UNIVERSIDAD DE GRANADA

Teoremas Tipo Hahn–Banach y Optimización no Lineal

Pablo Montiel López

Programa de Doctorado en Matemáticas

Editor: Universidad de Granada. Tesis Doctorales
Autor: Pablo Montiel López
ISBN: 978-84-9163-418-8
URI: http://hdl.handle.net/10481/47968

Índice general

El doctorando / *The doctoral candidate* **Pablo Montiel López** y el director de la tesis / *and the thesis supervisor* **Manuel Ruiz Galán**,

garantizamos, al firmar esta tesis doctoral, que el trabajo ha sido realizado por el doctorando bajo la dirección del director de la tesis y hasta donde nuestro conocimiento alcanza, en la realización del trabajo, se han respetado los derechos de otros autores a ser citados, cuando se han utilizado sus resultados o publicaciones.

/

guarantee, by signing this doctoral thesis, that the work has been done by the doctoral candidate under the direction of the thesis supervisor and, as far as our knowledge reaches, in the performance of the work, the rigths of the other authors to be cited (when their results or publications have been used) have been respected.

Lugar y fecha / *Place and date*: Granada a 29 de junio de 2017

Director de la Tesis / *Thesis supervisor;* Doctorando / *Doctoral candidate:*

Firma / *Signed* Firma / *Signed*

Agradecimientos

En primer lugar, querría agradecer al director de esta Memoria, profesor d. Manuel Ruiz Galán, su desinteresada ayuda en la elaboración de la misma, su guía científica y humana constante y su paciencia, especialmente estos últimos meses. En todo momento ha sabido entender mi situación familiar y profesional, animándome siempre a continuar el trabajo cuando me he sentido sobrepasado por las dificultades.

También ha sido muy generosa la ayuda y la acogida de los miembros del Grupo de Investigación FQM359 de la Junta de Andalucía, destacando la de la profesora dña. Ana Isabel Garralda, que me prestó el ánimo necesario para seguir en tiempos que ahora pueden ser recordados con una sonrisa pero que, entonces, fueron realmente difíciles.

Debo agradecer igualmente su hospitalidad al profesor Constantin Zalinescu, por haberme recibido en la Universidad Alexandru Ioan Cuza, en Iasi, Rumanía, donde realicé una provechosa estancia de investigación que fue financiada por el Centro de Magisterio La Inmaculada, al cual agradezco su generosidad y las facilidades que me ha brindado a lo largo de todos estos años.

Es especialmente emocionante poder darle también las gracias a mi mujer, Loli, y a mis hijas, Loli Jr. e Isabel, a las que he tenido que robar tiempo y atención, pero que siempre han sido el estímulo esencial para trabajar y el descanso necesario para no detenerme.

Terminar mencionando a mis padres, que aguantaron una tardía y prolongada adolescencia, y a Antoñita, sus croquetas han sido un gran consuelo durante la realización de esta tesis.

Abstract

The development of convex analysis has been in parallel to that of optimization from its beginnings, when seminal versions of the theorem of Hahn-Banach arose in the form of theorems of the alternative. And regarding this central result of convex and functional analysis, it is, at least, significant that in the almost century and a half since then, its study is still valid. And exactly the research we present in this thesis forms part of it, establishing new versions of the Hahn–Banach theorem and applying them to nonlinear programming.

The thesis consists of four main chapters, the first of which deals with collecting some concepts and results of that convex and functional worlds, which although well known, will serve to contextualize and address the problems that are treated. Special mention should be made of the Mazur–Orlicz theorem, an equivalent reformulation of the Hahn-Banach theorem that has become a powerful tool in various fields such as optimization, convex analysis, monotone operators or minimax theory.

The second chapter aims to establish from the theorem of Mazur–Orlicz a Hahn–Banach type theorem in finite dimension, sharp in a sense, as well as

to use it in obtaining results along the lines of those of Karush–Kuhn–Tucker, Fritz John and Lagrange multiplier in semi-infinite programming (a finite number of restrictions on arbitrary domains). Such result is an improvement of a reformulation of Hahn–Banach's theorem, König's maximum theorem, which is equivalent (a slight improvement), as we show at the beginning of the chapter, to Gordan's alternative theorem and a theorem of Simons on convex functions. In fact, it is proven that the class of convexity where it is valid is that of the infsup-convex (or affine weakly convexlike) functions, whose study goes back to that of the minimax inequalities. As a consequence, the mentioned results in programming not only improve the classic and even the more general ones in quasiconvex or convex in the sense of Fan frameworks, but also are optimal.

On the other hand, Chapter 3 begins with an optimal extension of the maximum theorem of König, König's supremum theorem, which, unlike that, allows us to work in a space of bounded functions. In fact, it is nothing more than another equivalent form of the Hahn–Banach theorem. We deduce from the Mazur–Orlicz theorem a generalization that is optimal once again, which provide us, because of its character not necessarily finite dimensional, with some optimal results for infinite programs with an arbitrary number of inequality-type constraints for Lagrange multiplier and Karush–Kuhn-Tucker or Fritz John conditions.

We conclude with a final chapter focused on stating the aforementioned results on nonlinear programming in the more general context: any domain and an arbitrary number of inequality or equality restrictions. The starting point now is the convex separation theorem, because it is more popular in the optimization context than the theorems of König. The results derived in infinite programming are again optimal in terms of infsup-convexity.

