	0.09	124	124	2	18	2
ADS18	22.4853	22.4853	0.07	0.09	124	124	2	18	2
ADS18	22.4853	22.4853	0.07	0.10	124	124	2	18	2

TABLE 5.1. Results for the 18 points artificial data set.

$$\mu_0 = 100, \lambda_0 = 10^{-6}, \sigma_1 = 1, \sigma_2 = 0.5$$

	Cost1	Cost2	Time1	Time2	Iter1	Iter2	k	m	n
EIL76	1125.48	1107.47	1.61	1.86	540	540	3	76	2
EIL76	1099.36	1099.36	1.52	1.81	540	540	3	76	2
EIL76	1099.36	1099.36	1.53	1.87	540	540	3	76	2

TABLE 5.2. Results for EIL76 data set.

$$\mu_0 = 1950, \lambda_0 = 10^{-6}, \sigma_1 = 7500, \sigma_2 = 0.5$$

	Cost1	Cost2	Time1	Time2	Iter1	Iter2	k	m	n
PR1002	1.63399e+06	1.63399e+06	22.16	25.37	330	330	6	1002	2
PR1002	1.63399e+06	1.63399e+06	22.30	25.54	330	330	6	1002	2
PR1002	1.63399e+06	1.63399e+06	23.54	26.08	330	330	6	1002	2

TABLE 5.3. Results for PR1002 data set.

For comparison purposes, both Cost1 and Cost2 are computed by the same way. First, we systematically reassign the k cluster centers returned by the respective algorithms by k real nodes that are close to them, i.e., for $\ell = 1, \dots, k$

$$\bar{x}^\ell = \operatorname{argmin}\{\|x^\ell - a^i\| \mid a^i \in A\}.$$

Then the total center x^* will be a real node, from the remaining nodes, whose sum of distances from the k reassigned centers is the minimal, i.e.,

$$x^* := \operatorname{argmin} \left\{ \sum_{\ell=1}^k \|\bar{x}^\ell - a^i\| \mid a^i \in A \right\}.$$

The total cost is computed by adding the distance of each real node to its closest center (including the total center), and the distances of the total center from the k cluster centers, i.e.,

$$\text{Cost} := \sum_{i=1}^m \min_{\ell=1, \dots, k+1} \|\bar{x}^\ell - a^i\| + \sum_{\ell=1}^k \|\bar{x}^\ell - x^{k+1}\|, \text{ where } x^{k+1} = x^*.$$

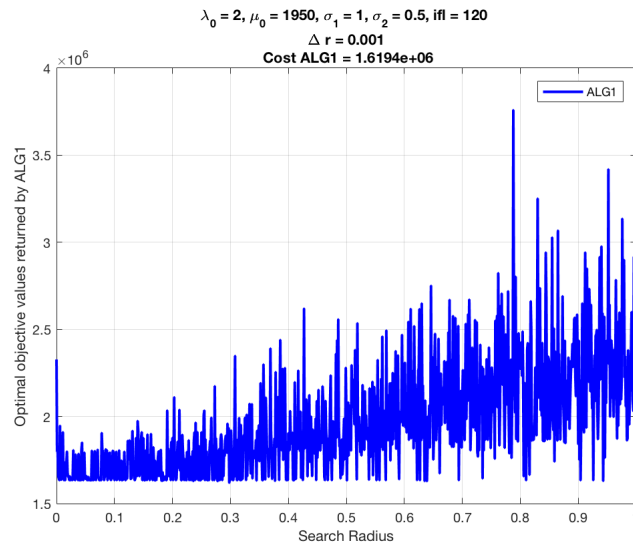


FIGURE 5.2. PR1002, The 1002 City Problem, with 6 Cluster Centers.

DC Programming for Multifacility Location via Mixed Integer Programming

6.1. Overview

In many practical applications, one needs to consider facility location problems in which more than one center must be found to serve a finite number of demand points. These problems are referred to as *multifacility location problems* (MFLPs). Given a finite number of demand points a_1, \dots, a_n in \mathbb{R}^d , we consider the facility location in which k centers v_1, \dots, v_k ($1 \leq k \leq n$) in \mathbb{R}^d need to be found to serve these demand points by assigning each demand point to its nearest center and minimizing the total distances from the centers to the assigned demand points. In the case where $k = 1$, this problem reduces to the generalized Fermat-Torricelli problem which asks for a point that minimizes the sum of the distances to a finite number of given points in \mathbb{R}^d . Motivated by recent work by An, Minh and Tao [5], we model this problem using *mixed integer programming* and develop a *DCA-based algorithm* to solve its approximate version from Nesterov's smoothing.

For convenience, we use a variable $k \times d$ matrix \mathbf{V} with v_i as its i th row to store the centers to be found. We also use another variable $k \times n$ matrix $\mathbf{U} = [u_{i,j}]$ with $u_{i,j} \in \{0, 1\}$ and $\sum_{i=1}^k u_{i,j} = 1$ for $j = 1, \dots, n$ to assign demand points to the centers. The set of all such matrices is denoted by \mathcal{U} . Note that $u_{i,j} = 1$ if the center v_i is assigned to the demand point a_j , and $\sum_{i=1}^k u_{i,j} = 1$ means that the demand point a_j is assigned to only one center. The goal is to solve the constrained optimization problem

$$(6.1.36) \quad \begin{aligned} & \text{minimize} && \mathcal{F}(\mathbf{U}, \mathbf{V}) := \sum_{i=1}^k \sum_{j=1}^n u_{i,j}^2 \|a_j - v_i\| \\ & \text{subject to} && \mathbf{U} \in \mathcal{U} \text{ and } \mathbf{V} \in \mathbb{R}^{k \times d}. \end{aligned}$$

Note that, since $u_{i,j} \in \{0, 1\}$, we use $u_{i,j}^2$ instead of $u_{i,j}$ in the definition of the objective function \mathcal{F} for convenience.

In their recent work [5], for the purpose of *clustering*, An, Minh and Tao studied a version of (6.1.36) which involves the squared Euclidean norm instead of the Euclidean norm as in our formulation. Note that the difference between the two models is significant because using the Euclidean norm allows us to model the total distance in *supply delivery*, while using the squared Euclidean norm is meaningful in clustering. In addition, in the case where $k = 1$ our problem becomes the generalized Fermat-Torricelli problem which does not have a closed-form solution, while the problem considered in [5] reduces to the problem of minimizing the sum of squares of the Euclidean distances to the demand points. The latter has a closed form solution given by the mean of the data points. Solving the clustering problem in [5] and solving our multifacility location problem (6.1.36) are difficult because of their discrete nature and nonconvexity, while the multifacility location problem (6.1.36) requires techniques to deal with its nondifferentiability/nonsmoothness. This chapter presents the work we have done in [7]. The reader can find more details in there.

