UNIVERSIDAD DE CONCEPCIÓN
DIRECCIÓN DE POSTGRADO
CONCEPCIÓN, CHILE

# Strong Duality in Non-convex Optimization and Related Properties

# Dualidad Fuerte en Optimización No Convexa y Propiedades Afines

**Gabriel Osvaldo Cárcamo Aravena**

*Tesis para optar al grado de*
*Doctor en Ciencias Aplicadas con mención en Ingeniería Matemática*

FACULTAD DE CIENCIAS FÍSICAS Y MATEMÁTICAS
DEPARTAMENTO DE INGENIERÍA MATEMÁTICA
NOVIEMBRE 2018

**Dualidad Fuerte en Optimización No Convexa y Propiedades Afines**

**Gabriel Osvaldo Cárcamo Aravena**

**Directores de Tesis** : Dr. Fabián Flores-Bazán, Universidad de Concepción, Chile.
Dr. Giandomenico Mastroeni, Universidad de Pisa, Italia.

**Director de Programa** : Dr. Rodolfo Rodriguez, Universidad de Concepción, Chile

**COMISIÓN EVALUADORA**

Dr. Nicolas Hadjisavvas, University of the Aegean, Greece.
Dr. Juan Peypouquet, Universidad de Chile, Chile.
Dr. Rubén López, Universidad de Tarapacá, Chile.

**COMISIÓN EXAMINADORA**

Firma : ______________________________
Dr. Cristián Vera, Universidad Arturo Prat, Chile.

Firma : ______________________________
Dr. Juan Peypouquet, Universidad de Chile, Chile.

Firma : ______________________________
Dr. Nguyen Dinh Hoang, Universidad de Concepción, Chile.

**Fecha Examen de Grado :** ______________________________

**Calificación :** ______________________________

*Concepción - Noviembre de 2018*

# Agradecimientos

Agradecer a todas las personas que me han brindado su apoyo este último tiempo. Quisiera comenzar agradeciendo a mi familia por su apoyo incondicional en este proceso. Gracias por darlo todo.

A mi director de tesis, el profesor Fabián Flores Bazán, por su apoyo académico al compartir su experiencia. A mi co-tutor, el profesor Giandomenico Mastroeni, por la gran hospitalidad que me brindó en mi estadía en Italia.

A todos y cada uno de los profesores del Programa que brindan su tiempo y conocimiento a cada generación.

A todas las personas que tuve la posibilidad de conocer, sin duda saque en limpio muchas enseñanzas.

Por el apoyo financiero, agradezco a la Comisión Nacional de Investigación Científica y Tecnológica (CONICYT) que a través de su programa de becas de doctorado nacional han financiado mis estudios, a la Red Doctoral (REDOC.CTA) y al Departamento de Ingeniería Matemática de la Universidad de Concepción por financiar las actividades académicas en las que participé, y a Becas Chile por financiar mi pasantía en el extranjero.

# Resumen

El objetivo principal de esta tesis doctoral es analizar y caracterizar la propiedad de dualidad fuerte Lagrangiana para un problema escalar no convexo sujeto a una restricción, más allá de aquellos resultados existentes en la literatura. Nuestros resultados son aplicados al caso cuadrático no convexo. En particular, obtenemos una versión relajada del teorema de Dines, asociado al problema de minimización cuadrática con una desigualdad, también del tipo cuadrática junto con varias igualdades del tipo afín. En ese sentido, discutimos el caso cuando el valor óptimal no es finito. A continuación, establecemos una caracterización del tipo geométrica de la propiedad de dualidad fuerte para el problema cuadrático no convexo, sin necesidad de asumir hipótesis de Slater, de ahí obtener condiciones necesarias y suficientes de optimalidad. Finalmente, a la luz de los resultados sobre no vacuidad del conjunto de solución, obtenidos por Frank y Wolfe, nuestra versión considera conjuntos asintóticamente lineales. En la segunda parte de esta tesis, establecemos una caracterización topológica y geométrica de la propiedad de dualidad fuerte, para un problema general no convexo sujeto a una restricción del tipo igualdad, junto con restricciones del tipo geométricas, de donde es revelada la convexidad de la envoltura cónica asociada a la imagen conjunta determinada por las funciones del problema original. Como aplicación, revisamos la validez de las condiciones de KKT sin asumir condición de regularidad estándar. En la parte final de esta tesis, estudiamos en detalle el problema cuadrático estándar, al sustituir al simplejo usual por un cono convexo, puntiagudo, no necesariamente poliédrico, que admita una base compacta, por lo cual asociamos a este problema tres duales distintos, en cada caso, caracterizamos la propiedad de dualidad fuerte en términos de la copositividad del Hessiano de la función objetivo, junto con algunas condiciones de optimalidad. En ese sentido, para el caso de dos diminsiones, caracterizamos cuando toda solución local es global.

# Abstract

The main objective of this doctoral thesis is to analyze and characterize the Lagrangian strong duality property for a non-convex optimization scalar problem subject to a single constraint, beyond those existing in the state of art. Our results are applied to the non-convex quadratic case. In particular, we obtained a relaxed version of Dine's theorem, associated to the quadratic minimization problem with an inequality, also the quadratic type together with several similarities of an affine type. In that sense, we discussed the case when the optimal value is not finite. Next, we established a characterization of the geometric type of strong duality property for the non convex quadratic problem, without the need to assume Slater's hypothesis, hence obtaining necessary and sufficient conditions of optimality. Finally, in light of the results on non-emptiness of the solution set, obtained by Frank Wolfe, our version sets asymptotically linear sets. In part two of the study, we established a topological and geometric characterization of the property of strong duality, for a general non-convex problem subject to a constraint of the equality type, together with constraints of the geometric type, from which the convexity of the conical envelope associated with the joint image determined by the functions of the original problem. As an application, we checked the validity of KKT conditions without assuming standard regularity condition. In the final part of this thesis, we study in detail the standard quadratic problem, by replacing the usual simplex with a convex cone, pointed, not necessarily polyhedral, that admits a compact base, for which we associate three different duals to this problem, in each case, we characterize the property of strong duality in terms of the Hessian's copositivity of the objective function, along with some conditions of optimality. In this sense, for the case of two dimensions, we characterized when every local solution is global.

# Contents

# List of Figures

# Chapter 1

# Introduction

## 1.1 Spanish version

En optimización, el concepto de dualidad en general, consiste en asignar a un problema de minimización[1], el primal, un problema de maximización, conocido como el dual, de modo tal que haya una manifiesta correspondencia entre los valores extremales, de manera directa o indirecta. Por ejemplo, el Teorema de dualidad débil, que asegura que el valor ínfimo del problema original, no puede ser menor al valor supremo del problema dual, tiene como consecuencia que el primal es infactible cuando el valor optimal del dual es $+\infty$, mientras que el dual es infactible cuando el valor optimal del primal es $-\infty$. El Teorema de Dualidad de *Fenchel* (ver por ejemplo Teorema 31.1 en [78]), ilustra también cómo se manifiesta el concepto de dualidad en optimización.

Debido al fuerte rol que cumple tanto en la formulación de condiciones de optimalidad, como también en optimización numérica, tiene que ver con la existencia de brechas entre los valores extremos del primal y el dual. A la diferencia entre los valores extremales del primal y el dual se le conoce como *gap* de dualidad. Dado que en general, dicha diferencia es no negativa, interesa determinar bajo qué condiciones se cumple la igualdad, o bien, bajo qué condiciones el primal satisface la propiedad de la brecha de dualidad cero (*zero duality gap*). Así, el gap de dualidad puede ser visto como la medida que cuantifica con qué precisión es posible aproximar el valor optimal del primal, por medio de estimadores inferiores obtenidos al resolver el problema dual. Asimismo, se dice que el problema satisface la propiedad de dualidad fuerte (*strong duality*), cuando su dual alcanza su valor óptimo, junto con satisfacer la propiedad de la brecha de dualidad cero.

De ahí que, al resolver un problema de optimización, se esta pensando en uno asociado que permita entre otras cosas, resolver el primal de una manera más sencilla, aprovechando las propiedades que el dual tiene, como por ejemplo, pasar de un problema con restricciones a uno sin restricciones. En general, la adecuada construcción del problema dual, tiene efectos en:

- dar condiciones que garanticen la existencia de soluciones optimales en el problema original;

[1] Cualquier problema en optimización, puede ser formulado como el de minimizar una función escalar (o bien, a valores en la recta real extendida) respecto a algún criterio determinado por el conjunto factible, cuyos elementos, las soluciones factibles adquieren la categoria de solución optimal, cuando la función objetivo alcanza su valor mínimo o valor optimal relativo al conjunto factible.

- la construcción a través del dual, del valor optimal del problema original;
- caracterizar el mejor estimador inferior del valor optimal.

Tales objetivos, típicamente son formulados bajo supuestos de convexidad, ingrediente clave tanto en el análisis como en la articulación de propiedades afínes en optimización, sólo por mencionar algunas: Condiciones de Optimalidad, Existencia y Unicidad de Soluciones Optimales, Convergencia de Algorítmos y Optimización Paramétrica.

En el caso convexo[2], la teoría de dualidad naturalmente tiene una interpretación desde la teoría de funciones conjugadas[3]. En un espacio vectorial topológico localmente convexo, los conjuntos cerrados y convexos son descritos como la intersección de todos los semiplanos cerrados que lo contienen. Así, cada conjunto cerrado y convexo tiene una reprensentación dual en su cono polar (no negativo), mientras que toda función semicontinua inferior y convexa, que cuenta con la propiedad de ser el supremo puntual entre todas las funciones afines que están por debajo de ésta, encuentra en la función polar o conjugada, su representante en el dual. De ahí que, lo fundamental en este caso subyace en los resultados de Fenchel-Moreau y el Teorema bipolar (ver por ejemplo Teoremas 2.3.3 y 1.1.9 en [94]), específicamente, en las hipótesis bajo las cuales una función coincide con su biconjugada, como también un conjunto coincide con su cono bipolar.

Como consecuencia, esta teoría puede ser adaptada a una amplia variedad de situaciones, generandose un vínculo fructífero con otras ramas de la optimización. Por ejemplo, la teoría de perturbación [28], subyacente a la teoría de dualidad, ha sido establecida en función de conectar propiedades entre dualidad y estabilidad en optimización [6], de modo que para el caso convexo, la validez de la brecha de dualidad cero, como también la propiedad de dualidad fuerte, dependen de la semicontinuidad inferior y la subdiferenciabilidad de la función valor en cero respectivamente. De esta manera, la relación entre dualidad y la cerradura del epígrafo de la función valor, permite visualizar la brecha de dualidad en términos de hiperplanos soportes del epígrafo de la función valor. El desarrollo de esta fructífera teoría de dualidad, se debe a las contribuciones de autores entre los que destacan W. Fenchel [29], J. J. Moreau y en especial a R. T. Rockafellar [78, 79].

En esta tesis doctoral, aplicamos el concepto de dualidad *Lagrangiana*. Mediante un proceso de escalarización, el Lagrangiano se construye por medio de una combinación lineal entre la función objetivo y las restricciones. Teniendo en cuenta esta estructura, la teoría de dualidad Lagrangiana, se basa en Teoremas del tipo alternativo, que junto a un adecuada condición de regularidad, son validados por medio de Teoremas de separación entre conjuntos convexos, que involucran variables auxiliares conocidas en la literatura como Multiplicadores de Lagrange.

En la práctica, interesa identificar aquellos casos que satisfacen la propiedad de dualidad fuerte, ya que permite dotar al problema de una estructura cuya complejidad computacional, se hace más tratable en el caso convexo. Por ejemplo, en modelos de programación matemática estándar[4], bajo la condición de factibilidad estricta o de *Slater*, la propiedad de dualidad fuerte se satisface cuando

[2]Se asume convexidad tanto para la función objetivo como también para el conjunto factible.

[3]La teoría funciones conjugada, fue desarrollada en sus inicios por Fenchel para luego ser generalizada por Rockafellar y Mourier.

[4]El conjunto factible es descrito a través de restricciones de desigualdad e igualdad. Se conoce como éstandar convexo, cuando las funciones que definen las desiguladades son del tipo convexo, mientras que las funciones respecto a las igualdades, son afínes.

el problema es del tipo convexo. De esta forma, hallar la solución del primal se reduce a resolver localmente un sistema de ecuaciones, cuyas incógnitas, corresponden a las variables del primal y el dual. Básicamente la idea detrás de métodos del tipo primal-dual [21]. Sin embargo, para puntos regulares, la condición de optimalidad de primer orden es en general solo una condición necesaria, salvo que se asuman ciertas hipótesis de convexidad, siendo inadecuado su uso como criterio de certificación de mínimo. Por el contrario, chequear de manera directa la condición de punto de silla, dificulta su uso como condición suficiente de optimalidad. Así, se presenta la dificultad de tratar con funciones no lineales, debido a que en general, bajo supuestos de convexidad es posible garantizar condiciones necesarias y suficientes, de optimalidad global, basadas en teoría de dualidad Lagrangiana.

Si la función objetivo o el conjunto factible vistos como objetos matemáticos no son convexos, el problema asociado es comúnmente identificado como del tipo *no convexo*, que a diferencia de su homólogo convexo, presenta serios inconvenientes en la construcción de la solución óptimal. Por un lado, resulta común detectar brechas de dualidad entre el problema (no convexo) y su dual. Mientras que, muchos problemas reales que son formulados como un modelo de optimización y que en la práctica resultan difíciles de resolver, tienen en común la falta de convexidad. Así, por ejemplo, ante una gran multiplicidad de mínimos locales, métodos de busqueda convencionales pierden eficacia localizando mínimos globales, limitando la busqueda a soluciones locales, o bien, aquellas realizables en la práctica para tamaños relativamente pequeños en las variables de entrada. Por tanto, de aquellos problemas difíciles de resolver, el énfasis esta en identificar condiciones de regularidad lo suficientemente mejoradas que permitan garantizar la existencia de estimadores inferiores para el valor óptimal. En ese sentido, nuestros resultados fueron aplicados al caso cuadrático[5] no convexo.

Pertenciente a la familia de programación matemática no lineal, cuenta con varias aplicaciones establecidas en términos de un problema cuadrático, que incluye al subproblema con una restricción del tipo cuadrática: métodos de región de confianza [42, 82], con aplicación en optimización robusta [64, 81] y problemas de mínimos cuadrados con restricción cuadrática [41], como también, los métodos de región de confianza generalizada [66, 74, 83]. La falta de convexidad en la forma cuadrática, se caracteriza por la presencia de valores propios negativos en la matriz Hessiana, convirtiendo al problema cuadrático asociado, en uno NP-completo[6], incluso, si las restricciones son del tipo afín [71, 68].

La relevancia que tiene estudiar la propiedad de dualidad fuerte en este tipo de problemas, apunta al desarrollo de métodos computacionales basados en condiciones de optimalidad, que permitan resolver un amplio espectro de problemas de optimización adaptados a un modelo cuadrático.

Por una parte, los algoritmos en optimización son iterativos, esto quiere decir que comenzando en un punto inicial, generan una sucesión de instancias, hasta que el algoritmo no pueda mejorar el valor de la función objetivo. La información de la función objetivo se utiliza para formar un modelo, de modo que cerca del punto actual, tenga en lo posible un comportamiento similar al de la función objetivo. Como el modelo puede no ser una buena aproximación de dicho objetivo, cuando se consideran puntos alejados del actual, se debe restringir la búsqueda de la solución a puntos pertencientes a una región

[5]En la versión general, las restricciones vienen representadas por medio de desigualdades del tipo cuadrático al igual que la función objetivo.

[6]NP denota la colección de todos los problemas de decisión los cuales tienen algoritmos de solución no-determinístico en tiempo real. Si además de ser NP, un problema tiene la característica de reducir en tiempo polinomial todo problema NP, se dice que es NP-completo.

de confianza, esto es, un entorno de este, típicamente representadas por esferas centradas en el punto actual, respecto de la norma usual. El modelo corresponde a uno del tipo cuadrático, cuando la función objetivo es la aproximación de Taylor de segundo orden.

En cuanto a la complejidad del problema, a veces es posible construir la solución del problema a través de un problema auxiliar o una *relajación* de este. Mediante técnicas de relajación semidefinida [20, 86], el problema cuadrático asociado, es relajado a un problema de programación semidefinida (SDP), cuyo objetivo es el de minimizar un funcional lineal, sobre el cono, convexo y cerrado, de todas las matrices simétricas y semidefinida positivas, que satisfacen un número finito de restricciones del tipo lineal, respecto al producto interno usual en el espacio de las matrices cuadradas, debido a esto, el problema (SDP) es del tipo convexo y puede ser resuelto en tiempo polinomial, mediante técnicas de punto interior [56, 93, 69]. En caso de que el problema (SDP) admita solución, por medio de un proceso de descomposición en matrices de rango uno [84], es posible reconstruir la solución óptimal del problema cuadrático propuesto, previo a esto, se debe garantizar que la relajación sea *exacta*, esto quiere decir, asegurar el emjor estimador inferior del valor óptimal del problema cuadrático, lo cual requieren de hipótesis adicionales, principalmente, condiciones que garanticen la validez de un sistema de inecuaciones cuadráticas, si y solamente si, es imposible resolver el sistema asociado de inecuaciones matriciales, obtenido en la etapa de relajación. Como el problema (SDP) se reformula de manera equivalente al dual del problema cuadrático original, entonces la validez de la propiedad de dualidad fuerte, garantiza que la relajación es exacta. Sin embargo, la falta de convexidad en general, restringe el uso de Teoremas alternativos. Sin embargo, para el caso de una restricción del tipo desigualdad, la validez del Teorema alternativo o más conocido como el *S-Lemma*[7], depende de la validez en la condición de Slater [88, 89], mientra que, en [87] dieron condiciones para garantizar la validez del S-Lemma, versión restricción de igualdad.

En este trabajo, apuntamos a debilitar las hipótesis que permiten garantizar la validez de la propiedad de dualidad fuerte, organizada como sigue. En el capítulo 2, introducimos las definiciones y notaciones que utilizaremos en los posteriores capítulos.

El capítulo 3 está estructurado de la siguiente manera. La sección 3.2.1 proporciona la formulación del problema que vamos a discutir, junto con la caracterización del Teorema de separación, entre un conjunto convexo y un cono abierto en términos de la convexidad de la envoltura cónica de conjuntos. También, incluye una versión relajada del Teorema de Dines, cuando el valor óptimo es $-\infty$. La sección principal 3.3, presenta una versión relajada del Teorema de Dines, cuando el valor óptimo es finito, junto con la caracterización geométrica de la propiedad de dualidad fuerte, para el problema de minimización sujeto a restricciones del tipo afín y una restricción de desigualdad del tipo cuadrático, sin asumir hipótesis de convexidad o supuestos de Slater. Esto permite obtener condiciones necesarias y suficientes de optimalidad, sin necesidad de asumir la condición de Slater. Relaciones con las condiciones empleadas en el Teorema de Finsler también son establecidas. La sección 3.4 presenta un refinamiento al Teorema de Frank y Wolfe, para conjuntos asintóticamente lineales. Los resultados contenidos en este capítulo fueron publicados en el artículo:

- Flores-Bazan, F.; Cárcamo, G., A geometric characterization of strong duality in nonconvex quadratic programming with linear and noncovex quadratic contraints, *Math. Programming, Ser. A* **145** (2014), 263–290.

[7] En [90], se discuten distintas versiones del S-Lemma, inicialmente propuesto en [30].

El capítulo 4 se divide en varias secciones. En la Sección 3, sin imponer ningún supuesto de diferenciabilidad establecemos primero, una completa descripción de la convexidad del cono de la imagen conjunta determinada por las funciones asociadas al problema (4.1). Posteriormente, establecemos algunas caracterizaciones topológicas, o de naturaleza geométrica de la propiedad de dualidad fuerte para el problema (4.1). En particular, bajo una condición del tipo Slater, probamos que una condición necesaria y suficiente, para obtener dualidad fuerte, es la convexidad del cono de la imagen. En la sección 4.4, bajo supuestos de diferenciabilidad, establecemos varias caracterizaciones de la validez de las condiciones de optimalidad KKT, aplicando el resultado principal de la Sección 4.3. Finalmente, en la sección 4.5 describe una aplicación concreta de nuestros resultados previos, a una generalización del problema de programación cuadrática estándar, donde el octante positivo es sustituido por un cono puntiagudo, cerrado y convexo. Los resultados contenidos en este capítulo fueron publicados en el artículo:

- Cárcamo, G.; Flores-Bazan, F., Strong duality and KKT conditions in nonconvex optimization with a single equality constraint and geometric constraint, *Math. Programming, Ser. B* **168** (2018), 369–400.

El capítulo 5 esta organizado como sigue. A la luz de [34], revisamos en la Sección 5.2, el esquema de dualidad Lagrangiana para el problema general con una restricción del tipo igualdad, en donde establecemos nuevas condiciones secuenciales, de la propiedad de la brecha de dualidad cero. Nuestros principales resultados, conectados con el problema (5.5), se presentan en la Sección 5.3. El interés esta en la caracterización de la propiedad de dualidad fuerte, respecto a distintos problemas duales, en términos de la copositividad del Hessiano de la función objetivo, revelando una convexidad oculta. En la sección 5.4, se analiza en detalle algunas condiciones de optimalidad, para el problema (5.5), mientras que, en la Sección 5.5 analizamos el caso $n = 2$. En particular, caracterizamos la copositividad del Hessiano de la función objetivo, y cuando toda solución local es global. Los resultados contenidos en este capítulo fueron publicados en el artículo:

- Flores-Bazan, F.; Cárcamo, G.; Caro, S., Extensions of the Standard Quadratic Optimization Problem: Strong Duality, Optimality, Hidden Convexity and S-Lemma, *Appl. Math. Optim.* (2018) https://doi.org/10.1007/s00245-018-9502-0.

## 1.2 English version

In optimization, the concept of duality consists of assigning to a minimization problem[8], the primal, a maximization problem, known as the dual, in such a way that there is a clear correspondence between their optimal values. For example, the weak duality theorem, which ensures that the minimum value of the original problem cannot be less than the maximum value of the dual problem, has as a consequence that the primal is infeasible when the optimal value of the dual is $+\infty$, while dual is infeasible when optimal value of the primal is $-\infty$. The Fenchel Duality Theorem (see example Theorem 31.1 in [78]) also illustrates how this symmetry is shown in optimization.

Due to the strong role fulfilled by both in the formulation of optimality conditions, as well as in numerical optimization, one has to do with the existence of gaps between the extreme values of primal and dual. The difference between the extreme values of primal and dual problems is known as the duality gap. Since, in general, this difference is non-negative, it is important to determine the conditions under which the equality is met, or under what conditions the primal satisfies the property of the *zero-duality gap*. Thus, the duality gap can be seen as the measure that quantifies with what precision it is possible to approximate the optimal value of the primal, by means of inferior estimators obtained by solving the dual problem. Likewise, it is said that the problem satisfies the property of *strong duality*, when its dual reaches its optimal value, along with satisfying the property of the zero duality gap.

Once the problem structure is known, studying its relationship with its dual facilitates a greater understanding of the nature of the objects involved, which has advantages from the theoretical point of view, as well as in the determination of numerical solutions.

Hence, when solving an optimization problem, is one that allows, among other things, to solve the primal in a simpler way, taking advantage of the properties that a dual has, such as, to pass from a restrictive problem to one without restrictions. Furthermore, the adequate construction of the dual problem has effects on:

- provide conditions that guarantee the existence of optimal solutions to the original problem;
- the construction through the dual, the optimal value of the original problem;
- characterize the best inferior estimator of the optimal value.

These objectives are typically formulated under certain convexity assumptions, a key ingredient in both analysis and the articulation of related properties in Optimization such as: optimality conditions, existence and uniqueness of solutions, convergence of algorithms and parametric optimization.

In the convex[9] case, the idea is to think of the convex sets in terms of hyperplanes supported, naturally, by the theory of conjugated functions[10]. In a locally convex topological vector space, the

[8] Any problem in optimization, can be formulated as to minimize a scalar function (or, an extended real valued function) with respect to some criterion determined by the feasible set, whose elements, the feasible solutions acquire the category of optimal solution, when the objective function reaches its minimum value or optimal value relative to the feasible set.

[9] Convexity is assumed for both the objective function and the feasible set.

[10] The conjugated functions theory, was developed in its beginnings by Fenchel to later be generalized by Rockafellar and Moreau.

closed and convex sets are described as the intersection of all closed half space that contain it. Thus, each closed and convex set has a dual representation in its positive polar cone, while every convex lower semicontinuous proper convex function, which has the property of being the supreme point among all affine functions bellow it, it finds in the polar or conjugated function, its representative in the dual. Hence, fundamentals lie in the results of Fenchel-Moreau and the bipolar Theorem (see for example Theorems 2.3.3 and 1.1.9 in [94]), Specifically, in the hypotheses under which a function coincides with its biconjugate as also in a set coincides with its bipolar cone.

As a consequence, this theory can be adapted to a wide variety of situations, generating a fruitful link with other branches of optimization. For example, the perturbational theory [28], Underlying the theory of duality, it has been established as a function of connecting properties between duality and stability in optimization [6], So that for the convex case, the validity of the zero duality gap, as well as the property of strong duality, depend on the lower-semicontinuity and the sub-differentiability of the value function at zero, respectively. In this way, the relationship between duality and the closedness of epigraph of value function allows to visualize the duality gap in terms of hyperplanes that support the epigraph of the value function. In addition, it is worth mentioning that the development of this fruitful theory of duality in convex optimization is due to the contributions of authors, among which W. Fenchel [29], J.J Moreau and especially R.T. Rockafellar [78, 79].

In this doctoral thesis, we apply the concept of *Lagrangian* duality. Through a process of scalarization, the Lagrangian is constructed by means of a linear combination between the objective function and the constraints. Taking this structure into account, the Lagrangian duality theory is based on theorems of the alternative type, which together with an adequate regularity condition, are validated by means of separation theorems between convex sets, which involve auxiliary variables known in Literature as Lagrange Multipliers.

In practice, it is interesting to identify those cases that satisfy the property of strong duality, since it allows the problem to be provided with a structure whose computational complexity becomes more treatable in the convex case. For example, in standard mathematical programming models[11], under the condition of strict feasibility or *Slater*, the strong duality property is satisfied when the problem is of the convex type. In this way, finding the solution of the primal is reduced to solving locally a system of equations, whose unknowns correspond to the variables of primal and dual. Basically the idea behind methods of the primal-dual type [21]. However, for regular points, the first order optimality condition is in general only a necessary condition, unless certain convexity hypothesis are assumed, doing inadequate its use as a certification of the optimal solution.

On the contrary, directly checking the condition of saddle point, makes difficult its use as a sufficient condition of optimality. Thus, the difficulty of dealing with non-linear functions is presented, because in general, under convexity assumptions it is possible to guarantee necessary and sufficient conditions of global optimality, based on Lagrangian duality theory.

If the objective function or the feasible set seen as mathematical objects are not convex, the associated problem is commonly identified as non-convex type, which, unlike its convex counterpart, presents serious drawbacks in the construction of the optimal solution. On the one hand, it is common to detect duality gaps between the problem (non-convex) and its dual. While, many real problems that are formulated as an optimization model and that in practice are difficult to solve, they have in

[11] The feasible set it is discribed by linear and non-linear constraint.

common the lack of convexity. Thus, for example, in the face of a large multiplicity of local minimums, conventional search methods lose efficiency by locating global minimums, limiting the search to local solutions, or those that are feasible in practice for relatively small input sizes. Therefore, of those problems that are difficult to solve, the emphasis is on identifying conditions of regularity that are sufficiently improved to guarantee the existence of inferior estimators for the optimal value. In that sense, our results were applied to the non-convex quadratic case[12]. Pertinent to the family of nonlinear mathematical programming, it has several applications established in terms of a quadratic problem, which includes the subproblem with a restriction of the quadratic type: methods of trusting region [42, 82], with application in robust optimization [64, 81] and least squares problems with quadratic restriction [41], as well as, generalized confidence region methods [66, 74, 83]. The lack of convexity in the quadratic form is characterized by the presence of negative eigenvalues in the Hessian matrix, converting the associated quadratic problem into an NP-complete[13], even if the constraints are affine [71, 68].

The relevance of studying the property of strong duality in this type of problems, points to the development of computational methods based on optimality conditions, which allow solving a wide spectrum of optimization problems adapted to a quadratic model. On the one hand, the algorithms in optimization are iterative, this means that starting at an initial point, they generate a succession of instances, until the algorithm can not improve the value of the objective function. The information of the objective function is used to form a model, so that close to the current point, it has, as far as possible, a behavior similar to that of the objective function. Since the model may not be a good approximation of this objective, when considering points away from the current one, the search of the solution must be restricted to points pertaining to a region of trust, that is, an environment of this, typically represented by spheres centered on the current point, with respect to the usual norm. The model corresponds to one of the quadratic type, when the objective function is the second order Taylor approximation.

Regarding the complexity of the problem, it is sometimes possible to construct a solution through an auxiliary problem or a relaxation of it. By means of semi-definite relaxation techniques [20, 86], the associated quadratic problem is relaxed to a semidefinite programming problem (SDP), whose objective is to minimize a linear function, on the closed convex cone, of all the symmetric and semi-definite positive matrices, which satisfy a finite number of constraints of the linear type, with respect to the usual internal product in the space of the square matrices, due to this, the problem (SDP) is of the convex type and can be solved in polynomial time, by means of internal point techniques [56, 93, 69]. In case the problem (SDP) admits solution, by means of a decomposition process in rank one matrices [84], it is possible to reconstruct the optimal solution of the proposed quadratic problem, prior to this, it must be guaranteed that the relaxation is exact, that is to say, assure the best inferior estimator of the optimal value of the quadratic problem, which requires additional hypothesis, mainly, conditions that guarantee the validity of a system of quadratic inequations, if and only if, it is impossible to solve the associated system of matrix inequalities, obtained in the relaxation stage. Since the problem

[12] In the general version, constraints are represented by inequalities of the quadratic type as well as the objective function.

[13] NP denotes the collection of all decision problems which have algorithms of non-deterministic solution in real time. If in addition to being NP, a problem has the characteristic of reducing polynomial time to all NP problems, it is said to be NP-complete.

(SDP) is reformulated in an equivalent way to the dual of the original quadratic problem, then the validity of the property of strong duality, guarantees that the relaxation is exact. However, the lack of convexity in general, restricts the use of alternative theorems, and there are only few general results. In the case of a single inequality constraint, the validity of the alternative Theorem or more known as the S-Lemma[14], depends on the validity of the Slater condition [88, 89], while in [87] gave conditions to guarantee the validity of S-Lemma for the equality constraint version.

In this work, we aim to weaken the hypothesis that allows us to guarantee the validity of the property of strong duality, organized as follows. In chapter 2, we introduce the definitions and notations that we will use in the later chapters.

Chapter 3 is structured as follows. Section 3.2.1 provides the formulation of the problem that we are going to discuss, together with the characterization of the separation theorem, between a convex set and an open cone in terms of the convexity of the complex envelope of sets. Also, it includes a relaxed version of Dines' Theorem, when the optimal value is $-\infty$. The main section **??** presents a relaxed version of Dines' Theorem, when the optimum value is finite, together with the geometric characterization of the strong duality property, for the minimization problem subject to restrictions with finitely many linear equality and a single quadratic inequality constraint, without assuming hypothesis of convexity or Slater assumptions. This allows obtaining necessary and sufficient conditions of optimality, without having to assume the Slater condition. Relations with the conditions used in Finsler's Theorem are also established. Section 3.4 presents a refinement to Frank and Wolfe's Theorem, for asymptotically linear sets. The results contained in this chapter were published in the article:

- Flores-Bazan, F.; Cárcamo, G., A geometric characterization of strong duality in nonconvex quadratic programming with linear and noncovex quadratic contraints, Math. Programming, Ser. A 145 (2014), 263-290.

Chapter 4 is divided into several sections. In Section 4.1, without imposing any assumption of differentiability, we first establish a complete description of the convexity of the conic hull of the joint-range of the pair of functions associated to problem (4.1). Subsequently, we establish some topological characterizations, or geometric nature of the strong duality property for the problem (4.1). In particular, under a Slater type condition, we prove that a necessary and sufficient condition, in order to obtain strong duality, is the convexity of the cone of the image. In section 4.4, under assumptions of differentiability, we establish several characterizations of the validity of the KKT optimality conditions, applying the main result of Section 4.3. Finally, in section 4.5 we describe a concrete application of our previous results, to a generalization of the standard quadratic programming problem, where the positive orthant is replaced by a pointed, closed and convex cone. The results contained in this chapter were published in the article:

- Cárcamo, G.; Flores-Bazan, F., Strong duality and KKT conditions in nonconvex optimization with a single equality constraint and geometric constraint, Math. Programming, Ser. B 168 (2018), 369-400.

[14] Various versions of the S-Lemma, originally proposed in [30], are discussed in [90].

Chapter 5 is organized as follows. In light of [34], we review in Section 5.2, the Lagrangian duality scheme for the general problem with a constraint of the equality type, where we establish new sequential conditions, of the property of the zero duality gap. Our main results, connected with problem (5.5), are presented in Section 5.3. The interest is in the characterization of the property of strong duality, with respect to different dual problems, in terms of the Hessian's copositiveness of the objective function, revealing a hidden convexity. In section 5.4, some optimality conditions are analyzed in detail for problem (5.5), while in Section 5.5 we analyze the case $n = 2$. In particular, we characterize the Hessian's copositivity of the objective function, and when every local solution is global. The results contained in this chapter were published in the article:

- Flores-Bazan, F.; Cárcamo, G.; Caro, S., Extensions of the Standard Quadratic Optimization Problem: Strong Duality, Optimality, Hidden Convexity and S-Lemma, Appl. Math. Optim. (2018) https://doi.org/10.1007/s00245-018-9502-0.

# Chapter 2

# Notations and Preliminaries

In this chapter we present the basic notations and the terminology that will be used throughout the thesis. This is a concise review, we only discuss the topics at the level necessary to follow the rest of the thesis.

## 2.1 Convex sets and cones

All the results in this section are presented in the context of the Euclidian vector space, although these results can mostly be generalized to infinite dimensional spaces like Banach or Hausdorff spaces.

By $\langle \cdot, \cdot \rangle$ we denote the inner or scalar product in $\mathbb{R}^n$ whose elements are considered column vectors. Thus, $\langle x, y \rangle = x^\top y$ for all $x,\ y \in \mathbb{R}^n$, where $x^\top$ means the transpose of the vector $x$. Given any set $M$ in $\mathbb{R}^n$, its closure, topological interior, boundary are denoted, respectively, by $\overline{M}$, int $M$, bd $M$. Similarly, ri $M$ denotes the relative interior of $M$, which is the interior with respect to the affine hull of $M$, denoted by aff $M$.

We recall that the set $M$ is convex if for every $x,\ y \in M$ and $\lambda \in [0, 1]$,

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

The convex hull set of $M$, denote by co $M$, is the smallest convex set containing $M$, i.e.,

$$\text{co } M \doteq \{x \in \mathbb{R}^n : \ x = \sum_{i=1}^{l} \lambda_i x_i,\ \sum_{i=1}^{l} \lambda_i = 1,\ 0 \le \lambda_i \le 1,\ x_i \in M\}.$$

An important aspect of convexity in duality theory is related to separation and support of sets, since they can be separated by an hyperplane. A hyperplane $H \subseteq \mathbb{R}^n$ is a set of the form

$$H(a, \alpha) \doteq \{y \in \mathbb{R}^n : \ \langle a, y \rangle = \alpha\} = H^- \cap H^+ \tag{2.1}$$

where $a \in \mathbb{R}^n \setminus \{0\}$, $\alpha \in \mathbb{R}$, whereas $H(a,\alpha)^- \doteq \{y \in \mathbb{R}^n : \langle a, y\rangle \leq \alpha\}$ and $H(a,\alpha)^+ \doteq \{y \in \mathbb{R}^n : \langle a, y\rangle \geq \alpha\}$ are closed half spaces in $\mathbb{R}^n$. Thus a hyperplane $H(a,\alpha)$ is said to separate two nonempty sets $M_1$ and $M_2$ if

$$\langle a, x\rangle \leq \alpha \leq \langle a, y\rangle, \ \forall\ x \in M_1, \ \forall\ y \in M_2. \tag{2.2}$$

The separation is said to be strict if

$$\langle a, x\rangle < \alpha < \langle a, y\rangle, \ \forall\ x \in M_1, \ \forall\ y \in M_2. \tag{2.3}$$

Whereas $H(a,\alpha)$ is said to support $M$ at $\bar{x} \in \text{bd } M$ if either

$$\langle a, y - \bar{x}\rangle \leq 0, \ \forall\ y \in M, \tag{2.4}$$

or else,

$$\langle a, y - \bar{x}\rangle \geq 0, \ \forall\ y \in M.$$

A separation theorem deals with the existence of an hyperplane that separates two given sets. The following theorem with the case when separates a point from a closed convex set which does not contain the point.

**Proposition 2.1.1.** ([*79, Theorem* 11.4]) *Let $M \subseteq \mathbb{R}^n$ be a closed convex cone such that $\bar{x} \notin M$. Then there exists $a \in \mathbb{R}^n \setminus \{0\}$ such that*

$$\langle a, \bar{x}\rangle < 0 \leq \langle a, x\rangle, \ \forall\ x \in M$$

Cones play a key role in establishing the characterization of strong duality. A cone is a set in $\mathbb{R}^n$ which is closed under nonnegative scalar multiplication. The smallest cone containing $M$ is denoted by

$$\text{cone}(M) \doteq \bigcup_{t \geq 0} tM.$$

Whereas $\overline{\text{cone}}(M)$ denotes the smallest closed cone containing $M$, i.e., $\overline{\text{cone}}(M) = \overline{\text{cone}(M)}$. Moreover $\text{cone}_+(M) \doteq \bigcup_{t>0} tM$.

Also we say that a (not necessarily convex) cone $M$ in $\mathbb{R}^n$, is pointed if

$$\text{co } M \cap (-\text{co } M) = \{0\}.$$

For boundary structure of a convex set the following notion is important.