Introducción

Sin lugar a dudas, el teorema de Hahn–Banach constituye uno de los resultados sobre los que se fundamenta gran parte de las estructuras del análisis convexo y la optimización. Teniendo esto presente, el objetivo de la presente memoria es bastante claro: establecer mejoras de dicho resultado con el fin de obtener perfeccionamientos de teoremas clásicos en programación no lineal. Nos encontramos pues ante un primer reto, el de determinar en qué sentido es mejorable el teorema de Hahn–Banach. Pensando en la versión que podría considerarse esencial (en un espacio vectorial real siempre es posible minorar un funcional sublineal por otro lineal), no parece en principio que ello sea posible. Sin embargo, centrándonos en otra de sus muchas versiones equivalentes, encontramos un hilo del que tirar. En concreto, se trata de la siguiente, el teorema del máximo de König:

Teorema 0.1 *Si E es un espacio vectorial real, X es un subconjunto no vacío y convexo de E, $n \geq 1$, $S_1, \ldots, S_n : E \longrightarrow \mathbb{R}$ son funcionales sublineales y $L : E \longrightarrow \mathbb{R}$ es un funcional lineal de forma que*

$$x \in X \;\Rightarrow\; L(x) \leq \max_{i=1,\ldots,n} S_i(x),$$

entonces existe $\mathbf{s} \in \Delta_n$ *tal que*

$$x \in X \;\Rightarrow\; L(x) \leq \sum_{i=1}^{n} s_i S_i(x).$$

Repárese en el hecho de que este resultado garantiza la existencia, bajo ciertas hipótesis, de un elemento del símplex de probabilidad Δ_n de $\mathbb{R}^n$ que satisface cierta desigualdad. Puesto que los vectores de este símplex,

$$\Delta_n := \left\{ \mathbf{s} \in \mathbb{R}^n : \; \mathbf{s} \in \mathbb{R}^n_+, \; \sum_{i=1}^{n} s_i = 1 \right\},$$

no son más que funcionales lineales y continuos de $\mathbb{R}^n$, y más específicamente, los que minoran al funcional lineal máximo, se asegura la existencia de un tal funcional.

Antes de plantearnos la posible generalización del teorema del máximo de König, probamos que equivale –de hecho, una ligera extensión suya– a otros teoremas tipo Hahn–Banach en dimensión finita, como son el clásico teorema de la alternativa de Gordan y un resultado sobre funciones convexas debido a S. Simons. Ello nos pone sobre la pista de cómo extender el teorema del máximo de König: debemos suavizar, en la medida de lo posible, la linealidad y la sublinealidad. En primer lugar, analizamos diversas clases de convexidad generalizada, todas ellas provenientes de la teoría minimax, como son la convexidad en sentido de Fan (o *convexlikeness*) y la quasiconvexidad, pero reparamos en que el concepto de convexidad que nos permite dar un teorema tipo König menos restrictivo es el de infsup-convexidad (o *affine weakly convexlikeness*), cuya utilidad en dicha teoría ha quedado patente recientemente. Más aún, hemos demostrado que este resultado es óptimo, pues la clase de convexidad considerada es justamente la que lo hace cierto:

Teorema 0.2 *Sean X un conjunto no vacío, $n \geq 1$ y $f, f_1, \cdots, f_n : X \longrightarrow \mathbb{R}$ funciones dadas. Entonces la familia $f_1 - f, \cdots, f_n - f$ es infsup-convexa en X si, y solo si, para todo α real se tiene que la validez de la desigualdad*

$$x \in X \;\Rightarrow\; f(x) + \alpha \leq \max_{i=1,\dots,n} f_i(x)$$

implica la existencia de un $\mathbf{s} \in \Delta_n$ tal que

$$x \in X \;\Rightarrow\; f(x) + \alpha \leq \sum_{i=1}^{n} s_i f_i(x).$$

Ello nos permite deducir mejoras óptimas de los teoremas de los multiplicadores de Lagrange, Karush–Kuhn–Tucker y Fritz John para programas semi-infinitos del siguiente tipo: si $N, M \geq 0$, X es un conjunto no vacío y $f, g_1, \dots, g_N, h_1, \dots, h_M : X \longrightarrow \mathbb{R}$ son funciones dadas, de forma que el conjunto factible

$$X_0 := \{x \in X : \; g_1(x) \leq 0, ..., g_N(x) \leq 0, h_1(x) = 0, ..., h_M(x) = 0\}$$

es no vacío, consideramos el programa semi-infinito

$$\inf_{x \in X_0} f(x). \tag{0.1}$$

En concreto, para el primero de los tres resultados clásicos mencionados deducimos esta caracterización de la dualidad entre soluciones del problema no lineal anterior y puntos de silla del correspondiente lagrangiano:

Teorema 0.3 *Sean X un conjunto no vacío, $f, g_1, ..., g_N, h_1, ..., h_M$ funciones reales definidas en X determinando un conjunto factible*

$$X_0 = \{x \in X : g_1(x) \leq 0, ..., g_N(x) \leq 0, h_1(x) = 0, ..., h_M(x) = 0\}$$

no vacío, y sea $x^0 \in X$ una solución del problema de programación semi-infinita (0.1). Supongamos además que existe $x^1 \in X$ tal que

$$g_1(x^1) < 0, ..., g_N(x^1) < 0, h_1(x^1) = 0, ..., h_M(x^1) = 0,$$

y dados $\boldsymbol{\alpha}, \boldsymbol{\beta} \in \mathbb{R}^M_+$ alguno distinto de cero, encontramos $x^2 \in X$ de forma que

$$\sum_{j=1}^{M} \alpha_j h_j(x^2) < \sum_{j=1}^{M} \beta_j h_j(x^2)$$

(condición de Slater). Entonces existe $(\boldsymbol{\lambda}^0, \boldsymbol{\mu}^0) \in \mathbb{R}^N_+ \times \mathbb{R}^M$ tal que $(x^0, \boldsymbol{\lambda}^0, \boldsymbol{\mu}^0)$ es un punto de silla del lagrangiano asociado a (0.1) *si, y solo si, la familia $g_1, ..., g_N, \pm h_1, ..., \pm h_M, f - f(x^0)$ es infsup-convexa en X.*