6.2. Smooth Approximation by Continuous DC Problems

Using Nesterov's smoothing technique allows us to approximate the objective function \mathcal{F} in (6.1.36) by a smooth DC function \mathcal{F}_μ as $\mu > 0$ defined as follows

$$\begin{aligned}\mathcal{F}_\mu(\mathbf{U}, \mathbf{V}) &:= \frac{1}{2\mu} \sum_{i=1}^k \sum_{j=1}^n u_{i,j}^2 \|a_j - v_i\|^2 - \frac{\mu}{2} \sum_{i=1}^k \sum_{j=1}^n u_{i,j}^2 \left[d\left(\frac{a_j - v_i}{\mu}; \mathbb{B}\right) \right]^2 \\ &= \mathcal{G}_\mu(\mathbf{U}, \mathbf{V}) - \mathcal{H}_\mu(\mathbf{U}, \mathbf{V}),\end{aligned}$$

where $\mathcal{G}_\mu, \mathcal{H}_\mu: \mathbb{R}^{k \times n} \times \mathbb{R}^{k \times d} \rightarrow \mathbb{R}$ are given by

$$\begin{aligned}\mathcal{G}_\mu(\mathbf{U}, \mathbf{V}) &:= \frac{1}{2\mu} \sum_{i=1}^k \sum_{j=1}^n u_{i,j}^2 \|a_j - v_i\|^2, \\ \mathcal{H}_\mu(\mathbf{U}, \mathbf{V}) &:= \frac{\mu}{2} \sum_{i=1}^k \sum_{j=1}^n u_{i,j}^2 \left[d\left(\frac{a_j - v_i}{\mu}; \mathbb{B}\right) \right]^2.\end{aligned}$$

This leads us to the construction of the following family of smooth approximations of the main problem (6.1.36) defined by

$$(6.2.37) \quad \begin{aligned} & \text{minimize} && \mathcal{F}_\mu(\mathbf{U}, \mathbf{V}) := \mathcal{G}_\mu(\mathbf{U}, \mathbf{V}) - \mathcal{H}_\mu(\mathbf{U}, \mathbf{V}) \text{ as } \mu > 0 \\ & \text{subject to} && \mathbf{U} \in \mathcal{U} = \Delta^n \cap \{0, 1\}^{k \times n} \text{ and } \mathbf{V} \in \mathbb{R}^{k \times d}, \end{aligned}$$

where Δ^n is the the n th Cartesian degree of the $(k - 1)$ -simplex

$$\Delta := \{y \in [0, 1]^k \mid \sum_{i=1}^k y_i = 1\},$$

which is a subset of \mathbb{R}^k .

Observe that for each $\mu > 0$ problem (6.2.37) is of *discrete optimization*, while our intention is to convert it to a family of problems of *continuous optimization* for which we are going to develop and implement a DCA-based algorithm in Section 6.3.

The rest of this section is devoted to deriving two results, which justify such a reduction. The first theorem allows us to verify the *existence* of optimal solutions to the constrained optimization problems that appear in this procedure. It is required for having well-posedness of the algorithm construction.

Theorem 6.2.1. *Let $(\bar{\mathbf{U}}, \bar{\mathbf{V}})$ be an optimal solution to problem (6.2.37). Then for any $\mu > 0$ we have $\bar{\mathbf{V}} \in \mathcal{B}$, where $\mathcal{B} := \prod_{i=1}^k B_i$ is the Cartesian product of the k Euclidean balls B_i centered at $0 \in \mathbb{R}^d$ with radius $r := \sqrt{\sum_{j=1}^n \|a_j\|^2}$ that contain the optimal centers \bar{v}_i for each index $i = 1, \dots, k$.*

Proof. We can clearly rewrite the objective function in (6.2.37) in the form

$$(6.2.38) \quad \mathcal{F}_\mu(\mathbf{U}, \mathbf{V}) = \frac{1}{2\mu} \sum_{i=1}^k \sum_{j=1}^n u_{i,j} \|a_j - v_i\|^2 - \frac{\mu}{2} \sum_{i=1}^k \sum_{j=1}^n u_{i,j} \left[d\left(\frac{a_j - v_i}{\mu}; \mathbb{B}\right) \right]^2$$

due to interchangeability between $u_{i,j}^2$ and $u_{i,j}$. Observe that $\mathcal{F}_\mu(\mathbf{U}, \mathbf{V})$ is differentiable on $R^{k \times n} \times \mathbb{R}^{k \times d}$. Employing the classical Fermat rule in (6.2.37) with respect to V gives us $\nabla_{\mathbf{V}} \mathcal{F}_\mu(\bar{\mathbf{U}}, \bar{\mathbf{V}}) = 0$. To calculate this partial gradient, we need some clarification for the second term in (6.2.38), which is differentiable as a whole while containing the nonsmooth distance function. The convexity of the distance function in the setting

of (6.2.38) allows us to apply the subdifferential calculation of convex analysis (see, e.g., [25, Theorem 2.39]) and to combine it with an appropriate chain rule to handle the composition in (6.2.38). Observe that the distance function square in (6.2.38) is the composition of the nondecreasing convex function $\varphi(t) := t^2$ on $[0, \infty)$ and the distance function to the ball \mathbb{B} . Thus the chain rule from [25, Corollary 2.62] is applicable. Thus, we can show that $d^2(\cdot; \mathbb{B})$ is differentiable with

$$(6.2.39) \quad \nabla d^2(x; \mathbb{B}) = 2[x - P(x; \mathbb{B})] \text{ for } x \in \mathbb{R}^d.$$

Using (6.2.39), we consider the following two cases:

Case 1: $(a_j - \bar{v}_i)/\mu \in \mathbb{B}$ for the fixed indices $i \in \{1, \dots, k\}$ and $j \in \{1, \dots, n\}$. Then

$$\nabla d^2\left(\frac{a_j - \bar{v}_i}{\mu}; \mathbb{B}\right) = \{0\},$$

which gives us

$$\frac{\partial \mathcal{F}_\mu}{\partial v_i}(\bar{\mathbf{U}}, \bar{\mathbf{V}}) = \frac{1}{\mu} \sum_{j=1}^n \bar{u}_{i,j} (\bar{v}_i - a_j), \quad i = 1, \dots, k,$$

for the corresponding partial derivatives of \mathcal{F}_μ .