We say the point $\bar{x} \in \text{bd } M$ is an extreme point of $M$, if it is not an interior point of any line segment from $M$. More precisely, if $\bar{x} = \lambda x + (1-\lambda) y$ for $x, y \in M$, and $\lambda \in (0,1)$, we have $\bar{x} = x = y$. If additionally $M$ is a cone, then $\bar{d} \in \text{bd } M \setminus \{0\}$ is an extreme direction of $M$, if it is not a positive combination of other directions $d_1$ and $d_2$ in $M$; that is, if $\bar{d} = \lambda_1 d_1 + \lambda_2 d_2$ for $\lambda_1,\ \lambda_2 > 0$, then $d_1,\ d_2 \in \mathbb{R}_+\bar{d}$.

The positive polar cone of $M$ is defined by

$$M^* \doteq \{y^* \in \mathbb{R}^n :\ \langle y^*, y\rangle \geq 0,\ \forall\, y \in M\}.$$

We denote by $M^{**} \doteq (M^*)^*$, the bipolar cone of $M$. Clearly if $M$ is a linear subspace then $M^* = M^\perp$, where $M^\perp$ stands for the orthogonal subspace of $M$.

We immediately obtains some properties from its definition.

**Proposition 2.1.2.** ([9, *Theorem* 1.3]) *Let $M$, $N$ be any sets in $\mathbb{R}^n$. Then*

(a) *$M^*$ is a closed convex set;*

(b) *$M \subset N$ then $N^* \subset M^*$;*

(c) *$M \subset M^{**}$;*

(d) *$M^* = M^{***}$;*

(e) *$M^* = (\overline{M})^* = (\text{co } M)^* = (\text{cone } M)^*$;*

(f) *$M^* \cap N^* \subset (M+N)^*$ and if $0 \in M \cap N$ then $(M+N)^* \subset M^* \cap N^*$;*

(g) *$M^* + N^* \subset (M \cap N)^*$.*

The well-known bipolar theorem is derived from Proposition 2.1.2 and Theorem 2.1.1.

**Proposition 2.1.3.** ([9, *Theorem* 1.5]) *Let $M \subseteq \mathbb{R}^n$. Then $M$ is a closed convex cone if and only if $M = M^{**}$., that is,*

$$x \in M \Longleftrightarrow \langle x^*, x\rangle \geq 0,\ \forall\ x^* \in M^*. \tag{2.5}$$

*Moreover, if* $\text{int } M \neq \emptyset$ *then*

$$x \in \text{int } M \Longleftrightarrow \langle x^*, x\rangle > 0,\ \forall\ x^* \in M^* \setminus \{0\}. \tag{2.6}$$

We also need to introduce the following cones.

Assume that $M \subseteq \mathbb{R}^n$ is a nonempty set. Given $\bar{x} \in \overline{M}$, we say that the contingent cone of $M$ (or tangent cone of Bouligand) at $\bar{x}$ is defined by

$$T(M;\bar{x}) \doteq \{v \in \mathbb{R}^n :\ \exists\ (t_k, x_k) \in \mathbb{R}_{++} \times M,\ x_k \longrightarrow \bar{x},\ t_k(x_k - \bar{x}) \longrightarrow v\},$$

which is always a closed cone. Moreover from its definition, if $M$ is convex then

$$T(M;\bar{x}) = \overline{\text{cone}}(M - \bar{x}). \tag{2.7}$$

Whereas, if $M$ is convex, $N_M(\bar{x})$ stands for the normal cone to $M$ at $\bar{x}$, is the set $N_M(\bar{x}) \doteq \{x^* : \langle x^*, x - \bar{x}\rangle \leq 0, \ \forall\, x \in M\}$. Thus

$$[T(M;\bar{x})]^* = -N_M(\bar{x}),$$

whenever $\bar{x} \in M$.

The asymptotic cone of $M$ is defined by

$$M^\infty \doteq \{v \in \mathbb{R}^n : \ \exists\ (t_k, x_k) \in \mathbb{R}_+ \times M, \ t_k \downarrow 0, \ t_k x_k \longrightarrow v\}.$$

It is not difficult to see that $M^\infty$ is always a closed cone. Additionally if $M$ is a closed convex set, then for each $\bar{x} \in M$

$$M^\infty = \{v \in \mathbb{R}^n : \ \bar{x} + tv \in K, \ \forall\ t > 0\}.$$

**Lemma 2.1.1.** ([79, *Corollary* 9.1.2.]) *Let $M_1$ and $M_2$ be nonempty closed convex sets in $\mathbb{R}^n$. Assume that $M_1^\infty \cap (-M_2^\infty) = \{0\}$. Then*

(*a*) $\overline{M_1 + M_2} = M_1 + M_2$;

(*b*) $[M_1 + M_2]^\infty = M_1^\infty + M_2^\infty$.

## 2.2 Convex and polar functions

We start in this section with the concept of convexity to functions taking values on the extended real line. Let $h : \mathbb{R}^n \longrightarrow \mathbb{R} \cup \{\pm\infty\}$ be an extended real valued function. The epigraph of $h$ is defined by

$$\text{epi } h \doteq \{(x, \alpha) \in \mathbb{R}^n \times \mathbb{R} : \ h(x) \leq \alpha\}.$$

Whereas, the effective domain of $h$ is the projection on $\mathbb{R}^n$ of the epigraph of $h$, i.e., a set of the form

$$\text{dom } h = \{x \in \mathbb{R}^n : \ h(x) < +\infty\}.$$

A function $h$ is said to be convex if its epigraph epi $h$ is a convex set of $\mathbb{R}^{n+1}$. If, furthermore, dom $h \neq \emptyset$ and $f(x) > -\infty$ for all $x \in \mathbb{R}^n$, then $h$ is a proper convex function. In case we know additionally that $h$ is proper we obtain that the function $h$ is convex, if and only if dom $h$ is convex and for all $x, \ y \in M$ and $\lambda \in [0, 1]$ we have

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

There are special functions, for instance the indicator function $\delta_M : \mathbb{R}^n \longrightarrow \{0, +\infty\}$ of $M$ that is defined by

$$\delta_M(x) \doteq \begin{cases} 0, & if\ x \in M \\ +\infty, & if\ x \notin M \end{cases} \tag{2.8}$$

which is convex if, and only if $M$ is convex.

A quadratic function defined by

$$h(x) \doteq \frac{1}{2} x^\top A x + a^\top x + \alpha, \tag{2.9}$$

where $A$ is a symmetric $n \times n$ matrix, $a \in \mathbb{R}^n$ and $\alpha \in \mathbb{R}$, is convex (see Theorem 4.5 in [79]), if, and only if the matrix $A$ is positive semi-definite, i.e., $x^\top A x \geq 0$ for all $x \in \mathbb{R}^n$. The space of all $(n \times n)$ symmetric matrices is denoted by $S^n$.

In addition, we need the following concepts of envelope representation of functions. We recall (see for instance section 7 in [79]) that a function $h : \mathbb{R}^n \longrightarrow \mathbb{R} \cup \{\pm\infty\}$ is called lower semicontinuous (lsc) at $\bar{x}$, if

$$h(\bar{x}) \leq \liminf_{k \to +\infty} h(x_k), \text{ for every sequence } \{x_k\}_{k \in \mathbb{N}} \text{ converging to } \bar{x}.$$

Whereas we said that $h$ is (lsc) if it is (lsc) at every point $x \in \mathbb{R}^n$. Just as the convexity of the epigraph characterizes the convexity of a function $h$, its closedness characterizes the lower semicontinuity of $h$, as shown by the next proposition.

**Proposition 2.2.1.** ([*79, Theorem* 7.1]) *Let* $h : \mathbb{R}^n \longrightarrow \mathbb{R} \cup \{\pm\infty\}$ *be an extended real valued function. The following assertions are equivalent:*

(*a*) $h$ *is* (lsc)*;*

(*b*) *the epigraph* epi $h$ *is a closed set in* $\mathbb{R}^{n+1}$*;*

(*c*) *the level set* $S_h(\lambda) \doteq \{x : \ h(x) \leq \lambda\}$ *is closed for every* $\lambda \in \mathbb{R}$.

Given a function $h : \mathbb{R}^n \longrightarrow \mathbb{R} \cup \{\pm\infty\}$, we stand by $\bar{h}$ for the greatest lower semicontinuous function not larger than $h$ (lsc hull $\bar{h}$) is given by

$$\bar{h}(x) \doteq \sup_{l \in \mathcal{S}_h} l(x), \tag{2.10}$$

where $\mathcal{S}_h \doteq \{l : \mathbb{R}^n \longrightarrow \mathbb{R} \cup \{\pm\infty\} : \ l \text{ is l.s.c and } l \leq h\}$. It is easy to check that since $-\infty \in \mathcal{S}_h$, then $\mathcal{S}_h$ is a nonempty set. Moreover, $\bar{h} \leq h$ and since $\bar{h}$ is the pointwise supremum of the familiy of functions indexed by $l \in \mathcal{S}_h$, the epigraph of $h$ is the intersection of the sets epi $l$. By virtue of Proposition 2.2.1, $\bar{h} \in \mathcal{S}_h$. The following result provides and interpretation of the l.s.c. hull of a function. See [79, 39].

**Proposition 2.2.2.** *Let* $h$ *be as above. Then*

(a)

$$\bar{h}(x) = \begin{cases} \inf\{\alpha : \ (x,\alpha) \in \overline{\text{epi } h}\}, & if \ [\{x\} \times \mathbb{R} \cap \overline{\text{epi } h}] \neq \emptyset \\ +\infty, & if \ [\{x\} \times \mathbb{R} \cap \overline{\text{epi } h}] = \emptyset \end{cases}$$

(b)

$$\text{epi } \bar{h} = \overline{\text{epi } h}. \tag{2.11}$$

The greatest convex and lower semicontinuous function not longer than $h$ given by $\overline{\text{co}}\ h : \mathbb{R}^n \longrightarrow \mathbb{R} \cup \{\pm\infty\}$ is defined as the function for which

$$\text{epi } \overline{\text{co}}\ h = \overline{\text{co}}(\text{epi } h). \tag{2.12}$$

The conjugate (or polar) of $h$ is defined by

$$h^*(y) \doteq \sup_{x \in \text{dom } h} \{\langle y, x\rangle - h(x)\}. \tag{2.13}$$

For instance the conjugate of indicator function defined by (2.8) is

$$\delta_M^*(y) = \sup_{x \in M} \langle y, x\rangle.$$

In the case that $M$ be a cone then one gets from above that $\delta_M^*(y) = 0$ if $y \in -M^*$, $\delta^*(y) = +\infty$ in other case, which implies that $\delta_M^*$ is in fact the indicator function of $-M^*$: $\delta_{(-M)^*}$, i.e., $(\delta_M)^* = \delta_{(-M)^*}$

The biconjugate (or bipolar) of $h$ is the conjugate of $h^*$, that is

$$h^{**}(y) = \sup_{x \in \text{dom } h^*} \{\langle y, x\rangle - h^*(x)\}.$$

Moreover,

$$\overline{\text{co}}\ h(y) > -\infty \ \ \forall\, y \in \mathbb{R}^m \Longrightarrow \overline{\text{co}}\ h(y) = h^{**}(y) \ \ \forall\, y \in \mathbb{R}^m, \tag{2.14}$$

# Chapter 3

# Nonconvex quadratic and constraints

## 3.1 Introduction

Given a subset $C$ of a finite dimensional space $\mathbb{R}^n$, and functions $f : \mathbb{R}^n \to \mathbb{R}$, $g : \mathbb{R}^n \to \mathbb{R}$, let us consider the following minimization problem:

$$\mu \doteq \inf\{f(x) : \; g(x) \leq 0, \; x \in C\}. \tag{3.1}$$

The Lagrangian dual problem associated to (3.1) is

$$\nu \doteq \sup_{\lambda \geq 0} \inf_{x \in C} [f(x) + \lambda g(x)]. \tag{3.2}$$

In general $\nu \leq \mu$. We say Problem (3.1) has a (Lagrangian) *zero duality gap* if the optimal values of (3.1) and (3.2) coincide, that is, $\mu = \nu$. Problem (3.1) is said to have *strong duality* if it has a zero duality gap and Problem (3.2) admits a solution.

Quadratic functions have proved to be very important in applications (e.g. in telecommunications, robust control [64, 81], trust region problems [42, 82] among others) and enjoy very nice properties. After the result ($C = \mathbb{R}^n$) due to Gay [42] and Sorensen [82] concerning a characterization of solutions for a special quadratic optimization problem without any convexity assumptions, several authors extended such a result for general quadratic optimization with a single inequality constraint. In particular, we mention the work by Moré [66] who considered the general case of a single equality constraint and then used it to cover the single inequality constraint under the standard Slater condition. Moré actually provided necessary and sufficient optimality conditions for a point to be optimal under no convexity conditions. Certainly, this may be seen as a strong duality-type result.

More recently, when $C = \mathbb{R}^n$ with $g$ being a quadratic function that is not identically zero, the authors in [49] prove that, (3.1) has strong duality for each quadratic function $f$ if, and only if there exists $\bar{x} \in \mathbb{R}^n$ such that $g(\bar{x}) < 0$, that is, the standard Slater condition holds. Unlike this result and many others established in [48, 50, 52, 51, 53], our approach allows us to derive conditions on the pair, $f$ and $g$ jointly, that ensure that (3.1) has strong duality without satisfying the Slater condition, and under no convexity assumptions on $f$ or $g$. This is carried out by further developing the geometric approach introduced in [31], where strong duality is characterized under a single inequality constraint for any (not necessarily quadratic) functions $f$ and $g$. We actually characterize completely the strong duality in the presence of finitely many linear equality and a single quadratic inequality constraints without convexity assumptions or Slater condition (Theorem 3.3.3), and derive necessary and sufficient optimality conditions.

Among the main results showing some of the nice properties of quadratic functions we mention two of them. The first one is due to Dines [24] (see also [76]) and it ensures the convexity of the set $\{(f(x), g(x)) \in \mathbb{R}^2 : x \in \mathbb{R}^n\}$ for any homogeneous quadratic functions $f$ and $g$. For general quadratic non homogeneous functions we provide a relaxed version of this result, see Theorem 3.3.1 when $\mu$ is finite, and when $\mu = -\infty$ it is provided a condition under which a Dines-type result holds. A complete characterization of convexity of the above set for any quadratic functions $f$ and $g$ is given in [36]. A second result showing another nice property of these functions is that due to Frank and Wolfe [37], which asserts that any quadratic function bounded from below on a nonempty (possibly unbounded) polyhedral set attains its infimum value. We establish several equivalences (including that due to Frank and Wolfe) for a larger family of sets than polyhedral, whose proof uses elementary analysis and it is related to that by Blum and Oettli [12], being suitable for expository purposes; whereas the original proof of Frank and Wolfe requires a decomposition theorem for convex polyhedra.

More precisely, in the present paper, we deal with the case where $f$ and $g$ are quadratic functions and $C = H^{-1}(d) = \{x \in \mathbb{R}^n : \ Hx = d\}$ where $H$ is a real matrix of order $m \times n$ and $d \in \mathbb{R}^m$, and the regularized Lagrangian dual problem is considered. It means that instead of considering the standard Lagrangian dual problem

$$\sup_{\lambda \geq 0, \gamma \in \mathbb{R}^m} \inf_{x \in \mathbb{R}^n} [f(x) + \lambda g(x) + \gamma(Hx - d)], \tag{3.3}$$

we choose the regularized Lagrangian dual problem

$$\sup_{\lambda \geq 0} \inf_{x \in H^{-1}(d)} [f(x) + \lambda g(x)], \tag{3.4}$$

which is more suitable for our purpose since there are instances, specially in trust-region problems, showing a non zero duality gap between (3.1) and (3.3) against with the zero duality gap between (3.1) and (3.4), even if the Slater condition holds, as stated in [53].

Apart from these characterizations several sufficient conditions of the zero duality gap for convex programs have been established in the literature, see [38, 3, 4, 94, 16, 18, 19, 85, 70].

## 3.2 Formulation of the problem

This section will provide the formulation of the problem in the non quadratic situation.

### 3.2.1 The general case with finite optimal value

Given a subset $C$ of a finite dimensional space $\mathbb{R}^n$, and functions $f:\mathbb{R}^n\to\mathbb{R}$, $g:\mathbb{R}^n\to\mathbb{R}$, let us consider the following minimization problem:

$$\mu \doteq \inf_{\substack{g(x)\le 0\\ x\in C}} f(x). \tag{3.5}$$

The Lagrangian dual problem associated to (3.1) is

$$\nu \doteq \sup_{\lambda\ge 0}\inf_{x\in C}[f(x)+\lambda g(x)], \tag{3.6}$$

We associate with Problem (3.5) the usual linear Lagrangian

$$L(\gamma,\lambda,x) \doteq \gamma f(x)+\lambda g(x),$$

where $\gamma\ge 0$ and $\lambda\ge 0$ are called the Lagrange multipliers. By setting $K\doteq\{x\in C:\ g(x)\le 0\}$, we obtain the trivial inequality

$$\gamma\inf_{x\in K} f(x)\ge \inf_{x\in K} L(\gamma,\lambda,x)\ge \inf_{x\in C} L(\gamma,\lambda,x),\quad \forall\ \gamma\ge 0,\ \forall\ \lambda\ge 0. \tag{3.7}$$

In order to get the equality, we need to find conditions under which the reverse inequality holds, that is, we must have:

$$\gamma(f(x)-\mu)+\lambda g(x)\ge 0\quad \forall\ x\in C. \tag{3.8}$$

This will imply strong duality once we get $\gamma>0$, and by recalling that $\mu=\inf_{x\in K} f(x)$.
By setting $F(x)\doteq(f(x),g(x))$ and so $F(C)\doteq\{(f(x),g(x))\in\mathbb{R}^2:\ x\in C\}$ along with $\rho\doteq(\gamma,\lambda)$, the previous inequality can be written as

$$\langle\rho,a\rangle\ge 0\quad \forall\ a\in F(C)-\mu(1,0). \tag{3.9}$$

The following result, which is important by itself, characterizes completely (3.9). Part of this result was established in [32, Theorem 4.1].

**Theorem 3.2.1.** *Let $P\subseteq\mathbb{R}^2$ be a convex closed cone such that* $\operatorname{int} P\neq\emptyset$, *and $A\subseteq\mathbb{R}^2$ be nonempty set. Then the following assertions are equivalent:*

*(a)* $\exists\ \lambda\in P^*\setminus\{0\},\ \forall\ a\in A:\ \langle\lambda,a\rangle\ge 0$;

*(b)* $A\cap(-\operatorname{int} P)=\emptyset$ *and* $\overline{\operatorname{cone}}(A+P)$ *is convex;*

*(c)* $A\cap(-\operatorname{int} P)=\emptyset$ *and* $\operatorname{cone}_+(A+\operatorname{int} P)$ *is convex;*

(*d*) $A \cap (-\text{int } P) = \emptyset$ *and* $\text{cone}(A + \text{int } P)$ *is convex;*

(*e*) $\text{cone}(A + \text{int } P)$ *is pointed;*

(*f*) $\text{co}(A) \cap (-\text{int } P) = \emptyset$.

*Proof.* Obviously $(c) \Longrightarrow (d) \Longrightarrow (b)$.
$(f) \Longrightarrow (a)$ It follows from a simple use of a separation result of convex sets.
$(a) \Longrightarrow (b)$: Clearly $\langle \lambda, x\rangle \geq 0$ for all $x \in \text{cone}(A+P)$. Choose $u \in \text{int } P$. Let $y, z \in A$. Then obviously

$$\text{cone}(\{y\}) + \text{cone}(\{u\}) = \{sy + tu : s,\ t \geq 0\}$$

is a closed convex cone containing $y$ and $u$ and contained in $\text{cone}(A + P)$. The same is true for the cone $\text{cone}(\{z\}) + \text{cone}(\{u\})$. The two cones have the line $\text{cone}(\{u\})$ in common and their union is contained in $\text{cone}(A + P)$, thus it is contained in the halfspace $\{x \in \mathbb{R}^2 :\ \langle \lambda, x\rangle \geq 0\}$. Hence, the set $B \doteq (\text{cone}(\{y\}) + \text{cone}(\{u\})) \cup (\text{cone}(\{z\}) + \text{cone}(\{u\}))$ is a convex cone. Since $y, z \in B$ we deduce that $[y, z] \subseteq B \subseteq \text{cone}(A + P)$. Thus $\text{co}(A) \subseteq \overline{\text{cone}}(A + P)$, from which we infer that $\overline{\text{cone}}(A + P)$ is convex since $P \subseteq \overline{\text{cone}}(A + P)$ holds as well.
$(b) \Longleftrightarrow (c)$: Obviously $(c)$ implies $(b)$. If $\overline{\text{cone}}(A + P)$ is convex then ([38])

$$\text{int}(\overline{\text{cone}}(A + P)) = \text{int}(\overline{\text{cone}_+(A) + P}) = \text{cone}_+(A) + \text{int } P = \text{cone}_+(A + \text{int } P)$$

is convex as well.
$(c) \Longrightarrow (e)$: Let $x, -x \in \text{cone}(A + \text{int } P)$. Then $x = t_1(a_1 + p_1,\ -x = t_2(a_2 + p_2)$ for some $t_i \geq 0$, $a_i \in A$, $p_i \in \text{int } P$ for $i = 1, 2$. Assuming $t_i > 0$, for $i = 1, 2$, we have $x, -x \in \text{cone}_+(A + \text{int } P)$. By convexity, $0 = x + (-x) \in \text{cone}_+(A + \text{int } P)$, which implies that $0 \in A + \text{int } P$, contradicting the first part of $(c)$.
$(e) \Longrightarrow (f)$: Assume on the contrary that $\text{co}(A) \cap (-\text{int } P) \neq \emptyset$. Then, there exist $a_i \in A$, $p_0 \in \text{int } P$, $\alpha_i \geq 0$, satisfying $\sum_{i=1}^m \alpha_i = 1$ and $0 = \sum_{i=1}^m \alpha_i a_i + p_0$. Thus, $0 = \sum_{i=1}^m \alpha_i(a_i + p_0)$. By pointedness, we get $\alpha_i(a_i + p_0) = 0$ for all $i = 1, \cdots, m$. Hence, $0 = a_j + p_0 \in A + \text{int } P$ for some $j$, which implies that $\text{cone}_+(A + \text{int } P) = \mathbb{R}^2$, contradicting $(e)$. $\square$

By virtue of the preceding result and following the reasoning developed in [31], we need to split the set $\text{cone}(F(C) - \mu(1, 0) + \mathbb{R}^2_{++})$. To that purpose, some notations are in order. By setting $K \doteq \{x \in C :\ g(x) \leq 0\}$, we get $K = S_g^-(0) \cup S_g^=(0)$, where

$$S_g^-(0) \doteq \{x \in C :\ g(x) < 0\},\ S_g^=(0) \doteq \{x \in C :\ g(x) = 0\}, S_g^+(0) \doteq \{x \in C :\ g(x) > 0\}.$$

Similarly, we define

$$S_f^-(\mu) \doteq \{x \in C :\ f(x) < \mu\},\ S_f^+(\mu) \doteq \{x \in C :\ f(x) > \mu\},$$

$$S_f^=(\mu) \doteq \{x \in C :\ f(x) = \mu\}.$$

Furthermore, whenever $S_g^-(0) \cap S_f^+(\mu) \neq \emptyset$ and $S_f^-(\mu) \neq \emptyset$, we set

$$r \doteq \inf_{x \in S_f^+(\mu) \cap S_g^-(0)} \frac{g(x)}{f(x) - \mu},\quad s \doteq \sup_{x \in S_f^-(\mu)} \frac{g(x)}{f(x) - \mu}.$$

Evidently, $-\infty \leq r < 0$, $-\infty < s \leq 0$. Notice that

$$x \in S_f^-(\mu) \Longrightarrow x \in S_g^+(0).$$

The latter and other basic facts about the previous sets are collected in the next proposition.

**Proposition 3.2.1.** *Let $\mu \in \mathbb{R}$, we have the following:*

(a) $C = K \iff S_g^+(0) = \emptyset$;

(b) $[\operatorname*{argmin}_K f \cap S_g^-(0) = \emptyset \text{ and } S_f^+(\mu) \cap S_g^-(0) = \emptyset] \iff S_g^-(0) = \emptyset$;

(c) $S_f^+(\mu) \cap S_g^-(0) = \emptyset \iff S_g^-(0) \subseteq \operatorname*{argmin}_K f$;

(d) $S_f^-(\mu) = \emptyset \iff \mu = \inf_{x \in C} f(x)$.

*Proof.* $(a)$ and $(c)$ are straightforward.
$(b)$: Suppose on the contrary that $S_g^-(0) \neq \emptyset$. Then, by assumption every $x \in C$ such that $g(x) < 0$ satisfies $f(x) \leq \mu$. Thus $f(x) = \mu$ yielding a contradiction. The other implication is obvious.
$(d)$: It follows by noticing that $S_f^-(\mu) = \emptyset$ if and only if $f(x) \geq \mu$ for all $x \in C$. $\square$

We now proceed to split the set $\text{cone}[F(C) - \mu(1,0) + \mathbb{R}^2_{++}]$ by writing $F(C) - \mu(1,0) + \mathbb{R}^2_{++} = \Omega_1 \cup \Omega_2 \cup \Omega_3$. This gives

$$\text{cone}(F(C) - \mu(1,0) + \mathbb{R}^2_{++}) = \text{cone}(\Omega_1) \cup \text{cone}(\Omega_2) \cup \text{cone}(\Omega_3), \tag{3.10}$$

where

$$\Omega_1 \doteq \mathbb{R}^2_{++} \cup \bigcup_{x \in \operatorname*{argmin}_K f \cap S_g^-(0)} [(0, g(x)) + \mathbb{R}^2_{++}];$$

$$\Omega_2 \doteq \bigcup_{x \in S_f^+(\mu) \cap S_g^-(0)} [(f(x) - \mu, g(x)) + \mathbb{R}^2_{++}] \cup \bigcup_{x \in S_f^+(\mu) \cap S_g^=(0)} [(f(x) - \mu, 0) + \mathbb{R}^2_{++}];$$

$$\Omega_3 \doteq \bigcup_{x \in S_f^-(\mu)} [(f(x) - \mu, g(x)) + \mathbb{R}^2_{++}] \cup \bigcup_{x \in S_f^=(\mu) \cap S_g^+(0)} [(0, g(x)) + \mathbb{R}^2_{++}] \cup$$
$$\cup \bigcup_{x \in S_f^+(\mu) \cap S_g^+(0)} [(f(x) - \mu, g(x)) + \mathbb{R}^2_{++}].$$

This decomposition will be used in Section 3.3.

### 3.2.2 The general case with unbounded optimal value

We continue by considering real-valued functions defined in a Hausdorff topological space $X$.
The case $\mu = -\infty$ deserves a special attention and it will be discussed in this subsection. First of all, it is not difficult to check that

$$\mu = -\infty \iff (F(C) + \mathbb{R}^2_+) \cap [(\rho, 0) - (\mathbb{R}_{++} \times \{0\})] \neq \emptyset, \quad \forall\, \rho \in \mathbb{R}. \tag{3.11}$$

We set

$$\gamma \doteq \inf_{x \in S^-_g(0)} \frac{g(x)}{f(x)},$$

whenever $S^-_g(0) \neq \emptyset$. Furthermore, set

$$W \doteq \{(u, v) \in \mathbb{R}^2 : \; v > \gamma u, \; v \leq 0\}, \quad \text{if} \quad \gamma \in \mathbb{R}.$$

The following theorem establishes the geometric structure of the set $\mathrm{cone}[F(C)+\mathbb{R}^2_{++}]$ in case $\mu = -\infty$.

**Theorem 3.2.2.** *Let $\mu = -\infty$. Then*

$$\mathbb{R} \times \mathbb{R}_+ \subseteq F(C) + \mathbb{R}^2_+. \tag{3.12}$$

*Furthermore,*

(a) *If $S^-_g(0) = \emptyset$ then $F(C) + \mathbb{R}^2_+ = \mathbb{R} \times \mathbb{R}_+$.*

(b) *If $S^-_g(0) \cap S^-_f(0) \neq \emptyset$ then $\mathrm{cone}_+[F(C) + \mathbb{R}^2_+] = \mathbb{R}^2$.*

(c) *If $S^-_g(0) \neq \emptyset$ and $S^-_g(0) \cap S^-_f(0) = \emptyset$ then $-\infty = \inf_{\substack{g(x)=0\\ x\in C}} f(x)$, $-\infty \leq \gamma < 0$ and*

  (c1) $\mathrm{cone}_+[F(C) + \mathbb{R}^2_{++}] = (\mathbb{R} \times \mathbb{R}_{++}) \cup W$ *if* $-\infty < \gamma < 0$;

  (c2) $\mathrm{cone}_+[F(C) + \mathbb{R}^2_{++}] = (\mathbb{R} \times \mathbb{R}_{++}) \cup (\mathbb{R}_{++} \times \mathbb{R})$ *if* $\gamma = -\infty$.

*Proof.* Let us prove (3.12). Take any $(\xi_1, \xi_2) \in \mathbb{R} \times \mathbb{R}_+$; by (3.11), there exist $x \in C$, $p \geq 0$, $q \geq 0$, $r > 0$, such that $f(x) + p = \xi_1 - r$ and $g(x) + q = 0$. It follows that

$$(\xi_1, \xi_2) = (f(x), g(x)) + (p + r, q + \xi_2) \in F(C) + \mathbb{R}^2_+.$$

$(a)$: Since $g(x) \geq 0$ for all $x \in C$, we obtain

$$F(C) + \mathbb{R}^2_+ \subseteq [f(C) \times g(C)] + \mathbb{R}^2_+ \subseteq (\mathbb{R} \times \mathbb{R}_+) + \mathbb{R}^2_+ = \mathbb{R} \times \mathbb{R}_+.$$

$(b)$: By assumption, there exists $x_0 \in C$ satisfying $g(x_0) < 0$ and $f(x_0) < 0$. This implies that $(0, 0) \in F(x_0) + \mathbb{R}^2_{++}$, which gives $\mathrm{cone}_+(F(x_0) + \mathbb{R}^2_{++}) = \mathbb{R}^2$ and therefore the conclusion follows.
$(c)$: By assumption, $f(x) \geq 0$ for all $x \in S^-_g(0)$, which implies that $-\infty = \inf_{\substack{g(x)=0\\ x\in C}} f(x)$ and $-\infty \leq \gamma < 0$.

$(c1)$: Let $(u_0, v_0) \in W$. Then, $u_0 > 0$ and there exists $x_0 \in C$ satisfying $g(x_0) < 0$ and

$$\gamma \leq \frac{g(x_0)}{f(x_0)} < \frac{v_0}{u_0}.$$

We choose $\varepsilon > 0$ satisfying $v_0 f(x_0) = u_0 g(x_0) + \varepsilon(u_0 - v_0)$ and write

$$u_0 = \frac{u_0}{f(x_0)+\varepsilon}(f(x_0)+\varepsilon), \quad v_0 = \frac{u_0}{f(x_0)+\varepsilon}(g(x_0)+\varepsilon).$$

This proves that $(u_0, v_0) \in \text{cone}_+[F(C) + \mathbb{R}^2_{++}]$. This result along with (3.12) prove one inclusion in $(c1)$.
For the other inclusion we reason as follows. Take any $(u_0, v_0) \in \text{cone}_+[F(C) + \mathbb{R}^2_{++}]$. Then, for some $(p, q) \in \mathbb{R}^2_{++}$, $t_0 > 0$, $x_0 \in C$, we have $u_0 = t_0(f(x_0) + p)$ and $v_0 = t_0(g(x_0) + q)$. If $(u_0, v_0) \notin \mathbb{R} \times \mathbb{R}_{++}$ then $v_0 \leq 0$. This implies that $g(x_0) < g(x_0) + q \leq 0$, and therefore, by assumption, $f(x_0) \geq 0$. Clearly $f(x_0) > 0$ since otherwise $\gamma = -\infty$. Hence $\gamma \leq \dfrac{g(x_0)}{f(x_0)}$, and so

$$\gamma u_0 = \gamma t_0(f(x_0) + p) \leq t_0 g(x_0) + \gamma t_0 p < t_0 g(x_0) < t_0(g(x_0) + q) = v_0,$$

showing that $(u_0, v_0) \in W$. Hence, the proof of $(c1)$ is completed.
$(c2)$: It is similar to $(c1)$. $\square$

## 3.3 The case with linear and quadratic constraints

In this section, we consider the case of quadratic functions defined in a finite dimensional space $\mathbb{R}^n$. Various problems arising in telecommunications, robust control [64, 81], trust region [42, 82], are modeled via quadratic non-homogeneous functions.
Consider the following quadratic optimization problem:

$$\mu \doteq \inf\Big\{\frac{1}{2}x^\top Ax + a^\top x + \alpha : \ \frac{1}{2}x^\top Bx + b^\top x + \beta \leq 0, \ Hx = d \ \Big\}, \tag{3.13}$$

where $A, B$ are symmetric matrices of order $n$; $a, \ b \in \mathbb{R}^n$; $d \in \mathbb{R}^m$; $\alpha, \beta \in \mathbb{R}$, and $H$ is a real matrix of order $m \times n$.
Setting, $C \doteq H^{-1}(d) \doteq \{x \in \mathbb{R}^n : \ Hx = d\}$, it is known that

$$C = x_0 + \ker\, H, \quad \forall \ x_0 \in C.$$

Let

$$f(x) = \frac{1}{2}x^\top Ax + a^\top x + \alpha, \ g(x) = \frac{1}{2}x^\top Bx + b^\top x + \beta.$$

One of the most important results concerning quadratic functions refers to Dine's theorem [24] (motivated by Finsler's theorem [30]), which ensures that

$$\Big\{(x^\top Ax, x^\top Bx) : \ x \in \mathbb{R}^n\Big\} \ \text{ is convex.}$$

This result does not hold in the non-homogeneous case as the next example shows.

**Example 3.3.1.** *Take* $f(x,y) = x + y - x^2 - y^2 - 2xy$, $g(x,y) = x^2 + y^2 + 2xy - 1$, *and consider the set* $M \doteq \{(f(x,y), g(x,y)) \in \mathbb{R}^2 : \ (x,y) \in \mathbb{R}^2\}$. *Clearly* $(0,0) = (f(0,1), g(0,1)) \in M$ *and* $(-2,0) = (f(-1,0), g(-1,0)) \in M$. *We claim that* $(-1,0) = \frac{1}{2}(0,0) + \frac{1}{2}(-2,0) \ \notin \ F(\mathbb{R}^2)$. *Indeed, if* $-1 = f(x,y)$ *and* $0 = g(x,y)$, *then* $|x+y| = 1$ *and* $x + y = 0$, *reaching a contradiction. Hence,* $(-1,0) \in \text{co } F(\mathbb{R}^2) \setminus F(\mathbb{R}^2)$, *showing that* $F(\mathbb{R}^2)$ *is nonconvex. More precisely, one can check that*

$$F(\mathbb{R}^2) = \{(t - t^2, t^2 - 1) : \ t \in \mathbb{R}\}.$$

*Let us consider the minimization problem*

$$\mu \doteq \inf_{\substack{g(x,y) \leq 0 \\ (x,y) \in \mathbb{R}^2}} f(x,y)$$

*We claim that* $\mu = -2$. *Indeed,*

$$\begin{aligned} f(x,y) + 2 + \frac{3}{2} g(x,y) &= x + y - (x+y)^2 + 2 + \frac{3}{2}((x+y)^2 - 1) \\ &= 2(x+y+1)^2 \geq 0, \ \forall \ (x,y) \ \in \ \mathbb{R}^2. \end{aligned}$$

*In particular, if* $(x,y)$ *is such that* $g(x,y) \leq 0$, *we obtain* $f(x,y) \geq -2 = f(-1,0)$, *proving our claim. We actually have* $\operatorname*{argmin}_K f = \{(x,y) \in \mathbb{R}^2 : \ x + y = -1\}$ *and* $\lambda = \frac{3}{2}$ *is a Lagrange multiplier. Furthermore,* $S_f^+(\mu) \cap S_g^-(0) = \{(x,y) \in \mathbb{R}^2 : \ |x+y| < 1\}$ *and* $S_f^-(\mu) = \{(x,y) \in \mathbb{R}^2 : \ x + y < -1\} \cup \{(x,y) \in \mathbb{R}^2 : \ 2 < x + y\}$. *Therefore*

$$r = \inf_{|x+y|<1} \frac{(x+y)^2 - 1}{x + y - (x+y)^2 + 2} = \inf_{|t|<1} \frac{t^2 - 1}{t - t^2 + 2} = \inf_{|t|<1} -\frac{t-1}{t-2} = -\frac{2}{3};$$

$$s = \sup_{(x,y) \in S_f^-(\mu)} \frac{(x+y)^2 - 1}{x + y - (x+y)^2 + 2} = -\frac{2}{3}.$$

However, we can prove a relaxed version of Dine's theorem. To that purpose the next result, valid for quadratic functions, will play an important role.

**Proposition 3.3.2.** [51, *Theorem* 3.6] *Let* $f, g : \mathbb{R}^n \to \mathbb{R}$ *be quadratic functions not necessarily homogeneous, let* $x_0 \in \mathbb{R}^n$, *and let* $S_0$ *be a subspace of* $\mathbb{R}^n$. *Then exactly one of the following statements holds:*

(*a*) $\exists \ x \in x_0 + S_0, \ f(x) < 0, \ g(x) < 0$;

(*b*) $\exists \ (\lambda_1, \lambda_2) \in \mathbb{R}^2_+ \setminus \{(0,0)\}, \ \lambda_1 f(x) + \lambda_2 g(x) \geq 0, \ \ \forall \ x \in x_0 + S_0$.

On combining the preceding result and Theorem 3.2.1, we obtain the following theorem which may be considered as a relaxed version of the Dines theorem and, according to the author's knowledge, it is new in the literature. Obviously our result is weaker than that provided by Dines when $f$ and $g$ are homogeneous quadratic functions (for the case $\mu = -\infty$ we refer Theorem 3.2.2).