Para abordar el estudio de programas infinitos, nos centramos en el tercer capítulo en aquellos que solo incluyen restricciones de desigualdad, esto es, problemas de la forma (0.1) donde, X y Λ son conjuntos no vacíos, $f : X \longrightarrow \mathbb{R}$ es una función y $\{g_\lambda\}_{\lambda\in\Lambda}$ es una familia de funciones reales definidas en X tal que para cualquier $x \in X$, $\{g_\lambda(x)\}_{\lambda\in\Lambda} \in \ell^\infty(\Lambda)$ y el conjunto factible

$$X_0 := \left\{ x \in X : \sup_{\lambda\in\Lambda} g_\lambda(x) \leq 0 \right\}$$

es no vacío. Para analizar este tipo de problema y establecer resultados que permitan estudiarlo a partir de otro sin restricciones (multiplicadores de Lagrange, Karush–Kuhn–Tucker, Fritz John), debemos retomar la cuestión inicial sobre el teorema de Hahn–Banach. Dado que el teorema del máximo de König es equivalente al teorema de Hahn–Banach solo en dimensión finita, debemos recurrir a otro resultado que no se vea limitado por la dimensión. Un tal enunciado es el teorema del supremo de König, análogo al teroema del máximo, pero reemplazando una familia finita por otra en un espacio de funciones acotadas y el funcional sublineal máximo por el supremo. Nuevamente la infsup-convexidad determina una perfecta equivalencia con la validez de los teoremas sin restricciones. A modo ilustrativo, hemos probado el siguiente teorema de Karush–Kuhn–Tucker:

Teorema 0.4 *Supongamos que X y Λ son conjuntos no vacíos, $f : X \longrightarrow \mathbb{R}$ es una función y que $\{g_\lambda\}_{\lambda\in\Lambda}$ es una familia de funciones reales definidas en X tal que*

$$x \in X \;\Rightarrow\; \{g_\lambda(x)\}_{\lambda\in\Lambda} \in \ell^\infty(\Lambda)$$

y de forma que el conjunto factible

$$X_0 = \left\{ x \in X :\; \sup_{\lambda\in\Lambda} g_\lambda(x) \leq 0 \right\}$$

es no vacío. Si además $\{g_\lambda\}_{\lambda\in\Lambda}, f - f(x^0)$ es una familia infsup-convexa en X que se verifica la condición de Slater: existe $x^1 \in X$ tal que

$$\sup_{\lambda\in\Lambda} g_\lambda(x^1) < 0,$$

entonces x^0 es una solución del correspondiente programa infinito si, y solo si, existe $\phi_0 \in \ell^\infty(\Lambda)^*_+$ verificando las condiciones tipo Karush–Kuhn–Tucker

(KKT1)

$$f + \phi_0(\{g_\lambda\}_{\lambda\in\Lambda}) \text{ alcanza su ínfimo en } X \text{ en } x^0$$

y

(KKT2)

$$\phi_0(\{g_\lambda(x^0)\}_{\lambda\in\Lambda}) = 0 \quad y \quad \sup_{\lambda\in\Lambda} g(x^0) \leq 0.$$

La condición de Slater considerada, que es uniforme, no puede ser impuesta puntualmente, tal y como se ilustra con un ejemplo concreto.

En el Capítulo 4 estudiamos un problema de programación no lineal general –dominio cualquiera, número arbitrario de restricciones de igualdad y desigualdad–, pero a diferencia de lo hecho en los casos anteriores, donde se deducen de una adecuada reformulación del teorema de Hahn–Banach –teoremas del máximo y supremo de König–, ahora se prueban directamente del teorema de Hahn–Banach en forma de teorema de separación de convexos. El motivo que nos ha llevado a ello es que esta versión del teorema de Hahn–Banach es de uso más extendido entre el mundo de la optimización. No obstante, los resultados derivados vuelven a ser óptimos, como por ejemplo ocurre con el teorema de Fritz John:

Teorema 0.5 *Sean X, Λ y Ω conjuntos tales que X es no vacío, $f : X \longrightarrow \mathbb{R}$ una función y $\{g_\lambda\}_{\lambda\in\Lambda}$ y $\{h_\omega\}_{\omega\in\Omega}$ dos familias de funciones reales definidas en X tales que para todo x en X, $\{g_\lambda(x)\}_{\lambda\in\Lambda} \in \ell^\infty(\Lambda)$ y $\{h_\omega(x)\}_{\omega\in\Omega} \in \ell^\infty(\Omega)$. Supongamos también que el conjunto factible*

$$X_0 = \left\{ x \in X : \sup_{\lambda\in\Lambda} g_\lambda(x) \leq 0, \ \sup_{\omega\in\Omega} \mid h_\omega(x) \mid = 0 \right\}$$

es no vacío y que x^0 es una solución del problema de programa infinito asociado. Entonces existen $\rho \geq 0$, $\phi_0 \in \ell^\infty(\Lambda)^*_+$, $\varphi_1 \in \ell^\infty(\Omega)^*_+$ y $\varphi_2 \in \ell^\infty(\Omega)^*_+$ tales que $\rho + \phi_0(\mathbb{1}) + \varphi_1(\mathbb{1}) + \varphi_2(\mathbb{1}) = 1$,