Case 2: $(a_j - \bar{v}_i)/\mu \notin \mathbb{B}$ for the fixed indices $i \in \{1, \dots, k\}$ and $j \in \{1, \dots, n\}$. In this case we have

$$\begin{aligned} \frac{\partial \mathcal{F}_\mu}{\partial v_i}(\bar{\mathbf{U}}, \bar{\mathbf{V}}) &= \frac{1}{2\mu} \sum_{j=1}^n \bar{u}_{i,j} 2(\bar{v}_i - a_j) + \sum_{j=1}^n \bar{u}_{i,j} \left[\frac{a_j - \bar{v}_i}{\mu} - P\left(\frac{a_j - \bar{v}_i}{\mu}; \mathbb{B}\right) \right] \\ &= \frac{1}{\mu} \sum_{j=1}^n \bar{u}_{i,j} (\bar{v}_i - a_j) + \sum_{j=1}^n \bar{u}_{i,j} \left[\frac{a_j - \bar{v}_i}{\mu} - \left(\frac{a_j - \bar{v}_i}{\|a_j - \bar{v}_i\|} \right) \right] \\ &= \frac{1}{\|a_j - \bar{v}_i\|} \sum_{j=1}^n \bar{u}_{i,j} (\bar{v}_i - a_j). \end{aligned}$$

Thus in both cases above it follows from the stationary condition $\nabla_{\mathbf{V}} \mathcal{F}_\mu(\bar{\mathbf{U}}, \bar{\mathbf{V}}) = 0$ that

$$\bar{v}_i = \frac{\sum_{j=1}^n \bar{u}_{i,j} a_j}{\sum_{j=1}^n \bar{u}_{i,j}} \text{ for all } i = 1, \dots, k,$$

since we have $\sum_{j=1}^n \bar{u}_{i,j} > 0$ due to the nonemptiness of the clusters. Then the classical Cauchy-Schwarz inequality leads us to the estimates

$$\|\bar{v}_i\|^2 \leq \frac{\left(\sum_{j=1}^n \bar{u}_{i,j} a_j\right)^2}{\left(\sum_{j=1}^n \bar{u}_{i,j}\right)^2} \leq \sum_{j=1}^n \|a_j\|^2 := r^2,$$

which therefore verify all the conclusions of this theorem. \square

Our next step is to enclose each discrete optimization problem (6.2.37) into the corresponding one of *continuous optimization*. For the reader's convenience if no confusion arises, we keep the same notation \mathbf{U} for *all* the $k \times n$ matrices without the discrete restrictions on their entries. Define now the function $\mathcal{P}: \mathbb{R}^{k \times n} \rightarrow \mathbb{R}$ by

$$\mathcal{P}(\mathbf{U}) := \sum_{i=1}^k \sum_{j=1}^n u_{i,j}(1 - u_{i,j}) \quad \text{for all } \mathbf{U} \in \mathbb{R}^{k \times n}$$

and observe that this function is concave on $\mathbb{R}^{k \times n}$ with $\mathcal{P}(\mathbf{U}) \geq 0$ whenever $\mathbf{U} \in \Delta^n$.

Furthermore, we have the representations

$$(6.2.40) \quad \mathcal{U} = \{\mathbf{U} \in \Delta^n \mid \mathcal{P}(\mathbf{U}) = 0\} = \{\mathbf{U} \in \Delta^n \mid \mathcal{P}(\mathbf{U}) \leq 0\}$$

for the set of feasible $k \times n$ -matrices \mathcal{U} in the original problem (6.1.36). Employing further the standard *penalty function* method allows us to eliminate the most involved constraint on \mathbf{U} in (6.2.40) given by the function \mathcal{P} . Taking the penalty parameter $\alpha > 0$ sufficiently large and using the smoothing parameter $\mu > 0$ sufficiently small, consider the following family of continuous optimization problems

$$(6.2.41) \quad \begin{array}{ll} \text{minimize} & \mathcal{F}_\mu(\mathbf{U}, \mathbf{V}) + \alpha \mathcal{P}(\mathbf{U}) = \mathcal{G}_\mu(\mathbf{U}, \mathbf{V}) - \mathcal{H}_\mu(\mathbf{U}, \mathbf{V}) + \alpha \mathcal{P}(\mathbf{U}) \\ \text{subject to} & \mathbf{U} \in \Delta^n \text{ and } \mathbf{V} \in \mathcal{B}. \end{array}$$

Observe that Theorem 6.2.1 ensures the existence of feasible solutions to problem (6.2.41) and hence optimal solutions to this problem by the Weierstrass theorem due to

the continuity of the objective functions therein and the compactness of the constraints sets Δ^n and \mathcal{B} .

Let us introduce yet another parameter $\rho > 0$ ensuring a DC representation of the objective function in (6.2.41) as follows

$$\begin{aligned}
\mathcal{F}_\mu(\mathbf{U}, \mathbf{V}) + \alpha\mathcal{P}(\mathbf{U}) &= \\
&= \frac{\rho}{2}\|\mathbf{U}, \mathbf{V}\|^2 - \left(\frac{\rho}{2}\|\mathbf{U}, \mathbf{V}\|^2 - \mathcal{F}_\mu(\mathbf{U}, \mathbf{V}) - \alpha\mathcal{P}(\mathbf{U}) \right) \\
&= \frac{\rho}{2}\|\mathbf{U}, \mathbf{V}\|^2 - \left(\frac{\rho}{2}\|\mathbf{U}, \mathbf{V}\|^2 - \mathcal{G}_\mu(\mathbf{U}, \mathbf{V}) + \mathcal{H}_\mu(\mathbf{U}, \mathbf{V}) - \alpha\mathcal{P}(\mathbf{U}) \right) \\
&=: \mathcal{G}(\mathbf{U}, \mathbf{V}) - \mathcal{H}(\mathbf{U}, \mathbf{V}),
\end{aligned}$$

where the function $\mathcal{G}(\mathbf{U}, \mathbf{V}) := \frac{\rho}{2}\|\mathbf{U}, \mathbf{V}\|^2$ is obviously convex, and

$$\mathcal{H}(\mathbf{U}, \mathbf{V}) := \frac{\rho}{2}\|\mathbf{U}, \mathbf{V}\|^2 - \mathcal{G}_\mu(\mathbf{U}, \mathbf{V}) + \mathcal{H}_\mu(\mathbf{U}, \mathbf{V}) - \alpha\mathcal{P}(\mathbf{U}).$$

Since $\mathcal{H}_\mu(\mathbf{U}, \mathbf{V}) - \alpha\mathcal{P}(\mathbf{U})$ is also convex as $\alpha > 0$, we are going to show that for any given number $\mu > 0$ it is possible to determine the values of the parameter $\rho > 0$ such that the function $\frac{\rho}{2}\|\mathbf{U}, \mathbf{V}\|^2 - \mathcal{G}_\mu(\mathbf{U}, \mathbf{V})$ is convex under an appropriate choice of ρ . This would yield the convexity of $\mathcal{H}(\mathbf{U}, \mathbf{V})$ and therefore would justify a desired representation of the objective function in (6.2.41). The following result gives us a precise meaning of this statement, which therefore verify the required reduction of (6.2.41) to *DC continuous optimization*.