**Theorem 3.3.1.** (*Relaxed Dine's theorem*) *Let $f : \mathbb{R}^n \to \mathbb{R}$, $g : \mathbb{R}^n \to \mathbb{R}$ be quadratic functions as above. If $\mu \in \mathbb{R}$, then*

$$\text{cone}(F(C) - \rho(1,0) + \mathbb{R}^2_{++}) \text{ is convex for all } \rho \leq \mu.$$

*Proof.* Since there is no $x$ satisfying $g(x) < 0$, $Hx = d$ and $f(x) - \rho < 0$, by Proposition 3.3.2 we obtain the existence of $(\gamma, \lambda) \in \mathbb{R}^2_+ \setminus \{(0,0)\}$ such that

$$\gamma(f(x) - \rho) + \lambda g(x) \geq 0 \quad \forall\, x \in x_0 + \ker H = C. \tag{3.14}$$

The desired result is a consequence of Theorem 3.2.1. □

In case $\mu = -\infty$ Theorem 3.2.2 provides a complete description of $\text{cone}_+(F(C)+\mathbb{R}^2_+)$; in particular, it establishes conditions under which $\text{cone}_+(F(C) + \mathbb{R}^2_+)$ is convex.

We next present an application of the previous theorem to derive the $S$-lemma for any (not necessarily homogeneous) quadratic functions already appeared in [75, Theorem 2.2]; [51, Corollary 3.7]). Some variants of the $S$-lemma may be found in [23].

**Theorem 3.3.2.** (*The $S$-lemma*) *Let $f, g : \mathbb{R}^n \to \mathbb{R}$ be quadratic functions and assume that there is $\bar{x} \in C$ such that $g(\bar{x}) < 0$. Then, (a) and (b) are equivalent:*

(*a*) *There is no $x \in C$ such that*

$$f(x) < 0, \;\; g(x) \leq 0.$$

(*b*) *There is $\lambda \geq 0$ such that*

$$f(x) + \lambda g(x) \geq 0, \quad \forall\, x \in C.$$

*Proof.* Obviously $(b) \Longrightarrow (a)$ always holds. Assume therefore that $(a)$ is satisfied. This means that $x \in C$, $g(x) \leq 0$ implies $f(x) \geq 0$, that is, $0 \leq \mu \doteq \inf_{x \in K} f(x)$. It follows that

$$\text{cone}[F(C) - \mu(1,0) + \mathbb{R}^2_{++}] \cap \mathcal{H} = \emptyset,$$

$\mathcal{H} \doteq \{(u,v) \in \mathbb{R}^2 \; : u < 0, v \leq 0\}$. By the previous theorem $\text{cone}[F(C) - \mu(1,0) + \mathbb{R}^2_{++}]$ is convex, and so by a separation theorem, there exist $(\gamma, \lambda) \in \mathbb{R}^2_+ \setminus \{(0,0)\}$ and $\alpha \in \mathbb{R}$ such that

$$\gamma(f(x) - \mu + p) + \lambda(g(x) + q) \geq \alpha \geq \gamma u + \lambda v, \; \forall\, x \in C, \; \forall\, (p,q) \in \mathbb{R}^2_{++}, \; \forall\, u < 0, \; \forall\, v \leq 0.$$

This implies $\alpha \geq 0$, $\gamma \geq 0$ and $\lambda \geq 0$. Thus

$$\gamma(f(x) - \mu) + \lambda g(x) \geq 0, \quad \forall\, x \in C,$$

that is, $\gamma f(x) + \lambda g(x) \geq \gamma\mu \geq 0, \;\; \forall\, x \in C$. The Slater condition yields $\gamma > 0$, completing the proof of the theorem. □

The important case, when $f$ and $g$ are quadratic, with $C = \mathbb{R}^n$, was studied by Yakubovich [88, 89], see the survey by Pólik and Terlaky in [75, Theorem 2.2]. Its proof uses the Dines theorem which asserts the convexity of the set $\{(f(x), g(x)) \in \mathbb{R}^2 : \ x \in \mathbb{R}^n\}$ when $f$ and $g$ are homogeneous quadratic functions.

We observe that (3.14) for $\rho = \mu$ amounts to writing that

$$(\gamma, \lambda) \in [\mathrm{cone}(F(C) - \mu(1,0) + \mathbb{R}^2_{++})]^*, \tag{3.15}$$

and we also get

$$\mathrm{cone}(F(C) - \mu(1,0) + \mathbb{R}^2_{++}) = \mathrm{cone}_+(F(C) - \mu(1,0) + \mathbb{R}^2_{++}) \cup \{(0,0)\}.$$

The slightly dark region in Figures 3.1 and 3.2 represents $\mathrm{cone}_+(F(C) - \mu(1,0) + \mathbb{R}^2_{++})$.

Taking into account the splitting (3.10) introduced in Subsection 3.2.1, we establish the main theorem in this section which is new in the literature and describes all the situations that may happen when considering quadratic minimization problems with finitely many linear equality and a single quadratic inequality constraints. It provides also the solution set of the regularized Lagrangian dual (3.4).

**Theorem 3.3.3.** *Let $f : \mathbb{R}^n \to \mathbb{R}$, $g : \mathbb{R}^n \to \mathbb{R}$ be the quadratic functions, and $C$ as above, $(\gamma, \lambda) \in \mathbb{R}^2_+ \setminus \{(0,0)\}$ and $\mu$ be finite. Then, exactly one of the following assertions holds:*

$(a1)$ *If either $\operatorname*{argmin}_K f \cap S^-_g(0) \neq \emptyset$ or $[S^+_f(\mu) \cap S^-_g(0) \neq \emptyset$ with $r = -\infty]$, then $S^-_f(\mu) = \emptyset$ and*

$$\mathrm{cone}_+(F(C) - \mu(1,0) + \mathbb{R}^2_{++}) = \{(u,v) \in \mathbb{R}^2 : \ u > 0\}.$$

*Hence,*

$$\gamma(f(x) - \mu) + \lambda g(x) \geq 0, \ \forall \ x \ \in C \Longleftrightarrow \gamma > 0, \ \lambda = 0.$$

$(a2)$ *If $\operatorname*{argmin}_K f \cap S^-_g(0) = \emptyset$, $S^+_f(\mu) \cap S^-_g(0) = \emptyset = S^-_f(\mu)$, then*

$$\mathrm{cone}_+(F(C) - \mu(1,0) + \mathbb{R}^2_{++}) = \{(u,v) \in \mathbb{R}^2 : \ u > 0, \ v > 0\}.$$

*Hence*

$$\gamma(f(x) - \mu) + \lambda g(x) \geq 0, \ \forall \ x \in C \Longleftrightarrow (\gamma, \lambda) \in \mathbb{R}^2_+ \setminus \{(0,0)\}.$$

$(a3)$ *If $S^+_f(\mu) \cap S^-_g(0) \neq \emptyset$ with $-\infty < r < 0$ and $S^-_f(\mu) \neq \emptyset$, then $\operatorname*{argmin}_K f \cap S^-_g(0) = \emptyset$, $s \leq r$ and*

$$\mathrm{cone}_+(F(C) - \mu(1,0) + \mathbb{R}^2_{++}) = \{(u,v) \in \mathbb{R}^2 : \ v > ru, \ v > su\}.$$

*Hence,*

$$\gamma(f(x) - \mu) + \lambda g(x) \geq 0, \ \forall \ x \ \in \ C \Longleftrightarrow \gamma > 0, \ -\frac{1}{s}\gamma \leq \lambda \leq -\frac{1}{r}\gamma.$$

$(a4)$ *If* $S_f^+(\mu) \cap S_g^-(0) = \emptyset$, $S_f^-(\mu) \neq \emptyset$ *with* $-\infty < s < 0$, *then* $\operatorname*{argmin}_K f \cap S_g^-(0) = \emptyset$ *and*

$$\text{cone}_+(F(C) - \mu(1,0) + \mathbb{R}^2_{++}) = \{(u,v) \in \mathbb{R}^2 : \ v > su, \ v > 0\}.$$

*Hence*

$$\gamma(f(x) - \mu) + \lambda g(x) \geq 0, \ \forall \ x \in C \Longleftrightarrow \gamma \geq 0, \ \lambda \geq -\frac{1}{s}\gamma, \ \lambda \neq 0.$$

$(a5)$ *If* $\operatorname*{argmin}_K f \cap S_g^-(0) = \emptyset$, $S_f^+(\mu) \cap S_g^-(0) \neq \emptyset$ *with* $-\infty < r < 0$ *and* $S_f^-(\mu) = \emptyset$, *then*

$$\text{cone}_+(F(C) - \mu(1,0) + \mathbb{R}^2_{++}) = \{(u,v) \in \mathbb{R}^2 : \ v > ru, \ u > 0\}.$$

*Hence,*

$$\gamma(f(x) - \mu) + \lambda g(x) \geq 0, \ \forall \ x \ \in C \Longleftrightarrow \gamma > 0, \ 0 \leq \lambda \leq -\frac{1}{r}\gamma.$$

$(a6)$ *If* $S_f^-(\mu) \neq \emptyset$, $s = 0$, *then* $\operatorname*{argmin}_K f \cap S_g^-(0) = \emptyset$ *and*

$$\text{cone}_+(F(C) - \mu(1,0) + \mathbb{R}^2_{++}) = \{(u,v) \in \mathbb{R}^2 : \ v > 0\}.$$

*Hence*

$$\gamma(f(x) - \mu) + \lambda g(x) \geq 0, \ \forall \ x \in C \Longleftrightarrow \gamma = 0, \ \lambda > 0.$$

*Proof.* Since the proof uses a frequent application of the convexity of $\text{cone}_+(F(C) - \mu(1,0) + \mathbb{R}^2_{++})$, the splitting (3.10) and (3.15) along with Figures 3.1, 3.2, we simply prove $(a1)$ and $(a2)$ just to give an idea of the reasoning to be employed.
$(a1)$: By assumptions and due to the convexity of $\text{cone}_+(F(C) - \mu(1,0) + \mathbb{R}^2_{++})$, looking at Figure 3.1$(a1)$, we immediately get that $S_f^-(\mu) = \emptyset$, and so

$$\text{cone}_+(F(C) - \mu(1,0) + \mathbb{R}^2_{++}) = \{(u,v) \in \mathbb{R}^2 : \ u > 0\}.$$

From this, the equivalence follows in view of (3.15).

$(a2)$: It is a consequence of the splitting (3.10) and (3.15)

As mentioned above all other assertions follow in similar way by taking into account the convexity of $\text{cone}(F(C) - \mu(1,0) + \mathbb{R}^2_{++})$, (3.10) and (3.15), see Figures 3.1, 3.2. $\square$

One can check that Example 3.3.1 satisfies $(a3)$, since

$$(0, -2) \in S_f^-(\mu) \ \text{ and } \ \operatorname*{argmin}_K f \cap S_g^-(0) = \emptyset.$$

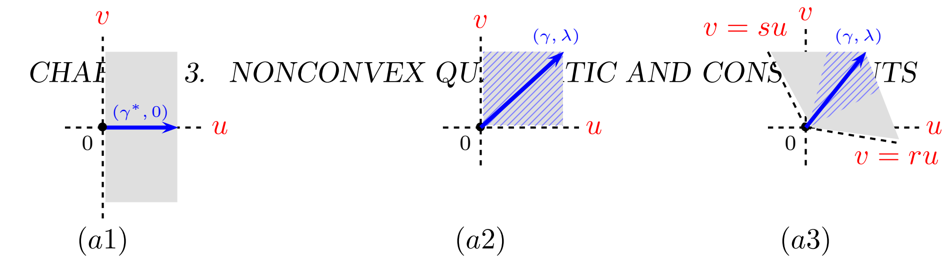


Figure 3.1: Theorem 3.3.3: $(a1)$, $(a2)$, $(a3)$ (figure produced by author)

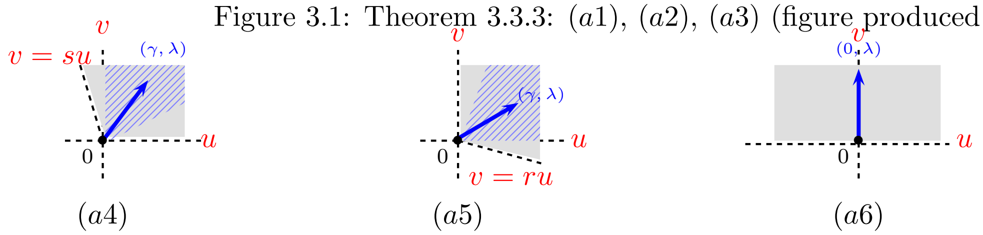


Figure 3.2: Theorem 3.3.3: $(a4)$, $(a5)$, $(a6)$ (figure produced by author)

We also obtain $r = s = -\dfrac{2}{3}$; therefore $\lambda = \dfrac{2}{3}\gamma$.

Before providing a characterization of strong duality, some preliminaries are necessary for linking the behaviour of the Hessians of $f$ and $g$ and the number $r$ and $s$. We first provide a necessary condition to have $\mu \in \mathbb{R}$.

**Proposition 3.3.3.** *Assume that $\mu$ is finite. Then,*

$$0 \neq v \in \ker\, H,\ v^\top B v \leq 0 \implies v^\top A v \geq 0. \tag{3.16}$$

*Proof.* Let $v \in \ker\, H,\ v \neq 0$, we distinguish the discussion into two cases: $v^\top Bv < 0$ and $v^\top Bv = 0$.

In the first case, given $x \in H^{-1}(d)$, we obtain $g(x + tv) = g(x) + t\nabla g(x)^\top v + \frac{t^2}{2} v^\top Bv \to -\infty$ as $|t| \to +\infty$ since $v^\top Bv < 0$. Thus, there exists $t_1 > 0$ such that $x + tv \in S_g^-(0)$ for all $|t| \geq t_1$, which gives $f(x) + t\nabla f(x)^\top v + \frac{t^2}{2} v^\top Av = f(x + tv) \geq \mu$ for all $|t| \geq t_1$ since $x + tv \in H^{-1}(d)$. On dividing by $t^2$ and letting $t \to +\infty$, we get $v^\top Av \geq 0$.

Now assume that $v^\top Bv = 0$, and suppose on the contrary that $v^\top Av < 0$. This yields, given any $x \in H^{-1}(d)$, $f(x + tv) \to -\infty$ for all $|t| \to +\infty$. Then $g(x + tv) = g(x) + t\nabla g(x)^\top v > 0$ for all $|t|$ sufficiently large, which implies that $\nabla g(x)^\top v = 0$, and therefore $g(x) = g(x + tv) > 0$ for all $t \in \mathbb{R}$ and all $x \in H^{-1}(d)$. This cannot happen if we choose $x$ satisfying in addition $g(x) \leq 0$.

□

The necessary condition (3.16) given in the previous proposition is stronger than the condition

$$0 \neq v \in \ker\, H,\ v^\top Bv = 0 \implies v^\top Av \geq 0. \tag{3.17}$$

This is related to a relaxed version of Finsler's theorem due to Moré [66, Theorem 2.3] and independently to Hamburger [45]: assume that $B$ be indefinite, then $(i)$ and $(ii)$ below are equivalent:

(*i*) $v \in \ker H,\ v^\top Bv = 0 \implies v^\top Av \geq 0.$

(*ii*) $\exists\, t \in \mathbb{R}$ such that $A + tB$ is positive semidefinite on $\ker H$.

**Proposition 3.3.4.** *Let $v \in \ker H$, $v^\top Av = 0$ and $v^\top Bv < 0$, and assume that $\mu$ is finite. Then,*

(*a*) $\exists\, t_1 > 0$ *such that* $x + tv \in S_g^-(0)$, $\forall\, x \in H^{-1}(d)$ *and* $\forall\, |t| \geq t_1$;

(*b*) $\nabla f(x)^\top v = 0$, $\forall\, x \in H^{-1}(d)$, *or equivalently,* $f(x + tv) = f(x)\ \forall\, x \in H^{-1}(d)$ *and* $\forall\, t \in \mathbb{R}$, *or equivalently,* $\exists\, y \in \mathbb{R}^m$ *such that* $Av = H^\top y$ *and* $d^\top y + a^\top v = 0$;

(*c*) $S_f^-(\mu) = \emptyset$, *and therefore* $\mu = \inf\limits_{x \in H^{-1}(d)} f(x)$ *with* $\operatorname*{argmin}\limits_{H^{-1}(d)} f \neq \emptyset$;

(*d*) $S_f^+(\mu) \neq \emptyset \implies r = -\infty.$

*Proof.* (*a*): Let $x \in H^{-1}(d)$. Then, $g(x + tv) = g(x) + t\nabla g(x)^\top v + \frac{t^2}{2} v^\top Bv \to -\infty$ as $|t| \to +\infty$ since $v^\top Bv < 0$. Thus, there exists $t_1 > 0$ such that $x + tv \in S_g^-(0)$ for all $|t| \geq t_1$.

(*b*): For the first equivalence; from (*a*), $f(x + tv) \geq \mu$ for all $|t| \geq t_1$ because of $x + tv \in H^{-1}(d)$ and $g(x + tv) < 0$. By writting $\mu \leq f(x + tv) = f(x) + t\nabla f(x)^\top v$, we conclude that $\nabla f(x)^\top v = 0$, and therefore $f(x + tv) = f(x)$ for all $t \in \mathbb{R}$.

One implication for the second equivalence is as follows. By noticing that $H^{-1}(d) = x_0 + \ker H$ for all $x_0 \in H^{-1}(d)$, the equality $(Ax + a)^\top v = 0$ for all $x \in H^{-1}(d)$ implies that $Av \in (\ker H)^\perp = H^\top(\mathbb{R}^m)$. Thus, there exists $y \in \mathbb{R}^m$ such that $Av = H^\top y$ and therefore

$$0 = x^\top Av + a^\top v = x^\top H^\top y + a^\top v = d^\top y + a^\top v.$$

The remaining implication is obvious.

(*c*): It follows from (*a*) and (*b*), along with Proposition 3.2.1 and Corollary 3.4.2.

(*d*): Take any $x_0 \in S_f^+(\mu)$. Then, from (*b*) it follows that $f(x_0 + tv) = f(x_0) > \mu$ for all $|t| \geq t_1$. For such $t$, (*a*) implies that $x_0 + tv \in S_f^+(\mu) \cap S_g^-(0)$. Hence, since

$$r \leq \frac{g(x_0) + t\nabla g(x_0)^\top v + \frac{t^2}{2} v^\top Bv}{f(x_0) - \mu}, \quad \forall\, |t| \geq t_1,$$

we infer that $r = -\infty$. ∎

**Proposition 3.3.5.** *Let $f, g : \mathbb{R}^n \to \mathbb{R}$ be quadratic functions as above with $\mu$ finite. Then,*

$$(a)\ \ r = -\infty \implies \begin{cases} \operatorname*{argmin}\limits_{K} f \neq \emptyset, \quad \text{or;} \\ \mu = \inf\limits_{x \in H^{-1}(d)} f(x) \text{ with } \operatorname*{argmin}\limits_{H^{-1}(d)} f \neq \emptyset, \quad \text{or;} \\ \exists\, x_k \in S_f^+(\mu) \cap S_g^-(0) : \|x_k\| \to +\infty,\ \frac{x_k}{\|x_k\|} \to v \in \ker H, \text{ and} \\ v^\top Av = 0,\ v^\top Bv = 0. \end{cases}$$

$$(b)\ \ s = 0 \Longrightarrow \begin{cases} \underset{K}{\operatorname{argmin}}\ f \cap S_g^{=}(0) \neq \emptyset, \qquad \text{or;} \\ \exists\ x_k \in\ S_f^{-}(\mu) : \|x_k\| \to +\infty,\ \frac{x_k}{\|x_k\|} \to v \in \ker\, H,\ \text{and} \\ v^\top Av = 0, \quad v^\top Bv = 0. \end{cases}$$

*Proof.* $(a)$: By assumption, there exists a sequence $x_k \in S_f^{+}(\mu) \cap S_g^{-}(0)$ such that

$$\lim_{k\to+\infty} \frac{g(x_k)}{f(x_k) - \mu} = -\infty.$$

We distinguish two cases.

**Case 1.** $\sup\limits_{k\in\mathbb{N}} \|x_k\| < +\infty$. Up to a subsequence we may assume that $x_k \to x_0$ as $k \to +\infty$. Thus, $g(x_0) \leq 0$ and $f(x_0) \geq \mu$. The case $g(x_0) = 0$, $f(x_0) = \mu$ (resp. $g(x_0) < 0$, $f(x_0) = \mu$) yields $\underset{K}{\operatorname{argmin}}\ f \cap S_g^{=}(0) \neq \emptyset$ (resp. $\underset{K}{\operatorname{argmin}}\ f \cap S_g^{-}(0) \neq \emptyset$). The other situations cannot occur since

$$\lim_{k\to+\infty} \frac{g(x_k)}{f(x_k) - \mu} = \begin{cases} \dfrac{g(x_0)}{f(x_0) - \mu} \neq -\infty & ,\ \text{if } g(x_0) < 0, \quad f(x_0) > \mu; \\ \\ 0 & ,\ \text{if } g(x_0) = 0, \quad f(x_0) > \mu. \end{cases}$$

**Case 2.** $\sup\limits_{k\in\mathbb{N}} \|x_k\| = +\infty$. Then, we can assume that

$$\|x_k\| \to +\infty, \qquad \frac{x_k}{\|x_k\|} \to v, \quad \text{as } k \to +\infty, \tag{3.18}$$

and therefore $v \in \ker\, H$, $v^\top Av \geq 0$ and $v^\top Bv \leq 0$. Moreover, we obtain, as in Case 1,

$$\lim_{k\to+\infty} \frac{g(x_k)}{f(x_k) - \mu} = \begin{cases} \dfrac{v^\top Bv}{v^\top Av} \neq -\infty & ,\ \text{if } v^\top Av > 0,\ v^\top Bv < 0; \\ \\ 0 & ,\ \text{if } v^\top Av > 0,\ v^\top Bv = 0. \end{cases}$$

Hence, we must have $v^\top Av = 0$ and $v^\top Bv \leq 0$. In case $v^\top Av = 0$ and $v^\top Bv < 0$, we apply Proposition 3.3.4$(c)$ to get the second posibility of $(a)$.

$(b)$: We have the existence of a sequence $x_k \in S_f^{-}(\mu)$ such that

$$\lim_{k\to+\infty} \frac{g(x_k)}{f(x_k) - \mu} = 0.$$

We likewise distinguish two cases.

**Case 1.** $\sup_{k\in\mathbb{N}} \|x_k\| < +\infty$. Up to a subsequence, we obtain $x_k \to x_0 \in H^{-1}(d)$, $g(x_0) \geq 0$ and $f(x_0) \leq \mu$. Since $g(x_0) = 0$ and $f(x_0) < \mu$ is impossible, and because of

$$\lim_{k\to+\infty} \frac{g(x_k)}{f(x_k)-\mu} = \begin{cases} \dfrac{g(x_0)}{f(x_0)-\mu} \neq 0 & \text{, if } g(x_0) > 0, \quad f(x_0) < \mu; \\ -\infty & \text{, if } g(x_0) > 0, \quad f(x_0) = \mu, \end{cases}$$

we must have $g(x_0) = 0$ and $f(x_0) = \mu$.

**Case 2.** $\sup_{k\in\mathbb{N}} \|x_k\| = +\infty$. Passing to a subsequence, if necessary, we have (3.18), and therefore $v \in \ker H$, $v^\top Av \leq 0$ and $v^\top Bv \geq 0$. As in $(a)$, we get necessarily $v^\top Av \leq 0$ and $v^\top Bv = 0$. The conclusion follows after noticing that $v^\top Av < 0$ and $v^\top Bv = 0$ cannot occur by Proposition 3.3.3. $\square$

In view of Propositions 3.3.4, 3.3.3 and 3.3.5, the following conditions arise:

- $[0 \neq v \in \ker H,\ v^\top Bv \leq 0] \implies v^\top Av > 0;$ (3.19)
- $[0 \neq v \in \ker H,\ v^\top Bv = 0] \implies v^\top Av > 0;$ (3.20)
- $[v \in \ker H,\ v^\top Av = 0 = v^\top Bv] \implies v = 0;$ (3.21)
- $[0 \neq v \in \ker H,\ v^\top Bv = 0] \implies v^\top Av \neq 0;$ (3.22)
- $0 \neq v \in \ker H \implies [v^\top Av \neq 0 \text{ or } v^\top Bv \neq 0].$ (3.23)

Clearly,

$$(3.19) \implies (3.20) \implies (3.21) \iff (3.22) \iff (3.23).$$

By Finsler's theorem [30] (see also [45]), condition (3.20) is equivalent to:

$$\exists\, t \in \mathbb{R}, \text{ such that } A + tB \text{ is positive definite on } \ker H. \tag{3.24}$$

When this condition is satisfied it is said that the Simultaneous Diagonalization Property holds, since it implies the existence of a nonsingular matrix $C$ such that both $C^\top AC$ and $C^\top BC$ are diagonal [46, Theorem 7.6.4]. Such an assumption allowed the authors in [8] to re-write the original problem in a more tractable one.

In [90] when $H = 0$ and $d = 0$, some relationships between (3.17), (3.20), (3.21) and the Yakuvobich $S$-lemma (with quadratic homogeneous functions) are established. They are related with the non-strict Finsler's, strict Finsler's and Finsler-Calabi's theorem, respectively.

Under assumption (3.22), $(b)$ of Proposition 3.3.5, implies the following corollary

**Corollary 3.3.6.** *Assume that $f$, $g$ are as above with $C = H^{-1}(d)$ and $\mu \in \mathbb{R}$. If $s = 0$ and (3.22) is satisfied then* $\operatorname{argmin}_K f \cap S_g^{=}(0) \neq \emptyset$ *and strong duality does not hold.*

*Proof.* If $s = 0$ then by $(b)$ of Proposition 3.3.5 we obtain that either $\operatorname*{argmin}_K f \cap S_g^{=}(0) \neq \emptyset$ or there exists $0 \neq v \in \ker H$ satisfying $v^\top Av = 0$ and $v^\top Bv = 0$. By assumption the second situation is not possible, and therefore the first holds proving the desired result. The lack of strong duality is a consequence of $(a6)$ in Theorem 3.3.3. $\square$

In contrast to a similar result due to Moré [66] where condition (3.19) (stronger than (3.22)) is imposed, our corollary applies to situations where Theorem 3.3 in [66] does not.

**Corollary 3.3.7.** *Assume that $f$, $g$ are as above with $C = H^{-1}(d)$ and $\mu \in \mathbb{R}$. If $r = -\infty$ and (3.22) is satisfied then strong duality holds and either $\operatorname*{argmin}_K f \neq \emptyset$ or $\mu = \inf_{x \in H^{-1}(d)} f(x)$ with $\operatorname*{argmin}_{H^{-1}(d)} f \neq \emptyset$.*

*Proof.* It is a direct consequence of $(a)$ in Proposition 3.3.5. $\square$

Next result, which is new, on one hand characterizes the regularized strong duality without requiring the nonemptiness of $\operatorname*{argmin}_K f$, and where the Slater condition may fail, and on the other, gives a sufficient or necessary condition in terms of inequality systems.

**Theorem 3.3.4.** *Let $\mu$ be finite with $C = H^{-1}(d)$. Let us consider the following assertions:*

- $(a)$ $\operatorname*{argmin}_K f \cap S_g^{=}(0) = \emptyset$ *and* (3.23) *holds;*
- $(b)$ *strong duality holds;*
- $(c)$ *either $S_f^-(\mu) = \emptyset$ or $[S_f^-(\mu) \neq \emptyset$ with $s < 0]$ holds;*
- $(d)$ *either $\inf_{x \in C} f(x) = \mu$ or* $[v \in \ker H,\ v^\top Av \leq 0 \implies v^\top Bv \geq 0]$ holds.

*Then, we have the following relationships:*

$$(a) \implies (b) \iff (c) \implies (d).$$

*Proof.* $(a) \implies (c)$: We have to check that $s < 0$. If on the contrary, $s = 0$, by using Proposition 3.3.5$(b)$ we get a contradiction.
$(b) \implies (c)$: Suppose that $S_f^-(\mu) \neq \emptyset$. Strong duality implies the existence of $\lambda_0 \geq 0$ such that $f(x) + \lambda_0 g(x) \geq \mu$ for all $x \in C$, which yields $\lambda_0 > 0$. Indeed, if $\lambda_0 = 0$, the previous inequality gives $f(x) - \mu \geq 0$ for all $x \in C$, which is impossible if $S_f^-(\mu) \neq \emptyset$.
Now, suppose that $s = 0$. Then, there exists $\bar{x} \in S_f^-(\mu) \neq \emptyset$ such that

$$\frac{g(\bar{x})}{f(\bar{x}) - \mu} > -\frac{1}{\lambda_0}.$$

It follows that $f(\bar{x}) + \lambda_0 g(\bar{x}) < \mu$, giving a contradiction; this proves that $s > 0$.
$(c) \implies (b)$: It is simply a consequence of Theorem 3.3.3 by looking at those items where $\gamma^* > 0$ is possible.
$(c) \implies (d)$ : If $S_f^-(\mu) = \emptyset$ then $f(x) \geq \mu$ for all $x \in C$ by Proposition 3.3.2. Assume now that $S_f^-(\mu) \neq$

$\emptyset$ and $s < 0$. Due to the convexity of $\text{cone}(F(C) - \mu(1,0) + \mathbb{R}^2_{++})$, we obtain $\underset{K}{\text{argmin}}\ f \cap S_g^-(0) = \emptyset$. We consider two cases: $S_g^-(0) = \emptyset$ or $S_g^-(0) \neq \emptyset$. Obviously in the first case, the implication in $(d)$ holds vacuously. If $S_g^-(0) \neq \emptyset$, it follows that $S_f^+(\mu) \cap S_g^-(0) \neq \emptyset$ since otherwise $S_g^-(0) = \emptyset$ by Proposition 3.2.1$(b)$. Thus, we must have $-\infty < s \leq r < 0$ again by the convexity of $\text{cone}(F(C) - \mu(1,0) + \mathbb{R}^2_{++})$. From Proposition 3.3.4$(d)$ it follows that $v \in \ker H$, $v^\top B v < 0 \Longrightarrow v^\top A v \neq 0$, which together with (3.16) yields the desired implication. $\square$

Example 3.3.1 shows that the implication $(c) \Longrightarrow (a)$ may be false, and the next instance shows the second part of $(d)$ does not necessarily imply the second part of $(c)$.

**Example 3.3.8.** *Let $C = \mathbb{R}^n$, $f(x_1, x_2) = x_1 + x_2$ and $g(x_1, x_2) = (x_1 + x_2)^2$. Clearly it satisfies the second part of $(d)$, but it holds $S_f^-(\mu) \neq \emptyset$ with $s = 0$. Indeed, $K = \{(0,0)\}$ and*

$$S_f^-(\mu) = \{(x_1, x_2) \in \mathbb{R}^2 :\ x_1 + x_2 < 0\}.$$

*Hence,*

$$s = \sup_{x_1 + x_2 < 0} \frac{(x_1 + x_2)^2}{x_1 + x_2} = 0,$$

*and the strong duality does not hold, since for any $\lambda > 0$, the inequality*

$$x_1 + x_2 + \lambda(x_1 + x_2)^2 \geq 0, \quad \forall\ (x_1, x_2) \in \mathbb{R}^2$$

*yields a contradiction. This agrees with $(a6)$ of Theorem 3.3.3.*

Next example illustrates a situation where our main Theorem 3.3.3 applies, exhibiting that strong duality holds without satisfying the Slater condition: there exists $x_0 \in H^{-1}(d)$ such that $g(x_0) < 0$.

**Example 3.3.9.** *Take $H(x_1, x_2) = x_1 - x_2$, $d = 0$, $f(x_1, x_2) = 2x_1^2 - x_2^2$, $g(x_1, x_2) = x_1^2 - x_2^2$. Here, $K = \{(x_1, x_2) \in \mathbb{R}^2 : x_1 = x_2,\ x_1^2 - x_2^2 \leq 0\} = \{(t,t) \in \mathbb{R}^2 :\ t \in \mathbb{R}\}$. Clearly, $S_g^-(0) = \emptyset = S_f^-(\mu)$ and $\mu = 0$ with $\underset{K}{\text{argmin}}\ f = \{(0,0)\}$. According to $(a2)$ of Theorem 3.3.3, we conclude that strong duality holds by choosing any $\lambda^* \geq 0$.*

The next theorem, which is new in the literature, considers non-convex situations.

**Theorem 3.3.5.** *Let $f$ and $g$ be quadratic functions as above, $\mu$ finite and $\bar{x}$ be feasible for (3.13). Set $C = H^{-1}(d)$. The following assertions are equivalent:*

- $(a)$ *$\bar{x}$ is a solution to (3.13) and either $S_f^-(\mu) = \emptyset$ or $[S_f^-(\mu) \neq \emptyset$ with $s < 0]$ holds;*
- $(b)$ *$\exists\ \lambda \geq 0\ \exists\ y \in \mathbb{R}^m$ such that $\nabla f(\bar{x}) + \lambda \nabla g(\bar{x}) + H^\top y = 0$, $\lambda g(\bar{x}) = 0$, $A + \lambda B$ is positive semidefinite on $\ker H$.*

*Proof.* $(a) \Longrightarrow (b)$: By Theorem 3.3.4, strong duality holds, thus, there exists $\lambda \geq 0$ such that

$$f(\bar{x}) + \lambda g(\bar{x}) \leq f(\bar{x}) = \inf_{x \in C}(f(x) + \lambda g(x)).$$

This implies that $\lambda g(\bar{x}) = 0$ and $\bar{x}$ is a minimum for $L(x) = f(x) + \lambda g(x)$ on $C$. The necessary optimality condition yields

$$\langle \nabla f(\bar{x}) + \lambda \nabla g(\bar{x}), x - \bar{x} \rangle \geq 0 \quad \forall\, x \in C.$$

Since $x - \bar{x} \in \ker H$ for all $x \in C$, we obtain $\nabla f(\bar{x}) + \lambda \nabla g(\bar{x}) \in (\ker H)^{\perp} = H^{\top}(\mathbb{R}^m)$. Thus, there exists $y \in \mathbb{R}^m$ such that $\nabla f(\bar{x}) + \lambda \nabla g(\bar{x}) + H^{\top} y = 0$. On the other hand, we also have $f(x) + \lambda g(x) \geq f(\bar{x})$ for all $x \in C$, which gives $\lambda g(\bar{x}) = 0$ and $v^{\top}(A + \lambda B)v \geq 0$ for all $v \in \ker H$, i.e., $A + \lambda B$ is positive semidefinite on $\ker H$.
$(b) \Longrightarrow (a)$: Setting $L(x) = f(x) + \lambda g(x)$, $x \in C$, we write

$$L(x) - L(\bar{x}) = \langle \nabla f(\bar{x}) + \lambda^* \nabla g(\bar{x}), x - \bar{x} \rangle + \frac{1}{2}\langle (A + \lambda B)(x - \bar{x}), x - \bar{x} \rangle.$$

By taking into account that $x - \bar{x} \in \ker H$ for all $x \in C$ and the assumptions, the previous equality implies that

$$f(x) \geq L(x) \geq L(\bar{x}) = f(\bar{x}) + \lambda g(\bar{x}) = f(\bar{x}), \quad \forall\, x \in C,\; g(x) \leq 0,$$

which yields $f(x) \geq f(\bar{x})$, proving that $\bar{x}$ is a solution to (3.13). □

By applying Theorem 3.3.1, we re-obtain Theorem 3.8 in [51] which generalizes the Moré theorem [66, Theorem 3.4].

**Corollary 3.3.10.** [*51, Theorem* 3.8] (*Under Slater condition*) *Let $f$, $g$ be quadratic functions with $\mu \in \mathbb{R}$. Assume that $Hx_0 = d$ and $g(x_0) < 0$ for some $x_0 \in \mathbb{R}^n$, and let $\bar{x} \in K$ (feasible for problem* (3.13)*). Then, the following assertions are equivalent:*

(a) $\bar{x} \in \operatorname*{argmin}_K f$;

(b) *$\exists\, \lambda \geq 0\; \exists\, y \in \mathbb{R}^m$ such that $\nabla f(\bar{x}) + \lambda \nabla g(\bar{x}) + H^{\top} y = 0$, $\lambda g(\bar{x}) = 0$ and $A + \lambda B$ is positive semidefinite on* $\ker H$.

*Proof.* In case $S_f^-(\mu) = \emptyset$ the result is a consequence of Theorem 3.3.5. If $S_f^-(\mu) \neq \emptyset$ we need to check that $s < 0$ and the result again is a consequence of Theorem 3.3.5. Suppose on the contrary that $s = 0$. Then, by the convexity of $\operatorname{cone}(F(\mathbb{R}^n) - \mu(1,0) + \mathbb{R}_{++})$ (see Theorem 3.3.1 or ($a6$) of Theorem 3.3.3), we must have

$$S_f^+(\mu) \cap S_g^-(0) = \emptyset \quad \text{and} \quad \operatorname*{argmin}_K f \cap S_g^-(0) = \emptyset.$$

This implies that $S_g^-(0) = \emptyset$ by Proposition 3.2.1, contradicting the Slater condition. Therefore, $s < 0$, and the conclusion follows. □

For completeness we establish a characterization of solutions when Slater condition fails, that is,

$$g(x) \geq 0, \quad \forall\, x \in H^{-1}(d). \tag{3.25}$$

Under this assumption,

$$K = \{x \in H^{-1}(d) :\ g(x) = 0\} = \operatorname*{argmin}_{H^{-1}(d)} g, \tag{3.26}$$

provided $K \neq \emptyset$. By Corollary 3.4.2, for $\bar{x} \in H^{-1}(d)$,

$$\bar{x} \in \operatorname*{argmin}_{H^{-1}(d)} g \Longleftrightarrow \left\{ \begin{array}{c} B \text{ is positive semidefinite on } \ker H \text{ and} \\ \exists\, y \in \mathbb{R}^m,\ B\bar{x} + b + H^\top y = 0. \end{array} \right. \tag{3.27}$$

Therefore, if $B$ is positive semidefinite on $\ker H$, then

$$\bar{x} \in \operatorname*{argmin}_{K} f \Longleftrightarrow \exists\, y \in \mathbb{R}^m,\ (\bar{x}, y) \in \operatorname*{argmin}_{\tilde{K}} \tilde{f}, \tag{3.28}$$

where

$$\tilde{K} = \left\{ (x, y) \in \mathbb{R}^n \times \mathbb{R}^m :\ \begin{pmatrix} B & H^\top \\ H & 0 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} -b \\ d \end{pmatrix} \right\} \text{ and}$$

$$\tilde{f}(x, y) = \frac{1}{2}(x \;\; y) \begin{pmatrix} A & 0 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} + (a \;\; 0) \begin{pmatrix} x \\ y \end{pmatrix} + \alpha = f(x).$$

Hence, an application of Corollary 3.4.2 to $\tilde{f}$ and $\tilde{K}$ instead of $f$ and $K$, respectively, leads to the following corollary.