(FJ1)

$$\rho f + \phi_0(\{g_\lambda\}_{\lambda\in\Lambda}) + (\varphi_1 - \varphi_2)(\{h_\omega\}_{\omega\in\Omega}) \ \textit{alcanza su ínfimo en } X \textit{ en } x^0$$

y

(FJ2)

$$\phi_0(\{g_\lambda(x^0)\}_{\lambda\in\Lambda}) = 0 \quad y \quad (\varphi_1 - \varphi_2)(\{h_\omega(x^0)\}_{\omega\in\Omega}) = 0,$$

si, y solo si, $\{g_\lambda\}_{\lambda\in\Lambda}, \{h_\omega\}_{\omega\in\Omega}, \{-h_\omega\}_{\omega\in\Omega}, f - f(x^0)$ *es una familia infsup-convexa en* X.

La optimalidad de todos las aportaciones originales en el ámbito de la programación no lineal es estricta en relación a otros resultados conocidos, clásicos o recientes, hecho que se ha ilustrado mediante convenientes ejemplos en la tesis.

Parte de los enunciados originales que aparecen en esta memoria han sido publicados en [40] y [41], y otros están recogidos en [42]. Además, algunos de ellos se han presentado en el *WorkShop on Banach Spaces*, (Granada, 2015) y en *17th International Conference on Computational and Mathematical Methods in Science and Engineering*, (Rota, 2017) [43].

Objetivos y Metodología

Los objetivos que en esencia nos marcamos al iniciar la labor que da su fruto en la elaboración de esta memoria son dos. En primer lugar, establecer teoremas tipo Hahn–Banach, en la línea de los del máximo y supremo de König, en los que el tipo de hipótesis de convexidad (linealidad y sublinealidad) puediese ser reemplazado por otro más general. Y por otro lado, aplicar los resultados obtenidos en relación al teorema de Hahn–Banach en la mejora de teoremas tipo multiplicadores de Lagrange, Karush–Kuhn–Tucker y Fritz John. El logro de ambos hitos, que de hecho ha generado resultados óptimos en términos de un concepto de convexidad –infsup-convexidad– proveniente de la teoría minimax, queda avalado por las publicaciones originales [40, 41].

Con relación a la metodología, partimos de un estudio de carácter recopilatorio de resultados conocidos del análisis convexo, clásicos o recientes, pero vinculados de algún modo al teorema de Hahn–Banach y la optimización no lineal. A partir de ahí, comenzamos el trabajo original de la tesis, avanzando gradualmente en la deducción de nuestras aportaciones. Dicho trabajo se vió beneficiado del contacto con otros investigadores, como los miembros del grupo de investigación *Análisis Numérico y sus Aplicaciones*, en cuyo seminario el autor ha participado a lo largo de este tiempo. Mención especial merece la estancia de investigación que, por un periodo de tres meses, llevó a cabo junto al profesor C. Zalinescu (Universidad Alexandru Ioan Cuza, Iasi, Rumanía). Hemos completado la formación elaborando los dos trabajos mencionados, además de otro sometido a publicación, [42], y presentando los resultados novedosos de esta tesis en dos congresos: *WorkShop on Banach Spaces*, (Granada, 2015) y *17th International Conference on Computational and Mathematical Methods in Science and Engineering*, (Rota, 2017), [43].

Capítulo 1

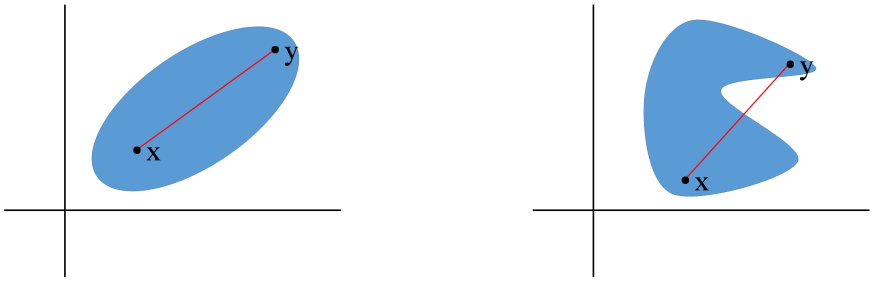
Nociones y Resultados Básicos

Comenzamos esta memoria con un capítulo de carácter introductorio, donde repasamos algunos conceptos propios del análisis convexo que, si bien son básicos, se usarán a lo largo de la misma. También recordamos algunos resultados bien conocidos, entre los que destaca el teorema de Mazur–Orlicz, una reformulación equivalente del teorema de Hahn–Banach que jugará un papel fundamental en gran parte de la tesis.

Introducimos en primer lugar la noción más elemental de todas, que necesariamente requiere un ambiente vectorial. Nosotros trabajaremos siempre con espacios vectoriales reales, que es el contexto donde nuestras aplicaciones en programación no lineal tienen sentido.

Definición 1.1 Un subconjunto C de un espacio vectorial real es *convexo* siempre que

$$x, y \in C, \quad t \in [0,1] \quad \Rightarrow \quad tx + (1-t)\,y \in C.$$



(a) Subonjunto convexo (b) Subconjunto no convexo

Figura 1.1

Mencionemos que, dados dos puntos x, y de un espacio vectorial, se denomina *segmento* de extremos x e y al conjunto

$$[x, y] := \{tx + (1-t)\, y : \ 0 \leq t \leq 1\}.$$

Con este concepto, tenemos una clara caracterización geométrica de un subconjunto convexo de un espacio vectorial real: contiene a todos los segmentos cuyos extremos pertenecen a él.