Theorem 6.2.2. *The function*

$$(6.2.42) \quad \mathcal{G}_1(\mathbf{U}, \mathbf{V}) := \frac{\rho}{2}\|\mathbf{U}, \mathbf{V}\|^2 - \mathcal{G}_\mu(\mathbf{U}, \mathbf{V})$$

is convex on $\Delta^n \times \mathcal{B}$ provided that

$$(6.2.43) \quad \rho \geq \frac{n}{2\mu} \left[\left(1 + \frac{1}{n}\xi^2 \right) + \sqrt{\left(1 + \frac{1}{n}\xi^2 \right)^2 + \frac{12}{n}\xi^2} \right],$$

where $\xi := r + \max_{1 \leq j \leq n} \|a_j\|$ and $r := \sqrt{\sum_{j=1}^n \|a_j\|^2}$.

Proof. Consider the function $\mathcal{G}_1(\mathbf{U}, \mathbf{V})$ defined in (6.2.42) for all $(\mathbf{U}, \mathbf{V}) \in \Delta^n \times \mathcal{B}$ and deduce by elementary transformations directly from its construction that

$$\begin{aligned}
\mathcal{G}_1(\mathbf{U}, \mathbf{V}) &= \frac{\rho}{2} \|(\mathbf{U}, \mathbf{V})\|^2 - \mathcal{G}_\mu(\mathbf{U}, \mathbf{V}) \\
&= \frac{\rho}{2} \|(\mathbf{U}, \mathbf{V})\|^2 - \frac{1}{2\mu} \sum_{i=1}^k \sum_{j=1}^n u_{i,j}^2 \|a_j - v_i\|^2 \\
&= \frac{\rho}{2} \|\mathbf{U}\|^2 + \frac{\rho}{2} \|\mathbf{V}\|^2 - \frac{1}{2\mu} \sum_{i=1}^k \sum_{j=1}^n u_{i,j}^2 \|a_j - v_i\|^2 \\
&= \frac{\rho}{2} \sum_{i=1}^k \sum_{j=1}^n u_{i,j}^2 + \frac{\rho}{2n} \sum_{i=1}^k \sum_{j=1}^n \|v_i\|^2 - \frac{1}{2\mu} \sum_{i=1}^k \sum_{j=1}^n u_{i,j}^2 \|a_j - v_i\|^2 \\
&= \sum_{i=1}^k \sum_{j=1}^n \frac{\rho}{2} u_{i,j}^2 + \frac{\rho}{2n} \|a_j - v_i\|^2 + \\
&\quad + \frac{\rho}{n} \langle a_j, v_i \rangle - \frac{\rho}{2n} \|a_j\|^2 - \frac{1}{2\mu} u_{i,j}^2 \|a_j - v_i\|^2.
\end{aligned}$$

Next we define the functions $\gamma_{i,j}: \mathbb{R} \times \mathbb{R}^d \rightarrow \mathbb{R}$ for all $i = 1, \dots, k$ and $j = 1, \dots, n$ by

$$(6.2.44) \quad \gamma_{i,j}(u_{i,j}, v_i) := \frac{\rho}{2} u_{i,j}^2 + \frac{\rho}{2n} \|a_j - v_i\|^2 - \frac{1}{2\mu} u_{i,j}^2 \|a_j - v_i\|^2$$

and show that each of these functions is convex on the set $\{u_{i,j} \in [0, 1], \|v_i\| \leq r\}$, where $r > 0$ is taken from Theorem 6.2.1.

To proceed, consider the Hessian matrix of each function in (6.2.44) given by

$$J_{\gamma_{i,j}}(u_{i,j}, v_i) := \begin{bmatrix} \rho - \frac{1}{\mu} \|a_j - v_i\|^2 & -\frac{2}{\mu} u_{i,j} (v_i - a_j) \\ -\frac{2}{\mu} u_{i,j} (v_i - a_j) & \frac{\rho}{n} - \frac{1}{\mu} u_{i,j}^2 \end{bmatrix}$$

and calculate its determinant $\det(J_{\gamma_{i,j}}(u_{i,j}, v_i))$ by

$$\begin{aligned}
\det(J_{\gamma_{i,j}}(u_{i,j}, v_i)) &:= \left(\rho - \frac{1}{\mu} \|a_j - v_i\|^2 \right) \left(\frac{\rho}{n} - \frac{1}{\mu} u_{i,j}^2 \right) - \frac{4}{\mu^2} u_{i,j}^2 (v_i - a_j)^T (v_i - a_j) \\
&= \frac{\rho^2}{n} - \rho \left(\frac{u_{i,j}^2}{\mu} + \frac{1}{n\mu} \|v_i - a_j\|^2 \right) - \frac{3u_{i,j}^2}{\mu^2} \|v_i - a_j\|^2.
\end{aligned}$$

It follows from the well-known second-order characterization of the convexity that the function $\gamma_{i,j}(u_{i,j}, v_i)$ is convex on $\{u_{i,j} \in [0, 1], \|v_i\| \leq r\}$ if $\det(J_{\gamma_{i,j}}(u_{i,j}, v_i)) \geq 0$. Using [6, Theorem 1] gives us the estimate

$$\det(J_{\gamma_{i,j}}(u_{i,j}, v_i)) \geq \frac{\rho^2}{n} - \rho \left(\frac{1}{\mu} + \frac{1}{n\mu} \|v_i - a_j\|^2 \right) - \frac{3}{\mu^2} \|v_i - a_j\|^2.$$

Then we get from the construction of \mathcal{B} in Theorem 6.2.1 that $0 < \|v_i - a_j\| \leq \|v_i\| + \|a_j\| \leq r + \max_{1 \leq j \leq n} \|a_j\| =: \xi$, and therefore

$$(6.2.45) \quad \det(J_{\gamma_{i,j}}(u_{i,j}, v_i)) \geq \frac{\rho^2}{n} - \frac{\rho}{\mu} \left(1 + \frac{1}{n} \xi^2 \right) - \frac{3}{\mu^2} \xi^2.$$

It allows us to deduce from the aforementioned condition for the convexity of $\gamma_{i,j}(u_{i,j}, v_i)$ that we do have this convexity if ρ satisfies the estimate (6.2.43). \square

6.3. Design and Implementation of the Solution Algorithm

Based on the developments presented in the previous sections and using the established smooth DC structure of problem (6.2.41) with the subsequent ρ -parameterization of the objective function therein as $\mathcal{G}(\mathbf{U}, \mathbf{V}) - \mathcal{H}(\mathbf{U}, \mathbf{V})$, we are now ready to propose and implement a new algorithm for solving this problem involving both DCA-2 and Nesterov's smoothing.