**Corollary 3.3.11.** (*Slater condition fails*) *Let $f$, $g$ be quadratic functions with $\mu \in \mathbb{R}$ and $\bar{x} \in K$. Assume that* (3.25) *holds. Then the following statements are equivalent:*

(*a*) $\bar{x} \in \operatorname*{argmin}_{K} f$;

(*b*) *$B$ is positive semidefinite on* $\ker H$*, $A$ is positive semidefinite on* $\ker H \cap B^{-1}[(\ker H)^\perp]$*, and* $\exists\, v \in \ker H$ *and* $\exists\, (y, z) \in \mathbb{R}^m \times \mathbb{R}^m$ *such that*

$$A\bar{x} + a + Bv + H^\top z = 0 \text{ and } B\bar{x} + b + H^\top y = 0.$$

## 3.4 The Frank-Wolfe and Eaves theorems revisited

In this section motivated by the form of the Lagrangian introduced in the previous section, we review the Frank Wolfe theorem [37], by providing several equivalences to the nonemptiness of the solution set, in contrast to the only equivalence between (*a*) and (*d*) (of Theorem 3.4.1) established by Frank Wolfe, or Blum Oettli (the latter authors use elementary analysis in their proof). We believe that our proof is still shorter than that by Blum Oettli [12], and it is suitable for expository purposes. The original proof of Frank Wolfe theorem requires a decomposition theorem for convex polyhedra. Furthermore, it is said that a subset $K \subseteq \mathbb{R}^n$ is *asymptotically linear* [4, Definition 2.3.1] if for all

$\rho > 0$ and all sequence $x_k \in K$, satisfying $\|x_k\| \to +\infty$, $\dfrac{x_k}{\|x_k\|} \to v \in K^\infty$, there exists $k_0 \in \mathbb{N}$, such that $x_k - \rho v \in K$ for all $k \geq k_0$. Here, $K^\infty$ is the asymptotic cone of $K$ defined as in Section 3.2.

Observe that polyhedral sets are asymptotically linear, but there are asymptotically linear sets that are not polyhedral, see after Definition 2.3.2 in [4]. For instance, convex sets without lines (see [4]).

The next theorem is a refinement of the Frank Wolfe theorem when the constraints set is asymptotically linear, and likewise it improves some of the main results of Section 3 in [26].
Other extensions in different directions of the Frank-Wolfe theorem may be found in [77, 7].

**Theorem 3.4.1.** *Let $K \subseteq \mathbb{R}^n$ be closed, convex and asymptotically linear; $h(x) = \frac{1}{2}x^\top Ax + a^\top x + \alpha$ with $A \in S^n$, $a \in \mathbb{R}^n$, $\alpha \in \mathbb{R}$. The assertions $(a)--(d)$ are equivalent, where*

- $(a)$ $-\infty < \nu \doteq \inf_{x\in K} h(x)$;
- $(b)$ *$A$ is copositive on $K^\infty$ and* $[v^\top Av = 0,\ v \in K^\infty \Longrightarrow (Ax + a)^\top v \geq 0 \ \ \forall\, x \in K]$;
- $(c)$ *$A$ is copositive on $K^\infty$ and* $[v^\top Av = 0,\ v \in K^\infty,\ (Ax + a)^\top v \leq 0,\ x \in K \Longrightarrow (Ax + a)^\top v = 0]$;
- $(d)$ $\operatorname*{argmin}_K h \neq \emptyset$.

*Furthermore, we have $(e) \Longrightarrow (f) \Longrightarrow (g) \Longrightarrow (h)$, where*

- $(e)$ *$h$ is coercive, i.e.,* $\lim_{\substack{\|x\|\to+\infty \\ x\in K}} h(x) = +\infty$;
- $(f)$ *$A$ is copositive on $K^\infty$ and* $[v^\top Av = 0,\ v \in K^\infty,\ (Ax + a)^\top v \leq 0,\ x \in K \Longrightarrow v = 0]$.
- $(g)$ $\operatorname*{argmin}_K h$ *is nonempty and bounded;*
- $(h)$ *$A$ is copositive on $K^\infty$ and* $[v^\top Av = 0,\ v \in K^\infty,\ (Ax + a)^\top v \leq 0,\ \forall\, x \in K \Longrightarrow v = 0]$.

*It is worth noticing that under convexity on $h$, i.e., positive semidefiniteness of $A$ (which infers that: $v^\top Av = 0$ if and only if $v \in \ker A$), one obtains $(h) \Longrightarrow (e)$, and therefore all of them are equivalent. In general, $(h)$ does not imply $(e)$ as Example 3.4.1 shows.*

*Proof.* $(a) \Longrightarrow (b)$: Let us prove first that $A$ is copositive on $K^\infty$. For $x_0 \in K$ and $v \in K^\infty$, we obtain by assumption,

$$h(x_0 + tv) = h(x_0) + t\langle \nabla h(x_0), v\rangle + \frac{1}{2}t^2 v^\top Av \geq \nu \quad \forall\, t \in \mathbb{R}. \tag{3.29}$$

Thus,

$$\frac{1}{t^2}h(x_0) + \frac{1}{t}\langle \nabla h(x_0), v\rangle + \frac{1}{2}v^\top Av \geq \frac{\nu}{t^2} \quad \forall\, t \in \mathbb{R},\ t \neq 0.$$

Letting $t \to +\infty$, we get $v^\top Av \geq 0$ for all $v \in K^\infty$, proving that $A$ is copositive on $K^\infty$. Take $v \in K^\infty$ such that $v^\top Av = 0$, then from (3.29) we obtain, $(Ax_0 + a)^\top v \geq 0$, concluding that $(b)$ holds.
$(b) \implies (c)$: It is straightforward.
$(b) \implies (d)$: For every $k \in \mathbb{N}$, setting $B_k \doteq \{x \in K : \|x\| \leq k\}$, we may assume that $B_k \neq \emptyset$ for all $k \in \mathbb{N}$. Let us consider the problem

$$\inf_{x \in B_k} h(x), \tag{3.30}$$

which always has solution. Let $x_k$ be such that

$$\|x_k\| = \min\{\|x\| : x \in \operatorname*{argmin}_{B_k} h\}.$$

**Case 1.** $\sup_{k\in\mathbb{N}} \|x_k\| < \infty$. One can check that any limit point of $(x_k)$ belongs to $\operatorname*{argmin}_K h$.
**Case 2.** $\sup_{k\in\mathbb{N}} \|x_k\| = +\infty$. We can assume that $\|x_k\| \to +\infty$ and $\dfrac{x_k}{\|x_k\|} \to v$ as $k \to +\infty$, thus $v \in K^\infty$. Since $K$ is asymptotically linear given $\rho > 0$ there exists $k_0$ such that $x_k - \rho v \in K$ for all $k \geq k_0$. We can also assume that $\left\|\dfrac{x_k}{\|x_k\|} - v\right\| < 1$ and $\dfrac{\rho}{\|x_k\|} < 1$ for all $k \geq k_0$. Then, by writing

$$x_k - \rho v = \Big(1 - \frac{\rho}{\|x_k\|}\Big)x_k + \frac{\rho}{\|x_k\|}\Big(x_k - \|x_k\|v\Big), \tag{3.31}$$

we get $\|x_k - \rho v\| < \|x_k\|$. On the other hand, given any $x \in K$, there exists $k_1 \in \mathbb{N}$ such that

$$h(x_k) = \frac{1}{2}x_k^\top Ax_k + a^\top x_k + \alpha \leq h(x), \ \forall\ k \geq k_1.$$

It follows that $v^\top Av \leq 0$ and so by the copositive assumption $v^\top Av = 0$. Again by assumption we have $(Ax + a)^\top v \geq 0$ for all $x \in K$.
Set $u_k \doteq x_k - \rho v$. Then, for all $k \geq k_0$, $u_k \in K$, $\|u_k\| < \|x_k\|$ and

$$h(u_k) = h(x_k - \rho v) = h(x_k) - \rho(Ax_k + a)^\top v + \rho^2 v^\top Av \leq h(x_k).$$

This means that $u_k \in \operatorname*{argmin}_{B_k} h$ for all $k$ sufficiently large, contradicting the choice of $x_k$.
Consequently, Case 2 cannot happen, and hence $\operatorname*{argmin}_K h \neq \emptyset$.
$(d) \implies (a)$: It is straightforward.
$(e) \implies (f)$: Evidently the coercive property of $h$ implies the first part of $(f)$, and the second part easily follows as well.
$(f) \implies (g)$: That $\operatorname*{argmin}_K h \neq \emptyset$ follows from $(c)$ implies $(d)$. Suppose there exists a sequence of minimizers $x_k$ such that $\|x_k\| \to +\infty$. Up to a subsequence we may assume that $\dfrac{x_k}{\|x_k\|} \to v \in K^\infty \backslash \{0\}$. From the equality $h(x_k) = \nu$ it follows that $v^\top Av = 0$. On the other hand, by the classical optimality condition, $\nabla h(x_k)^\top (x - x_k) \geq 0$ for all $x \in K$. Given $\rho > 0$, as above, we choose $k$ sufficiently large such that $x_k - \rho v \in K$. Thus $(Ax_k + a)^\top v \leq 0$, which by assumption yields $v = 0$, giving a

contradiction.
$(g) \Longrightarrow (h)$: The first part of $(h)$ is a consequence of $(a)$ implies $(b)$. Take $v \in K^\infty$ satisfying $v^\top Av = 0$ and $(Ax+a)^\top v \leq 0$ for all $x \in K$. We suppose on the contrary that $v \neq 0$. From the equality in (3.29) for $x_0$ to be a minimizer, we deduce that $h(x_0 + tv) \leq h(x_0)$ for all $t > 0$, which says that $x_0 + tv \in \operatorname*{argmin}_K h$ for all $t > 0$, which is not possible if $\operatorname*{argmin}_K h$ is bounded and $v \neq 0$. □

The following example show that the reverse implications in the preceding theorem need not to be true in general.

**Example 3.4.1.** *This example shows, that in general $(h)$ does not imply $(g)$. Take $h_1(x_1,x_2) = x_1^2 - x_2^2$, $K_1 = \{(x_1,x_2) \in \mathbb{R}^2 : \; |x_1 - x_2| \leq 1\}$. Thus $K_1^\infty = \{(x,x) \in \mathbb{R}^2 : \; x \in \mathbb{R}\}$. One can easily check that $(h)$ holds but $\operatorname*{argmin}_{K_1} h_1 = \emptyset$.*

The case $K = H^{-1}(d)$ deserves a special attention, and it is in connection with the Lagrangian appeared in Section 3.3.

**Corollary 3.4.2.** *Let $h$ be as above. The following assertions are equivalent:*

- $(a)$ $-\infty < \nu \doteq \inf_{x \in H^{-1}(d)} h(x)$;
- $(b)$ *$A$ is positive semidefinite on* $\ker H$ *and* $[v^\top Av = 0, \; v \in \ker H \Longrightarrow (Ax+a)^\top v = 0 \;\; \forall\, x \in H^{-1}(d)]$;
- $(c)$ *$A$ is positive semidefinite on* $\ker H$ *and* $[v^\top Av = 0, \; v \in \ker H \Longrightarrow \exists\, y \in \mathbb{R}^m : \;\; Av = H^\top y$ and $d^\top y + a^\top v = 0]$;
- $(d)$ $\operatorname*{argmin}_{H^{-1}(d)} h \neq \emptyset$;
- $(e)$ *$A$ is positive semidefinite on* $\ker H$ *and there exist $\bar{x} \in H^{-1}(d)$, $y \in \mathbb{R}^m$ such that $A\bar{x} + a + H^\top y = 0$.*

*Proof.* By virtue of the previous theorem we need only to check $(b) \Longleftrightarrow (c)$ and $(d) \Longrightarrow (e) \Longrightarrow (a)$. The equivalence between $(b)$ and $(c)$ follows as in $(b)$ of Proposition 3.3.4.
$(d) \Longrightarrow (e)$: Let $\bar{x} \in \operatorname*{argmin}_{H^{-1}(d)} h$. Then by the usual necessary optimality condition, we have $\langle \nabla h(\bar{x}), x - \bar{x} \rangle \geq 0$ for all $x \in H^{-1}(d)$. Since $x - \bar{x} \in \ker H$ for all $x \in H^{-1}(d)$, we get $A\bar{x} + a = \nabla h(\bar{x}) \in (\ker H)^\perp = H^\top(\mathbb{R}^m)$. Hence there exists $y \in \mathbb{R}^m$ such that $A\bar{x} + a + H^\top y = 0$, which is the desired result.
$(e) \Longrightarrow (a)$: it is straightforward, once we notice that $H^{-1}(d) = \bar{x} + \ker H$ and

$$h(x + \bar{x}) = h(\bar{x}) + \langle \nabla h(\bar{x}), x \rangle + \frac{1}{2} x^\top Ax, \quad x \in \ker H.$$

□

When $H$ is the null matrix and $d = 0$, the previous result admits a more precise formulation as expressed in the following corollary. Recall that when $A$ is positive semidefinite ($A \succcurlyeq 0$), we have

$$v^\top A v = 0 \iff v \in \ker A.$$

**Corollary 3.4.3.** *Let $h$ be as above. The following assertions are equivalent:*

(*a*) $-\infty < \nu \doteq \inf_{x \in \mathbb{R}^n} h(x)$;

(*b*) $A \succcurlyeq 0$ *and* $[v \in \ker A \Longrightarrow a^\top v = 0]$;

(*c*) $\operatorname*{argmin}_{\mathbb{R}^n} h \neq \emptyset$;

(*d*) $A \succcurlyeq 0$ *and there exists* $\bar{x} \in \mathbb{R}^n$ *such that* $A\bar{x} + a = 0$.

# Chapter 4

# Single equality and geometric constraints

## 4.1 Introduction

Let $X$ be a real Hausdorff topological space and $C$ be nonempty subset of $X$. Given two functions $f : X \to \mathbb{R} \cup \{+\infty\}$ and $g : X \to \mathbb{R}$, with $C$ being a subset of the effective domain of $f$, dom $f$, let us consider the following minimization problem with a single equality and geometric constraints:

$$\mu \doteq \inf\{f(x) : \quad g(x) = 0,\ x \in C\}, \tag{4.1}$$

and its (Lagrangian) dual problem

$$\nu \doteq \sup_{\lambda \in \mathbb{R}} \inf_{x \in C} \{f(x) + \lambda g(x)\}. \tag{4.2}$$

We assume that the feasible set $K \doteq \{x \in C : \quad g(x) = 0\}$ is nonempty. It is known that several important models in real-life problems can be formulated as (4.1), and there are some devices allowing us to re-write problems with a finite number of inequality constraints (or more generally a cone constrained minimization problem) into one with a single inequality constraint. Observe that the latter problem may be considered as a specialization of one with a single equality constraint, and that problem (4.1) encompasses that of trust-region subproblems well-known in quadratic programming, in addition to the Portfolio problem.

To be more precise, we will discuss the validity of the zero duality gap and strong duality properties, where by zero duality gap, we mean $\mu = \nu$; whereas strong duality means that problem (4.2) admits solution and $\mu = \nu$.

One of our main tasks is to characterize the strong duality property without convexity assumptions, from geometrical and topological points of view.

Since

$$\inf_{x\in C}\{f(x)+\lambda g(x)\}\leq \mu,\ \ \forall\ \lambda\in\mathbb{R}, \tag{4.3}$$

in case $\mu=-\infty$, there is no duality gap because of $\nu=-\infty$ as well; and from (4.3), we conclude that any real number is a solution for the problem (4.2), and therefore, strong duality always holds for (4.1) whenever $\mu=-\infty$. Thus, we assume throughout that $\mu$ is finite.

When our approach applies to a quadratic programming problem under a single inequality constraint, that is, $f$ and $g$ are quadratic functions (setting $C=\mathbb{R}^n$), it provides new results without using the semidefinite programming approach to the dual (4.2), which is based on a Schur complement, see [21, Appendix B], [84, 55, 95]. For instance, under a Slater-type condition, and for general $f$, $g$ and $C$, we establish that strong duality is valid if and only if the conic hull of $(f,g)(C)-\mu(1,0)+\mathbb{R}_+(1,0)$ is convex, a first characterization of this kind, see Corollary 4.3.8.

Once the characterization of strong duality is established in a general context, it will be applied to find equivalent formulations to the validity of KKT optimality conditions for the same problem, since such a validity is characterized in terms of strong duality for a suitable linear approximation of the given problem. Thus, our results will be useful to deal with models where some theorems based on exact penalization techniques ([91, 65]) cannot be applied, see Example 4.4.3; or where classical constraint qualifications fail (see Example 4.4.4), which occurs, for instance, in structural optimization [2].

Just to point out one concrete application, let us consider one of the simplest but important problems (with a single equality constraint), which is well-known as the standard quadratic problem

$$\min_{x\in\Delta}\frac{1}{2}x^\top Ax, \tag{4.4}$$

where, $\Delta$ is the simplex $\{x\in\mathbb{R}^n:\ \mathbf{1}^\top x=1,\ x\geq 0\}$ with $A$ being any real symmetric matrix with positive entries and $\mathbf{1}^\top \doteq (1,\ldots,1)\in\mathbb{R}^n$ is the vector with components equal to one. Such a problem retains, as asserted in [14], most of the complexity of the general quadratic case having a polyhedron as a feasible set. As applications of (4.4), we mention quadratic allocation problems [47], portfolio optimization problems [61, 62], the maximum weight clique problem [67, 43], among others. In Section 4.5, our results will apply to a more general problem than (4.4) where $\Delta$ is substituted by a convex and compact base of any pointed, closed, convex (possibly non polyhedral) cone $C$. In particular, we will prove the validity of strong duality and provide further qualitative and quantitative information, in addition to the uniqueness of the KKT multiplier in our sense.

Our line of reasoning follows previous ones and is based on a careful analysis of the structure of the problem. Nevertheless, the method employed in [31] (for a single inequality constraint) cannot be applied to our problem (4.1), since equality constraints are more difficult to handle. We actually explore the hidden convexity of the underlying problem under the presence of strong duality. Thus, unlike some of the results appeared elsewhere, which involve conditions on $g$ and $C$ that guarantee (4.1) has strong duality for every $f$ in a certain class of functions, our approach allows us to derive conditions on $f$, $g$ and $C$, jointly, that ensure (4.1) has strong duality under no convexity assumption. Thus, we provide results where none of those in [38, 16, 50, 48, 17, 18, 19, 51, 49, 52] is applicable.

Various sufficient conditions for the zero duality gap have been also established in the literature, see [38, 3, 4, 94, 16, 18, 19, 85, 27]. By using a Shor' scheme and the semidefinite programming relaxation a characterization of zero duality gap in quadratic programming under inequality constraints may be found in [95]. A different approach is applied in [55] to establish conditions ensuring zero duality gap.

## 4.2 Some preliminaries

Denote $F \doteq (f, g)$. Assuming that $\mu \in \mathbb{R}$, we obtain

$$(F(C) - \mu(1,0)) \cap -(\mathbb{R}_{++} \times \{0\}) = \emptyset, \tag{4.5}$$

which amounts to writing

$$(F(C) + \mathbb{R}_{+}(1,0) - \mu(1,0)) \cap -(\mathbb{R}_{++} \times \{0\}) = \emptyset, \tag{4.6}$$

or, equivalently,

$$\text{cone}(F(C) + \mathbb{R}_{+}(1,0) - \mu(1,0)) \cap -(\mathbb{R}_{++} \times \{0\}) = \emptyset. \tag{4.7}$$

Consequently, $\text{cone}(F(C) + \mathbb{R}_{+}(1,0) - \mu(1,0)) \neq \mathbb{R}^2$.

The following result, is a particular case of Theorem 3.2 in [35]

**Proposition 4.2.1.** *Assume that $\mu$ is finite. The following assertions are equivalent*

(*a*) *Strong duality holds for* (4.1), *that is*

$$\exists \lambda^* \in \mathbb{R}: \ f(x) + \lambda^* g(x) \geq \mu;$$

(*b*) $\overline{\text{cone}}\ \text{co}(F(C) + \mathbb{R}_{+}(1,0) - \mu(1,0)) \cap (-\mathbb{R}_{++} \times \{0\}) = \emptyset.$

We first recall a bidimensional optimal alternative theorem valid for convex cones possibly with empty interior proved in [33].

**Theorem 4.2.1.** ([*33, Theorem* 2.4] *Let $P \subseteq \mathbb{R}^2$ be a convex cone and $A \subseteq \mathbb{R}^2$ be such that* $\text{int}(\text{cone}_{+}(A+P)) \neq \emptyset$. *The following assertions are equivalent:*

(*a*) $0 \notin \text{int}(\overline{\text{cone}}(A+P))$ *and* $\overline{\text{cone}}(A+P)$ *is convex;*

(*b*) $\exists\ p^* \in P^* \setminus \{0\}$ *such that* $\langle p^*, a\rangle \geq 0 \ \ \forall\ a \in A.$

A remark is in order when we particularize to our minimization problem (4.1).

**Remark 4.2.2.** *Let us analyze the assumption of the previous theorem when it is applied to problem* (4.1)*, that is,* $A = F(C) - \mu(1,0)$ *and* $P = \mathbb{R}_+(1,0)$*. It is easy to see that*

$$\begin{aligned} \mathrm{int}[\mathrm{cone}_+(F(C) - \mu(1,0) + \mathbb{R}_+(1,0))] = \emptyset &\iff F(C) - \mu(1,0) \subseteq \mathbb{R}_+(1,0) \\ &\iff g(x) = 0 \ \ \forall\, x \in C \\ &\iff K = C. \end{aligned}$$

*Hence, Theorem 4.2.1 is applicable to problem* (4.1) *when* $K \neq C$*.*

## 4.3 Strong duality: Geometric and topological characterization

We first describe the Lagrangian duality approach for the minimization problem (4.1) under one single equality and geometric constraints in a general setting, and then it will be particularized to quadratic nonconvex programming.

We associate to problem (4.1) the usual linear Lagrangian

$$L(\gamma, \lambda, x) \doteq \gamma f(x) + \lambda g(x),$$

with $\gamma \geq 0$ and $\lambda \in \mathbb{R}$.

We obtain the trivial inequality

$$\inf_{x \in C} L(\gamma, \lambda, x) \leq \inf_{x \in K} L(\gamma, \lambda, x) = \gamma \inf_{x \in K} f(x), \quad \forall\, \gamma \geq 0,\ \forall\, \lambda \in \mathbb{R}. \tag{4.8}$$

In order to get the equality, we need to find conditions under which the reverse inequality holds, that is, we must have:

$$\gamma(f(x) - \mu) + \lambda g(x) \geq 0 \quad \forall\, x \in C. \tag{4.9}$$

This will imply strong duality once we get $\gamma > 0$.

The sets

$$\mathcal{F} \doteq F(C) + \mathbb{R}_+(1,0),\ \ \mathcal{F}_\mu \doteq \mathcal{F} - \mu(1,0), \tag{4.10}$$

will play an important role in our analysis.

Set $\rho \doteq (\gamma, \lambda)$. Then, (4.9) can be written equivalently as

$$\langle \rho, a \rangle \geq 0 \quad \forall\, a \in \mathcal{F}_\mu,$$

which means by virtue of Theorem 2.1.2 $(e)$,

$$(\gamma, \lambda) \in [\overline{\mathrm{cone}}\ \mathcal{F}_\mu]^*. \tag{4.11}$$

Set

$$\mathcal{L}_{SD} \doteq \Big\{\lambda \in \mathbb{R}:\ (1,\lambda) \in [\text{cone } \mathcal{F}_\mu]^*\Big\}. \tag{4.12}$$

Then, (4.1) has the strong duality property if, and only if $\mathcal{L}_{SD} \neq \emptyset$. Hence

$$\mathcal{L}_{SD} \subseteq \mathcal{S}_D, \tag{4.13}$$

where $\mathcal{S}_D$ is the solution set to the dual problem (4.2). Notice that $\mathcal{L}_{SD} = \mathcal{S}_D$, whenever $\mathcal{L}_{SD} \neq \emptyset$.

We will split the set cone $\mathcal{F}_\mu$ in order to describe its convexity. To that purpose some notation in Section 3.4 is needed, i.e., $S_g^-(0)$, $S_g^=(0)$, $S_g^+(0)$, $S_f^-(\mu)$, $S_f^=(\mu)$ and $S_f^+(\mu)$.

We notice that

$$C = \operatorname*{argmin}_K\ f \cup (K \setminus \operatorname*{argmin}_K\ f) \cup (C \setminus K).$$

Obviously

$$K \setminus \operatorname*{argmin}_K\ f = S_f^+(\mu) \cap S_g^=(0);$$

and

$$\begin{aligned} C \setminus K &= \Big(S_f^-(\mu) \cap S_g^-(0)\Big) \cup \Big(S_f^-(\mu) \cap S_g^+(0)\Big) \cup \big(S_f^=(\mu) \cap S_g^-(0)\big) \\ &\cup\ \big(S_f^=(\mu) \cap S_g^+(0)\big) \cup \Big(S_f^+(\mu) \cap S_g^-(0)\Big) \cup \Big(S_f^+(\mu) \cap S_g^+(0)\Big). \end{aligned}$$

We set

$$\Omega_-^- \doteq S_f^-(\mu) \cap S_g^-(0);\ \ \Omega_+^- \doteq S_f^-(\mu) \cap S_g^+(0);\ \ \Omega_-^= \doteq S_f^=(\mu) \cap S_g^-(0).$$

Similarly for $\Omega_+^=$, $\Omega_-^+$, $\Omega_+^+$. Thus, $F(C) - \mu(1,0) + \mathbb{R}_+(1,0) = \Omega_1 \cup \Omega_2 \cup \Omega_3$, where

$$\begin{aligned} \Omega_1 &= F(\operatorname*{argmin}_K\ f) - \mu(1,0) + \mathbb{R}_+(1,0) = \mathbb{R}_+(1,0),\ \text{provided } \operatorname{argmin}_K\ f \neq \emptyset; \\ \Omega_2 &= F(\Omega_=^+) - \mu(1,0) + \mathbb{R}_+(1,0); \\ \Omega_3 &= [F(\Omega_-^-) - \mu(1,0) + \mathbb{R}_+(1,0)] \cup [F(\Omega_+^-) - \mu(1,0) + \mathbb{R}_+(1,0)] \\ &\cup\ [F(\Omega_-^=) - \mu(1,0) + \mathbb{R}_+(1,0)] \cup [F(\Omega_+^=) - \mu(1,0) + \mathbb{R}_+(1,0)] \\ &\cup\ [F(\Omega_-^+) - \mu(1,0) + \mathbb{R}_+(1,0)] \cup [F(\Omega_+^+) - \mu(1,0) + \mathbb{R}_+(1,0)]. \end{aligned} \tag{4.14}$$

Evidently

$$\overline{\text{cone}}(\mathcal{F}_\mu) = \overline{\text{cone}}(\Omega_1) \cup \overline{\text{cone}}(\Omega_2) \cup \overline{\text{cone}}(\Omega_3),$$

and

$$\overline{\text{cone}}(\Omega_1) \cup \overline{\text{cone}}(\Omega_2) = \mathbb{R}_+(1,0).$$

Thus,

$$\overline{\text{cone}}(\mathcal{F}_\mu) = \overline{\text{cone}}(\Omega_3) \cup \mathbb{R}_+(1,0). \tag{4.15}$$

The next result provides simple expressions for $\overline{\text{cone}}(\mathcal{F}_\mu)$, and will be exploited along the proof of our main theorem in this section.

**Proposition 4.3.1.** *Let $f, g$ be as above with $\mu$ being finite. The following hold:*

(*a*) *if $C \neq K$ then $\overline{\text{cone}}(\mathcal{F}_\mu) = \overline{\text{cone}}(\Omega_3)$;*

(*b*) *if $C = K$ then $F(C) - \mu(1,0) + \mathbb{R}_{++}(1,0) = \mathbb{R}_{++}(1,0)$, and so*

$$\overline{\text{cone}}(\mathcal{F}_\mu) = \mathbb{R}_+(1,0).$$

*Proof.* ($a$): Notice that (see ($g$) of Proposition 2.1 in [33])

$$\mathbb{R}_+(1,0) \subseteq \overline{\text{cone}}(F(C \setminus K) - \mu(1,0) + \mathbb{R}_+(1,0)),$$

since every $z \in \mathbb{R}_+(1,0)$ is the limit of $\frac{1}{k}(a + kz)$ for fixed $a \in F(C \setminus K) - \mu(1,0)$. But $\overline{\text{cone}}(F(C \setminus K) - \mu(1,0) + \mathbb{R}_+(1,0)) = \overline{\text{cone}}(\Omega_3)$. Thus $\overline{\text{cone}}(\mathcal{F}_\mu) = \overline{\text{cone}}(\Omega_3)$ by (4.15).

($b$): In case $C = K$, by Remark 4.2.2,

$$F(C) - \mu(1,0) + \mathbb{R}_{++}(1,0) \subseteq \mathbb{R}_{++}(1,0).$$

Let $(u,v) \in \mathbb{R}_{++} \times \{0\}$. Then, there exists $x_0 \in C$ satisfying $f(x_0) < \mu + u$. Thus, $(u,v) \in F(x_0) - \mu(1,0) + \mathbb{R}_{++}(1,0) \subseteq F(C) - \mu(1,0) + \mathbb{R}_{++}(1,0)$, and the conclusion follows. □

Furthermore, we need the following numbers:

- if $\Omega_+^- \neq \emptyset$,

$$s \doteq \sup_{x \in \Omega_+^-} \frac{g(x)}{f(x) - \mu} \in \,]-\infty, 0];$$

- if $\Omega_-^+ \neq \emptyset$,

$$r \doteq \inf_{x \in \Omega_-^+} \frac{g(x)}{f(x) - \mu} \in [-\infty, 0[;$$

- if $\Omega_-^- \neq \emptyset$,

$$l \doteq \inf_{x \in \Omega_-^-} \frac{g(x)}{f(x) - \mu} \in [0, +\infty[;$$

- if $\Omega_+^+ \neq \emptyset$,

$$m \doteq \sup_{x \in \Omega_+^+} \frac{g(x)}{f(x) - \mu} \ \in \, ]0, +\infty].$$

Observe that, if $\nu \in \mathbb{R}$ and $\lambda^* \in \mathcal{S}_D$, then

$$\frac{\nu - f(x)}{|g(x)|} \leq |\lambda^*| \ \ \forall\, x \in C, \ g(x) \neq 0. \tag{4.16}$$

Thus, if $\mu = \nu$, one gets

$$\max\{s, -l\} \leq -\frac{1}{|\lambda^*|} \ \ \forall\, \lambda^* \in \mathcal{S}_D, \ \lambda^* \neq 0. \tag{4.17}$$

We will see that under further assumptions (see Remark 4.3.7) exactly one of the sets $\Omega_+^-$ or $\Omega_-^-$ is nonempty.

The following proposition, whose proof is straightforward, collects some easy useful facts to be used in what follows.

**Proposition 4.3.2.** *For the above data with $\mu$ being finite, we have:*

(*a*) *if $S_f^-(\mu) = \emptyset$ then $\mathcal{F}_\mu \subseteq \mathbb{R}_+ \times \mathbb{R}$, and so $\Omega_-^- = \emptyset = \Omega_+^-$;*

(*b*) *if $m = +\infty$ then $\overline{\text{cone}}(F(\Omega_+^+) - \mu(1,0) + \mathbb{R}_+(1,0)) = \mathbb{R}_+^2$;*

(*c*) *if $r = -\infty$ then $\overline{\text{cone}}(F(\Omega_-^+) - \mu(1,0) + \mathbb{R}_+(1,0)) = \mathbb{R}_+ \times \mathbb{R}_-$;*

(*d*) *if $\Omega_-^= \neq \emptyset \neq \Omega_+^=$ then,*

(*d*1) $\text{cone}(F(\Omega_-^=) - \mu(1,0) + \mathbb{R}_+(1,0)) = (\mathbb{R}_+ \times \mathbb{R}_{--}) \cup \{(0,0)\}$;

(*d*2) $\text{cone}(F(\Omega_+^=) - \mu(1,0) + \mathbb{R}_+(1,0)) = (\mathbb{R}_+ \times \mathbb{R}_{++}) \cup \{(0,0)\}$.

The next result, is important for the main theorem of this section.

**Proposition 4.3.3.** *For the above data with $\mu$ being finite, we have:*

(*a*) *if the sets $\Omega_-^-$, $\Omega_-^=$ and $\Omega_-^+$ are nonempty, then,*

(*a*1) $F(\Omega_-^+) - \mu(1,0) + \mathbb{R}_+(1,0) \subseteq \text{cone}(F(\Omega_-^=) - \mu(1,0) + \mathbb{R}_+(1,0))$;

(*a*2) $F(\Omega_-^=) - \mu(1,0) + \mathbb{R}_+(1,0) \subseteq \text{cone}(F(\Omega_-^-) - \mu(1,0) + \mathbb{R}_+(1,0))$.

(*b*) *if the sets $\Omega_+^-$, $\Omega_+^=$ and $\Omega_+^+$ are nonempty, then,*

(*b*1) $F(\Omega_+^+) - \mu(1,0) + \mathbb{R}_+(1,0) \subseteq \text{cone}(F(\Omega_+^=) - \mu(1,0) + \mathbb{R}_+(1,0))$;

(*b*2) $F(\Omega_+^=) - \mu(1,0) + \mathbb{R}_+(1,0) \subseteq \text{cone}(F(\Omega_+^-) - \mu(1,0) + \mathbb{R}_+(1,0))$.

*Proof.* $(a)$: For the first inclusion let $(u,v) \in F(\Omega_-^+) - \mu(1,0) + \mathbb{R}_+(1,0)$. Then, there exist $x_0 \in \Omega_-^+$ and $p_0 \geq 0$, such that $u = f(x_0) - \mu + p_0$, $v = g(x_0)$. On the other hand, given any $x_1 \in \Omega_-^=$, one may find $t_1 > 0$ satisfying $g(x_0) = t_1 g(x_1)$; thus,

$$u = t_1[(f(x_1) - \mu) + \frac{1}{t_1}(f(x_0) - \mu + p_0)],\ v = t_1 g(x_1),$$

showing the first inclusion. For the second one, take $(u,v) \in F(\Omega_-^=) - \mu(1,0) + \mathbb{R}_+(1,0)$. Then, there exist $x_0 \in \Omega_-^=$ and $p_0 \geq 0$, such that $u = p_0$, $v = g(x_0)$. Since for any $x_1 \in \Omega_-^-$, $g(x_0) = t_1 g(x_1)$ for some $t_1 > 0$, we can express

$$u = t_1[(f(x_1) - \mu) + \frac{p_0}{t_1} - (f(x_1) - \mu)],\ v = t_1 g(x_1),$$

proving the second inclusion.
$(b)$: The first inclusion is a consequence of the following reasoning. Take any $(u,v) \in F(\Omega_+^+) - \mu(1,0) + \mathbb{R}_+(1,0)$, then there exist $x_0 \in \Omega_+^+$ and $p_0 \geq 0$ such that $u = f(x_0) - \mu + p_0$, $v = g(x_0)$. On the other hand, every $x_1 \in \Omega_+^=$ yields $g(x_0) = t_1 g(x_1)$ for some $t_1 > 0$. Thus,

$$u = t_1[(f(x_1) - \mu) + \frac{1}{t_1}(f(x_0) - \mu + p_0)],\ v = t_1 g(x_1),$$

showing the first inclusion. For the second inclusion we follow a similar reasoning. In fact, take any $(u,v) \in F(\Omega_+^=) - \mu(1,0) + \mathbb{R}_+(1,0)$. Then, there exist $x_0 \in \Omega_+^=$ and $p_0 > 0$, such that $u = p_0$, $v = g(x_0)$. Moreover, for any $x_1 \in \Omega_+^-$, we can find $t_1 > 0$ satisfying $g(x_0) = t_1 g(x_1)$; thus

$$u = t_1[(f(x_1) - \mu) + \frac{p_0}{t_1} - (f(x_1) - \mu)],\ v = t_1 g(x_1).$$

This proves $(u,v) \in \operatorname{cone}(F(\Omega_+^-) - \mu(1,0) + \mathbb{R}_+(1,0))$.