Ejemplos 1.2 1. Los subconjuntos convexos de $\mathbb{R}$ son sus intervalos.

2. En la Figura 1.1 podemos ver un ejemplo de un subconjunto convexo de $\mathbb{R}^2$ y otro que no lo es.

3. Si adoptamos el convenio de que $\mathbf{s}$ representa al vector $(s_1, ..., s_n) \in \mathbb{R}^n$, entonces el *n-símplex de probabilidad* de $\mathbb{R}^n$ definido como

$$\Delta_n := \left\{ \mathbf{s} \in \mathbb{R}^n : \ \mathbf{s} \in \mathbb{R}^n_+ \quad \text{y} \quad \sum_{i=1}^{n} s_i = 1 \right\},$$

es un subconjunto convexo de dicho espacio vectorial.

■

Resulta inmediato comprobar que la intersección arbitraria de subconjuntos convexos de un espacio vectorial real E es un conjunto convexo. Igualmente es fácil probar por inducción que un subconjunto C de E es convexo si, y solo si, para cualesquiera $m \geq 1$, $x_1, ..., x_m \in C$ y $\mathbf{t} \in \Delta_m$, se tiene que $\sum_{j=1}^{m} t_j x_j \in C$. El siguiente concepto hace referencia precisamente a los vectores de este tipo.

Definición 1.3 Dados un espacio vectorial real E y $x_1, ..., x_m \in E$, llamamos *combinación convexa* de los vectores $x_1, ..., x_m$ a un vector de la forma

$$\sum_{j=1}^{m} t_j x_j,$$

siendo $m \geq 1$ y $\mathbf{t} \in \Delta_m$.

Ejemplos 1.4

1. En la Figura 1.2 se muestran en color naranja todos los puntos que son combinación convexa de A, B, C y D, como por ejemplo, el punto P.

2. En $\mathbb{R}^n$, el conjunto de las combinaciones convexas de los vectores de la base usual coincide con el n-símplex de probabilidad.

■

El conjunto de todas las combinaciones convexas de los vectores de un subconjunto de un espacio vectorial tiene interés en sí mismo:

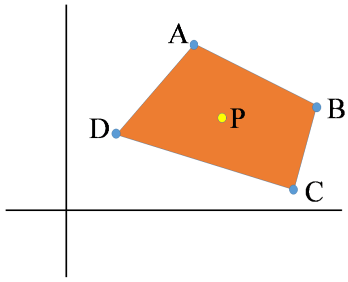


Figura 1.2: Combinación convexa

Definición 1.5 Sean E un espacio vectorial real y A un subconjunto suyo. La *envolvente convexa* de A, que se denota por $\mathrm{co}(A)$, es el subconjunto de E constituido por todas las combinaciones convexas de los vectores de A, esto es,

$$\mathrm{co}(A) := \left\{ \sum_{j=1}^{m} t_j x_j : \ m \geq 1, x_1, ..., x_m \in A, \mathbf{t} \in \Delta_m \right\}.$$

Podemos deducir sin esfuerzo que la envolvente convexa de A coincide con el menor (relación de orden determinada por la inclusión conjuntista) subconjunto convexo de E que contiene a A, es decir,

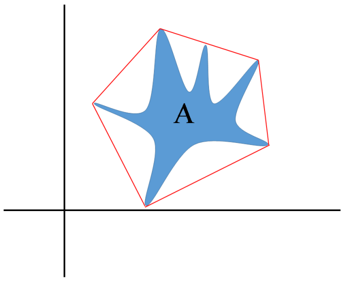
$$\mathrm{co}(A) = \bigcap \{C : \ A \subset C \subset E, \ C \text{ convexo}\}.$$

Ejemplos 1.6

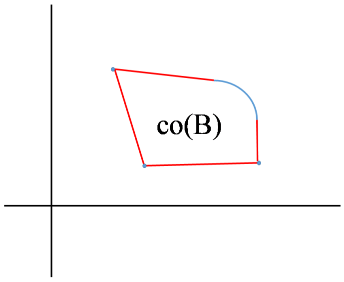
1. En la Figura 1.3 hemos dibujado la envolvente convexa de dos subconjuntos de $\mathbb{R}^2$. Los dos subconjuntos, A y B, están representados en azul, y sus respectivas envolventes convexas son las delimitadas por el contorno rojo.

2. En la Figura 1.4 observamos como en $\mathbb{R}^3$ una circunferencia y punto no coplanario a esta generan un cono como envolvente convexa.

■



(a) Envolvente convexa de A



(b) Envolvente convexa de B

Figura 1.3

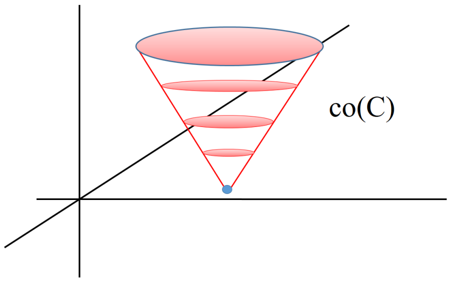


Figura 1.4: Envolvente convexa de C

Pasamos a definir otro concepto elemental, el de función convexa, que desempeña un papel importante en optimización, tanto en sus aspectos teóricos como prácticos: véase, por ejemplo, [2, Chapter 1].