To proceed, let us present the problem under consideration in the equivalent *unconstrained* format by using the infinite penalty via the indicator function

$$(6.3.46) \quad \begin{aligned} & \text{minimize} && \frac{\rho}{2} \|\mathbf{U}, \mathbf{V}\|^2 - \mathcal{H}(\mathbf{U}, \mathbf{V}) + \delta_{\Delta \times \mathcal{B}}(\mathbf{U}, \mathbf{V}) \\ & \text{subject to} && (\mathbf{U}, \mathbf{V}) \in \mathbb{R}^{k \times n} \times \mathbb{R}^{k \times d}, \end{aligned}$$

where \mathcal{B} , Δ , and ρ are taken from Section 6.2.

We first explicitly compute the gradient of the convex function $\mathcal{H}(\mathbf{U}, \mathbf{V})$ in (6.3.46).

Denoting

$$\begin{aligned} [\mathcal{Y}, \mathcal{Z}] &:= \nabla \mathcal{H}(\mathbf{U}, \mathbf{V}) = \nabla \left(\frac{\rho}{2} \|\mathbf{U}, \mathbf{V}\|^2 - \frac{1}{2\mu} \sum_{i=1}^k \sum_{j=1}^n u_{i,j}^2 \|a_j - v_i\|^2 \right. \\ &\quad \left. + \frac{\mu}{2} \sum_{i=1}^k \sum_{j=1}^n u_{i,j}^2 \left[d \left(\frac{a_j - v_i}{\mu}; \mathbb{B} \right) \right]^2 - \alpha \sum_{i=1}^k \sum_{j=1}^n u_{i,j} (1 - u_{i,j}) \right), \end{aligned}$$

we have $\mathcal{Y} = \nabla \mathcal{H}_{\mathbf{U}}(\mathbf{U}, \mathbf{V})$ and $\mathcal{Z} = \nabla \mathcal{H}_{\mathbf{V}}(\mathbf{U}, \mathbf{V})$. Thus for each $i = 1, \dots, k$ and $j = 1, \dots, n$ the (j, i) -entry of the matrix \mathcal{Y} and the i th row of the matrix \mathcal{Z} are

$$\begin{aligned} \mathcal{Y}_{j,i} &:= \rho u_{i,j} - \frac{u_{i,j}}{\mu} \|a_j - v_i\|^2 + \mu u_{i,j} \left[d \left(\frac{a_j - v_i}{\mu}; \mathbb{B} \right) \right]^2 + 2\alpha u_{i,j} - \alpha, \\ \mathcal{Z}_i &:= \rho v_i - \frac{1}{\mu} \sum_{j=1}^n u_{i,j}^2 (v_i - a_j) - \sum_{j=1}^n u_{i,j}^2 \left[\frac{a_j - v_i}{\mu} - P \left(\frac{a_j - v_i}{\mu}; \mathbb{B} \right) \right], \end{aligned}$$

respectively. Let us now describe the proposed algorithm for solving the DC program (6.3.46) and hence the original problem (6.1.36) of multifacility location. The symbols $\mathcal{Y}_{[j,:]}^l$ and \mathcal{Z}_i^l in this description represents the j th row of the matrix \mathcal{Y} and the i th row of the matrix \mathcal{Z} at the l th iteration, respectively. Accordingly we use the symbols $\mathbf{U}_{[:j]}^{(l+1)}$ and $\mathbf{V}_i^{(l+1)}$.

Next we employ Algorithm 9 to solving several multifacility location problems of some practical meaning. By trial and error we verify that the values chosen for μ determine the performance of the algorithm for each data set. It can be seen that very small values of the smoothing parameter μ may prevent the algorithm from clustering, and thus we gradually decrease these values. This is done via multiplying μ by some number $0 < \beta < 1$ and stopping when $\mu < \mu_f$. Note also that in the implementation of our algorithm we use the standard approach of choosing \mathbf{U}^0 by computing the distance between the point in question and each group center \mathbf{V}^0 and then by classifying this point to be in the group whose center is the closest to it by assigning the value of 1, while otherwise we assign the value of 0.

Let us now present several numerical examples, where we compute the optimal centers by using Algorithm 9 via MATLAB calculations. Fix in what follows the values of

Algorithm 9 Solving Multifacility Location Problems

INPUT: \mathbf{X} (the dataset), \mathbf{V}^0 (initial centers), ClusterNum (number of clusters), $\mu > 0$, β (scaling parameter) > 0 .

INITIALIZATION \mathbf{U}^0 , $\varepsilon > 0$, μ_f (minimum threshold for μ) > 0 , $\alpha > 0$, $\rho > 0$, tol (tolerance parameter) = 1, $l = 0$

while tol $> \varepsilon$ and $\mu > \mu_f$ **do**

Repeat

for $1 \leq i \leq k$ and $1 \leq j \leq n$ **do**

$$\mathcal{Y}_{j,i}^l := \rho u_{i,j}^l - \frac{u_{i,j}^l}{\mu} \|a_j - v_i^l\|^2 + \mu u_{i,j}^l \left[d\left(\frac{a_j - v_i^l}{\mu}; \mathbb{B}\right) \right]^2 + 2\alpha u_{i,j}^l - \alpha,$$

$$\mathcal{Z}_i^l := \rho v_i^l - \frac{1}{\mu} \sum_{j=1}^n (u_{i,j}^l)^2 (v_i^l - a_j) - \sum_{j=1}^n (u_{i,j}^l)^2 \left[\frac{a_j - v_i^l}{\mu} - P\left(\frac{a_j - v_i^l}{\mu}; \mathbb{B}\right) \right].$$

end for

for $1 \leq i \leq k$ and $1 \leq j \leq n$ **do**

$$\mathbf{U}_{[i,j]}^{(l+1)} := P\left(\frac{\mathcal{Y}_{j,i}^l}{\rho}; \Delta\right),$$

$$\mathbf{V}_i^{(l+1)} := P\left(\frac{\mathcal{Z}_i^l}{\rho}; B_i\right) = \begin{cases} \frac{\mathcal{Z}_i^l}{\rho} & \text{if } \|\mathcal{Z}_i^l\| \leq \rho r, \\ \frac{\rho}{r} \frac{\mathcal{Z}_i^l}{\|\mathcal{Z}_i^l\|} & \text{if } \|\mathcal{Z}_i^l\| > \rho r. \end{cases}$$

end for

$l := l + 1$

Update:

$$\text{tol} := \left\| [\mathbf{U}^{(l+1)}, \mathbf{V}^{(l+1)}] - [\mathbf{U}^l, \mathbf{V}^l] \right\|_F$$
$$\mu := \beta \mu$$

end while

OUTPUT: $[\mathbf{U}^{(l+1)}, \mathbf{V}^{(l+1)}]$

$\mu = 0.5$, $\beta = 0.85$, $\varepsilon = 10^{-6}$, $\mu_f = 10^{-6}$, $\alpha = 30$, and $\rho = 30$ unless otherwise stated. The objective function is the *total distance* from the centers to the assigned data point. Note that this choice of the objective function seems to be natural from practical aspects in, e.g., airline and other transportation industries, where the goal is to reach the destination via the best possible route available. This reflects minimizing the transportation cost.