□

**Proposition 4.3.4.** *For the above data with $\mu$ being finite, we have:*

(a) *if $-\infty < r < 0$ then*

$$\overline{\operatorname{cone}}(F(\Omega_-^+) - \mu(1,0) + \mathbb{R}_+(1,0)) = \{(u,v) \in \mathbb{R}^2 :\ ru \leq v \leq 0,\ u \geq 0\}; \tag{4.18}$$

(b) *if $0 < m < +\infty$, then*

$$\overline{\operatorname{cone}}(F(\Omega_+^+) - \mu(1,0) + \mathbb{R}_+(1,0)) = \{(u,v) \in \mathbb{R}_+^2 :\ v \leq mu,\ \}; \tag{4.19}$$

(c) *if $0 < l < +\infty$ then*

$$\overline{\operatorname{cone}}(F(\Omega_-^-) - \mu(1,0) + \mathbb{R}_+(1,0)) = \{(u,v) \in \mathbb{R}^2 :\ v \leq lu,\ v \leq 0\}; \tag{4.20}$$

(d) *if $-\infty < s < 0$ then*

$$\overline{\operatorname{cone}}(F(\Omega_+^-) - \mu(1,0) + \mathbb{R}_+(1,0)) = \{(u,v) \in \mathbb{R}^2 :\ v \geq su,\ v \geq 0\}. \tag{4.21}$$

*Proof.* We will prove only $(a)$ and $(b)$, being the others entirely similar.
$(a)$: Obviously

$$F(\Omega^+_-) - \mu(1,0) \subseteq \{(u,v) \in \mathbb{R}^2 : \; ru \le v \le 0, \; u \ge 0\},$$

implying one inclusion. For the other inclusion, let $(u,v)$ be in the set on the right-hand side of (4.18). There exists a sequence $x_k \in \Omega^+_-$ such that $\dfrac{g(x_k)}{f(x_k)-\mu} \to r$. Then, $\dfrac{1}{f(x_k)-\mu}(f(x_k)-\mu, g(x_k)) \to (1,r)$, which implies

$$(1,r) \in \overline{\text{cone}}(F(\Omega^+_-) - \mu(1,0) + \mathbb{R}_+(1,0)).$$

Consequently,

$$(\frac{v}{r}, v) \in \overline{\text{cone}}(F(\Omega^+_-) - \mu(1,0) + \mathbb{R}_+(1,0)).$$

It follows that

$$\begin{aligned}(u,v) = (\frac{v}{r}, v) + (u - \frac{v}{r}, 0) \;\; &\in \;\; \overline{\text{cone}}(F(\Omega^+_-) - \mu(1,0) + \mathbb{R}_+(1,0)) + \mathbb{R}_+(1,0) \\ &\subseteq \;\; \overline{\text{cone}}(F(\Omega^+_-) - \mu(1,0) + \mathbb{R}_+(1,0)).\end{aligned}$$

$(b)$: Similarly as above, it is obvious that

$$F(\Omega^+_+) - \mu(1,0) \subseteq \{(u,v) \in \mathbb{R}^2_+ : \; v \le mu \; \}.$$

For the other inclusion, take any $(u,v)$ in the set on the right-hand side of (4.19). Then, there exists $x_k \in \Omega^+_+$ such that $\dfrac{g(x_k)}{f(x_k)-\mu} \to m$. Thus

$$\frac{1}{f(x_k)-\mu}(f(x_k)-\mu, g(x_k)) \to (1,m),$$

which implies $(1,m) \in \overline{\text{cone}}(F(\Omega^+_+) - \mu(1,0) + \mathbb{R}_+(1,0))$. Hence

$$(\frac{v}{m}, v) \in \overline{\text{cone}}(F(\Omega^+_+) - \mu(1,0) + \mathbb{R}_+(1,0)).$$

By the choice of $(u,v)$, we get

$$\begin{aligned}(u,v) = (\frac{v}{m}, v) + (u - \frac{v}{m}, 0) \;\; &\in \;\; \overline{\text{cone}}(F(\Omega^+_+) - \mu(1,0) + \mathbb{R}_+(1,0)) + \mathbb{R}_+(1,0) \\ &\subseteq \;\; \overline{\text{cone}}(F(\Omega^+_+) - \mu(1,0) + \mathbb{R}_+(1,0)).\end{aligned}$$

□

Before establishing our complete description of the convexity of $\overline{\text{cone}}\,\mathcal{F}_\mu$, we need some additional preliminary results.

**Proposition 4.3.5.** *Let us consider problem* (4.1) *with* $\mu \in \mathbb{R}$ *and* $\overline{\text{cone}}(\mathcal{F}_\mu)$ *being convex.*

(a) *Assume that* $\Omega^-_+ \neq \emptyset$ *and* $s < 0$. *Then,*

($a1$) $\Omega_-^- = \Omega_-^= = \emptyset$;

($a2$) *if additionaly* $\Omega_-^+ \neq \emptyset$ *one has* $s \leq r$.

($b$) *Assume that* $\Omega_-^- \neq \emptyset$ *and* $l > 0$. *Then,*

($b1$) $\Omega_+^- = \Omega_+^= = \emptyset$;

($b2$) *if additionaly* $\Omega_+^+ \neq \emptyset$ *one has* $l \leq m$.

*Proof.* We will prove only ($a$), since ($b$) is similar. By ($d$) of Proposition 4.3.4,

$$\{(u, v) \in \mathbb{R}^2 : \ v \geq su, \ v \geq 0\} \subseteq \overline{\text{cone}}(\mathcal{F}_\mu).$$

Thus, if $\overline{\text{cone}}(\mathcal{F}_\mu)$ is convex then immediately $\Omega_+^- = \Omega_+^= = \emptyset$, proving ($a1$). In case $\Omega_-^+ \neq \emptyset$, one gets $s \leq r$ due to convexity again. □

We now proceed to describe all the situations that may occur under which $\overline{\text{cone}}\ \mathcal{F}_\mu$ is convex. In what follows the slightly dark regions in Figures 4.1 to 4.5, represent $\overline{\text{cone}}(\mathcal{F}_\mu)$, and its polar cone is the darker part. Recall that $\mathcal{L}_{SD} = \{\lambda \in \mathbb{R} : (1, \lambda) \in [\overline{\text{cone}}\ \mathcal{F}_\mu]^*\}$.

**Theorem 4.3.1.** *Let us consider problem (4.1) with* $\mu \in \mathbb{R}$. *Assume that* $\overline{\text{cone}}\ \mathcal{F}_\mu \neq \mathbb{R}^2$. *Then,* $\overline{\text{cone}}\ \mathcal{F}_\mu$ *is convex if, and only if exactly one of the following assertions holds:*

($a1$) $C = K$. *In which case* $F(C) - \mu(1, 0) + \mathbb{R}_{++}(1, 0) = \mathbb{R}_{++}(1, 0)$. *Hence*

$$[\text{cone}\ \mathcal{F}_\mu]^* = \mathbb{R}_+ \times \mathbb{R},$$

*and* $\mathcal{L}_{SD} = \mathcal{S}_D = \mathbb{R}$.

($a2$) $S_f^-(\mu) = \emptyset$ *and either* $\Omega_-^= \neq \emptyset$ *or* $[\Omega_-^+ \neq \emptyset, \ r = -\infty]$ *and either* $[\Omega_+^+ \neq \emptyset, \ m = +\infty]$ *or* $\Omega_+^= \neq \emptyset$. *In which case* $\overline{\text{cone}}(\mathcal{F}_\mu) = \mathbb{R}_+ \times \mathbb{R}$. *Hence*

$$[\text{cone}\ \mathcal{F}_\mu]^* = \mathbb{R}_+ \times \{0\},$$

*and* $\mathcal{L}_{SD} = \mathcal{S}_D = \{0\}$.

($a3$) $S_f^-(\mu) = \Omega_-^= = \emptyset, \ \Omega_-^+ \neq \emptyset, \ -\infty < r < 0$ *and either* $[\Omega_+^+ \neq \emptyset, \ m = +\infty]$ *or* $\Omega_+^= \neq \emptyset$. *In which case* $S_g^-(0) = \Omega_-^+$ *and* $\overline{\text{cone}}(\mathcal{F}_\mu) = \{(u, v) \in \mathbb{R}^2 : \ v \geq ru, \ u \geq 0\}$. *Hence*

$$[\text{cone}\ \mathcal{F}_\mu]^* = \{(\gamma, \lambda) \in \mathbb{R}^2 : \ \gamma \geq 0, \ 0 \leq \lambda \leq -\frac{1}{r}\},$$

*and* $\mathcal{L}_{SD} = \mathcal{S}_D = \{\lambda \in \mathbb{R} : \ 0 \leq \lambda \leq -\frac{1}{r}\}$.

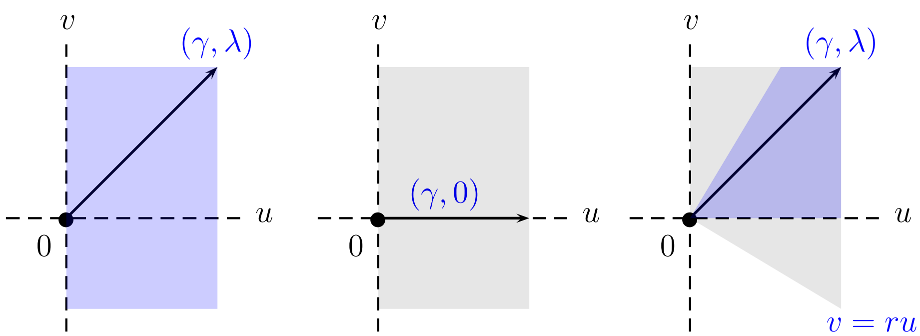


Figure 4.1: Theorem 4.3.1: $(a1)$, $(a2)$, $(a3)$ (figure produced by author)

$(a4)$ *$S_f^-(\mu) = \Omega_+^= = \emptyset,\ \Omega_+^+ \neq \emptyset,\ 0 < m < +\infty$ and either $[\Omega_-^+ \neq \emptyset,\ r = -\infty]$ or $\Omega_-^= \neq \emptyset$. In which case $S_g^+(0) = \Omega_+^+$ and $\overline{\text{cone}}(\mathcal{F}_\mu) = \{(u,v) \in \mathbb{R}^2 :\ v \leq mu,\ u \geq 0\}$. Hence,*

$$[\text{cone } \mathcal{F}_\mu]^* = \{(\gamma, \lambda) \in \mathbb{R}^2 :\ \gamma \geq 0,\ -\frac{1}{m} \leq \lambda \leq 0\},$$

*and $\mathcal{L}_{SD} = \mathcal{S}_D = \{\lambda \in \mathbb{R} :\ -\frac{1}{m} \leq \lambda \leq 0\}$.*

$(a5)$ *$S_f^-(\mu) = \Omega_+^= = \Omega_-^= = \emptyset,\ \Omega_+^+ \neq \emptyset,\ 0 < m < +\infty,\ \Omega_-^+ \neq \emptyset,\ -\infty < r < 0$. In which case $S_g^-(0) = \Omega_-^+$, $S_g^+(0) = \Omega_+^+$ and $\overline{\text{cone}}(\mathcal{F}_\mu) = \{(u,v) \in \mathbb{R}^2 :\ ru \leq v \leq mu,\ u \geq 0\}$. Hence,*

$$[\text{cone } \mathcal{F}_\mu]^* = \{(\gamma, \lambda) \in \mathbb{R}^2 :\ \gamma \geq 0,\ -\frac{1}{m}\gamma \leq \lambda \leq -\frac{1}{r}\gamma\},$$

*and $\mathcal{L}_{SD} = \mathcal{S}_D = \{\lambda \in \mathbb{R} :\ -\frac{1}{m} \leq \lambda \leq -\frac{1}{r}\}$.*

$(a6)$ *$S_g^-(0) = S_f^-(\mu) = \emptyset$ and either $\Omega_+^= \neq \emptyset$ or $[\Omega_+^+ \neq \emptyset,\ m = +\infty]$. In which case $\overline{\text{cone}}(\mathcal{F}_\mu) = \mathbb{R}_+^2$. Hence,*

$$[\text{cone } \mathcal{F}_\mu]^* = \mathbb{R}_+^2,$$

*and $\mathcal{L}_{SD} = \mathcal{S}_D = \{\lambda \in \mathbb{R} :\ 0 \leq \lambda\}$.*

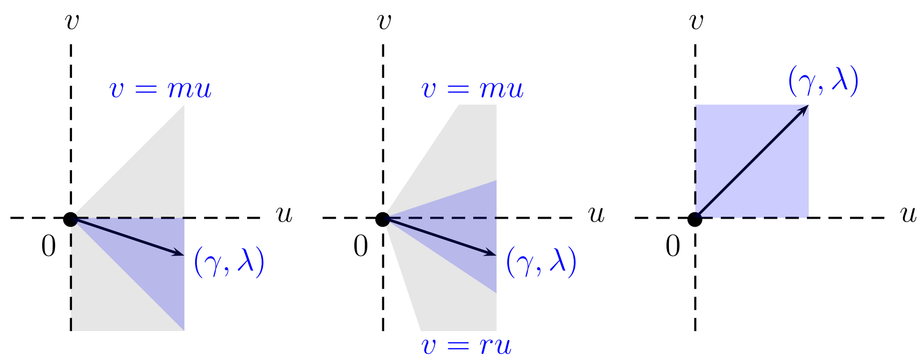


Figure 4.2: Theorem 4.3.1: $(a4)$, $(a5)$, $(a6)$ (figure produced by author)

(a7) *Si* $S_g^-(0) = S_f^-(\mu) = \Omega_+^= = \emptyset$, $\Omega_+^+ \neq \emptyset$, $0 < m < +\infty$. *In which case* $S_g^+(0) = \Omega_+^+$ *and* $\overline{\text{cone}}(\mathcal{F}_\mu) = \{(u,v) \in \mathbb{R}^2 : \; 0 \le v \le mu\}$. *Hence*

$$[\text{cone } \mathcal{F}_\mu]^* = \{(\gamma,\lambda) \in \mathbb{R}^2 : \; \gamma \ge 0, \; -\frac{1}{m}\gamma \le \lambda \le m\gamma\},$$

*and* $\mathcal{L}_{SD} = \mathcal{S}_D = \{\lambda \in \mathbb{R} : \; -\frac{1}{m} \le \lambda \le m\}$.

(a8) $\Omega_+^= = \Omega_+^- = \emptyset$, $\Omega_+^+ \neq \emptyset$, $\Omega_-^- \neq \emptyset$, $m \le l$. *In which case* $S_g^+(0) = \Omega_+^+$ *and* $\overline{\text{cone}}(\mathcal{F}_\mu) = \{(u,v) \in \mathbb{R}^2 : \; v \le mu, \; v \le lu\}$. *Hence,*

$$[\text{cone } \mathcal{F}_\mu]^* = \{(\gamma,\lambda) \in \mathbb{R}^2 : \; \gamma \ge 0, \; -\frac{1}{m}\gamma \le \lambda \le -\frac{1}{l}\gamma\},$$

*and* $\mathcal{L}_{SD} = \mathcal{S}_D = \{\lambda \in \mathbb{R} : \; -\frac{1}{m} \le \lambda \le -\frac{1}{l}\}$.

(a9) $\Omega_-^= = \Omega_-^- = \emptyset$, $\Omega_-^+ \neq \emptyset$, $\Omega_+^- \neq \emptyset$, $s \le r$. *In which case* $S_g^-(0) = \Omega_-^+$ *and* $\overline{\text{cone}}(\mathcal{F}_\mu) = \{(u,v) \in \mathbb{R}^2 : \; v \ge su, \; v \ge ru\}$. *Hence,*

$$[\text{cone } \mathcal{F}_\mu]^* = \{(\gamma,\lambda) \in \mathbb{R}^2 : \; \gamma \ge 0, \; -\frac{1}{s}\gamma \le \lambda \le -\frac{1}{r}\gamma\},$$

*and* $\mathcal{L}_{SD} = \mathcal{S}_D = \{\lambda \in \mathbb{R} : \; -\frac{1}{s} \le \lambda \le -\frac{1}{r}\}$.

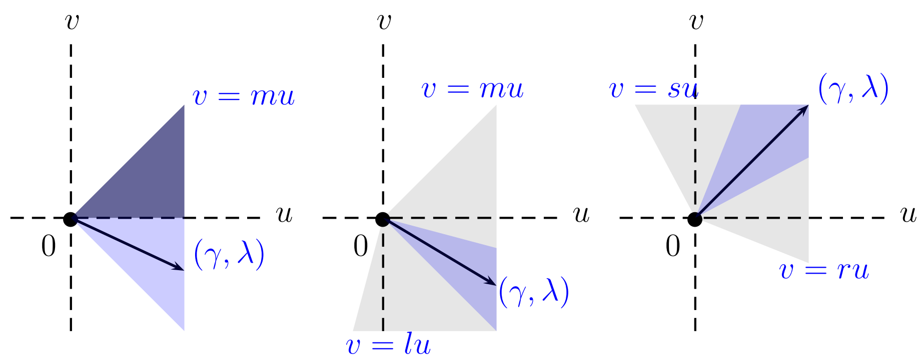


Figure 4.3: Theorem 4.3.1: ($a7$), ($a8$), ($a9$) (figure produced by author)

($a10$) $S_g^-(0) = \emptyset,\ \Omega_+^- \neq \emptyset,\ -\infty < s < 0$. *In which case* $\overline{\text{cone}}(\mathcal{F}_\mu) = \{(u,v) \in \mathbb{R}^2 :\ v \geq su,\ v \geq 0\}$. *Hence*

$$[\text{cone } \mathcal{F}_\mu]^* = \{(\gamma,\lambda) \in \mathbb{R}^2 :\ \gamma \geq 0,\ -\frac{1}{s}\gamma \leq \lambda\},$$

*and* $\mathcal{L}_{SD} = \mathcal{S}_D = \{\lambda \in \mathbb{R} :\ -\frac{1}{s} \leq \lambda\}$.

($a11$) $S_g^-(0) = \emptyset$, $\Omega_+^- \neq \emptyset,\ s = 0$. *In which case* $\overline{\text{cone}}(\mathcal{F}_\mu) = \{(u,v) \in \mathbb{R}^2 :\ v \geq 0\}$. *Hence,*

$$[\text{cone } \mathcal{F}_\mu]^* = \{0\} \times \mathbb{R}_+,$$

*and so* $\mathcal{L}_{SD} = \emptyset$.

($a12$) $S_g^+(0) = \emptyset$, $\Omega_-^- \neq \emptyset,\ 0 < l < +\infty$. *In which case* $\overline{\text{cone}}(\mathcal{F}_\mu) = \{(u,v) \in \mathbb{R}^2 :\ v \geq lu,\ v \leq 0\}$. *Hence*

$$[\text{cone } \mathcal{F}_\mu]^* = \{(\gamma,\lambda) \in \mathbb{R}^2 :\ \gamma \geq 0,\ \lambda \leq -\frac{1}{l}\gamma\},$$

*and* $\mathcal{L}_{SD} = \mathcal{S}_D = \{\lambda \in \mathbb{R} :\ \lambda \leq -\frac{1}{l}\}$.

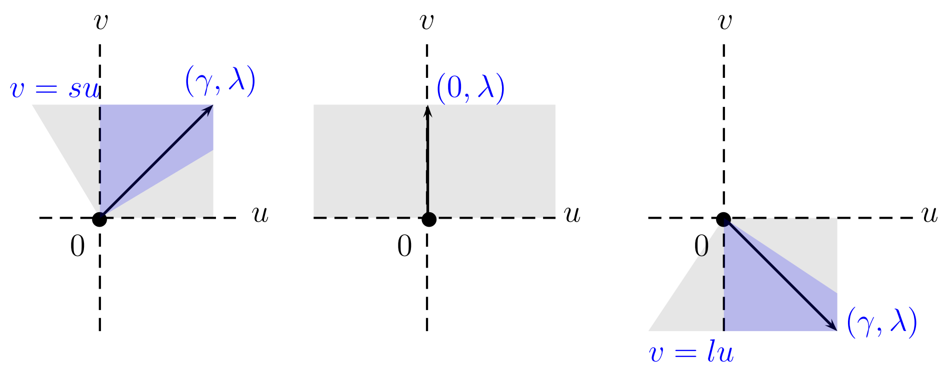


Figure 4.4: Theorem 4.3.1: $(a10)$, $(a11)$, $(a12)$ (figure produced by author)

$(a13)$ $S_g^+(0) = \emptyset,\ \Omega_-^- \neq \emptyset,\ l = 0$. *In which case* $\overline{\text{cone}}(\mathcal{F}_\mu) = \{(u, v) \in \mathbb{R}^2 :\ v \leq 0\}$. *Hence*

$$[\text{cone } \mathcal{F}_\mu]^* = \{0\} \times \mathbb{R}_-,$$

*and so* $\mathcal{L}_{SD} = \emptyset$.

$(a14)$ $S_g^+(0) = S_f^-(\mu) = \emptyset$ *and either* $[\Omega_-^+ \neq \emptyset,\ r = -\infty]$ *or* $\Omega_-^= \neq \emptyset$. *In which case* $\overline{\text{cone}}(\mathcal{F}_\mu) = \mathbb{R}_+ \times \mathbb{R}_-$. *Hence,*

$$[\text{cone } \mathcal{F}_\mu]^* = \mathbb{R}_+ \times \mathbb{R}_-,$$

*and* $\mathcal{L}_{SD} = \mathcal{S}_D = \{\lambda \in \mathbb{R} :\ \lambda \leq 0\}$.

$(a15)$ $S_g^+(0) = S_f^-(\mu) = \emptyset,\ \Omega_-^= = \emptyset,\ \Omega_-^+ \neq \emptyset,\ -\infty < r < 0$. *In which case* $S_g^-(0) = \Omega_-^+$ *and* $\overline{\text{cone}}(\mathcal{F}_\mu) = \{(u, v) \in \mathbb{R}^2 :\ ru \leq v \leq 0\}$. *Hence,*

$$[\text{cone } \mathcal{F}_\mu]^* = \{(\gamma, \lambda) \in \mathbb{R}^2 :\ \gamma \geq 0,\ r\gamma \leq \lambda \leq -\frac{1}{r}\gamma\},$$

*and* $\mathcal{L}_{SD} = \mathcal{S}_D = \{\lambda \in \mathbb{R} :\ r \leq \lambda \leq -\frac{1}{r}\}$.

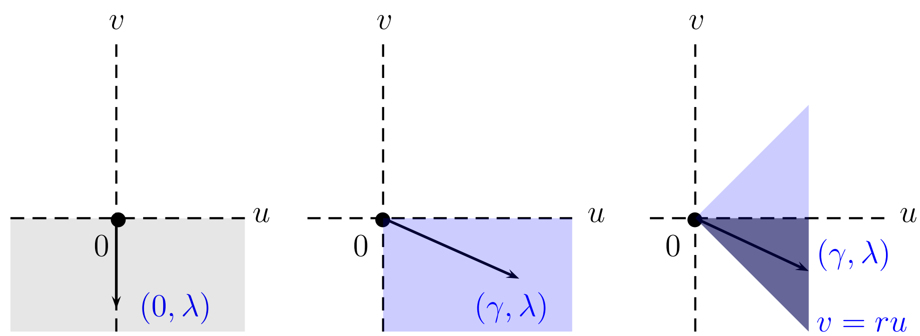


Figure 4.5: Theorem 4.3.1: $(a13)$, $(a14)$, $(a15)$ (figure produced by author)

*Proof.* The proof consists of two parts: the first part is devoted to find the precise expression for $\overline{\text{cone}}\ \mathcal{F}_\mu$, and the second part will prove the "only if" issue.
**1st Part**:
$(a1)$: See Proposition 4.3.1.
$(a2)$ : It follows from Proposition 4.3.2.
$(a3)$ : By Proposition 4.3.2, $\overline{\text{cone}}(F(\Omega_+^=) - \mu(1,0) + \mathbb{R}_+(1,0)) = \mathbb{R}^2_+$, and if $m = +\infty$ then

$$\overline{\text{cone}}(F(\Omega_+^+) - \mu(1,0) + \mathbb{R}_+(1,0)) = \mathbb{R}^2_+.$$

Both results along with the equalities $S_f^-(\mu) = \Omega_-^= = \emptyset$ give (see Proposition 4.3.1)

$$\overline{\text{cone}}(\mathcal{F}_\mu) = \overline{\text{cone}}(\Omega_3) = \overline{\text{cone}}(F(\Omega_-^+) - \mu(1,0) + \mathbb{R}_+(1,0)) \cup \mathbb{R}^2_+.$$

Then, the result follows from $(a)$ of Proposition 4.3.4.
$(a4)$ By Proposition 4.3.4, $\Omega_-^= \neq \emptyset$ and $r = -\infty$ imply

$$\overline{\text{cone}}(F(\Omega_-^=) - \mu(1,0) + \mathbb{R}_+(1,0)) = \overline{\text{cone}}(F(\Omega_-^+) - \mu(1,0) + \mathbb{R}_+(1,0)) = \mathbb{R}_+ \times \mathbb{R}_-.$$

Both results along with $S_f^-(\mu) = \Omega_+^= = \emptyset$, allow us to infer that (see Proposition 4.3.1)

$$\overline{\text{cone}}(\mathcal{F}_\mu) = \overline{\text{cone}}(\Omega_3) = \overline{\text{cone}}(F(\Omega_+^+) - \mu(1,0) + \mathbb{R}_+(1,0)) \cup \mathbb{R}_+ \times \mathbb{R}_-.$$

Thus, the conclusion follows from $(b)$ of Proposition 4.3.4.
$(a5)$: Since $S_f^-(\mu) = \Omega_+^= = \Omega_-^= = \emptyset$, one obtains (see Proposition 4.3.1)

$$\overline{\text{cone}}(\mathcal{F}_\mu) = \overline{\text{cone}}(\Omega_3) \ = \overline{\text{cone}}(F(\Omega_+^+) - \mu(1,0) + \mathbb{R}_+(1,0)) \cup \overline{\text{cone}}(F(\Omega_-^+) - \mu(1,0) + \mathbb{R}_+(1,0)).$$

Thus, the conclusion follows from $(a)$ and $(b)$ of Proposition 4.3.4.
$(a6)$: Since $S_f^-(\mu) = S_g^-(0) = \emptyset$, we get (Proposition 4.3.1)

$$\overline{\text{cone}}(\mathcal{F}_\mu) = \overline{\text{cone}}(\Omega_3) = \overline{\text{cone}}(F(\Omega_+^+) - \mu(1,0) + \mathbb{R}_+(1,0)) \cup \overline{\text{cone}}(F(\Omega_+^=) - \mu(1,0) + \mathbb{R}_+(1,0)) = \mathbb{R}^2_+,$$

where the last equality is obtained by Propositions 4.3.2 and 4.3.4.
($a7$): The conclusion of the first part follows since $S_f^-(\mu) = S_g^-(0) = \Omega_+^= = \emptyset$ yield (Proposition 4.3.1)

$$\overline{\text{cone}}(\mathcal{F}_\mu) = \overline{\text{cone}}(\Omega_3) = \overline{\text{cone}}(F(\Omega_+^+) - \mu(1,0) + \mathbb{R}_+(1,0)) = \{(u,v) \in \mathbb{R}_+^2 : \ v \leq mu \},$$

because of Proposition 4.3.4.
($a8$): Taking into account ($a$) of Proposition 4.3.3, we get, as above,

$$\overline{\text{cone}}(\mathcal{F}_\mu) = \overline{\text{cone}}(\Omega_3) = \overline{\text{cone}}(F(\Omega_+^+) - \mu(1,0) + \mathbb{R}_+(1,0)) \cup \overline{\text{cone}}(F(\Omega_-^-) - \mu(1,0) + \mathbb{R}_+(1,0)).$$

So, the conclusion follows as a consequence of Proposition 4.3.4.
($a9$): It is analogous to ($a8$). Taking into account ($b$) of Proposition 4.3.3, we obtain

$$\begin{aligned}\overline{\text{cone}}(\mathcal{F}_\mu) &= \overline{\text{cone}}(\Omega_3)\\ &= \overline{\text{cone}}(F(\Omega_+^-) - \mu(1,0) + \mathbb{R}_+(1,0)) \cup \overline{\text{cone}}(F(\Omega_-^+) - \mu(1,0) + \mathbb{R}_+(1,0)).\end{aligned}$$

($a10$): As above, we get

$$\overline{\text{cone}}(\mathcal{F}_\mu) = \overline{\text{cone}}(F(\Omega_+^-) - \mu(1,0) + \mathbb{R}_+(1,0)) = \{(u,v) \in \mathbb{R}^2 : \ v \geq su, \ v \geq 0\}.$$

($a11$): It is similar to ($a10$).
($a12$): In this case, we get

$$\overline{\text{cone}}(\mathcal{F}_\mu) = \overline{\text{cone}}(F(\Omega_-^-) - \mu(1,0) + \mathbb{R}_+(1,0)) = \{(u,v) \in \mathbb{R}^2 : \ v \leq lu, \ v \leq 0\}.$$

($a13$): It is similar to ($a12$).
($a14$): As above, we obtain

$$\begin{aligned}\overline{\text{cone}}(\mathcal{F}_\mu) &= \overline{\text{cone}}(\Omega_3)\\ &= \overline{\text{cone}}(F(\Omega_-^+) - \mu(1,0) + \mathbb{R}_+(1,0)) \cup \overline{\text{cone}}(F(\Omega_-^=) - \mu(1,0) + \mathbb{R}_+(1,0))\\ &= \mathbb{R}_+ \times \mathbb{R}_-.\end{aligned}$$

($a15$): Finally, as above, we get

$$\begin{aligned}\overline{\text{cone}}(\mathcal{F}_\mu) &= \overline{\text{cone}}(\Omega_3) = \overline{\text{cone}}(F(\Omega_-^+) - \mu(1,0) + \mathbb{R}_+(1,0))\\ &= \{(u,v) \in \mathbb{R}^2 : \ ru \leq v \leq 0, \ u \geq 0\}.\end{aligned}$$

**2nd Part**: Let us check that the convexity of $\overline{\text{cone}}(\mathcal{F}_\mu)$ implies exactly one of the fifteen cases ($ai$), $i = 1, 2, \ldots, 15$. First of all, when $C = K$, it follows that $\overline{\text{cone}}(\mathcal{F}_\mu) = \mathbb{R}_+(1,0)$ by Proposition 4.3.1, and so ($a1$) holds. All the situations occur when $C \neq K$. Notice also that from (4.7), $\text{cone}(\mathcal{F}_\mu) \neq \mathbb{R}^2$.

From $C \setminus K = S_g^-(0) \cup S_g^+(0)$, we identify all the remaining fourteen cases.

1⌋ $S_g^-(0) \neq \emptyset$, $S_g^+(0) = \emptyset$: from (4.14), one gets,

$$\begin{aligned}\overline{\text{cone}}(\Omega_3) &= \overline{\text{cone}}(F(\Omega_-^-) - \mu(1,0) + \mathbb{R}_+(1,0)) \cup \overline{\text{cone}}(F(\Omega_-^=) - \mu(1,0) + \mathbb{R}_+(1,0)) \\ &\cup \overline{\text{cone}}(F(\Omega_-^+) - \mu(1,0) + \mathbb{R}_+(1,0)).\end{aligned}$$

Then, depending if either $\Omega_-^-$ or $\Omega_-^=$ or $\Omega_-^+$ is nonempty, we obtain any of the cases $(a12) - -(a15)$.

2⌋ $S_g^-(0) = \emptyset$, $S_g^+(0) \neq \emptyset$: again from (4.14) one gets

$$\begin{aligned}\overline{\text{cone}}(\Omega_3) &= \overline{\text{cone}}(F(\Omega_+^-) - \mu(1,0) + \mathbb{R}_+(1,0)) \cup \overline{\text{cone}}(F(\Omega_+^=) - \mu(1,0) + \mathbb{R}_+(1,0)) \\ &\cup \overline{\text{cone}}(F(\Omega_+^+) - \mu(1,0) + \mathbb{R}_+(1,0)),\end{aligned}$$

obtaining any of the cases $(a6)$, $(a7)$, $(a10)$, $(a11)$.

3⌋ $S_g^-(0) \neq \emptyset$, $S_g^+(0) \neq \emptyset$, $\Omega_+^- \neq \emptyset$: First, assume that $s = 0$. Then

$$\mathbb{R} \times \mathbb{R}_+ \subseteq \overline{\text{cone}}(\mathcal{F}_\mu),$$

which, along with the convexity, yields $S_g^-(0) = \emptyset$, giving a contradiction. Hence $-\infty < s < 0$. Thus, by $(a1)$ of Proposition 4.3.5, $\Omega_-^- = \emptyset = \Omega_-^=$. It follows that $\emptyset \neq S_g^-(0) = \Omega_-^+$. Using $(a2)$ of Proposition 4.3.5 again, one obtains $s \leq r$. Hence $(a9)$ holds.

4⌋ $S_g^-(0) \neq \emptyset$, $S_g^+(0) \neq \emptyset$, $\Omega_-^- \neq \emptyset$: This case is analogous to 3⌋, so we conclude that $(a8)$ holds.

5⌋ $S_g^-(0) \neq \emptyset$, $S_g^+(0) \neq \emptyset$, $\Omega_+^- = \Omega_-^- = \emptyset$: Since $S_f^-(\mu) = \Omega_-^- \cup \Omega_+^- \cup \Omega_=^- = \emptyset$, one obtains

$$\begin{aligned}\overline{\text{cone}}(\Omega_3) &= \overline{\text{cone}}(F(\Omega_-^=) - \mu(1,0) + \mathbb{R}_+(1,0)) \cup \overline{\text{cone}}(F(\Omega_+^=) - \mu(1,0) + \mathbb{R}_+(1,0)) \\ &\cup \overline{\text{cone}}(F(\Omega_+^=) - \mu(1,0) + \mathbb{R}_+(1,0)) \cup \overline{\text{cone}}(F(\Omega_+^+) - \mu(1,0) + \mathbb{R}_+(1,0)).\end{aligned}$$

Thus each of the cases $(a2) - -(a5)$, is obtained.

6⌋ The remaining case $S_g^-(0) \neq \emptyset$, $S_g^+(0) \neq \emptyset$, $\Omega_+^- \neq \emptyset$, $\Omega_-^- \neq \emptyset$ is not possible since $\overline{\text{cone}}(\mathcal{F}_\mu) \neq \mathbb{R}^2$. $\square$

Now, we establish our geometric and topological characterizations of strong duality.

**Theorem 4.3.2.** *Let us consider problem* (4.1) *with* $\mu \in \mathbb{R}$. *Then, the following are equivalent:*

(a) *Strong Duality holds for* (4.1)*, that is*

$$\exists\ \lambda^* \in \mathbb{R}:\ f(x) + \lambda^* g(x) \geq \mu,\ \forall\ x \in C; \tag{4.22}$$

(b) $\overline{\text{cone}}(\mathcal{F}_\mu) \cap (-\mathbb{R}_{++} \times \{0\}) = \emptyset$ *and* $\overline{\text{cone}}(\mathcal{F}_\mu)$ *is convex;*

(*c*) $\mathrm{cone}(\mathcal{F}_\mu)$ *is convex and exactly one of the following assertions holds:*

(*c*1) $S_f^-(\mu) = \emptyset$, *in which case* $0 \in \mathcal{L}_{SD}$;

(*c*2) $\Omega_+^- \neq \emptyset$, $s < 0$, *in which case* $\min \mathcal{L}_{SD} = -\dfrac{1}{s}$;

(*c*3) $\Omega_-^- \neq \emptyset$, $l > 0$, *in which case* $\max \mathcal{L}_{SD} = -\dfrac{1}{l}$.

*Consequently, under condition* (*a*), *one obtains:* $\mathcal{L}_{SD} = \mathcal{S}_D$;

$$\inf_{x \in K} f(x) = \inf_{\substack{\lambda^* g(x) \leq 0 \\ x \in C}} f(x); \tag{4.23}$$

$$\bar{x} \textit{ is a solution to } (4.1) \iff \begin{cases} \bar{x} \in C, \quad g(\bar{x}) = 0, \\ f(\bar{x}) = \inf_{x \in C}[f(x) + \lambda^* g(x)]. \end{cases} \tag{4.24}$$

*Proof.* (*b*) $\Leftrightarrow$ (*a*) By applying Theorem 3.2 in [35], it remains only to check that (*a*) implies the convexity of $\overline{\mathrm{cone}}(\mathcal{F}_\mu)$. This is a consequence of Theorem 2.4 in [33] because of (4.22).
(*a*) $\Rightarrow$ (*c*) As shown above, (*a*) yields the convexity of $\overline{\mathrm{cone}}(\mathcal{F}_\mu)$. Assume that $S_f^-(\mu) \neq \emptyset$. Then $\lambda^* \neq 0$, $C \neq K$, and at least one of the sets $\Omega_-^-$, $\Omega_+^-$ is nonempty. We claim that exactly one of such sets is nonempty. Indeed, if there exist $x_1 \in \Omega_+^-$ and $x_2 \in \Omega_-^-$, we get $(f(x_1) - \mu, g(x_1))$, $(f(x_2) - \mu, g(x_2)) \in F(C) - \mu(1,0) + \mathbb{R}_+(1,0) \subseteq \overline{\mathrm{cone}}(\mathcal{F}_\mu)$. Setting $t_0 \doteq \dfrac{-g(x_2)}{g(x_1) - g(x_2)} \in \,]0,1[$, we get by convexity

$$\begin{aligned} & t_0(f(x_1) - \mu, g(x_1)) + (1 - t_0)(f(x_2) - \mu, g(x_2)) \\ & \quad = (t_0(f(x_1) - \mu) + (1 - t_0)(f(x_2) - \mu), 0) \in \overline{\mathrm{cone}}(\mathcal{F}_\mu) \cap -\mathbb{R}_{++} \times \{0\}, \end{aligned}$$

which contradicts (*b*).
Consider the case $\Omega_+^- \neq \emptyset$. Let us prove that $s < 0$. We first check that $\lambda^* > 0$. If $\lambda^* < 0$ then (4.22) implies

$$g(x) \leq \frac{-(f(x) - \mu)}{\lambda^*} < 0, \ \forall\, x \in S_f^-(\mu),$$

showing that $\Omega_-^- \neq \emptyset$, which is impossible. Hence $\lambda^* > 0$, and by (4.22), we get

$$\frac{g(x)}{f(x) - \mu} \leq -\frac{1}{\lambda^*} < 0, \ \forall\, x \in \Omega_+^-,$$

proving that $s \leq -\dfrac{1}{\lambda^*} < 0$, and so $\lambda^* \geq -\dfrac{1}{s}$.
The case $\Omega_-^- \neq \emptyset$ is treated similarly, concluding that $\Omega_+^- = \emptyset$ and $\lambda^* < 0$. Thus, from (4.22), we obtain

$$0 < -\frac{1}{\lambda^*} \leq \frac{g(x)}{f(x) - \mu}, \ \forall\, x \in \Omega_-^-,$$

implying that $0 < -\dfrac{1}{\lambda^*} \leq l$, and so $\lambda^* \leq -\dfrac{1}{l}$.
Then, from Theorem 4.3.1, we conclude that cone $\mathcal{F}_\mu$ is convex.

$(c1) \Rightarrow (a)$ In this case we obtain that $f(x) - \mu \geq 0$, $\forall\, x \in C$. Thus, by choosing $\lambda^* = 0$, (4.22) holds.
$(c2) \Rightarrow (a)$ From $\Omega_+^- \neq \emptyset$, $s < 0$, it follows that

$$f(x) - \mu + \left(-\frac{1}{s}\right) g(x) \geq 0,\ \forall\, x \in \Omega_+^-. \tag{4.25}$$

We claim the inequality in (4.25) holds for all $x \in C$. In fact, since $C = (\mathrm{argmin}_K f) \cup (K \setminus \mathrm{argmin}_K f) \cup (C \setminus K)$, we obtain:
if $x \in \mathrm{argmin}_K f$ then $0 = f(x) - \mu = f(x) - \mu + \left(-\frac{1}{s}\right) g(x)$;
if $x \in K \setminus \mathrm{argmin}_K f = \Omega_=^{\pm}$ then $0 < f(x) - \mu = f(x) - \mu + \left(-\frac{1}{s}\right) g(x)$;
finally, assume that $x \in C \setminus K = S_g^+(0) \cup S_g^-(0)$. The case $x \in S_g^+(0)$, which may be splitted as $\Omega_+^= \cup \Omega_+^+ \cup \Omega_+^-$, allows us to prove that (4.22) holds from (4.25) directly. The case $x \in S_g^-(0) = \Omega_-^= \cup \Omega_-^+ \cup \Omega_-^-$, allows us to prove that (4.22) holds by assumption and $(a1)$ of Proposition 4.3.5. In fact, $\Omega_-^- = \Omega_-^= = \emptyset$, and if $x \in \Omega_-^+$ then $f(x) - \mu + \left(-\frac{1}{s}\right) g(x) \geq 0$, since otherwise, $s > \frac{g(x)}{f(x)-\mu} \geq r$, contradicting $(a2)$ of Proposition 4.3.5.
$(c3) \Rightarrow (a)$ We apply $(b)$ of Proposition 4.3.5 in a similar way as in the previous case to conclude with the expected result $f(x) - \mu + \left(-\frac{1}{l}\right) g(x) \geq 0$ for all $x \in C$.
Assertions (4.23) and (4.24) are straightforward.