Definición 1.7 Una función real f definida en un subconjunto convexo y no vacío C de un espacio vectorial real es *convexa* si para cada $x, y \in C$ y para todo $t \in [0, 1]$, se verifica

$$f(tx + (1 - t)y) \leq tf(x) + (1 - t)f(y).$$

Diremos que f es *cóncava* si $-f$ es convexa.

Ejemplos 1.8 1. La función $f : \mathbb{R} \longrightarrow \mathbb{R}$, definida como

$$f(x) := x^2, \quad (x \in \mathbb{R}),$$

es convexa y no cóncava, y la función $f : (0, +\infty) \longrightarrow \mathbb{R}$

$$f(x) := \log x, \quad (x \in (0, +\infty)),$$

es cóncava pero no convexa.

2. Dado un espacio vectorial real E, una función $S : E \longrightarrow \mathbb{R}$ se llama funcional *sublineal* si

$$x, y \in E \;\Rightarrow\; S(x + y) \leq S(x) + S(y)$$

y

$$x \in E, \; \rho > 0 \;\Rightarrow\; S(\rho x) = \rho S(x).$$

Es claro que un funcional sublineal es una función convexa. En un espacio vectorial, todo funcional lineal es sublineal y, si además está dotado de una seminorma (en particular, una norma), dicha función es un funcional sublineal. También es fácil comprobar que la función $H : \mathbb{R}^n \longrightarrow \mathbb{R}$ definida para cada $\mathbf{u} = (u_1, ..., u_n) \in \mathbb{R}^n$ como

$$H(\mathbf{u}) := \max_{i=1,\ldots,n} u_i,$$

es sublineal.

■

Es inmediato probar por inducción la *desigualdad de Jensen*, que caracteriza la convexidad de una función: sean E un espacio vectorial real y C

un subconjunto convexo y no vacío suyo; entonces la función $f : C \longrightarrow \mathbb{R}$ es convexa si, y solo si, para cada $m \geq 1$, $x_1, ..., x_m \in E$ y $\mathbf{t} \in \Delta_m$ se verifica

$$f\left(\sum_{j=1}^{m} t_j x_j\right) \leq \sum_{j=1}^{m} t_j f(x_j).$$

Una propiedad muy útil de las funciones convexas, a pesar de su enorme simplicidad, se enuncia en estos términos:

Lema 1.9 *Si E es un espacio vectorial real, C es un subconjunto no vacío y convexo de E y $\varphi : \mathrm{co}(C) \longrightarrow \mathbb{R}$ es una función convexa, entonces*

$$\sup_{x \in C} \varphi(x) = \sup_{x \in \mathrm{co}(C)} \varphi(x).$$

Muchos de los problemas propios del análisis convexo resultan especialmente potentes cuando se debilita la noción de convexidad. Dos de las más populares, que son independientes, aparecen recogidas a continuación. La primera, introducida esencialmente por J. von Neumman ([58, 15, 14, 23, 45]), se refiere a una función, mientras que la segunda, debida a K. Fan ([11, p. 42]), involucra una familia de funciones.

Definición 1.10 Dados un espacio vectorial real E y un subconjunto convexo no vacío C de E, diremos que una función $f : C \longrightarrow \mathbb{R}$ es *quasiconvexa* en C si para cualquier $\alpha \in \mathbb{R}$ se tiene que el subconjunto de E

$$\{y \in C : \ f(y) \leq \alpha\}$$

es convexo.

Un cálculo directo nos convence de que toda función convexa también es quasiconvexa, y un ejemplo simple, como la función logaritmo, de que esta generalización es estricta.

Definición 1.11 Si X y Λ son conjuntos no vacíos, una familia $\{f_\lambda\}_{\lambda\in\Lambda}$ de funciones reales definidas en X es *Fan-convexa* en X cuando para cualesquiera $t \in [0,1]$ y $x_1, x_2 \in X$, existe x_0 en X de forma que

$$\lambda \in \Lambda \ \Rightarrow \ f_\lambda(x_0) \leq t f_\lambda(x_1) + (1-t) f_\lambda(x_2).$$

Es claro que una familia de funciones reales y convexas definidas en un subconjunto convexo de un espacio vectorial real, también es Fan-convexa (o *convexlike*, término de uso extendido) en dicho dominio común y que este tipo de convexidad generalizada no exige estructura vectorial alguna. En el siguiente capítulo necesitaremos trabajar con un concepto de convexidad aún menos restrictivo que el de Fan-convexidad, infsup-convexidad, que se revelará como esencial para establecer la optimalidad de algunas de nuestras aportaciones.

Una vez que hemos recordado algunas debilitaciones de la convexidad, nos centramos en presentar ciertos resultados convexos conocidos y que determinarán muchas de las aportaciones de esta memoria. Algunos de ellos se establecen en el contexto de los espacios normados, y varios de estos espacios se utilizarán explícitamente: ℓ^∞, c_0 y ℓ^1. Por un lado,

$$\ell^\infty := \{\{x_n\}_{n\in\mathbb{N}} \in \mathbb{R}^{\mathbb{N}} : \ \{x_n\}_{n\in\mathbb{N}} \text{ es acotada}\};$$

dotado de la siguiente norma, ℓ^∞ es un espacio de Banach:

$$\|x\|_\infty := \sup_{n\in\mathbb{N}} |x_n|, \qquad (x \in \ell^\infty)\,.$$

De forma más general, dado Λ no vacío,

$$\ell^\infty(\Lambda) := \left\{ x \in \mathbb{R}^\Lambda : \; x \; \text{ es acotada} \right\},$$

que se convierte en un espacio de Banach con la norma del supremo,

$$\|x\| := \sup_{\lambda \in \Lambda} |x(\lambda)|, \qquad (x \in \ell^\infty(\Lambda)),$$