In the following examples we implement the standard k -means algorithm in MATLAB using the built-in function `kmeans()`.

Example 6.3.1. Let us consider a data set with 14 entries in \mathbb{R}^2 given by

$$\mathbf{X} := \begin{bmatrix} 0 & 2 & 7 & 2 & 3 & 6 & 5 & 8 & 8 & 9 & 1 & 7 & 0 & 0 \\ 3 & 2 & 1 & 4 & 3 & 2 & 3 & 1 & 3 & 2 & 1 & 4 & 4 & 1 \end{bmatrix}^T$$

with the initial data defined by

$$\mathbf{V}^0 := \begin{bmatrix} 7.1429 & 2.2857 \\ 1.1429 & 2.5714 \end{bmatrix} \text{ is obtained from the } k\text{-means algorithm; see Table 6.1,}$$

ClusterNum := 2.

Employing Algorithm 9, we obtain the optimal centers as depicted in Table 6.1 and Figure 6.1.

TABLE 6.1. Comparison between Algorithm 3 and k -means.

Method	Optimal Center (\mathbf{V}^N)	Cost Function
k -means	$\begin{bmatrix} 7.1429 & 2.2857 \\ 1.1429 & 2.5714 \end{bmatrix}$	22.1637
Algorithm 9	$\begin{bmatrix} 7.2220 & 2.1802 \\ 1.1886 & 2.5069 \end{bmatrix}$	22.1352

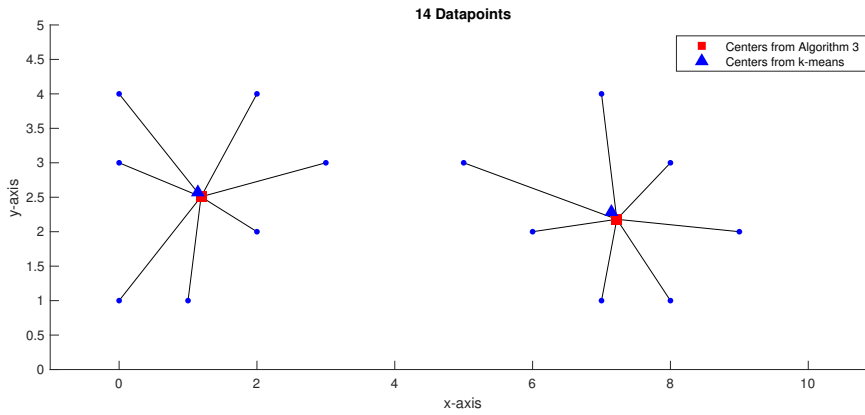


FIGURE 6.1. MFLP with 14 demand points and 2 centers.

Table 6.1 shows that the proposed Algorithm 9 is marginally better for the given data in comparison to the classical k -means approach in terms of the objective function.

In the next example we choose the initial cluster by the process of *random selection* and see its effect on the optimal centers. Then the results obtained in this way by Algorithm 9 are compared with those computed by the *k*-means approach.

Example 6.3.2. Let \mathbf{X} be 200 standard normally distributed random datapoints in \mathbb{R}^2 , and let the initial data be given by

$$\mathbf{V}^0 := \text{randomly permuting and selecting 2 rows of } \mathbf{X},$$

$$\text{ClusterNum} := 2.$$

We obtain the optimal centers as outlined in Table 6.2.

TABLE 6.2. Comparison between Algorithm 3 and *k*-means.

Method	Optimal Center (\mathbf{V}^N)	Cost Function
<i>k</i> -means	$\begin{bmatrix} 2.1016 & 1.2320 \\ -1.3060 & -1.0047 \end{bmatrix}$	403.3966
Algorithm 9	$\begin{bmatrix} 1.4902 & 0.7406 \\ -1.3464 & -1.0716 \end{bmatrix}$	401.7506

Observe from Table 6.2 that the proposed Algorithm 9 is better for the given data in comparison to the standard *k*-means approach. In addition, our approach gives a better approximation for the optimal center as shown in Figure 6.2.

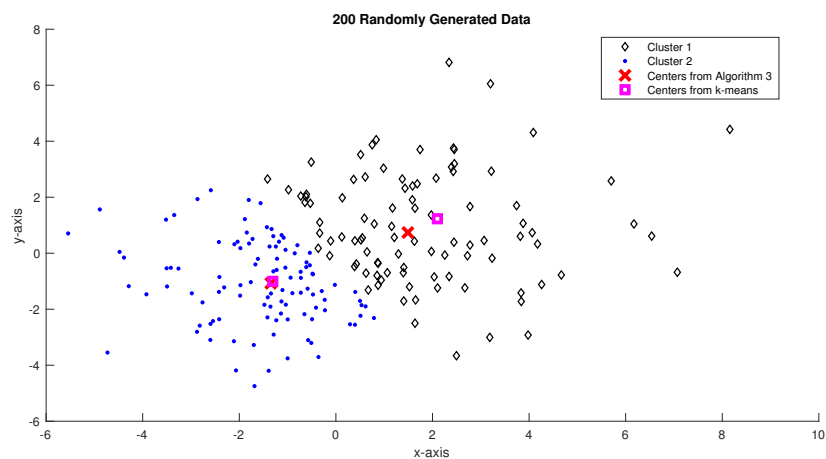


FIGURE 6.2. MFLP with 200 demand points and 2 centers.

Note that a real-life data may not be as efficiently clustered as in Example 6.3.2. Thus a suitable selection of the initial cluster \mathbf{V}^0 is vital for the convergence of the DCA based algorithms. In the next Example 6.3.3 we select \mathbf{V}^0 in Algorithm 9 by using the standard k -means method. The results achieved by our Algorithm 9 are again compared with those obtained by using the k -means approach.