$\square$

We immediately point out that the convexity of $\overline{\mathrm{cone}}(\mathcal{F}_\mu)$ does not imply the convexity of $\mathrm{cone}(\mathcal{F}_\mu)$ without any additional assumption. This is illustrated by Example 4.3.6. However, under strong duality that implication holds, as shown by the previous theorem.

**Example 4.3.6.** *Let* $f(x_1, x_2) = 2x_1x_2$, $g(x_1, x_2) = x_1$ *and* $C = \mathbb{R}^2$. *Then,* $\mu = 0$, $F(\mathbb{R}^2) = \{(0,0)\} \cup (\mathbb{R}^2 \setminus \mathbb{R} \times \{0\})$, *and so*

$$\mathrm{cone}(F(\mathbb{R}^2) - \mu(1,0) + \mathbb{R}_+(1,0)) = \mathbb{R}^2 \setminus (-\mathbb{R}_{++} \times \{0\}),$$

*which is nonconvex, but* $\overline{\mathrm{cone}}(\mathcal{F}_\mu) = \mathbb{R}^2$. *It is easy to check that strong duality does not hold.*

**Remark 4.3.7.** *From the proof of* $(a) \Rightarrow (c)$ *of Theorem 4.3.2 we conclude that if strong duality holds and* $S_f^-(\mu) \neq \emptyset$, *then exactly one of the sets* $\Omega_+^-$ *or* $\Omega_-^-$ *is nonempty.*

The following result, which is new in the literature, provides a characterization of strong duality under a Slater-type condition.

**Corollary 4.3.8.** *Let* $\mu \in \mathbb{R}$ *and assume that there exist* $x_1, x_2 \in C$ *such that* $g(x_1) < 0 < g(x_2)$. *Then,* $\mathrm{cone}(\mathcal{F}_\mu)$ *is convex if and only if strong duality holds for* (4.1).

*Proof.* Since $\mathrm{cone}(\mathcal{F}_\mu) \cap -(\mathbb{R}_{++} \times \{0\}) = \emptyset$ (see (4.7)), a standard convex separation theorem yields the existence of $\gamma, \lambda \in \mathbb{R}$ such that

$$\gamma(f(x) - \mu + r) + \lambda g(x) \geq 0, \quad \forall\, x \in C,\ r \geq 0.$$

Then $\gamma \geq 0$, and the Slater-type condition gives $\gamma > 0$, which ensures that strong duality holds. The other implication follows from Theorem 4.3.2. $\square$

**Remark 4.3.9.** (*The quadratic case*) *When $f$ and $g$ are quadratic functions, it is proven in [36] that, under the same Slater condition imposed in Corollary 4.3.8, strong duality holds if, and only if $F(\mathbb{R}^n) + \mathbb{R}_+(1,0)$ is convex. Necessary and sufficient conditions for the convexity are provided in Theorem 4.12 of [36]. Such a characterization encompasses the case when the Hessian of $g$ is not the zero matrix, or when $g$ is strictly concave (or strictly convex).*

Next example shows that a Slater-type condition: $g(x_1) < 0 < g(x_2)$ for some $x_1, x_2 \in C$, is needed.

**Example 4.3.10.** *Let us consider $f(x_1,x_2) = x_1 + x_2$, $g(x_1,x_2) = (x_1+x_2)^2$ and $C = \mathbb{R}^2$. One can deduce that $\mu = 0$, $\operatorname{argmin}_K f = \{(x_1,x_2) : \ x_1 + x_2 = 0\}$ and there is no duality gap. Moreover, $F(\mathbb{R}^2) = \{(u,v) \in \mathbb{R}^2 : \ v = u^2\}$, $S_g^-(0) = \emptyset$ and $\Omega_+^- = S_f^-(\mu) = \{(x_1,x_2) \in \mathbb{R}^2 : \ x_1 + x_2 < 0\}$, and therefore* cone $\mathcal{F}_\mu$ *is convex ((a11) of Theorem 4.3.1 holds) and*

$$s = \sup_{x_1+x_2<0} \frac{(x_1+x_2)^2}{x_1+x_2} = 0.$$

*Hence, by virtue of (c) of Theorem 4.3.2, strong duality does not hold, which may be checked directly.*

Now we exhibit an example illustrating a situation where our main Theorem 4.3.2 applies: strong duality is present without satisfying the Slater-type condition.

**Example 4.3.11.** *Take $f(x_1,x_2) = 2x_1^2 - x_2^2$, $g(x_1,x_2) = x_1^2 + x_2^2$ and $C = \mathbb{R}^2$. Here, $K = \{(0,0)\}$, $\mu = 0$, $S_g^-(0) = \emptyset$. One gets $s = -1$, and according to (a10) of Theorem 4.3.1 the convexity of* cone $\mathcal{F}_\mu$ *follows (which is also a consequence of Dine's theorem). Thus, by (c2) of Theorem 4.3.2 strong duality holds, with $\mathcal{L}_{SD} = [1, +\infty[$.*

## 4.4 Characterizing KKT optimality conditions

This section deals with some characterizations of the validity of the KKT optimality conditions for problem (4.1), see [57, 58]. For simplicity, take $X$ to be $\mathbb{R}^n$, and $f$ and $g$ to be differentiable on $\mathbb{R}^n$. Such characterizations will be derived as a consequence of Theorem 4.3.2 applied to the linearized approximation problem defined, given $\bar{x} \in C$, by

$$\mu_L \doteq \inf\{\nabla f(\bar{x})^\top v : \ v \in G'(\bar{x})\}, \tag{4.26}$$

where

$$G'(\bar{x}) \doteq \Big\{ v \in T(C;\bar{x}) : \ \nabla g(\bar{x})^\top v = 0 \Big\}.$$

Set $F_L(v) \doteq (\nabla f(\bar{x})^\top v, \nabla g(\bar{x})^\top v)$. Since $G'(\bar{x})$ is a cone, then $\mu_L \leq 0$. Thus, if $\mu_L < 0$, then there exist $v_0 \in G'(\bar{x})$ such that $\nabla f(\bar{x})^\top v_0 < 0$. Then, we obtain for all $t > 0$:

$$\mu_L < t\nabla f(\bar{x})^\top v_0 \longrightarrow -\infty, \text{ as } t \to +\infty.$$

That is, $\mu_L \in \{-\infty, 0\}$. Additionally

$$\begin{aligned}\mu_L = 0 \quad &\Longleftrightarrow \quad [v \in T(C;\bar{x}),\ \nabla f(\bar{x})^\top v < 0 \Longrightarrow \nabla g(\bar{x})^\top v \neq 0] \\ &\Longleftrightarrow \quad F_L(T(C;\bar{x})) \cap -(\mathbb{R}_{++} \times \{0\}) = \emptyset \\ &\Longleftrightarrow \quad [F_L(T(C;\bar{x})) + \mathbb{R}_+(1,0)] \cap (-\mathbb{R}_{++} \times \{0\}) = \emptyset. \end{aligned} \tag{4.27}$$

In view of Theorem 4.3.2, we introduce the following sets:

$$\widehat{S}_f^-(0) \doteq \{v \in T(C;\bar{x}) :\ \nabla f(\bar{x})^\top v < 0\},\ \widehat{S}_g^+(0) \doteq \{v \in T(C;\bar{x}) :\ \nabla g(\bar{x})^\top v > 0\},$$

$$\widehat{\Omega}_+^- \doteq \widehat{S}_f^-(0) \cap \widehat{S}_g^+(0),\quad \widehat{\Omega}_-^- \doteq \widehat{S}_f^-(0) \cap \widehat{S}_g^-(0).$$

Furthermore, whenever $\widehat{\Omega}_+^- \neq \emptyset \neq \widehat{\Omega}_-^-$, we put

$$\widehat{s} \doteq \sup_{v \in \widehat{\Omega}_+^-} \frac{\nabla g(\bar{x})^\top v}{\nabla f(\bar{x})^\top v},\quad \widehat{l} \doteq \inf_{v \in \widehat{\Omega}_-^-} \frac{\nabla g(\bar{x})^\top v}{\nabla f(\bar{x})^\top v}.$$

We denote by $\mathcal{L}(\bar{x})$ the set of Lagrange multipliers to problem (4.1) associated to a (not necessarily feasible) point $\bar{x} \in C$, i.e., the set of $\lambda^* \in \mathbb{R}$ satisfying (4.29). It obvious that

$$\lambda^* \in \mathcal{L}(\bar{x}) \Longleftrightarrow (1,\lambda^*) \in [F_L(T(C;\bar{x})) + \mathbb{R}_+(1,0)]^*. \tag{4.28}$$

When $\mathcal{L}(\bar{x}) \neq \emptyset$, we say that $\bar{x}$ is a KKT point.

Next remark describes the case $\nabla g(\bar{x}) = 0$.

**Remark 4.4.1.** *Assume that $\bar{x} \in C$ and $\nabla g(\bar{x}) = 0$. Observe that, if $\nabla g(\bar{x}) = 0$, then $G'(\bar{x}) = T(C;\bar{x})$ and*

$$\mu_L = \inf\{\nabla f(\bar{x})^\top v :\ v \in T(C;\bar{x})\}.$$

*Thus, if $\mathcal{L}(\bar{x}) \neq \emptyset$, one gets from (4.28) for some $\lambda^* \in \mathbb{R}$ that*

$$\big\langle (1,\lambda^*), (\nabla f(\bar{x})^\top v + u, 0) \big\rangle = \nabla f(\bar{x})^\top v + u \geq 0,\ \forall\ v \in T(C;\bar{x}),\ \forall\ u \geq 0.$$

*Then, it is not difficult to check that:*

- $\mu_L = 0$ *if and only if* $\mathcal{L}(\bar{x}) = \mathbb{R}$.
- $\mu_L = -\infty$ *if and only if* $\mathcal{L}(\bar{x}) = \emptyset$.

We are now ready to describe some equivalent formulations of the validity of the KKT optimality conditions when $\nabla g(\bar{x}) \neq 0$.

**Theorem 4.4.1.** *Assume that $\bar{x} \in C$. The following assertions are equivalent:*

(*a*) $\exists\ \lambda^* \in \mathbb{R}$ *such that*

$$\nabla f(\bar{x}) + \lambda^* \nabla g(\bar{x}) \in [T(C;\bar{x})]^*. \tag{4.29}$$

(*b*) $\mu_L = 0$ *and strong duality holds for the problem* (4.26)*.*

(*c*) $\overline{F_L(T(C;\bar{x})) + \mathbb{R}_+(1,0)}$ *is convex and*

$$\overline{[F_L(T(C;\bar{x})) + \mathbb{R}_+(1,0)]} \cap (-\mathbb{R}_{++} \times \{0\}) = \emptyset. \tag{4.30}$$

(*d*) $F_L(T(C;\bar{x})) + \mathbb{R}_+(1,0)$ *is convex and exactly one of the following assertions holds:*

(*d*1) $\widehat{S}_f^-(0) = \emptyset$*, in which case* $0 \in \mathcal{L}(\bar{x})$*;*

(*d*2) $\widehat{\Omega}_+^- \neq \emptyset$*,* $\widehat{s} < 0$*, in which case* $\min \mathcal{L}(\bar{x}) = -\dfrac{1}{\widehat{s}}$*;*

(*d*3) $\widehat{\Omega}_-^- \neq \emptyset$*,* $\widehat{l} > 0$*, in which case* $\max \mathcal{L}(\bar{x}) = -\dfrac{1}{\widehat{l}}$*.*

(*e*) $\overline{F_L(T(C;\bar{x})) + \mathbb{R}_+(1,0)}$ *is convex,* $\mu_L = 0$ *and*

$$\left.\begin{array}{l} v_k \in T(C;\bar{x}), \|v_k\| \to +\infty, \\ \nabla g(\bar{x})^\top v_k \to 0, \nabla f(\bar{x})^\top v_k < 0 \end{array}\right\} \Longrightarrow \limsup_k \nabla f(\bar{x})^\top v_k = 0. \tag{4.31}$$

*Proof.* It is not difficult to check that (*a*) is equivalent to (*b*). The equivalence between (*c*) and (*d*) follows from Theorem 4.3.2.

(*c*) $\Longleftrightarrow$ (*e*): We now prove that (4.30) is equivalent to $\mu_L = 0$ and (4.31). Assume that (4.30) holds. Then, (4.27) implies immediately that $\mu_L = 0$. Take any $v_k \in T(C;\bar{x})$, $\lim_k \nabla g(\bar{x})^\top v_k = 0$, $\nabla f(\bar{x})^\top v_k < 0$. Suppose, on the contrary, that $\limsup_k \nabla f(\bar{x})^\top v_k = \xi < 0$. Up to a subsequence, $\nabla f(\bar{x})^\top v_k \to \xi \in\ [-\infty, 0[$, which says that $\nabla f(\bar{x})^\top v_k < 0$ for all $k$ sufficiently large. Due to linearity

$$\begin{pmatrix} \nabla f(\bar{x})^\top \\ \nabla g(\bar{x})^\top \end{pmatrix} v'_k \in F_L(T(C;\bar{x})) + \mathbb{R}_+(1,0),$$

where $v'_k = -\dfrac{v_k}{\nabla f(\bar{x})^\top v_k}$. Thus

$$(-1,0) \in \overline{[F_L(T(C;\bar{x})) + \mathbb{R}_+(1,0)]} \cap -(\mathbb{R}_{++} \times \{0\}),$$

yielding a contradiction, which proves one implication.
Assume that (4.31) and $\mu_L = 0$ hold. Take $(a,0) \in \overline{F_L(T(C;\bar{x})) + \mathbb{R}_+(1,0)}$, and suppose that $a < 0$. Then, there exist $v_k \in T(C;\bar{x})$, $r_k \geq 0$ such that $\nabla f(\bar{x})^\top v_k + r_k \to a$. Assume first that $\sup_k \|v_k\| < +\infty$. Up to a subsequence, we get $v_k \to v \in T(C;\bar{x})$. Thus $r_k \to a - \nabla f(\bar{x})^\top v \geq 0$, which implies that $0 > a \geq \nabla f(\bar{x})^\top v$ yielding a contradiction since $\nabla g(\bar{x})^\top v = 0$ and $\mu_L = 0$.
Assume now that $\sup_k \|v_k\| = +\infty$. Up to a subsequence, we have $\|v_k\| \to +\infty$. Clearly $\nabla f(\bar{x})^\top v_k < 0$ for all $k$ sufficiently large. Thus, by applying (4.31), we obtain a subsequence $v_{k_l}$ such that $\nabla f(\bar{x})^\top v_{k_l} \to 0$ as $l \to +\infty$. It means that $r_{k_l} \to a$, yielding a contradiction, since $a < 0$.

□

A simple sufficient condition for a minimum to be a KKT point under strong duality is expressed in the following result, which is important by itself. Its importance lies on the fact that it may be applied to situations where results based either on exact penalization techniques or where Abadie's constraint qualification fails.

**Proposition 4.4.2.** *Assume that strong duality holds for* (4.1). *Then, every solution to* (4.1) *is a KKT point, that is,* $\mathcal{L}_{SD} \subseteq \mathcal{L}(\bar{x})$ *for all* $\bar{x} \in \underset{K}{\operatorname{argmin}}\, f$.

*Proof.* Let $\bar{x}$ be a solution to (**??**). By (4.24), $\bar{x}$ is a minimum for $f + \lambda^* g$ on $C$ for some $\lambda^* \in \mathbb{R}$. Thus, the standard optimality condition yields $\nabla f(\bar{x}) + \lambda^* \nabla g(\bar{x}) = \nabla(f + \lambda^* g)(\bar{x}) \in [T(C;\bar{x})]^*$, which is the desired result.

□

The previous proposition has its own merit. First of all, there are instances (see Example 4.3.10) where no minimizer is a KKT point, if strong duality is not satisfied. On the other hand, as we already pointed out, such a proposition applies to situations where, for instance, Theorem 3.1 from [91] based on exact penalization techniques cannot be applied. This is shown by the following example.

**Example 4.4.3.** *Take*

$$0 = \mu \doteq \min\{f(x_1, x_2) \doteq x_2 : \ g(x_1, x_2) \doteq x_2 - x_1^2 = 0, \ (x_1, x_2) \in \mathbb{R}^2\}.$$

*Clearly, here* $C = \mathbb{R}^2$, $\operatorname{argmin}_K f = \{\bar{x} = (0,0)\}$ *is a solution to that minimization problem,* $\mathcal{F}_\mu$ *is convex* (*see* (a8) *of Theorem 4.3.1*) *and* (c3) *of Theorem 4.3.2 is satisfied* ($l = 1 = m$). *Thus, strong duality holds, and by Proposition 4.4.2,* $\bar{x}$ *is a KKT point, actually* $\mathcal{L}_{SD} = \mathcal{L}(\bar{x}) = \{-1\}$, *and therefore Theorem 3.1 in [91] is not applicable, since such a result yields KKT multipliers which are nonnegative. Notice that* $g^\infty(\bar{x}; u) = -u_1^2$ *and* $C(\bar{x}) = \{(u_1, 0) : \ u_1 \in \mathbb{R}\}$, *and so the assumptions of Theorem 3.1 in [91] are not verified. Moreover, our result also applies to the problem with the same* $f$ *but with* $-g$ *instead of* $g$, *yielding positive KKT multipliers: in such a case,* $s = r = -1$.

We now exhibit an instance where Abadie's constraint qualification fails but still our Proposition 4.4.2 applies.

**Example 4.4.4.** (*without Abadie's CQ*) *Take* $f(x_1, x_2) = x_2$, $g(x_1, x_2) = (x_1 - 1)^2 + (x_2 - 1)^2 - 1$ *and* $C \doteq \{(x_1, x_2) \in \mathbb{R}^2 : g_0(x_1, x_2) \leq 0\}$ *with* $g_0(x_1, x_2) \doteq (x_1 - 1)^2 + (x_2 + 1)^2 - 1$, *and consider*

$$0 = \mu \doteq \min\{f(x_1, x_2) : \ g(x_1, x_2) = 0, \ (x_1, x_2) \in C\}.$$

*Clearly the feasible set reduces to* $K = \{\bar{x} \doteq (1, 0)\}$, *and since* $-2 \leq x_2 \leq 0$ *for all* $(x_1, x_2) \in C$, *we get* $f(x) + \lambda g(x) \geq 0$ *for all* $x \in C$, *and all* $\lambda \geq \dfrac{1}{2}$. *We actually get* $\mathcal{L}_{SD} = [\dfrac{1}{2}, +\infty[$ *since* $s = -2$, *so* (a10) *of Theorem 4.3.1 is satisfied. This implies that strong duality holds. Then, by Proposition 4.4.2, there exists* $\lambda \in \mathbb{R}$ *such that*

$$\nabla f(\bar{x}) + \lambda \nabla g(\bar{x}) \in [T(C;\bar{x})]^* = \mathbb{R}_+(0, -1),$$

*which finally yields the existence of* $\lambda \in \mathbb{R}$, $\lambda_0 \geq 0$ *satisfying*

$$\nabla f(\bar{x}) + \lambda \nabla g(\bar{x}) + \lambda_0 \nabla g_0(\bar{x}) = 0,$$

*which says that* $\bar{x}$ *is a standard KKT point. However, we will see that Abadie's constraint qualification condition fails. Indeed,* $T(K;\bar{x}) = \{(0,0)\}$, *whereas*

$$G'(\bar{x}) \doteq \{v = (v_1, v_2) \in \mathbb{R}^2 : \ \nabla g_0(\bar{x})^\top v \leq 0\} = \{(v_1, v_2) : \ v_2 \leq 0\},$$

$$H_0(\bar{x}) \doteq \{v = (v_1, v_2) \in \mathbb{R}^2 : \ \nabla g(\bar{x})^\top v = 0\} = \{(v_1, v_2) : \ v_2 = 0\}.$$

*Thus,* $T(K;\bar{x}) \neq G'(\bar{x}) \cap H_0(\bar{x})$, *and therefore Theorem 5.3.1 of* [5] *is not applicable.*

We now show a situation where strong duality is not satisfied (so Proposition 4.4.2 is not applicable), but our Theorem 4.4.1 applies.

**Example 4.4.5.** (*without strong duality where every minimizer is a KKT point*) *Let us revise Example 4.3.6. Consider*

$$0 = \mu \doteq \min\{f(x_1, x_2) \doteq 2x_1x_2 : \ g(x_1, x_2) \doteq x_1 = 0, \ (x_1, x_2) \in \mathbb{R}^2\}.$$

*Clearly, here* $C = \mathbb{R}^2$, $K = \mathrm{argmin}_K f = \{(0, x_2) : \ x_2 \in \mathbb{R}\}$. *By virtue of Remark 4.3.7, strong duality does not hold, since* $S_f^-(0) = \{(x_1, x_2) : \ x_1x_2 < 0\}$, $S_g^-(0) = \{(x_1, x_2) : \ x_1 < 0\}$, $S_g^+(0) = \{(x_1, x_2) : \ x_1 > 0\}$, *and therefore*

$$\Omega_+^- = \mathbb{R}_{++} \times (-\mathbb{R}_{++}), \ \Omega_-^- = (-\mathbb{R}_{++}) \times \mathbb{R}_{++}.$$

*On the other hand, given any* $\bar{x} = (0, \bar{x}_2)$, $\bar{x}_2 \in \mathbb{R}$, *one gets* $\nabla g(\bar{x}) = (1, 0)$, $\nabla f(\bar{x}) = 2\bar{x}_2 \nabla g_(\bar{x}) = 2\bar{x}_2(1, 0)$. *Let us check that* $\bar{x}$ *is a KKT point. To that purpose we apply* (*d*) *of Theorem 4.4.1. Obviously* $F_L(\mathbb{R}^2)$ *is convex. We distinguish three cases: if* $\bar{x}_2 = 0$, $\widehat{S}_f^-(0) = \emptyset$ *and so* $\mathcal{L}(\bar{x} = 0) = \{0\}$; *if* $\bar{x}_2 < 0$, *then*

$$\widehat{\Omega}_+^- = \mathbb{R}_{++} \times \mathbb{R}, \ \widehat{\Omega}_-^+ = -\widehat{\Omega}_+^-, \ \widehat{s} = \widehat{r} = \frac{1}{2\bar{x}_2}, \quad \widehat{\Omega}_-^- = \emptyset;$$

*in case* $\bar{x}_2 > 0$, *then*

$$\widehat{\Omega}_-^- = (-\mathbb{R}_{++}) \times \mathbb{R}, \ \widehat{\Omega}_+^+ = -\widehat{\Omega}_-^-, \ \widehat{l} = \widehat{m} = \frac{1}{2\bar{x}_2}, \quad \widehat{\Omega}_+^- = \emptyset.$$

*To be more precise, by applying* (*a*9) *of Theorem 4.3.1, we conclude that* $\mathcal{L}(\bar{x}) = \{-2\bar{x}_2\}$, *that is, there is uniqueness of multipliers at* $\bar{x} = (0, \bar{x}_2)$ *for all* $\bar{x}_2 \in \mathbb{R}$.

The next result provides a sufficient condition for a KKT point to be a strict local minimum. Recall that $K \doteq \{x \in C : \ g(x) = 0\}$.

**Proposition 4.4.6.** *Assume that* $\bar{x} \in C$, $g(\bar{x}) = 0$, *satisfies* (4.29) *for some* $\lambda^* \in \mathbb{R}$, *and*

$$[v \in T(C;\bar{x}), \nabla g(\bar{x})^\top v = \nabla f(\bar{x})^\top v = 0] \Longrightarrow v = 0. \tag{4.32}$$

*Then* $\bar{x}$ *is a strict local solution to problem* (4.1).

*Proof.* By the choice of $\lambda^*$, we obtain

$$\nabla f(\bar{x})^\top v \geq -\lambda^* \nabla g(\bar{x})^\top v, \quad \forall\, v \in T(C;\bar{x}).$$

Since

$$T(K;\bar{x}) \subseteq \{v \in T(C;\bar{x}) :\ \nabla g(\bar{x})^\top v = 0\},$$

we get $\nabla f(\bar{x})^\top v \geq 0$, for all $v \in T(K;\bar{x})$. By (4.32), it follows that

$$\nabla f(\bar{x})^\top v > 0, \quad \forall\, v \in T(K;\bar{x}) \setminus \{0\},$$

which implies that $\bar{x}$ is a strict local solution to (4.1). $\square$

Next instance shows that the previous proposition may be applied to non-pseudoconvex objective functions.

**Example 4.4.7.** *Let us consider* $f(x_1,x_2) = \sin x_1$, $g(x_1,x_2) = x_2 - x_1$, $C \doteq \{(x_1,0) \in \mathbb{R}^2 :\ x_1 \geq 0\} \cup \{(x_1,x_2) \in \mathbb{R}^2_+ :\ x_1 = x_2\}$ *and* $\bar{x} = (0,0)$. *Clearly* $f$ *is not pseudoconvex,* $T(C;\bar{x}) = C$ *and* (4.32) *holds. Furthermore,* (4.29) *is satisfied for* $\lambda^* = 1$, $\mu = -1$ *and* $\operatorname{argmin}_K f = \left\{\dfrac{4k-1}{2}\pi(1,1) :\ k \in \mathbb{N}\right\}$. *By Proposition 4.4.6,* $\bar{x} = (0,0)$ *is a strict local solution (in any open ball with positive radius* $\delta < \pi$*) but not global. Note that* $K \neq C$ *and* $S^-_f(\mu) = \emptyset$ *and clearly* $S^+_g(0) = \{(x_1,x_2) \in C :\ x_2 - x_1 > 0\} = \emptyset$.

*On the other hand,*

$$S^+_f(\mu) = \bigcup_{k\in\mathbb{N}} \left\{(x_1,0) :\ x_1 > 0,\ x_1 \neq \frac{4k-1}{2}\pi\right\} \cup \left\{(x_1,x_1) :\ x_1 \geq 0,\ x_1 \neq \frac{4k-1}{2}\pi\right\},$$

*and* $S^-_g(0) = \{(x_1,0) :\ x_1 > 0\}$. *Thus* $\Omega^+_- = \{(x_1,0) :\ x_1 > 0,\ x_1 \neq \dfrac{4k-1}{2}\pi,\ k \in \mathbb{N}\}$. *Moreover,* $x_k \doteq (k,0) \in \Omega^+_-$ *for all* $k \in \mathbb{N}$, *and*

$$\frac{g(x_k)}{f(x_k)-\mu} = \frac{-k}{\sin k + 1} = \frac{-1}{\frac{1}{k}(\sin k + 1)} \longrightarrow -\infty, \text{ as } k \to +\infty.$$

*Hence* $r = -\infty$, *and by Theorem 4.3.1, part (a14), we obtain* $\overline{\operatorname{cone}}(\mathcal{F}_\mu) = \mathbb{R}_+ \times \mathbb{R}_-$, *which implies that* $\mathcal{L}_{SD} = \,]-\infty, 0]$. *Notice that* $\mathcal{L}(\bar{x}) = \,]-\infty, 1]$ *and that the Hessain of* $f + \lambda^* g$ *is the null matrix.*

The next proposition shows that under suitable assumptions, and starting from an infeasible sequence, one may obtain either an optimal solution or an infeasible KKT-point, as one of its limit points (parts ($a1$) or ($b1$) in the first case, and ($a2$) or ($b2$) in the second one). This result is in connection to enhanced KKT conditions as described, for instance, in [10, 92]. Although the cases $l > 0$ and $s < 0$ are considered, similar expressions can be obtained for $m > 0$ and $r < 0$.

**Proposition 4.4.8.** *Let* $f, g$ *be differentiable with* $\mu \in \mathbb{R}$ *and* $C$ *be closed.*

($a$) *Let* $l > 0$ *and* $x_k \in \Omega^-_-$ *such that* $x_k \to \bar{x}$, $\dfrac{g(x_k)}{f(x_k)-\mu} \to l$ *and* $\dfrac{x_k - \bar{x}}{\|x_k - \bar{x}\|} \to v$, $(v \in T(C;\bar{x}))$. *Then, at least one of the following assertions hold:*

(*a*1) $\bar{x} \in \underset{K}{\operatorname{argmin}}\, f$ *and either* $v \in \widehat{\Omega}_-^-$ (*so* $\mathcal{L}(\bar{x}) \subseteq\, ]-\infty, 0[$) *and* $l = \dfrac{\nabla g(\bar{x})^\top v}{\nabla f(\bar{x})^\top v}$ (*so* $l \geq \widehat{l}$), *or* $\nabla f(\bar{x})^\top v = 0 = \nabla g(\bar{x})^\top v$;

(*a*2) $\bar{x} \in \Omega_-^-$ *and* $l = \dfrac{g(\bar{x})}{f(\bar{x}) - \mu} > 0$; *in which case,*

$$\nabla f(\bar{x}) - \frac{1}{l}\nabla g(\bar{x}) \in [T(C;\bar{x})]^*. \tag{4.33}$$

*Consequently* $F_L(T(C;\bar{x})) + \mathbb{R}_+(1,0)$ *is convex. Hence, if* $\widehat{\Omega}_-^- \neq \emptyset$ *then* $l \leq \widehat{l}$.

(*b*) *Let* $s < 0$ *and* $x_k \in \Omega_+^-$ *such that* $x_k \to \bar{x}$, $\dfrac{g(x_k)}{f(x_k) - \mu} \to s$ *and* $\dfrac{x_k - \bar{x}}{\|x_k - \bar{x}\|} \to v$. *Then, at least one of the following assertions hold:*

(*b*1) $\bar{x} \in \underset{K}{\operatorname{argmin}}\, f$ *and either* $v \in \widehat{\Omega}_+^-$ (*so* $\mathcal{L}(\bar{x}) \subseteq\, ]0, +\infty[$) *and* $s = \dfrac{\nabla g(\bar{x})^\top v}{\nabla f(\bar{x})^\top v}$ (*so* $s \leq \widehat{s}$), *or* $\nabla f(\bar{x})^\top v = 0 = \nabla g(\bar{x})^\top v$;

(*b*2) $\bar{x} \in \Omega_+^-$ *and* $s = \dfrac{g(\bar{x})}{f(\bar{x}) - \mu} < 0$; *in which case,*

$$\nabla f(\bar{x}) - \frac{1}{s}\nabla g(\bar{x}) \in [T(C;\bar{x})]^*. \tag{4.34}$$

*Consequently* $F_L(T(C;\bar{x})) + \mathbb{R}_+(1,0)$ *is convex. Hence, if* $\widehat{\Omega}_+^- \neq \emptyset$ *then* $s \geq \widehat{s}$.

*Proof.* We only check (*a*), since the other are entirely similar. Obviously $\bar{x} \in C$, $g(\bar{x}) \leq 0$ and $f(\bar{x}) \leq \mu$. Certainly, none of the conditions: $f(\bar{x}) < \mu$, $g(\bar{x}) = 0$, or $f(\bar{x}) = \mu$, $g(\bar{x}) < 0$ are possible.
Assume first that $f(\bar{x}) = \mu$ and $g(\bar{x}) = 0$. Then $\bar{x} \in \underset{K}{\operatorname{argmin}}\, f$. Since $0 > f(x_k) - \mu = f(x_k) - f(\bar{x}) = \nabla f(\bar{x})^\top (x_k - \bar{x}) + o(\|x_k - \bar{x}\|)$ with $o(t)/t \to 0$ as $t \downarrow 0$, we obtain $\nabla f(\bar{x})^\top v \leq 0$; similarly for $g$, $\nabla g(\bar{x})^\top v \leq 0$. Since

$$\frac{g(x_k)}{f(x_k) - \mu} = \frac{[g(\bar{x}) + \nabla g(\bar{x})^\top (x_k - \bar{x}) + o(\|x_k - \bar{x}\|)]/\|x_k - \bar{x}\|}{[\nabla f(\bar{x})^\top (x_k - \bar{x}) + o(\|x_k - \bar{x}\|)]/\|x_k - \bar{x}\|} \to l,$$

one concludes that, in fact, either $\nabla f(\bar{x})^\top v < 0$ and $\nabla g(\bar{x})^\top v < 0$ or $\nabla f(\bar{x})^\top v = 0 = \nabla g(\bar{x})^\top v = 0$. In the first case, $v \in \widehat{\Omega}_-^-$ and $l = \dfrac{\nabla g(\bar{x})^\top v}{\nabla f(\bar{x})^\top v}$. This completes the proof of (*a*1).

Assume now that $f(\bar{x}) < \mu$ and $g(\bar{x}) < 0$, that is, $\bar{x} \in \Omega_-^-$, and therefore $l = \dfrac{g(\bar{x})}{f(\bar{x}) - \mu} > 0$. It remains to check (4.33). By the first order necessary optimality condition

$$\nabla \left(\frac{g}{f - \mu}\right)(\bar{x}) \in [T(\Omega_+^-;\bar{x})]^*.$$

Since $T(\Omega_-^-;\bar{x}) = T(C;\bar{x})$, the last expression reduces to (4.33), and so the proof of (*a*2) is completed. $\square$

The previous proposition has important consequences related to the KKT optimality conditions. In that direction some remarks are in order.

**Remark 4.4.9.** *(i) Notice that, in the situation (a1), i.e., $\bar{x} \in \operatorname*{argmin}_K f$, if one additionally assumes the convexity of $\operatorname{cone}(\mathcal{F}_\mu)$, then strong duality for problem (4.1) holds and $\mathcal{L}_{SD} = \mathcal{S}_D$. This implies $\nabla f(\bar{x}) + \lambda_0 \nabla g(\bar{x}) \in [T(C;\bar{x})]^*$ for all $\lambda_0 \in \mathcal{L}_{SD}$ (so $\mathcal{L}_{SD} \subseteq \mathcal{L}(\bar{x})$), because of (4.24). A similar reasoning applies in case of (b1).*

*(ii) Assume the convexity of $F_L(T(C;\bar{x})) + \mathbb{R}_+(1,0)$. Under (a), if $\bar{x} \in \operatorname*{argmin}_K f$ and $l = \widehat{l}$, then $-\dfrac{1}{\widehat{l}} \in \mathcal{L}(\bar{x})$, i. e., $\bar{x}$ is a KKT point; in case $l > \widehat{l}$, $-\dfrac{1}{l} \notin \mathcal{L}(\bar{x})$; and in case $\widehat{l} = 0$, we have $\mathcal{L}(\bar{x}) = \emptyset$, i. e., $\bar{x}$ is not a KKT point, as a consequence of Theorem 4.4.1.*

Next, the critical values $l = 0$ and $s = 0$, under which strong duality is not satisfied, are considered.

**Proposition 4.4.10.** *Let $f, g$ be differentiable with $\mu \in \mathbb{R}$ and $C$ be closed.*

*(a) Let $l = 0$ and $x_k \in \Omega_-^-$ such that $x_k \to \bar{x}$, $\dfrac{g(x_k)}{f(x_k) - \mu} \to 0$ and $\dfrac{x_k - \bar{x}}{\|x_k - \bar{x}\|} \to v$, $(v \in T(C;\bar{x}))$. Then, $\bar{x} \in \operatorname*{argmin}_K f$ and either $\nabla f(\bar{x})^\top v < 0 = \nabla g(\bar{x})^\top v$ (so $\mathcal{L}(\bar{x}) = \emptyset$), or $\nabla f(\bar{x})^\top v = 0 = \nabla g(\bar{x})^\top v$.*

*(b) Let $s = 0$ and $x_k \in \Omega_+^-$ such that $x_k \to \bar{x}$, $\dfrac{g(x_k)}{f(x_k) - \mu} \to 0$ and $\dfrac{x_k - \bar{x}}{\|x_k - \bar{x}\|} \to v$, $(v \in T(C;\bar{x}))$. Then, $\bar{x} \in \operatorname*{argmin}_K f$ and either $\nabla f(\bar{x})^\top v < 0 = \nabla g(\bar{x})^\top v$ (so $\mathcal{L}(\bar{x}) = \emptyset$), or $\nabla f(\bar{x})^\top v = 0 = \nabla g(\bar{x})^\top v$.*

*Proof.* The proof follows the same line of reasoning of the preceding proposition.

□

Now let us go back to Example 4.3.10, which satisfies $s = 0$. With the same data, consider

$$\bar{x} = (\bar{x}_1, \bar{x}_2) \in \operatorname*{argmin}_K f = \{(x_1, x_2) \in \mathbb{R}^2 : \; x_1 + x_2 = 0\}, \quad x_k = (x_{1k}, x_{2k}) = (\bar{x}_1, \bar{x}_2 - \frac{1}{k}).$$

Then, $x_k$ satisfies the conditions in $(b)$ of Proposition 4.4.10 with $v = (0, -1)$. We see that

$$\nabla f(\bar{x})^\top v = -1 < 0, \quad \nabla g(\bar{x}) = 0.$$

Hence, it is possible to have a solution without being a KKT-point when strong duality fails.

## 4.5 The non-polyhedral standard quadratic optimization problem

Let us consider the standard quadratic problem:

$$\min_{x\in\Delta} \frac{1}{2}x^\top Ax, \tag{4.35}$$

where $A = (a_{ij})$ is a real symmetric matrix with positive entries $a_{ij} > 0$, and $\Delta = \{x \in \mathbb{R}^n_+ : \mathbf{1}^\top x = 1\}$.

Although this model is very special, it retains, as asserted in [14], most of the complexity of the general quadratic case, where a polyhedron $P$, instead of $\Delta$, is considered. As applications of (4.35), we mention quadratic allocation problems [47], portfolio optimization problems [61, 62], the maximum weight clique problem [67, 43], among others. Due to the structure of $\Delta$, any quadratic objective function may be reduced to an homogeneous one.