(con lo que $\ell^\infty = \ell^\infty(\mathbb{N})$). El espacio c_0 de las sucesiones reales convergentes a cero es un subespacio cerrado de ℓ^∞. Así pues, con la norma heredada de ℓ^∞, c_0 también es un espacio de Banach. Por su parte, el subespacio vectorial de c_0

$$\ell^1 := \{\{x_n\}_{n \in \mathbb{N}} \in \mathbb{R}^\mathbb{N} : \; \sum_{n=1}^{\infty} |x_n| < \infty\}$$

es un espacio de Banach con la norma

$$\|x\|_1 := \sum_{n=1}^{\infty} |x_n|, \qquad \left(x \in \ell^1\right),$$

que obviamente no es la inducida por ℓ^∞. Es bien conocido que el espacio dual topológico de c_0 coincide con ℓ^1, y el dual topológico de ℓ^1 es ℓ^∞, es decir

$$(c_0)^* = \ell^1 \quad \text{y} \quad (\ell^1)^* = \ell^\infty.$$

El primero de los resultados a enunciar debe ser, por su enorme potencia en las aplicaciones, simplicidad y carga geométrica, sin lugar a dudas, la versión geométrica del teorema de Hahn–Banach ([3, Theorem 1.6]). Se trata del teorema tipo Hahn–Banach más conocido en el ambiente de la optimización, y además lo usaremos para probar uno de nuestros resultados en el Capítulo 4.

Teorema 1.12 *Supongamos que E es un espacio normado real y que A y B son dos subconjuntos suyos no vacíos, disjuntos y convexos, de forma que A tiene interior no vacío. Entonces existen un funcional lineal y continuo $f : E \longrightarrow \mathbb{R}$ y un escalar α tales que*

$$x \in A,\ y \in B \ \Rightarrow\ f(x) < \alpha \leq f(y).$$

Nos ocupamos ahora de otro teorema de Hahn–Banach, en este caso analítico –lo que no impide que posea un fuerte carácter geométrico–, conocido como teorema de Mazur–Orlicz (véanse [39, Théorème 2.41], [44, Theorem, p. 365], [33, Satz, p. 482] y [50, Theorem 28], y sus generalizaciones [33, Satz, p. 482 y Zusatz, p.483], [35, Theorem 1.1], [52, Theorem 2.9], [38, Theorem 2], [8, Theorem 5.1], [22, Theorem 12], [9, Theorem 3.1]), [55, Theorem 3.5] y [53, Theorem 3.5 y Theorem 6.1]). Este teorema es fundamental en esta memoria, ya que constituye una herramienta que genera parte de nuestros principales enunciados.

Teorema 1.13 *Sean E un espacio vectorial real, C un subconjunto no vacío y convexo de E y $S : E \longrightarrow \mathbb{R}$ un funcional sublineal. Entonces existe un funcional lineal $L : E \longrightarrow \mathbb{R}$ verificando*

$$x \in E \ \Rightarrow\ L(x) \leq S(x)$$

e

$$\inf_{x \in C} L(x) = \inf_{x \in C} S(x).$$

Aun suponiendo una reformulación equivalente del teorema de Hahn–Banach, donde no se considera ningún subconjunto convexo C y únicamente se encuentra un funcional lineal L que minora a S, el control sobre C de los ínfimos no es trivial y genera numerosas aplicaciones en áreas como el análisis funcional, teoría minimax, optimización y análisis variacional ([52, 38, 9, 18]).

Para ilustrar el hecho de que un funcional sublineal mayore un funcional lineal, consideremos el funcional sublineal recogido en el Ejemplo 1.8, lo que supone una caracterización dual de los n-símplex de probabilidad de $\mathbb{R}^n$. La prueba no presenta dificultad alguna.

Lema 1.14 *Sea $H : \mathbb{R}^n \longrightarrow \mathbb{R}$ el funcional sublineal definido en cada $\mathbf{u} \in \mathbb{R}^n$ como*

$$H(\mathbf{u}) := \max_{i=1,\ldots,n} u_i.$$

Entonces, cualquier funcional lineal $L : \mathbb{R}^n \longrightarrow \mathbb{R}$ tal que $L \leq H$ es de la forma

$$L(\mathbf{u}) = \sum_{i=1}^{n} s_i u_i, \qquad (\mathbf{u} \in \mathbb{R}^n),$$

para cierto $\mathbf{s} \in \Delta_n$.

Concluimos el capítulo con un teorema atribuido a Dixmier, que aplicaremos a los espacios de Banach de sucesiones ya presentados, lo que será esencial en un ejemplo del Capítulo 3. Dado un espacio normado E, su *inyección canónica* en su bidual es la aplicación $J_E : E \longrightarrow E^{**}$ que a cada x le hace corresponder $J_E(x) \in E^{**}$, que viene dado por

$$J_E(x)(x^*) := x^*(x), \qquad (x^* \in E^*).$$

Además, notaremos por $J_E(E)^{\perp}$ el *anulador* de dicho subespacio de E^{**}.