Example 6.3.3. Consider the dataset \mathbf{X} consisting of the latitudes and longitudes of 50 most populous cities in the USA¹ with

$$\mathbf{V}^0 := \begin{bmatrix} -80.9222 & 37.9882 \\ -97.8273 & 35.3241 \\ -118.3121 & 36.9535 \end{bmatrix} \text{ is obtained from the } k\text{-means algorithm; see Table 6.3,}$$

ClusterNum := 3.

By using Algorithm 9 we obtain the following optimal centers as given in Table 6.3.

TABLE 6.3. Comparison between Algorithm 9 (combined with k -means) and standard k -means.

Method	Optimal Center (\mathbf{V}^N)	Cost Function
k -means	$\begin{bmatrix} -80.9222 & 37.9882 \\ -97.8273 & 35.3241 \\ -118.3121 & 36.9535 \end{bmatrix}$	288.8348
Algorithm 9 (combined with k -means)	$\begin{bmatrix} -81.0970 & 38.3092 \\ -97.4138 & 35.3383 \\ -119.3112 & 36.5410 \end{bmatrix}$	286.6523

We see that Algorithm 3 (combined with k -means) in which the initial cluster \mathbf{V}^0 is selected by using k -means method performs better in comparison to the standard k -means approach (Table 6.3). Moreover, it gives us optimal centers as depicted in Figure 6.3.

¹Available at https://en.wikipedia.org/wiki/List_of_United_States_cities_by_population

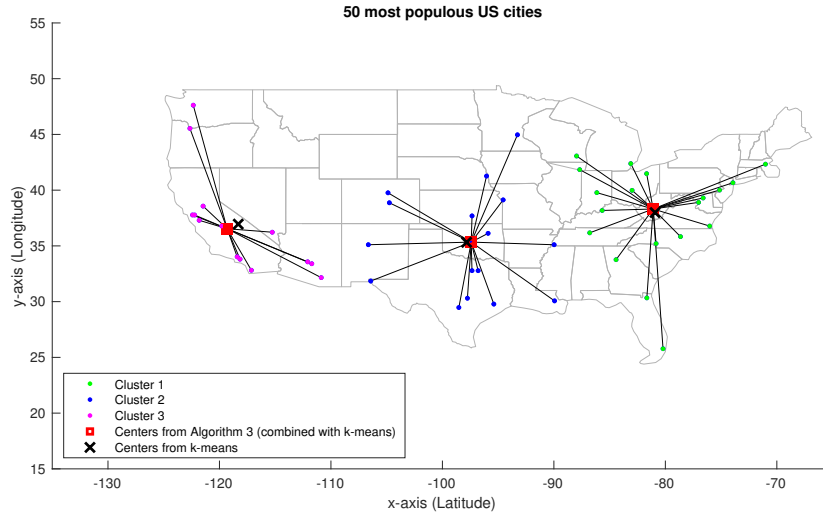


FIGURE 6.3. MFLP with 50 demand points and 3 centers.

In the last example presented in this section we efficiently solve a higher dimensional multifacility location problem by using Algorithm 9 and compare it's value of the cost function with the standard k -means algorithm.

Example 6.3.4. Let X in \mathbb{R}^{13} be the *wine dataset* from the UCI Machine Learning Repository [50] consisting of 178 demand points. We apply Algorithm 3 with

V^0 is obtained from the k -means algorithm,

ClusterNum := 3.

The total costs using Algorithm 9 and the k -means algorithm are obtained in Table 6.4 showing that former algorithm is better than the latter.

TABLE 6.4. Cost comparison between Algorithm 3 (combined with k -means) and k -means.

Method	Cost Function
k -means	16556
Algorithm 9 (combined with k -means)	16460

Conclusion

In chapter 1 and 2, we presented the basic tools of convex analysis and optimization which built the mathematical foundation for later chapters, and gave numerical techniques and algorithms, namely Nesterov's smoothing technique and DCA.

In chapter 3 we represented new results of convex analysis as well as their unified and simplified proofs based on the geometric variational analysis approach for the case of convex objects in locally convex topological and Banach spaces.

Throughout this thesis we saw that all the algorithms were simple, which is one of the key advantages of this thesis. In chapter 4, 5 and 6 we demonstrated the successful usage of DCA and Nesterov's smoothing technique to solve facility locations and clustering numerically in many different settings. However, it should be noted that most of the algorithms were implemented in two dimensions with nonnegative weights. So, extending them to higher dimensions with arbitrary weights is another direction towards future works.

Finally, we see that extensive research could be to improve convergence rates by applying other numerical algorithms such as boosted/stochastic DCA. Furthermore, it is possible to study the problems in which the distances are generated by different norms, for instance l_1 and l_∞ . Additionally, interesting areas to explore include using Nesterov's Accelerated version, refining the initial cluster selection, for instance, using derivative-free methods for initializations in the DCA, the stopping criterion and the enhancing the effectiveness of the gradient/subgradient-based nonconvex algorithms.

References

- [1] L.T.H. An, M.T. Belghiti, P.D. Tao: A new efficient algorithm based on DC programming and DCA for clustering, *J. Glob. Optim.*, **37**(4) (2007) 593–608
- [2] L. T. H. An, and L. H. Minh, Optimization based DC programming and DCA for hierarchical clustering. *European J. Oper. Res.* **183** (2007), 1067–1085.
- [3] N.T. An, N.M. Nam: Convergence analysis of a proximal point algorithm for minimizing differences of functions, *Optim.*, **66**(1) (2017) 129–147
- [4] L.T.H. An, H.V. Ngai, P.D. Tao: Convergence Analysis of Difference-of-Convex Algorithm with Subanalytic Data, *J. Optim. Theory Appl.*, **179**(1) (2018) 103–126
- [5] L.T.H. An, L.H. Minh, P.D. Tao, New and efficient DCA based algorithms for minimum sum-of-squares clustering, *Pattern Recogn.*, **47**(1) (2014) 388–401
- [6] L.T.H. An, P.D. Tao, Minimum Sum-of-Squares Clustering by DC Programming and DCA, *Int. Conf. Intell. Comp.*, (2009) 327–340
- [7] A. Bajaj, B. Mordukhovich, N.M. Nam, and T. Tran, *Solving multifacility location problems by DC Algorithms*, (2019), *J. Optim Method Softw.*(arXiv:1906.10331v1).
- [8] H. Attouch and H. Brézis, Duality of the sum of convex functions in general Banach spaces, in *Aspects of Mathematics and Its Applications* (J. A. Barroso, ed.), vol. 34, pp. 125-133, North-Holland, Amsterdam, 1986.
- [9] H. H. Bauschke, J. M. Borwein and W. Li, Strong conical hull intersection property, bounded linear regularity, Jameson’s property (G), and error bounds in convex optimization, *Math. Program.* **86** (1999), 135–160.
- [10] A. M. Bagirov, Derivative-free methods for unconstrained nonsmooth optimization and its numerical analysis. *Investigacao Operacional.* **19** (1999), 75–93.