It was established in [14, Theorem 4], via the existence theorem due to Frank Wolfe and Theorem 5 in [13], that strong duality holds for (4.35).

We will prove the validity of strong duality by applying our Theorem 4.3.1, providing further qualitative and quantitative information for the more general problem where $\Delta$ is substituted by a convex and compact base of any pointed, closed, convex (possibly ice-cream, or more general circular) cone $C$ (so, the Frank-Wolfe theorem is not applicable in this case):

$$\mu_q \doteq \min\{\frac{1}{2}x^\top Ax : \ e^\top x = 1, \ x \in C\}, \tag{4.36}$$

where $C \subseteq \mathbb{R}^n$ is as above with $e \in \text{int } C^*$, and $A$ is a symmetric copositive matrix on $C$, i. e., $x^\top Ax \geq 0$ for all $x \in C$, satisfying

$$\mu_q > 0 (= \min_{x\in C} \frac{1}{2}x^\top Ax). \tag{4.37}$$

This requirement holds, for instance, if $A$ is strictly copositive on $C$, i. e., $x^\top Ax > 0$ for all $x \in C$, $x \neq 0$. Thus, the dual to (4.36) is

$$\sup_{\lambda\in\mathbb{R}} \inf_{x\in C} L(\lambda, x) \doteq \frac{1}{2}x^\top Ax + \lambda(e^\top x - 1), \tag{4.38}$$

whose solution set is $\mathcal{S}_D = \{-2\mu_q\}$ as we shall see next. Setting

$$f(x) \doteq \frac{1}{2}x^\top Ax, \quad g(x) \doteq e^\top x - 1, \quad K \doteq \{x \in C : \ g(x) = 0\},$$

the following proposition collects all the properties satisfied by problem (4.36). In particular, Part $(f)$ shows that every infeasible sequence admits limit points which are either optimal solutions and KKT-points, or infeasible but still KKT points.

Notice that, because of the choice of $e$, $C = \text{cone } K$, i. e., $K$ is a convex and compact base of $C$. We point out that next result improves and extends that in [14, Theorem 4] (valid only for polyhedra) to non-polyhedral cones $C$, and in this context, the result is new. We will use the notations of Sections 4.3 and 4.4.

**Proposition 4.5.1.** *Assume that (4.37) is satisfied. For the problem (4.36) with the data as above, we have that* $\operatorname*{argmin}_K f$ *is nonempty and compact. Furthermore the following is true:*

(a) $\Omega_+^- = \emptyset = \Omega_+^=$, *and so* $S_g^+(0) = \Omega_+^+ \neq \emptyset$, $\emptyset \neq S_f^-(\mu_q) = \Omega_-^-$;

(b) $m = l = \dfrac{1}{2\mu_q}$, *so* $l > 0$ *and* $\mathcal{L}_{SD} = \mathcal{S}_D = \{-2\mu_q\}$, *and so*

$$\operatorname{cone}(F(C) + \mathbb{R}_+(1,0) - \mu_q(1,0)) = \{(u,v) \in \mathbb{R}^2 : \ v \leq \frac{1}{2\mu_q}u\};$$

(c) *strong duality holds for* (4.36)*;*

(d) $\mathcal{L}(\bar{x}) = \{-2\mu_q\}$ *for all* $\bar{x} \in \operatorname*{argmin}_K f$*;*

(e) *Let* $x_k \in S_f^-(\mu_q)$, $\dfrac{g(x_k)}{f(x_k) - \mu_q} \to l = \dfrac{1}{2\mu_q}$. *Then, there exists a subsequence (still indexed by* $k$*) and* $\bar{x}$ *such that* $x_k \to \bar{x}$, $\dfrac{x_k - \bar{x}}{\|x_k - \bar{x}\|} \to v \in T(C;\bar{x})$, *and at least one of the following assertions holds:*

   (e1) $\bar{x} \in \operatorname*{argmin}_K f$ *(so* $-2\mu_q \in \mathcal{L}(\bar{x})$*) and either* $[v \in \widehat{\Omega}_-^-$ *and* $\dfrac{1}{2\mu_q} = \dfrac{e^\top v}{(A\bar{x})^\top v}]$ *or* $e^\top v = 0 = (A\bar{x})^\top v$.

   (e2) $\bar{x} \in S_f^-(\mu_q)$ *and* $\dfrac{1}{2\mu_q} = \dfrac{g(\bar{x})}{f(\bar{x}) - \mu_q} > 0$, *in which case,* $-2\mu_q \in \mathcal{L}(\bar{x})$.

*Proof.* (a): Take any $y \in \Omega_+^-$; then there exists $t > 0$, $x \in C$ such that $y = tx$ and $g(x) = 0$. Then $0 < g(y) = e^\top y - 1 = t - 1$ and $0 > f(y) - \mu_q \geq \mu_q(t^2 - 1)$, implying that $\mu_q < 0$, yielding a contradiction. This proves that $\Omega_+^- = \emptyset$. The same argument also shows $\Omega_+^= = \emptyset$. From (4.37), we get $S_f^-(\mu_q) \neq \emptyset$, and so the last equality follows from (a).

(b): We claim that any $x \in \Omega_+^+$ satisfies $\dfrac{g(x)}{f(x) - \mu_q} \leq \dfrac{1}{2\mu_q}$. Indeed, writting $x = ty$ for some $0 < t < 1$ and $y \in K$, one obtains

$$\mu_q(2t - 1) \leq t^2\mu_q \leq t^2\frac{1}{2}y^\top Ay.$$

From which the claim is proved, implying $m \leq \dfrac{1}{2\mu_q}$. A similar argument also shows $\dfrac{1}{2\mu_q} \leq l$.

On the other hand, take any $\overline{x} \in \operatorname{argmin}_K f$ and consider the sequences in $C$, $x_k \doteq (1 + \dfrac{1}{k})\overline{x}$ and $y_k \doteq (1 - \dfrac{1}{k})\overline{x}$. Then, for all $k \in \mathbb{N}$, $g(x_k) = e^\top x_k - 1 = \dfrac{1}{k} > 0$ and $f(y_k) = \dfrac{1}{2}y_k^\top Ay_k = (1 - \dfrac{1}{k})^2\dfrac{1}{2}\overline{x}^\top A\overline{x} < \mu_q$, which proves that $x_k \in S_g^+(0)$ and $y_k \in S_f^-(\mu_q)$. Thus, by (a), we get $x_k \in \Omega_+^+$ and $y_k \in \Omega_-^-$. Hence, for all $k \in \mathbb{N}$,

$$\frac{g(x_k)}{f(x_k) - \mu_q} = \frac{1/k}{(1 + \frac{1}{k})^2\mu_q - \mu_q} = \frac{1/k}{[(1 + \frac{1}{k})^2 - 1]\mu_q} = \frac{1/k}{\frac{1}{k}(2 + \frac{1}{k})\mu_q} = \frac{1}{(2 + \frac{1}{k})\mu_q}.$$

This implies that $m \geq \dfrac{1}{2\mu_q}$ and so $m = \dfrac{1}{2\mu_q}$. Similarly, by using the sequence $y_k$, one concludes that $l \leq \dfrac{1}{2\mu_q}$, and so $l = \dfrac{1}{2\mu_q}$, proving the first part. This means that $(a8)$ in Theorem 4.3.1 is satisfied, obtaining the expression for the set $\mathrm{cone}(F(C) + \mathbb{R}_+(1,0) - \mu_q(1,0))$.
$(c)$: By $(c)$ and Theorem 4.3.2, it follows that strong duality holds.
$(d)$: By $(b)$ and Proposition 4.4.2 one has $-2\mu_q \in \mathcal{L}(\bar{x})$, and so, by Theorem 4.4.1 the cone

$$\overline{F_L(T(C;\bar{x})) + \mathbb{R}_+(1,0)} = \{(x,y) \in \mathbb{R}^2 : \ y \leq \widehat{m}x, \ y \leq \widehat{l}x\}$$

is convex. This means that $\widehat{m} \leq \widehat{l}$. On the other hand, by taking $x_k \doteq (1 - \dfrac{1}{k})\bar{x}$ or $x_k \doteq (1 + \dfrac{1}{k})\bar{x}$, one shows, respectively, that $-\bar{x} \in T(C;\bar{x})$ and $\bar{x} \in T(C;\bar{x})$. Thus $-\bar{x} \in \widehat{\Omega}^-_-$ and $\bar{x} \in \widehat{\Omega}^+_+$. It follows that $\widehat{l} \leq \dfrac{1}{2\mu_q}$ and $\widehat{m} \geq \dfrac{1}{2\mu_q}$, which, along with a previous inequality, allows us to conclude that $\widehat{l} = \widehat{m} = \dfrac{1}{2\mu_q}$. Hence $\mathcal{L}(\bar{x}) = \{-2\mu_q\}$, since, by (4.28), one gets

$$\gamma \in \mathcal{L}_(\bar{x}) \Longleftrightarrow (1,\gamma) \in [F_L(T(C;\bar{x})) + \mathbb{R}_+(1,0)]^* = \mathbb{R}_+(1, -2\mu_q).$$

$(e)$: This follows from $(a)$ of Proposition 4.4.8.

$\square$

We emphasize that when $C = \mathbb{R}^n_+$ and $e = (1,\ldots,1)$, every feasible KKT point, to (4.35), $\bar{x}$, is characterized by the existence of $\lambda \in \mathbb{R}$ and $y \in \mathbb{R}^n_+$ such that

$$\nabla f(\bar{x}) + \lambda e - y = 0, \quad y^\top \bar{x} = 0.$$

In other words, every feasible KKT point is a KKT point in the usual sense. However, the strong duality property associated to the Lagrangian

$$L(x,\lambda,y) \doteq f(x) + \lambda g(x) - y^\top x, \ \lambda \in \mathbb{R}, \ y \in \mathbb{R}^n_+,$$

holds if and only if $A$ is positive semidefinite as was observed in [14, Section 3].

# Chapter 5

# Strong duality reveals hidden convexity

## 5.1 Introduction

The standard quadratic optimization problem:

$$\min_{x\in\Delta} \frac{1}{2} x^\top A x, \tag{5.1}$$

where, $\Delta$ is the simplex $\{x \in \mathbb{R}^n : \ \mathbf{1}^\top x = 1, \ x \geq 0\}$ with $A$ being any real symmetric matrix having positive entries.

In connection to (5.1) three main dual problems may be considered:

$$\sup_{\lambda_0\in\mathbb{R}} \inf_{x\in\mathbb{R}^n_+} \left\{\frac{1}{2} x^\top A x + \lambda_0(\mathbf{1}^\top x - 1)\right\}; \tag{5.2}$$

$$\sup_{\lambda\in\mathbb{R}^n_+} \inf_{x\in X} \left\{\frac{1}{2} x^\top A x - \lambda^\top x\right\}, \quad X \doteq \{x \in \mathbb{R}^n : \ \mathbf{1}^\top x = 1\}; \tag{5.3}$$

$$\sup_{(\lambda_0,\lambda)\in\mathbb{R}\times\mathbb{R}^n_+} \inf_{x\in\mathbb{R}^n} \left\{\frac{1}{2} x^\top A x - \lambda_0(\mathbf{1}^\top x - 1) - \lambda^\top x\right\}. \tag{5.4}$$

Because of many real concrete applications, for instance in allocation problems, $A$ is simply copositive on $\mathbb{R}^n_+$. Thus, one cannot expect that strong duality with respect to (5.4) (resp. with respect to (5.3)) holds, as stated partly in [14], since this property requires the positive semidefiniteness of $A$ (resp. copositivity of $A$ on the orthogonal subspace to $\mathbf{1}$.

Motivated by the previous considerations, we propose to analyze the same issues, and beyond, for the generalized standard quadratic optimization problem, where $\Delta$ is substituted by a convex and compact base of any pointed, closed, convex (possibly circular) cone $C \subseteq \mathbb{R}^n$. More precisely, we will deal with the following problem

$$\mu_q \doteq \min\left\{\frac{1}{2} x^\top A x : \ e^\top x = 1, \ x \in C\right\}, \tag{5.5}$$

where $A$ is a real symmetric matrix, $e \in \text{int } C^*$, and $C \subseteq \mathbb{R}^n$ is a pointed, closed, convex cone having non-empty interior. The feasible set to (5.5), that is, $K \doteq \{x \in C: \ e^\top x = 1\}$, becomes a convex and compact base of $C$. We analyze the cases $\mu_q = 0$ and $\mu_q > 0$.

In order to discuss the validity of strong duality of (5.5) with respect to the dual problems analogous to (5.3) and (5.4), we formulate (5.5) as a semi-infinite optimization problem, and will apply some of the main results from [25]. To be more precise, we will establish that strong duality with respect to (5.2) (respectively, (5.3) and (5.4)) holds if and only if $A$ is copositive on $\mathbb{R}^n_+$ (respectively, $A$ is copositive on $\mathbf{1}^\perp$, $A$ is positive semidefinite), see Section 5.3.

## 5.2 The general case with one single equality and geometric constraints

We now deal with the abstract minimization problem under one single equality and a geometric constraints. Let $f, g : C \subseteq X \to \mathbb{R}$ be any finite-valued functions, with $X$ to be a normed vector space. Let us consider the problem

$$\mu \doteq \inf\{f(x): \ \ g(x) = 0, \ x \in C\}, \tag{5.6}$$

whose (Lagrangian) dual problem is defined by

$$\nu \doteq \sup_{\lambda \in \mathbb{R}} \inf_{x \in C} [f(x) + \lambda g(x)]. \tag{5.7}$$

Set $F(x) \doteq (g(x), f(x))$. Assuming that $\mu \in \mathbb{R}$, we obtain

$$(F(C) - \mu(0,1)) \cap -(\{0\} \times \mathbb{R}_{++}) = \emptyset, \tag{5.8}$$

which amounts to writing

$$(F(C) + \{0\} \times -\mu(0,1)) \cap -(\{0\} \times \mathbb{R}_{++}) = \emptyset, \tag{5.9}$$

or, equivalently,

$$\text{cone}(F(C) + \mathbb{R}_+(0,1) - \mu(0,1)) \cap -(\{0\} \times \mathbb{R}_{++}) = \emptyset. \tag{5.10}$$

We will show, next, that the zero duality gap and strong duality can be characterized by reinforcing (5.10).

The optimal value function $\psi : \mathbb{R} \to \mathbb{R} \cup \{\pm\infty\}$ to problem (5.6) is defined by

$$\psi(a) = \begin{cases} \inf\{f(x): \ x \in K(a)\} & \text{if} \quad K(a) \neq \emptyset; \\ +\infty & \text{otherwise}, \end{cases}$$

where

$$K(a) \doteq \{x \in C: \ g(x) = a\}. \tag{5.11}$$

Notice that $K = K(0)$, and $K(a) \neq \emptyset$ if and only if $a \in g(C)$, that is,

$$\text{dom } \psi \doteq \{a \in \mathbb{R}: \ \psi(a) < +\infty\} = g(C).$$

The sets

$$\mathcal{F} \doteq F(C) + \mathbb{R}_+(0,1), \;\; \mathcal{E}_\rho \doteq \mathcal{F} - \rho(0,1), \quad \rho \in \mathbb{R}. \tag{5.12}$$

will play an important role in our analysis.

Some topological and geometrical properties of $\psi$ are shown in the following theorem, which is a particular case of Theorem 3.2 in [34].

**Proposition 5.2.1.** *Let $f, g, F$ be as above. The following assertions hold.*

(a) $(a, r) \in \text{epi } \psi \Longleftrightarrow (a, r + \frac{1}{k}) \in F(C) + \mathbb{R}_+(0,1), \quad \forall\, k \in \mathbb{N}$.
*As a consequence, if $F(C) + \mathbb{R}_+(0,1)$ is convex then $\psi$ is convex.*

(b) $F(C) + \mathbb{R}_+(0,1) \subseteq \text{epi } \psi \subseteq \overline{F(C) + \mathbb{R}_+(0,1)}$.
*Consequently,*

$$\overline{\mathcal{E}_\mu} = \overline{\text{epi } \psi} - \mu(0,1) = \text{epi } \overline{\psi} - \mu(0,1); \quad \overline{\text{co}}\; \mathcal{E}_\mu = \overline{\text{co}}(\text{epi } \psi) - \mu(0,1) = \text{epi}(\overline{\text{co}}\; \psi) - \mu(0,1).$$

Recall that (see for instance [34, Theorem 3.1]) if $\mu = \psi(0) \in \mathbb{R}$ then,

$$\nu = \psi^{**}(0) \tag{5.13}$$

Proposition 5.2.1 leads to the following characterization of lower semicontinuity of $\psi$ at 0 and to zero duality gap. Similar results may be found in Lemma 3.1 and Theorem 4.1 of [63]. Notice that (5.14) reinforces (5.10).

**Theorem 5.2.1.** *Assume that $\mu = \psi(0)$ is finite. Then,*

(a) $\overline{\psi}(0) = \psi(0)$ *if and only if*

$$\overline{\mathcal{E}_\mu} \cap (-\{0\} \times \mathbb{R}_{++}) = \emptyset. \tag{5.14}$$

(b) $\nu = \overline{\text{co}}\; \psi(0)$.

*Proof.* (a): Since $\mu = \psi(0)$, Theorem 5.2.1 asserts that (5.14) implies $\overline{\psi}(0) \geq \psi(0)$, which along with

$$\overline{\psi}(a) \leq \psi(a), \quad \forall\, a \in \mathbb{R}, \tag{5.15}$$

yield $\overline{\psi}(0) = \psi(0)$. Conversely, from (5.15) and Theorem 5.2.1, we immediately obtain that $\overline{\psi}(0) = \psi(0)$ implies (5.14).
(b) By virtue of (5.13), we must check that $\psi^{**}(0) = \overline{\text{co}}\; \psi(0)$. If $\overline{\text{co}}\; \psi(0) = -\infty$ then $\psi^{**}(0) = -\infty$ since $\psi^{**} \leq \overline{\text{co}}\; \psi$. In case $\overline{\text{co}}\; \psi(0) \in \mathbb{R}$, we get $\overline{\text{co}}\; \psi$ never takes the value $-\infty$, since $\overline{\text{co}}\; \psi$ is convex and lsc, and so $\overline{\text{co}}\; \psi = \psi^{**}$ by (2.14) (see Chapter 2). $\square$

We now characterize the zero duality gap for problem (5.6), which reduces to the lower semicontinuity of $\psi$ at 0, under convexity of $\overline{\mathcal{E}}_\mu$. Under this latter assumption, a similar results was obtained in [54] for a semi-infinite optimization problem.

**Proposition 5.2.2.** *Assume that* $\mu = \psi(0)$ *is finite. Then,*
$\mu = \nu$ *if and only if*

$$\overline{\text{co}}\ \mathcal{E}_\mu \cap (-\{0\} \times \mathbb{R}_{++}) = \emptyset. \tag{5.16}$$

*Proof.* By using the last part of Proposition 5.2.1 we infer from (5.16) that $\overline{\text{co}}\ \psi(0) \geq \mu = \psi(0)$. Since $\overline{\text{co}}\ \psi \leq \psi$ always holds, we obtain $\overline{\text{co}}\ \psi(0) = \psi(0)$. Conversely, again from Proposition 5.2.1 it follows that $\overline{\text{co}}\ \psi(0) = \psi(0)$ implies (5.16).

□

## 5.3 The circular standard quadratic optimization problem

Let us go back to our problem

$$\mu_q \doteq \min\left\{\frac{1}{2}x^\top Ax : \ e^\top x = 1, \ x \in C\right\}, \tag{5.17}$$

where $C \subseteq \mathbb{R}^n$ is a pointed, closed, convex cone having non-empty interior. Obviously the feasible set to (5.17) is $K \doteq \{x \in C : \ e^\top x = 1\}$, which becomes a convex and compact base of $C$ provided $e \in \text{int}\ C^*$ (it is non empty since $C$ is pointed); $A$ is a real symmetric matrix. We say that $A$ is strictly copositive on $P$ if $x^\top Ax > 0$ for all $x \in P$, $x \neq 0$. We discuss the cases $\mu_q = 0$ and $\mu_q > 0$.

It is easy to check that

- $\mu_q \ = \ 0 \Longleftrightarrow A$ is copositive but not strictly copositive on $C$;
- $\mu_q \ > \ 0 \Longleftrightarrow A$ is strictly copositive on $C$.

The specialization $C = \mathbb{R}^n_+$; $A$ having positive entries; $e = (1, 1, \ldots, 1) \in \mathbb{R}^n$, discussed at the introduction, is termed the standard quadratic optimization problem and was studied in many papers, and models quadratic allocation problems, portfolio optimization problems, the maximum weight clique problem, among others.

We will describe the three main dual problems associated to (5.17). To that end, we formulate problem (5.17) as a semi-infinite optimization problem:

$$\mu_q \doteq \min\{f(x) : \ x \in X, \ g_i(x) \leq 0, \ \forall\ i \in I\}, \tag{5.18}$$

where

$$I \doteq -C^*, \ f(x) \doteq \frac{1}{2}x^\top Ax, \ g_0(x) \doteq e^\top x - 1, \quad g_i(x) \doteq i^\top x, \ i \in I,$$

and $X \doteq \{x \in \mathbb{R}^n : \ e^\top x = 1\}$. Thus, we consider the following three dual problems:

$$\nu_0 \ \doteq \ \sup_{\lambda \in \mathbb{R}} \inf_{x \in C} \Big\{f(x) + \lambda g_0(x)\Big\}; \tag{5.19}$$

$$\nu_1 \ \doteq \ \sup_{\lambda \in \mathbb{R}^{(I)}_+} \inf_{x \in X} \Big\{f(x) + \sum_{i \in I} \lambda_i g_i(x)\Big\}; \tag{5.20}$$

$$\nu_2 \ \doteq \ \sup_{(\lambda_0, \lambda) \in \mathbb{R} \times \mathbb{R}^{(I)}_+} \inf_{x \in \mathbb{R}^n} \Big\{f(x) + \lambda_0 g_0(x) + \sum_{i \in I} \lambda_i g_i(x)\Big\}, \tag{5.21}$$

analogous to (5.2), (5.3) and (5.4) respectively. Here, $\mathbb{R}^{(I)}$ is the topological dual of $\mathbb{R}^I$ (it stands for the set of real-valued functions defined on $I$, endowed with the usual product topology), which is the space of generalized sequences $\lambda = (\lambda_i)_{i\in I}$ such that $\lambda_i \in \mathbb{R}$, for each $i \in I$, and with only finitely many $\lambda_i$ different from zero. The supporting set of $\lambda$ is supp $\lambda \doteq \{i \in I : \lambda_i \neq 0\}$. Thus

$$\langle \lambda, z\rangle = \lambda(z) = \sum_{i\in I} \lambda_i z_i \doteq \sum_{i\in \text{supp } \lambda} \lambda_i z_i, \quad \forall\ z \in \mathbb{R}^I,\ \forall\ \lambda \in \mathbb{R}^{(I)}.$$

If $\lambda = 0$ then supp $\lambda = \emptyset$, and so we have $\sum_\emptyset = 0$. In addition, $\mathbb{R}^{(I)}_+$ denotes the non-negative cone in $\mathbb{R}^{(I)}$.

We will discuss first the validity of strong duality for (5.17) with respect to the duals (5.20) and (5.21), and we will see that strong duality with respect to (5.19) is suitable, since in most applications the positive semidefiniteness of $A$ fails. To that purpose, we need some preliminaries.

Set

$$M \doteq \text{cone co}\left(\bigcup_{i\in I} \text{epi } g_i^* \cup \text{epi } \delta_X^*\right).$$

Here, $g_i^*$ (resp. $\delta_X^*$), denotes the conjugate function of $g_i$ (resp. $\delta_X$). For a problem formulated as in (5.18) for general $f$, $g_i$ and $X$, with optimal value $\mu$ instead of $\mu_q$, the following two conditions arise:

$$M \ \text{ is closed;} \tag{5.22}$$

$$\text{epi } f^* + \overline{M} \ \text{ is closed.} \tag{5.23}$$

Notice that

$$M = \text{cone co}\left(\bigcup_{i\in I} \text{epi } g_i^*\right) + \text{epi } \delta_X^*.$$

It is proved, in Theorem 5 of [25], that, given any proper lsc and convex functions $f$ and $g_i$ ($i \in I$), $X$ a non-empty convex closed set, under (5.22) and (5.23) and assuming $\mu$ finite, one gets

$$\exists\ \lambda^* \in \mathbb{R}^{(I)}_+ : \quad f(x) + \sum_{i\in I} \lambda_i^* g_i(x) \geq \mu,\ \forall\ x \in X, \tag{5.24}$$

or, equivalently, there exists $\lambda^* \in \mathbb{R}^{(I)}_+$ such that

$$\mu = \sup_{x\in X}\{f(x) + \sum_{i\in I} \lambda_i^* g_i(x)\}. \tag{5.25}$$

**Remark 5.3.1.** *Theorem 1 in [25] asserts that if $f$ is either linear or continuous at some point of the feasible set of* (5.18)*, $K$, then, the fulfillment of* (5.22) *implies that* (5.23) *is also satisfied.*

### 5.3.1 Characterizing strong duality with respect to (5.21)

We are ready to apply Theorem 5 in [25] to our model (5.17). The following theorem establishes that standard strong duality for problem (5.18) or equivalently, (5.17), with respect to (5.20) holds if, and only if $A$ is positive semidefinite. So, this dual is not suitable when considering the porfolio optimization problem since $A$ is only copositive.

**Theorem 5.3.1.** *Let us consider problem* (5.17) *with $C$ being any pointed closed convex cone with nonempty interior, and $e \in \operatorname{int} C^*$. The following assertions are equivalent:*

(a) $A \succcurlyeq 0$;

(b) $A \succcurlyeq 0$ *and there exists* $\lambda^* \in \mathbb{R}^{(I)}_+$ *such that*

$$f(x) + \sum_{i \in I} \lambda_i^* g_i(x) \geq \mu_q, \ \forall\ x \in X; \tag{5.26}$$

(c) *there exist* $(\lambda_0^*, \lambda^*) \in \mathbb{R} \times \mathbb{R}^{(I)}_+$ *such that*

$$f(x) + \lambda_0^* g_0(x) + \sum_{i \in I} \lambda_i^* g_i(x) \geq \mu_q, \ \forall\ x \in \mathbb{R}^n, \tag{5.27}$$

*or, equivalently,*

$$\mu_q = \inf_{x \in \mathbb{R}^n} \Big\{ f(x) + \lambda_0^* g_0(x) + \sum_{i \in I} \lambda_i^* g_i(x) \Big\}.$$

*Proof.* (c) $\Rightarrow$ (a): It is immediate.
(a) $\Rightarrow$ (b): By virtue of Remark 5.3.1, we need to check only that condition (5.22) holds. Since

$$g_i^*(u) = \sup_{x \in \mathbb{R}^n} \{u^\top x - i^\top x\} = \delta_{\{i\}}(u), \tag{5.28}$$

one obtains epi $g_i^* = \{i\} \times \mathbb{R}_+$, and so

$$\operatorname{co} \left( \bigcup_{i \in I} \operatorname{epi} g_i^* \right) = \operatorname{co}[(-C^*) \times \mathbb{R}_+] = (-C^*) \times \mathbb{R}_+.$$

Hence

$$\operatorname{cone} \operatorname{co} \left( \bigcup_{i \in I} \operatorname{epi} g_i^* \right) = (-C^*) \times \mathbb{R}_+.$$

On the other hand, by writing $X = e^\perp + \bar{x}$ with $\bar{x} \in X$, we get

$$\delta_X^*(u) = \sup_{x \in \mathbb{R}^n} \{u^\top x - \delta_X(x)\} = \sup_{x \in X} u^\top x = \sup_{v \in e^\perp} u^\top v + u^\top \bar{x} = \delta_{\mathbb{R}e}(u) + u^\top \bar{x}.$$

Thus,

$$\operatorname{epi} \delta_X^* = \mathbb{R}(e, 1) + \mathbb{R}_+(0, 1).$$

Consequently,

$$\begin{aligned} M &= \text{cone co}\left(\bigcup_{i\in I}\text{epi } g_i^* \cup \text{epi } \delta_X^*\right) = \text{cone co}\left(\bigcup_{i\in I}\text{epi } g_i^*\right) + \text{epi } \delta_X^* && (5.29)\\ &= (-C^*)\times\mathbb{R}_+ + \mathbb{R}(e,1) + \mathbb{R}_+(0,1) = (-C^*)\times\mathbb{R}_+ + \mathbb{R}(e,1), && (5.30)\end{aligned}$$

which is closed, i. e., (5.22) is satisfied. Here we use the result: given two closed sets $M$ and $N$, $M-N$ is closed provided $M^\infty \cap N^\infty = \{0\}$. Then, (5.24) holds, and so (5.25) holds as well, proving $(b)$.
$(b) \Rightarrow (c)$: By setting $\varphi(x) = f(x) + \sum_{i\in I}\lambda_i^* g_i(x)$, it is not difficult to check that: $(\varphi, g_0)(\mathbb{R}^n) + \mathbb{R}_+(1,0)$ is convex (since $\varphi$ is convex and $g_0$ is affine), and there exist $x_0, x_1 \in \mathbb{R}^n$ satisfying $g_0(x_1) < 0 < g_0(x_1)$. Thus, strong duality holds for problem (5.25) (a usual application of a convex separation theorem yields the conclusion), that is, there exists $\lambda_0^* \in \mathbb{R}$ such that

$$\varphi(x) + \lambda_0^* g_0(x) \geq \mu_q, \quad \forall\, x \in \mathbb{R}^n,$$

which is nothing else than (5.27). □

**Remark 5.3.2.** *One can check easily that if* (5.26) *holds for some* $\lambda^* \in \mathbb{R}_+^{(I)}$*, then* $A$ *is copositive on the hyperplane* $e^\perp$*. Such a notion arises naturally in the next subsection.*

### 5.3.2 Characterizing strong duality with respect to (5.20)

In this subsection, we deal with the second dual problem (5.20).

**Theorem 5.3.2.** *Let us consider problem* (5.17) *with* $C$ *being any pointed closed convex cone with nonempty interior, and* $e \in \text{int } C^*$*. The following assertions are equivalent:*

$(a)$ *$A$ is copositive on $e^\perp$;*

$(b)$ *there exists $\lambda^* \in \mathbb{R}_+^{(I)}$ such that*

$$f(x) + \sum_{i\in I}\lambda_i^* g_i(x) \geq \mu_q, \ \forall\, x \in X, \tag{5.31}$$

*or, equivalently,*

$$\mu_q = \inf_{x\in X}\Big\{f(x) + \sum_{i\in I}\lambda_i^* g_i(x)\Big\}.$$

$(c)$ *$f$ is convex on $X$.*

$(d)$ *$f$ is convex on $K$.*

*Hence, under any of the conditions $(a)--(d)$, every local solution to problem* (5.17) *is global.*

*Proof.* $(b) \Rightarrow (a)$: It is straightforward.
$(a) \Rightarrow (b)$: Take any $\bar{x} \in \operatorname*{argmin}_K f$. By writing $X = \bar{x} + e^{\perp}$, one obtains for all $x \in X$,

$$
\begin{aligned}
x^\top Ax &= (x-\bar{x})^\top A(x-\bar{x}) + 2\bar{x}^\top Ax - \bar{x}^\top A\bar{x} && (5.32)\\
&\geq 2\bar{x}^\top Ax - \bar{x}^\top A\bar{x}, && (5.33)
\end{aligned}
$$

where the inequality was obtained by copositivity of $A$ on $e^{\perp}$. Thus,

$$\frac{1}{2}x^\top Ax \geq \bar{x}^\top Ax - \frac{1}{2}\bar{x}^\top A\bar{x},\ \forall\, x \in X. \tag{5.34}$$

Let us consider the convex problem:

$$\mu' \doteq \min\Big\{\bar{x}^\top Ax :\ g_i(x) \leq 0,\ i \in I,\ x \in X\Big\}. \tag{5.35}$$

It is easy to check that $\mu' = 2\mu_q$. Indeed, obviously $\mu' \leq \bar{x}^\top A\bar{x} = 2\mu_q$. On the other hand, by the first order optimality condition,

$$(A\bar{x})^\top (x - \bar{x}) \geq 0,\ \forall\, x \in K,$$

that is, $\mu' \geq 2\mu$, proving the claim.
We now check that strong duality holds for (5.35) with respect to the dual

$$\sup_{\lambda \in \mathbb{R}^{(I)}_+} \inf_{x \in X} \Big\{\bar{x}^\top Ax + \sum_{i \in I} \lambda_i g_i(x)\Big\}. \tag{5.36}$$

It will be a consequence, as before, of Theorem 5 of [25], with objective function given by $\bar{x}^\top Ax$. In fact, such a theorem is applicable since the assumptions are verified, see also Remark 5.3.1. Hence, there exists $\lambda^* = (\lambda_i^*) \in \mathbb{R}^{(I)}_+$ such that

$$\bar{x}^\top Ax + \sum_{i \in J} \lambda_i^* g_i(x) \geq 2\mu_q,\ \forall\, x \in X.$$

This along with inequality (5.34) yield that

$$f(x) + \sum_{i \in J} \lambda_i^* g_i(x) \geq (A\bar{x})^\top x + \sum_{i \in J} \lambda_i^* g_i(x) - \mu_q \geq 2\mu_q - \mu_q = \mu_q,\ \forall\, x \in X.$$

$(a) \Leftrightarrow (c)$: First observe that $A$ is copositive on $e^{\perp}$ if, and only if $(x-y)^\top A(x-y) \geq 0$ for all $x, y \in X$. Given $t \in\, ]0,1[$, and $x, y \in X$, on combining the two identities:

$$f(x) = f(y) + \nabla f(y)^\top (x-y) + \frac{1}{2}(x-y)^\top A(x-y);$$

$$f(y + t(x-y)) = f(y) + t\nabla f(y)^\top (x-y) + \frac{t^2}{2}(x-y)^\top A(x-y),$$

one obtains

$$f(y + t(x-y)) = f(y) + t(f(x) - f(y)) - \frac{t}{2}(1-t)(x-y)^\top A(x-y).$$

From which the desired inequality follows.
$(c) \Leftrightarrow (d)$: It follows from the same arguments as in $(a) \Leftrightarrow (c)$.

□

**Remark 5.3.3.** *In case $C = \mathbb{R}^n_+$ and $e = \mathbf{1}$, $(d)$ is equivalent to ([14, Lemma 6])*

$$P_{\bar{x}}^\top A P_{\bar{x}} \succcurlyeq 0 \ \text{ for any } \bar{x} \in X, \text{ with } P_{\bar{x}} \doteq Id - \bar{x}\mathbf{1}^\top, \tag{5.37}$$

*where $Id$ is the identity matrix of order $n$.*

### 5.3.3 Characterizing strong duality with respect to (5.19)

We now analize the strong duality property in connection to (5.19). It will be suitable for the portfolio optimization problem, since copositivity of $A$ on $C$ arises naturally in such a problem. Denote, given $a \in \mathbb{R}$,

$$K(a) \doteq \{x \in C : \ g_0(x) = a\}.$$

The following proposition, whose proof is straightforward, collects some basic facts on the perturbed problem

$$\min_{x \in K(a)} f(x).$$

**Proposition 5.3.4.** *Let $A$ be a real symmetric matrix; $C$ be a pointed closed convex cone with nonempty interior, and $e \in \operatorname{int} C^*$. The following assertions hold.*

(a) *$K(a) \neq \emptyset$ if and only if $a \geq -1$; $K(-1) = \{0\}$.*

(b) *Let $a > -1$. Then, $x \in K(a)$ if and only if $\frac{1}{1+a}x \in K$. Hence*

$$\min_{x \in K(a)} f(x) = \mu_q(1+a)^2, \quad \forall\ a \geq -1.$$

(c) *Let $\mu_q > 0$. Then, $f(x) > \mu_q$ for all $x \in K(a)$ and all $a > 0$.*

Now, denote the objective function of the dual problem (5.19) by

$$\theta(\lambda) \doteq \inf_{x \in C} L(\lambda, x) \doteq f(x) + \lambda g_0(x).$$

We now describe the main properties shared by this problem, which appear here for the first time about non-polyhedral cones $C$. In particular, it reveals a hidden convexity of the general standard quadratic optimization problem.

Notice that for problem (5.1), $(c1)$ below was obtained in [14, Theorem 4] by using the Frank-Wolfe theorem, which is not applicable here (see also [13]).

**Theorem 5.3.3.** *Let $A$, $C$ and $e$ be as in the preceding proposition. Then,*

(a) *the optimal value function is given by*

$$\psi(a) = \begin{cases} \mu_q(1+a)^2 & \text{if } \ a \geq -1; \\ +\infty & \text{if } \ a < -1. \end{cases}$$

*So it is strictly convex if* $\mu_q > 0$.

(b) $F(C)$ *and* $F(C) + \mathbb{R}_+(0,1)$ *are closed, so* $\text{epi}\ \psi = F(C) + \mathbb{R}_+(0,1)$.

(c) *Let* $\mu_q > 0$. *Then,*

(c1) *The objective function* $\theta$ *is given by*

$$\theta(\lambda) = \begin{cases} -\dfrac{\lambda^2}{4\mu_q} - \lambda & \text{if } \ \lambda < 0; \\ -\lambda & \text{if } \ \lambda \geq 0. \end{cases}$$

*Hence* $\mathcal{S}_D = \{-2\mu_q\}$.

(c2) *One obtains*

$$\text{cone}(F(C) + \mathbb{R}_+(0,1) - \mu_q(0,1)) = \Big\{(u,v) \in \mathbb{R}^2 : \ 2\mu_q u \leq v\Big\}.$$

*Hence, strong duality holds for* (5.17) *and* (5.19).

(d) *Let* $\mu_q = 0$. *Then,*

(d1) *The objective function* $\theta$ *is given by*

$$\theta(\lambda) = \begin{cases} -\infty & \text{if } \ \lambda < 0; \\ -\lambda & \text{if } \ \lambda \geq 0. \end{cases}$$

*Hence* $\mathcal{S}_D = \{0\}$.

(d2) *One obtains*

$$\text{cone}(F(C) + \mathbb{R}_+(0,1)) = \Big\{(u,v) \in \mathbb{R}^2 : \ v \geq 0\Big\}.$$

*Hence, strong duality holds for* (5.17) *with respect to* (5.19).