- [11] A. Bagirov, Long Jia, I. Ouveysi, and A.M. Rubinov, Optimization based clustering algorithms in Multicast group hierarchies, in: Proceedings of the Australian Telecommunications, Networks and Applications Conference (ATNAC), 2003, Melbourne Australia (published on CD, ISBN 0-646-42229-4).
- [12] R. I. Boş, *Conjugate Duality in Convex Optimization*, Springer, Berlin, 2010.
- [13] J. M. Borwein and Q. J. Zhu, *Techniques of Variational Analysis*, Springer, New York, 2005.
- [14] R. S. Burachik and V. Jeyakumar, A dual condition for the convex subdifferential sum formula with applications, *J. Convex Anal.* **12** (2005), 279–290.
- [15] E. Chi, H. Zhou, K. Lange, Distance majorization and its applications, *Math. Program. Series A* **146**, (2014), 409–436.
- [16] N. Dunford and J. T. Schwartz, *Linear Operators, Part I*, Interscience, New York, 1964
- [17] E. Ernst and M. Théra, Boundary half-strips and the strong CHIP, *SIAM J. Optim.* **18** (2007), 834–852.
- [18] D. Fisher, Iterative optimization and simplification of hierarchical clusterings, *Journal of Artificial Intelligence Research* **4** (1996) 147–180.
- [19] A.K. Jain, M.N. Murty, P.J. Flynn, Data clustering: A review, *ACM Computing Surveys* **31** (3) (1999) 264–323.
- [20] P. Hartman, On functions representable as a difference of convex functions. *Pacific J. Math.* **9**, (1959), 707–713.
- [21] J. B. Hiriart-Urruty, C. Lemaréchal: *Fundamentals of Convex Analysis*, Springer-Verlag (2001)
- [22] R. Horst, P. M. Pardalos, N. Van Thoai, *Introduction to global optimization*, Springer Science & Business Media (2000).
- [23] C. Li, K. F. Ng and T. K. Pong, The SECQ, linear regularity, and the strong CHIP for an infinite system of closed convex sets in normed linear spaces, *SIAM. J. Optim.* **18** (2007), 643–665.

- [24] B. S. Mordukhovich, *Variational Analysis and Generalized Differentiation, I: Basic Theory, II: Applications*, Springer, Berlin, 2006.
- [25] B.S. Mordukhovich, N.M. Nam: *An Easy Path to Convex Analysis and Applications*. Morgan & Claypool Publishers (2014)
- [26] B.S. Mordukhovich, N.M. Nam: *Convex Analysis with Applications to Optimization and Location Problems*, Springer, 2020.
- [27] B. S. Mordukhovich and N. M. Nam, Extremality of convex sets with some applications, submitted (2016); <http://arxiv.org/abs/1608.07726>.
- [28] B. S. Mordukhovich and N. M. Nam, Geometric approach to convex subdifferential calculus, *Optim.* **66**, (2017), 839–873.
- [29] B. S. Mordukhovich, N. M. Nam, R. B. Rector, T. Tran, Variational geometric approach to generalized differential and conjugate calculi in convex analysis. *Set-Valued and Variational Analysis*, **25** (4) (2017), 731-755.
- [30] B. S. Mordukhovich and H. M. Phan, Tangential extremal principle for finite and infinite systems II: Applications to semi-infinite and multiobjective optimization, *Math. Program.* **136** (2012), 31–63.
- [31] J. J. Moreau, Étude locale d’une fonctionnelle convexe, Université de Montpellier, Montpellier, France, 1963.
- [32] J. J. Moreau: Fonctionnelles sous-différentiables. *C.R.Acad.Sci.* **257**(1963) 4117-4119
- [33] F. Murtagh, A survey of recent advances in hierarchical clustering algorithms, *The Computer Journal* **26** (4) (1983).
- [34] N. M. Nam, N. T. An, S. Reynolds, T. Tran, Clustering and multifacility location with constraints via distance function penalty methods and DC programming, *Optimization.* **67** (2018), 1869–1894.
- [35] N. M. Nam, W. Geremew, S. Reynolds and T. Tran, Nesterov’s smoothing technique and minimizing differences of convex functions for hierarchical clustering, *Optim. Lett.*, **12**(3), (2018) 455–473

- [36] N.M. Nam, R.B. Rector, D. Giles, Minimizing differences of convex functions with applications to facility location and clustering, *J. Optim. Theory Appl.*, **173**(1) (2017) 255–278
- [37] Y. Nesterov: Smooth minimization of nonsmooth functions. *Math. Program.* **103**(1) (2005) 127–152
- [38] J. Nocedal, S.J. Wright, *Numerical Optimization*, Springer, New York, 2nd Edition, 2006.
- [39] R. R. Phelps, *Convex Functions, Monotone Operators and Differentiability*, 2nd edition. Springer, Berlin, 1993.
- [40] G. Reinelt, TSPLIB: A Traveling Salesman Problem Library, *ORSA Journal of Computing.* **3**, 376–384, 1991.
- [41] R. T. Rockafellar: *Convex Analysis*. Princeton University Press, Princeton, NJ (1970)
- [42] R. T. Rockafellar, *Conjugate Duality and Optimization*, SIAM, Philadelphia, PA, 1974.
- [43] R. T. Rockafellar and R. J-B. Wets, *Variational Analysis*, Springer, Berlin, 1998.
- [44] S. Simons, *From Hahn-Banach to Monotonicity*, 2nd edition, Springer, Berlin, 2008.
- [45] P.D. Tao, L.T.H. An: A D.C. optimization algorithm for solving the trust-region subproblem. *SIAM J. Optim.* **8**(2) (1998) 476–505
- [46] P.D. Tao, L.T.H. An: Convex analysis approach to D.C. programming: Theory, Algorithms and Applications. *Acta Math. Vietnam.* **22**(1) (1997) 289–355
- [47] L. Thibault, On subdifferentials of optimal value functions, *SIAM J. Control Optim.* **29** (1991), 1019–1036.
- [48] H. Tuy, et al *Convex analysis and global optimization*, Dordrecht: Kluwer (1998).
- [49] United States Cities Database. *Simple Maps: Geographic Data Products, 2017*, <http://simplemaps.com/data/us-cities>.

- [50] UCI Machine Learning Repository, <https://archive.ics.uci.edu/ml/datasets/wine>.
- [51] C. Zălinescu, *Convex Analysis in General Vector Spaces*, World Scientific, Singapore, 2002.