*Proof.* $(a)$ is a consequence of the previous proposition.
$(b)$:By virtue of $(b)$ of Proposition 5.2.1, we need only to check the closedness of $F(C) + \mathbb{R}_+(0,1)$. The same argument also shows that $F(C)$ is closed. Let $(a,r) \in \overline{F(C) + \mathbb{R}_+(0,1)}$. Then, there exist sequences $x_k \in C$, $q_k \geq 0$ satisfying $f(x_k) + q_k \to r$ and $g_0(x_k) \to a$. From the second relation, we deduce that $\|x_k\|$ is bounded. Thus, up to a subsequence, $x_k \to \bar{x} \in C$, implying that $q_k = f(x_k) + q_k - f(x_k) \to r - f(\bar{x})$. Setting $q \doteq r - f(\bar{x})$, we get $q \geq 0$, and so $(r,a) = (g_0(\bar{x}), f(\bar{x}) + q) \in$

$F(C) + \mathbb{R}_+(0,1)$.
$(c1)$: From the inequality $(x^\top Ax + \lambda e^\top x)^2 \geq 0$, one obtains

$$L(\lambda, x) \geq -\frac{\lambda^2 (e^\top x)^2}{2x^\top Ax} - \lambda, \ \forall\, x \in C,\ x \neq 0.$$

Since $K$ is a base for $C$, we conclude

$$\frac{(e^\top x)^2}{2x^\top Ax} \leq \frac{1}{4\mu_q}, \ \forall\, x \in C,\ x \neq 0.$$

Hence

$$L(\lambda, x) \geq -\frac{\lambda^2}{4\mu_q} - \lambda, \ \forall\, x \in C. \tag{5.38}$$

In case $\lambda \geq 0$, it is easy to see that

$$\theta(\lambda) = \min_{x \in C} L(\lambda, x) = L(\lambda, 0) = -\lambda.$$

If $\lambda < 0$ and $\bar{x} \in \operatorname{argmin}_K f$, then by taking $x_0 = -\dfrac{\lambda}{2\mu_q}\bar{x} \in C$, we get

$$L(\lambda, x_0) = -\frac{\lambda^2}{4\mu_q} - \lambda.$$

Thus, from (5.38),

$$\theta(\lambda) = -\frac{\lambda^2}{4\mu_q} - \lambda.$$

$(c2)$: The first part follows from $(a)$, and strong duality for (5.17) with respect to (5.19) is a consequence of the fact that

$$\overline{\text{cone}}\,(\mathcal{E}_\mu) \cap -(\{0\} \times \mathbb{R}_{++}) = \overline{\text{cone}}(F(C) + \mathbb{R}_+(0,1) - \mu_q(0,1)) \cap -(\{0\} \times \mathbb{R}_{++}) = \emptyset.$$

$(d1)$: We consider only the case $\lambda < 0$ (if $\lambda \geq 0$ is exactly as in $(c1)$), and check that

$$\inf\Big\{L(\lambda, x) :\ x \in C,\ x^\top Ax = 0\Big\} = -\infty.$$

Indeed, since there exists $x_0 \in C$, $x_0 \neq 0$, such that $x_0^\top Ax_0 = 0$ by assumption, we obtain

$$\begin{aligned} \inf\Big\{L(\lambda, x) :\ x \in C,\ x^\top Ax = 0\Big\} &= \inf\Big\{\lambda(e^\top x - 1) :\ x \in C,\ x^\top Ax = 0\Big\} \\ &\leq \inf_{t>0} \lambda(te^\top x_0 - 1) = -\infty. \end{aligned}$$

Thus, $\theta(\lambda) = -\infty$ in case $\lambda < 0$.
$(d2)$: The proof is as $(c2)$. $\square$

From the previous theorem, one characterizes the copositivity of $A$ on $C$ by means of the convexity of $F(C)+\mathbb{R}_+(0,1)$. This result is new and corresponds to the nice challenge of proving convexity of joint-range for a pair of quadratic functions.

The equivalence between $(a)$ and $(c)$ is a consequence of $(a)$ and $(b)$ of Theorem 5.3.3.

**Theorem 5.3.4.** *Let us consider problem* (5.17) $C$, $e$ *as above. The following assertions are equivalent:*

$(a)$ *$A$ is copositive on $C$;*

$(b)$ *there exists $\lambda_0^* \in \mathbb{R}$ such that*

$$f(x)+\lambda_0^* g_0(x) \geq \mu_q, \;\; \forall\, x \in C, \tag{5.39}$$

*or, equivalently,*

$$\mu_q = \inf_{x\in C} \Big\{ f(x)+\lambda_0^* g_0(x) \Big\}.$$

$(c)$ *$F(C)+\mathbb{R}_+(1,0)$ is convex.*

## 5.4 Local vs global optimality

Very recently, second-order necessary and sufficient conditions for local (resp. global) optimality for a quadratic optimization problem on a polyhedron were established in Theorem 1.2 (resp. Theorem 2.3) of [15]. In this section, due to the special structure of the circular standard quadratic programming problem (5.17), we derive second-order sufficient and/or necessary conditions for local or global optimality. We refer to [11] for a method locating some particular local minima.

Set, as before, $f(x) \doteq \dfrac{1}{2}x^\top Ax$.
Due to the assumptions on $C$ and $C^*$, we can write

$$C^* \setminus \{0\} = \mathrm{co}(\mathrm{extrd}\; C^*),$$

where extrd $C^*$ stands for the extremal directions of $C^*$. Recall that $d \in \mathrm{extrd}\; C^*$ if and only if $d \in C^* \setminus \{0\}$ and for all $d_1, d_2 \in C^*$ such that $d = d_1 + d_2$, we have $d_1, d_2 \in \mathbb{R}_+ d$. Thus, for every $\lambda \in C^* \setminus \{0\}$, one has

$$\lambda = \sum_{i=1}^{k} \lambda_i d_i, \;\;\; \lambda_i > 0, \; d_i \in \mathrm{extrd}\; C^*, \; i = 1,2,\ldots,k. \tag{5.40}$$

Firstly, we establish first and second-order necessary conditions for local optimality some of the main consequences of local optimality to problem (5.17).

**Theorem 5.4.1.** *Let $A$ be any real symmetric matrix, and $C, e$ be as before. Let $\bar{x}$ be any local solution to problem* (5.17) *with $\bar{\mu} \doteq f(\bar{x})$. Then*

$(a)$ *$\lambda \doteq A\bar{x} - 2\bar{\mu}e \in \mathrm{bd}\; C^*$ and so also $A\bar{x} - 2\mu_q e \in C^*$.*

(*b*) *If* $\lambda = 0$ *in* (*a*), *then* $A$ *is copositive on* $T(K;\bar{x})$ *and so* $\bar{x}$ *is a global solution to* (5.17).

(*c*) *If* $\lambda \neq 0$, *then* $\bar{x} \in \mathrm{bd}\ C$.

*Proof.* (*a*): Let $\bar{x} \neq 0$ be a local solution to problem (5.17), i. e., $\bar{x} \in \operatorname*{argmin}_{K\cap U_0} f$ for some open neighborhood of $\bar{x}$, $U_0$. Clearly $\bar{\mu} \geq \mu_q$. By the first order optimality condition, $\nabla f(\bar{x})^\top v \geq 0$ for all $v \in T(K \cap U_0;\bar{x}) = T(K;\bar{x})$. In other words,

$$\begin{aligned}\nabla f(\bar{x}) \in [T(X\cap C;\bar{x})]^* &= [T(X;\bar{x}) \cap T(C;\bar{x})]^* = \overline{[T(X;\bar{x})]^* + [T(C;\bar{x})]^*}\\ &= [T(X;\bar{x})]^* + [T(C;\bar{x})]^*\\ &= \mathbb{R}e + (\overline{C + \mathbb{R}\bar{x}})^* = \mathbb{R}e + (\bar{x}^\perp \cap C^*),\end{aligned}$$

where the first equality follows from Table 4.3 in [1] (since $0 \in \mathrm{int}(X - C)$) and the second one is a consequence of Corollary 16.4.2 in [79] (since $\mathrm{ri}(e^\perp) \cap \mathrm{ri}(C + \mathbb{R}\bar{x}) \neq \emptyset$). Thus there exists $\lambda \in \mathbb{R}$ satisfying $A\bar{x} - \lambda e \in \bar{x}^\perp \cap C^*$. It follows that $\lambda = 2\bar{\mu}$ and so $A\bar{x} - 2\bar{\mu}e \in C^*$. This implies that

$$A\bar{x} - 2\mu_q e = A\bar{x} - 2\bar{\mu} + 2(\bar{\mu} - \mu_q)e \in C^* + C^* = C^*.$$

(*b*): Let $v \in T(K;\bar{x})$. Then, there exist $t_k > 0$, $x_k \in K$, $x_k \to \bar{x}$ such that $t_k(x_k - \bar{x}) \to v$. Thus for all $k$ sufficiently large,

$$\begin{aligned}0 \leq f(x_k) - 2\bar{\mu}e^\top x - f(\bar{x}) + 2\bar{\mu}e^\top \bar{x} &= (\nabla f(\bar{x}) - 2\bar{\mu}e)^\top (x_k - \bar{x}) + \frac{1}{2}(x_k - \bar{x})^\top A(x_k - \bar{x})\\ &= \frac{1}{2}(x_k - \bar{x})^\top A(x_k - \bar{x}).\end{aligned}$$

Hence $v^\top Av \geq 0$, proving the copositivity on $T(K;\bar{x})$. Therefore, given any $x \in K$, the equality

$$\begin{aligned}f(x) - f(\bar{x}) &= f(x) - 2\mu_q g_0(x) - f(\bar{x}) + 2\mu_q g_0(\bar{x})\\ &= (A\bar{x} - 2\mu_q e)^\top (x - \bar{x}) + \frac{1}{2}(x - \bar{x})^\top A(x - \bar{x}),\end{aligned}$$

yields the desired result, since $K - \bar{x} \subseteq T(K;\bar{x})$.
(*c*): It is obvious. □

Next, we derive a sufficient condition for global optimality.

**Proposition 5.4.1.** *Let* $C$, $e$ *be as above. If* $\bar{x}$ *feasible for* (5.17), $A$ *is copositive on* $C - \bar{x}$ *or equivalently on* $T(C;\bar{x})$, *and* $A\bar{x} - 2\bar{\mu}e \in C^*$ *holds with* $\mu = f(\bar{x})$, *then* $\bar{x} \in \operatorname*{argmin}_K f$.

*Proof.* For all $x \in K$, one obtains

$$\begin{aligned}f(x) - f(\bar{x}) &= f(x) - 2\bar{\mu}g_0(x) - f(\bar{x}) + 2\bar{\mu}g_0(\bar{x})\\ &= (A\bar{x} - 2\bar{\mu}e)^\top (x - \bar{x}) + \frac{1}{2}(x - \bar{x})^\top A(x - \bar{x}),\end{aligned}$$

from which the result follows. □

We now deal with the standard quadratic optimization problem, that is, when $C = \mathbb{R}^n_+$ and $e = \mathbf{1}$. Denote, given $\bar{x} \in K$,

$$I(\bar{x}) \doteq \{i : \ \bar{x}_i = 0\}, \ I_+ \doteq \{i \in I(\bar{x}) : \ \lambda_i > 0\}.$$

$$Z(\bar{x}) \doteq \Big\{v \in \mathbb{R}^n : \ v_i = 0, \ i \in I_+; \ v_i \geq 0, \ i \in I(\bar{x}) \setminus I_+; \ \sum_{i \in I(\bar{x}) \setminus I_+} v_i + \sum_{i \notin I(\bar{x})} v_i = 0\Big\}.$$

It is not difficult to check that

$$T(K; \bar{x}) = \Big\{v \in \mathbb{R}^n : \ v_i \geq 0, \ i \in I(\bar{x}); \ \sum_{i=1}^{n} v_i = 0\Big\}.$$

Hence, if $I_+ = \emptyset$, then $Z(\bar{x}) = T(K; \bar{x})$. The following result is a consequence of Theorem 4.4.3 in [5] and the above remark.

**Theorem 5.4.2.** *Let $A$, $C, e$ be as just mentioned. Let $\bar{x}$ be any local solution to problem* (5.1) *with $\bar{\mu} \doteq f(\bar{x})$. Then*

(*a*) *$\bar{x}$ is a KKT point:*

$$A\bar{x} - 2\bar{\mu}\mathbf{1} - \lambda = 0, \ \lambda \geq 0, \ \lambda_i \bar{x}_i = 0, \ i = 1, \ldots, n,$$

*and $A$ is copositive on $Z(\bar{x})$.*

(*b*) *If additionally $I_+ = \emptyset$, then $\bar{x}$ is a global solution.*

**Example 5.4.2.** *Take the matrix*

$$A = \begin{pmatrix} 1 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix}.$$

By computing, one obtains ($x = (x_1, x_2, x_3)$),

$$x^\top Ax = x_1^2 + 2x_1x_2 + 2x_2x_3, \quad f(x_1, x_2, 1 - x_1 - x_2) = \frac{1}{2}x^\top Ax = \frac{1}{2}(x_1^2 - 2x_2^2 + 2x_2).$$

Then: $\mu_q = 0$, that is, $A$ is copositive on $\mathbb{R}^n_+$; $A$ is not copositive on $\mathbf{1}^\perp$ since $f(-1, 1, 0) = -\dfrac{1}{2}$. Moreover, the associated StQOP has two solutions, namely $\mathrm{argmin}_K f = \{\bar{x}^1 \doteq (0, 0, 1), \bar{x}^2 \doteq (0, 1, 0)\}$. One can also check that no local-nonglobal solution exists. Moreover, $\bar{\lambda}^1 = A\bar{x}^1 = (0, 1, 0)$, $\bar{\lambda}^2 = A\bar{x}^2 = (1, 0, 1)$, so ($b$) of Theorem 5.4.2 is not satisfied at $\bar{x}^1$ or $\bar{x}^2$. Notice that

$$Z(\bar{x}^1) = \{t(1, 0, -1) : \ t \geq 0\}, \ T(K; \bar{x}^1) = \{t(1, 0, -1) + s(0, 1, -1) : \ t \geq 0, \ s \geq 0\},$$

and

$$Z(\bar{x}^2) = \{(0, 0, 0)\}, \ T(K; \bar{x}^2) = \{t(1, -1, 0) + s(0, -1, 1) : \ t \geq 0, \ s \geq 0\}.$$

## 5.5 The bidimensional case

Just for illustration, let us consider $n = 2$, $\mathbf{1} = (1, 1)$ and $A = A^\top \in \mathbb{R}^{2\times 2}$ with

$$A = \begin{pmatrix} a & b \\ b & c \end{pmatrix}.$$

The following proposition is easily obtained, see also [44, 73], and from it, one infers that there is no relationship between copositivity on $\mathbf{1}^\perp$ and on $\mathbb{R}^2_+$.

**Proposition 5.5.1.** *Let $A$ be as above. Then*

(*a*) *$A$ is positive semidefinite if and only if $a \geq 0,\ ac \geq b^2$.*

(*b*) *$A$ is copositive on $\mathbf{1}^\perp$ if and only if $a - 2b + c \geq 0$.*

(*c*) *$A$ is copositive on $\mathbb{R}^2_+$ if and only if $a \geq 0$, $c \geq 0$, $b \geq -\sqrt{ac}$.*

(*d*) *Let $A$ be any real symmetric matrix with real entries. Then, every local solution to*

$$\min\{f(x_1, x_2) \doteq \frac{1}{2}x^\top Ax :\ x_1 + x_2 = 1,\ x_1 \geq 0,\ x_2 \geq 0\}, \tag{5.41}$$

*is global if and only if any of the following assertions holds:*

(*d*1) $a + c \geq 2b$;

(*d*2) $a + c < 2b$, $a \geq b$;

(*d*3) $a + c < 2b$, $b \leq c$;

(*d*4) $a + c < 2b$, $a = c$.

*Proof.* We have $x^\top Ax = ax_1^2 + 2bx_1x_2 + cx_2^2$.
($a$): This is a consequence of the fact that $A$ is positive semidefinite if and only if all the principal minors are nonnegative.
($b$): Let $x \in \mathbf{1}^\perp$. Then, $x^\top Ax = (a - 2b + c)x_1^2$, and so the result follows.
($c$): The condition $x^\top Ax \geq 0$ for all $x \in \mathbb{R}^2_+$, implies $a \geq 0$ and $c \geq 0$. Thus

$$x^\top Ax = (\sqrt{a}x_1 - \sqrt{c}x_2)^2 + 2(\sqrt{ac} + b)x_1x_2, \tag{5.42}$$

which is nonnegative if $b \geq 0$. In case $b < 0$ and $a > 0,\ c > 0$, we require $b \geq -\sqrt{ac}$, since otherwise, by choosing $x_1 = \dfrac{\sqrt{c}}{\sqrt{a}}x_2$, $x_2 > 0$, in (5.42), one gets $x^\top Ax < 0$.
($d$): It follows from the equivalent formulation to (5.41)

$$\min\left\{f(x_1, 1 - x_1) = \frac{1}{2}x_1^2(a - 2b + c) + x_1(b - c) + \frac{c}{2} :\ 0 \leq x_1 \leq 1\right\}, \tag{5.43}$$

and by noticing that the function $\varphi(x_1) \doteq f(x_1, 1 - x_1)$ satisfies

$$\varphi'(0) = b - c,\ \varphi'(1) = a - b,\ \varphi(0) = \frac{c}{2},\ \varphi(1) = \frac{a}{2}.$$

□

It worth noticing that, from the previous result, there are bi-dimensional quadratic problems with $A$ being non-copositive on $\mathbb{R}^2_+$, having local solutions without being global, for instance $a = -1$, $b = 1$, $c = 1$.

# Chapter 6

# Conclusions

## 6.1 Spanish version

En esta tesis se propuso establecer condiciones necesarias y suficientes que garantizan la validez de la propiedad de dualidad fuerte para un problema no convexo con una restricción, junto con estudiar conexiones con otras propieades afines. Obtuvimos condiciones del tipo geométricas y topológicas de la envoltura cónica del conjunto imagen $\{(f(x), g(x)) : \ x \in C\}$ asociada al problema original. A continuación, presentamos las principales conclusiones en este trabajo.

En el Capitulo 3 se estudió el problema cuadrático no convexo con una restricción cuadrática del tipo desigualdad y junto a varias restricciones lineales del tipo igualdad, es decir, el conjunto $C$ viene dado por un subespacio afín y $f(x) = \dfrac{1}{2}x^\top Ax + a^\top x + \alpha$, $g(x) = \dfrac{1}{2}x^\top Bx + b^\top x + \beta$, $C = H^{-1}(d)$. El resultado principal dado por el Teorema 3.3.4, caracteriza la validez de la dualidad fuerte, que depende de las matrices Hessianas de las funciones $f$ y $g$, sin hipótesis de regularidad como la condición de Slater. En cambio, bajo la condición de Slater el Teorema 3.3.2 permite derivar condiciones de optimalidad necesarias y suficientes sin supuesto alguno de convexidad. Finalmente, obtuvimos una caracterización en el Teorema 3.4.1 de la no vacuidad del conjunto solución para el problema de minimizar una forma cuadrática no necesariamente convexa, sobre un conjunto convexo, cerrado y asintóticamente lineal.

En el Capitulo 4, estudiamos el caso general de un restricción de igualdad estableciendo una completa caracterización, en virtud del Teorema 4.3.2. De acuerdo al Colorario 4.3.8 la convexidad de la envoltura cónica del conjunto imagen $\{(f(x), g(x)) : \ x \in C\}$ es condición necesaria y suficiente para la validez de la propiedad de dualidad fuerte, bajo una condición del tipo Slater. Nuestros resultados permiten asegurar que bajo la propiedad de dualidad fuerte, toda solución del problema original (4.1) es un punto KKT, sin imponer hipótesis de convexidad.

En el Capitulo 5, el enfoque estuvo en el problema cuadrático estándar no-poliédrico. En particular, para cada cono no vacío, convexo y puntiagudo, con interior no vacío $C \subseteq \mathbb{R}^n$, $f(x) = \frac{1}{2}x^\top Ax$ y $g(x) = e^\top x - 1$, establecimos un nuevo resultado que caracteriza la validez de la propiedad de dualidad fuerte para el problema (5.17) es equivalente a la copositividad en el cono, cerrado, convexo y puntiagudo $C$ de interior topológico no vacío. A raíz de lo cual establecemos un nuevo resultado que caracteriza que la copositividad de una matriz con respecto a $C$, es equivalente a la convexidad del conjunto imagen $\{(f(x), g(x)) : \ x \in C\}$.

## 6.2 English version

In this thesis it was proposed to establish necessary and sufficient conditions to guarantee the validity of the property of strong duality for a non-convex problem with a single restriction, along with the study of connections with other related properties. We obtained conditions of geometrical and topological type on the conic hull of $\{(f(x), g(x)) : x \in C\}$ associated with the original problem. Next, we present the main conclusions of this work.

In Chapter 3 the non-convex quadratic problem was studied with a quadratic inequality constraint type and with various linear type restrictions, namely the set $C$ is given by an affine sub-space, and, $f(x) = \frac{1}{2}x^\top Ax + a^\top x + \alpha$, $g(x) = \frac{1}{2}x^\top Bx + b^\top x + \beta$, $C = H^{-1}(d)$. The main result given by Theorem 3.3.4, characterizes the validity of the strong duality in terms of the Hessian matrices of the functions $f$ and $g$, without any regularity condition as the Slater condition. On the other hand, under the condition of Slater, Theorem 3.3.2 allows to derive necessary and sufficient optimality conditions without any assumption of convexity. Finally, we obtained a characterization in Theorem 3.4.1 of the non-emptiness of the solution set for the problem of minimizing a quadratic form not necessarily convex, on a convex, closed and asymptotically linear set.

In Chapter 4, we study the general case of an equality constraint establishing a complete characterization, by virtue of Theorem 4.3.2. According to Corollary 4.3.8 the convexity of the conical envelope of the image set $\{(f(x), g(x)) : x \in C\}$ is a necessary and sufficient condition for the validity of the property of strong duality, under a Slater type condition. Our results allow us assuring that under the property of strong duality, every solution of the original problem (4.1) is a KKT point, without imposing convexity hypothesis.

Chapter 5 was focused on the standard non-polyhedral quadratic problem. In particular, for any non-empty subset $C \subset \mathbb{R}^n$ being a closed, convex and pointed cone having non-empty interior, $f(x) = \frac{1}{2}x^\top Ax$ and $g(x) = e^\top x - 1$, we establish a new result that characterizes the copositivity of a matrix with respect to $C$, it is equivalent to the convexity of the image set $\{(f(x), g(x)) : x \in C\}$.

# Bibliography


[1] Aubin, J. P.; Frankowska, H., Set-Valued Analysis, Birkhäuser, Boston, Basel, Berlin, 1990.

[2] Achtziger, W.; Kanzow, Ch., Mathematical programs with vanishing constraints: optimality conditions and constraint qualifications, *Math. Program. A*, **114** (2008), 69-99.

[3] Auslender, A., Existence of optimal solutions and duality results under weak conditions, *Math. Program., Ser. A*, **88** (2000), 45–59.

[4] Auslender, A; Teboulle, M., Asymptotic cones and functions in optimization and variational inequalities, Springer Monograph. Math., Springer-Verlag, New york, 2003.

[5] Bazaraa, M.S.; Sherali, H. D.; Shetty, C. M., Nonlinear programming, theory and algorithms, Wiley-Interscience, A John Wiley & Sons, Inc., Hoboken, New Jersey, 2006.

[6] Bank, B.; Guddat, J.; Klatte, D.; Kummer, B. and Tammer, K., Non-linear Parametric Optimization, Springer Fachmedien Wiesbaden, Berlin, 1982.

[7] Belousov, E. G.; Klatte, D., A Frank-Wolfe type theorem for convex polynomial programs, *Comput. Optim. Appl.*, **22** (2002), 37–48.

[8] Ben-Tal, A.; Teboulle, M., Hidden convexity in some nonconvex quadratically constrained quadratic programming, *Math. Program., Ser. A*, **72** (1996), 51–63.

[9] Ben-Israel A., Linear equations and inequalities on finite dimensional, real or complex, vector spaces: A unified theory, J. Math Anal. Appl. (1969), 367-389.

[10] Bertsekas, D. P.; Ozdaglar, A. E.; Tseng, P., Enhanced Fritz John conditions for convex programming, *SIAM J. Optim.*, **16** (2006), 766–797.

[11] Best, M. J.; Ding, B., Global and local quadratic minimization, *J. Global Optim.*, **10** (1997), 77–90.

[12] Blum, E.; Oettli, W., Direct proof of the existence theorem for quadratic programming, *Operation Research*, **20** (1972), 165–167.

[13] Bomze, I. M., On standard quadratic optimization problems, *J. Global Optim.*, **13** (1998), 369–387.

[14] Bomze, I. M.; Locatelli, M.; Tardella, F., New and old bounds for standard quadratic optimization: dominance, equivalence and incomparability, *Math. Program., Ser. A* **115** 31–64 (2008).

[15] Bomze, I. M., Copositivity for second-order optimality conditions in general smooth optimization problems, *Optimizaion*, **65** (2016), 779–795.

[16] Bot, R. I.; Wanka, G., An alternative formulation for a new closed cone constraint qualification, *Nonlinear Analysis*, **64** (2006), 1367–1381.

[17] Bot, R. I.; Grad, S.-M.; Wanka, G., New regularity conditions for strong and total Fenchel-Lagrange duality in infinite dimensional spaces, *Nonlinear Analysis*, **69** (2008), 323–336.

[18] Bot, R. I.; Csetnek, E. R.; Moldovan, A., Revisiting some duality theorems via the quasirelative interior in convex optimization, *J. Optim. Theory Appl.*, **139** (2008), 67–84.

[19] Bot, R. I.; Csetnek, E. R.; Wanka, G., Regularity conditions via quasi-relative interior in convex programming, *SIAM J. Optim.*, **19** (2008), 217–233.

[20] Boyd, S.; Vandenberghe, L., Semidefinite Programming, *SIAM Rev.*, **38** (1996), 49–95.

[21] Boyd, S.; Vandenberghe, L., Convex Optimization, Cambridge University Press, Cambrigde, 2004.

[22] Jia-Ming Cao, Necessary and sufficient condition for local minima of a class of nonconvex quadratic programs, *Math. Program., Ser. A.*, **69** (1995), 403–411.

[23] Derinkuyu, K.; Pinar, M. C., On the S-procedure and some variants, *Math. Meth. Oper. Res.* **64** (2006), 55–77.

[24] Dines, L. L., On the mapping of quadratic forms, *Bull. Am. Math. Soc.*, **47** (1941), 494–498.

[25] Dinh, N,; Goberna, M. A.; Lopez, M. A.; Son, T. Q., New Farkas-type constraint qualifications in convex infinite programming, *ESAIM: Control, Optim. and Calculus of Variations*, **13** (2007), 580–597.

[26] Eaves, B. C., On quadratic programming, *Management Science: Theory Series*, **17** (1971), 698–711.

[27] Ernst, E.; Volle, M., Zero duality gap for convex programs: a generalization of the Clark-Duffin theorem, *J. Optim. Theory Appl.*, **158** (2013), 668–686.

[28] Ekeland, I.; Temam, R., Analyse convexe et problemes variationnels, Dunod-Gauthier-Villars, Paris, 1974.

[29] Fenchel, W., On conjugate convex functions, *Canadian Journal of Mathematics*, **1** (1949), 73–77.

[30] Finsler, P., Über das Vorkommen definiter und semi-definiter Formen in scharen quadratische Formen, *Commentarii Matematici Helvetici*, **9** (1937), 188–192.

[31] Flores-Bazán, F.; Flores-Bazán, Fernando; Vera, C., A complete characterization of strong duality in nonconvex optimization with a single constraint, *J. Global Optim.*, **53** (2012), 185–201.

[32] Flores-Bazán, F.; Hadjisavvas, N; Vera, C., An optimal alternative theorem and applications to mathematical programming, *J. Global Optim.*, **37** (2007), 229–243.

[33] Flores-Bazán, F.; Flores-Bazán, Fernando; Vera, C., Gordan-type alternative theorems and vector optimization revisited, in *Recent Developments in Vector Optimization*, pp 29–59 (2012).

[34] Flores-Bazán, F.; Jourani, A.; Mastroeni, G. On the convexity of the value function for a class of nonconvex variational problems: existence and optimality conditions, *SIAM J. Control and Optim.*, **56** (2014), 3673–3693.

[35] Flores-Bazán, F.; Mastroeni, G., Strong duality in cone constrained nonconvex optimization, *SIAM J. Optim.* , **23** (2013), 153–169.

[36] Flores-Bazán, F.; Opazo, F., Characterizing the convexity of joint-range for a pair of inhomogeneous quadratic functions and strong duality. *Minimax Theory and its Applications,* **2** (2016), 257–290.

[37] Frank, M.; Wolfe, P., An algorithm for quadratic programming, *Naval Res. Log. Quart.*, **3** (1956), 95–110.

[38] Frenk, J. B. G.; Kassay, G., On classes of generalized convex functions. Gordan-Farkas type theorems, and Lagrangian duality *J. Optim. Theory and Appl.*, **102** (1999), 315–343.

[39] Frenk J. B. G. and Kassay G., Introduction to Convex and Quasiconvex Analysis, (17 2004 9,). ERIM Report Series Reference No. ERS-2004-075-LIS.

[40] Fritz J., *Extremum problems with inequalities as subsidiary conditions*, in: Studies and Essays, Presented to R. Courant on his 60th Birthday, Interscience, New York (1948) 187–204.

[41] Gander, W., Least Squares with a Quadratic Constraint, *Numer. Math.*, **36** (1981), 291–307.

[42] Gay, D. M., Computing optimal locally constrained steps, *SIAM J. Sci. Statist. Comput.*, **2** (1981), 186–197.

[43] Gibbons, L. E.; Hearn, D. W.; Pardalos, P. M.; Ramana, M. V., Continuous characterizations of the maximum clique problem, *Math. Oper. Res.*, **22** (1997), 754–768.

[44] Hadeler, K. P., On copositive matrices, *Linear Alg. Appl.*, **49** (1983), 79–89.

[45] Hamburger, C., Two extensions to Finsler's recurring theorem, *Appl. Math. Optim.*, **40** (1999), 183–190.

[46] Horn, R. E.; Johnson, C. R., Matrix Analysis, Cambridge University Press, New york, 1985.

[47] Ibaraki, T.; Katoh, N., Resource allocations problems: algorithm approaches. MIT Press, Cambridge, 1988.

[48] Jeyakumar, V., Constraint qualifications characterizing Lagrangian duality in convex optimization, *J. Optim. Theory and Appl.*, **136** (2008), 31–41.

[49] Jeyakumar, V.; Huy, N. Q.; Li, G. Y., Necessary and sufficient conditions for S-lemma and nonconvex quadratic optimization, *Optim. Eng.*, **10** (2009), 491–503

[50] Jeyakumar, V.; Lee, G. M., Complete characterizations of stable Farkas'lemma and cone-convex programming duality, *Math. Program.* **A**, **114** (2008), 335–347.

[51] Jeyakumar, V.; Lee, G. M.; Li, G. Y., Alternative theorems for quadratic inequality systems and global quadratic optimization, *SIAM J. Optim.*, **20** (2009), 983–1001.

[52] Jeyakumar, V.; Li, G. Y., Stable zero duality gaps in convex programming: Complete dual characterizations with applications to semidefinite programs, *J. Math. Anal. and Appli.*, **360** (2009), 156–167.

[53] Jeyakumar, V.; Li, G. Y., Regularized Lagrangian duality for linearly constrained quadratic optimization and the trust-region problems, *J. Global Optim.*, **49** (2011), 1–14.

[54] Jeyakumar, V.; Wolkowicz, H., Zero duality gaps in infinite-dimensional programming, *J. Optim. Theory Appl.*, **67** (1990), 87–108.

[55] Jin, Q; Fang, A. Ch.; Xing, W., On the global of generalized trust region subproblems, *Optim.* **58** (2010), 1139–1151.

[56] Karmarkar, K., A new polynomial-time algorithm for linear programming, *Combinatorica*, **4** (1984), 373–395.

[57] Karush, K., Minima of functions of several variables with inequalities as side conditions. Master thesis. Department of Mathematics, University of Chicago, 1939.

[58] Kuhn, H. W.; Tucker, A. W., Nonlinear programming, in: J. Neyman, ed., Proceedings of the second Berkeley simposium on mathematical statistics and probability, University of California Press, Berkeley (1950) 481–492.

[59] Mangasarian, O. L., Locally unique solutions of quadratic programs, linear and nonlinear complementarity problems, *Math. Program., A*, **19** (1980), 200–212.

[60] Martínez, J. M., & Santos, S. A. A trust-region strategy for minimization on arbitrary domains, *Math. Program., B*, **68** (1995), 3–39.

[61] Markowitz, H. M., Portfolio selection, *J. Finance*, **7** (1952), 77–91.

[62] Markowitz, H. M., The general mean-variance portfolio selection problem, in Howison, S. D. et al (eds) Mathematical models in finance; **93-99**. Chapman & Hall, London (1995).

[63] Mastroeni, G., Some applications of the image space analysis to the duality theory for constrained extremum problems, *J. Global Optim.*, **46** (2010), 603–614.

[64] Matskani, E.; Sidiropoulos, N. D.; Luo, Z. Q.; Tassiulas, L., Convex approximations techniques for joint multiuser downlink beamforming and admission control, *IEEE Trans. Wireless Comm.*, **7** (2008), 2682–2693.

[65] Meng, K. W.; Yang, X. Q., Optimality conditions via exact penalty functions, *SIAM J. Optim.*, **20** (2010), 3208–3231.

[66] Moré, J. J., Generalizations of the trust region problem, *Optim. Methods, Softw.*, **2** (1993), 189–209.

[67] Motzkin, T. S.; Strauss, E. G., Maxima for graphs and a new proff of a theorem of Turán, *Can. J. Math.*, **17** (1965), 533–540.

[68] Murty, K. G.; Kabadi S. N, Some NP-complete problems in quadratic and nonlinear programming *Math. Program.*, **2** (1987), 117–129.

[69] Nesterov, Y.; Nemirowskii, A., Interior-point Polynomial Methods in Convex Programming, SIAM, (1994).

[70] Ozdaglar, A.; Tseng, P., Existence of global minima for constrained optimization, *J. Optim. Theory Appl.* **128** (2006), 523–546.

[71] Pardalos, P. M.; Vavasis, S. A., Quadratic programming with one negative eigenvalues is NP-hard, *J. Glob. Optim.* **1**, (1991) 15–22.

[72] Peng, J.; Yuan, Y;, Optimality Conditions for the Minimization of a quadratic with two quadratic constraints, *SIAM, J. Optim.,* **2** (1997), 579–594.

[73] Ping, L.; Yu, F. Y., Criteria for copositive matrices of order four, *Linear Alg. Appl.*, **194** (1993), 109–124.

[74] Pong, T. K.; Wolkowicz, H., The generalized trust region subproblem, *Computational Optimization and Applications*, **58(2)**, (2014), 273–322.

[75] Pólik, I.; Terlaky, T., A survey of the S-Lemma, *SIAM Review*, **49** (2007), 371–418.

[76] Polyak, B. T., Convexity of quadratic transformations and its use in control and optimization, *J. Optim. Theory Appl.*, **99** (1998), 553–583.

[77] Luo, Z. Q.; Zhang, S., On extensions of the Frank-Wolfe Theorems, *Computational Optimization and Applications*, **13** (1999) 87–110.

[78] Rockafellar, R. T., Convex Analysis, Princeton University Press, Princeton, 1970.

[79] ROCKAFELLAR, R. T., Conjugate Duality and Optimization, SIAM, (1974).

[80] SHOR, N., Quadratic optimization problems. *Sov. J. Comput. Syst. Sci.*, **25** (1987) 1–11.

[81] SIDIROPOULOS, N. D.; DAVIDSON, T. N.; LUO, Z. Q., Transmit beamforming for physical layer multicasting, *IEEE Trans. Signal Process*, **54** (2006), 2239–2251.

[82] SORENSEN, D. C., Newton's method with a model trust region modification, *SIAM Numer. Anal.*, **19** (1982), 409–426.

[83] STERN, R. J.; WOLKOWICZ, H., Indefinite trust region subproblems and nonsymmetric eigenvalue perturbations. SIAM Journal on Optimization, 5(2):286313, 1995.

[84] STURM, J. F.; ZHANG, S., On cones of nonnegative quadratic functions, *Math. Oper. Res.*, **28** (2003), 246–267.

[85] TSENG, P., Some convex programs without a duality gap, *Math. Program., B*, **116** (2009), 553–578.

[86] Wolkowicz, H., Saigal, R., Vandenberghe, L.: Handbook of Semidefinite Programming. Theory, Algorithms, and Applications. Kluwer (2000).

[87] XIA, Y.; WANG, S.; SHEU, R. L., S-Lemma with equality and its applications, *Math. Program. A*, **156** (2015), 513–547.

[88] YAKUBOVICH, V. A., S-procedure in nonlinear control theory, Vestnik Leningrad. Univ., **1** (1971), 62–77 (in Russian).

[89] YAKUBOVICH, V. A., S-procedure in nonlinear control theory, Vestnik Leningrad. Univ., **4** (1977), 73–93 (English translation).

[90] YAN, Z. Z.; GUO, J. H., Some equivalent results with Yakuvobich $S$-lemma, *SIAM J. Control Optim.*, **48** (2010), 4474–4480.

[91] YANG, X. Q.; MENG, Z. Q., Lagrange multipliers and calmness conditions of order $p$, *Math. Oper. Res.*, **32** (2007), pp. 95–101.

[92] YE, J. J.; ZHANG, J., Enhanced Karush-Kuhn-Tucker condition and weaker constraint qualifications, *Math. Program. B*, **139** (2013), 353–381.

[93] YE, Y., Interior Point Algorithms. Theory and Analysis, John Wiley & Sons, (1997).

[94] ZALINESCU, C., "Convex analysis in general vector spaces", World Scientific, Singapore (2002).

[95] ZHENG, X.J.; SUN, X. L.; LI, D.; XU, Y. F., On zero duality gap in nonconvex quadratic programming problems, *J. Global Optim.*, **52** (2012), 229–242.