Teorema 1.15 *Si E es un espacio normado, entonces su tercer dual topológico admite una descomposición en suma topológica de la forma*

$$E^{***} = J_E(E)^{\perp} \oplus J_{E^*}(E^*).$$

En particular, el tercer dual del espacio de sucesiones nulas como la suma topológica

$$(\ell^\infty)^* = (c_0)^\perp \oplus \ell^1.$$

Capítulo 2

Hahn–Banach en Dimensión Finita y Optimización

En este segundo capítulo presentamos aportaciones novedosas con relación al teorema de Hahn–Banach para espacios vectoriales reales de dimensión finita y sus aplicaciones a programación no lineal semi-infinita, y más concretamente, a programas con un número finito de restricciones –tanto igualdades como desigualdades– definidos en conjuntos arbitrarios, que incluso pueden carecer de estructura lineal. Para ello, vamos a obtener una extensión de una reformulación equivalente del teorema de Hahn–Banach finito dimensional, conocida como teorema del máximo de König, que de hecho es óptima. Dicha optimalidad se establece en términos de una noción de convexidad poco restrictiva que surge en la teoría minimax, y que es conocida, entre otros, por los términos de infsup-convexidad o affine weakly convexlikeness.

La versión generalizada del teorema del máximo de König es la herramienta fundamental que permite probar la dualidad entre la existencia de solución para un programa no lineal y la de punto de silla del correspondiente lagrangiano. Este resultado no solo mejora el clásico teorema de los multiplicadores de Lagrange para programas convexos, sino que, de hecho, es óptimo nuevamente en términos de infsup-convexidad. Además, de nuestra versión del teorema del máximo de König se derivan generalizaciones de los teoremas de Karush–Kuhn–Tucker y de Fritz John.

2.1. Teorema del Máximo de König

En esta primera sección presentamos un resultado que, sin ser óptimo aún –de ello nos ocupamos en la Sección 2.2–, generaliza el teorema del máximo de König. Así mismo, se pone en equivalencia dicha extensión inicial del teorema del máximo de König con el teorema de la alternativa de Gordan y otro sobre funciones convexas debido a S. Simons.

Comenzamos recordando el teorema del máximo de König ([34, Maximumssatz p. 500], [35, Theorem 1.2.1]). A la vista del Lema 1.14, se trata de un resultado de existencia de funcionales lineales en $\mathbb{R}^n$, como ocurre con el teorema de Hahn–Banach en ese espacio vectorial finito dimensional.

Teorema 2.1 *Sean E un espacio vectorial real, X un subconjunto no vacío y convexo de E, $n \geq 1$, $S_1, \ldots, S_n : E \longrightarrow \mathbb{R}$ funcionales sublineales y sea*

$L : E \longrightarrow \mathbb{R}$ un funcional lineal tal que

$$L \leq \max_{i=1,\ldots,n} S_i \quad \textit{en } X.$$

Entonces, existe $\mathbf{s} \in \Delta_n$ tal que

$$L \leq \sum_{i=1}^{n} s_i S_i \quad \textit{en } X.$$

Reparemos en el hecho de que la demostración que H. König hace en [35] sigue siendo válida si reemplazamos cada funcional sublineal S_i por una función convexa f_i definida en un subconjunto convexo y no vacío X de E y el funcional lineal L por una función cóncava f definida también en X. De hecho, la única dificultad estriba en comprobar la convexidad del subconjunto de $\mathbb{R}^n$

$$\{\mathbf{u} \in \mathbb{R}^n : \text{ existe } \ x \in X \ \text{ con } \ u_l \geq f_l(x) - f(x), \text{ para todo } \ l = 1, ..., n\}.$$

Introduciendo esta variación en la demostración de [35, Theorem 1.2.1], deducimos una versión ligeramente más general del teorema del máximo de König:

Teorema 2.2 *Dados E un espacio vectorial real, X un subconjunto no vacío y convexo de E, $n \geq 1$, $f_1, \ldots, f_n : X \longrightarrow \mathbb{R}$ funciones convexas y $f : X \longrightarrow \mathbb{R}$ una función cóncava de forma que*

$$f \leq \max_{i=1,\ldots,n} f_i \ \textit{ en } \ X,$$

entonces existe $\mathbf{s} \in \Delta_n$ tal que

$$f \leq \sum_{i=1}^{n} s_i f_i \ \textit{ en } \ X.$$

Por su parte, S. Simons estableció en [51, Lemma 9], como consecuencia del teorema de Hahn–Banach, el siguiente resultado sobre funciones convexas.

Teorema 2.3 *Supongamos que E es un espacio vectorial real y que X es un subconjunto no vacío y convexo de E. Dados $n \geq 1$ y $f_1, \ldots, f_n : X \longrightarrow \mathbb{R}$ funciones convexas, existe $\mathbf{s} \in \Delta_n$ tal que*

$$\inf_{x \in X} \max_{i=1,\ldots,n} f_i(x) = \inf_{x \in X} \sum_{i=1}^{n} s_i f_i(x).$$

Evidentemente este teorema no se sigue del teorema del máximo de König, Teorema 2.1, ya que su hipótesis de sublinealidad es más fuerte que la de convexidad. Pero el teorema de Simons, Teorema 2.3, sí que es equivalente a la generalización del teorema del máximo de König dada en el Teorema 2.2:

Observación 2.4 El Teorema 2.2 es equivalente al Teorema 2.3.

Probemos en primer lugar que la versión del teorema del máximo de König, Teorema 2.2, es consecuencia del teorema de Simons, Teorema 2.3. Sean E un espacio vectorial real, X un subconjunto no vacío y convexo de E, $n \geq 1$, $f_1, \ldots, f_n : X \longrightarrow \mathbb{R}$ funciones convexas y $f : X \longrightarrow \mathbb{R}$ una función cóncava que verifica

$$f \leq \max_{i=1,\ldots,n} f_i \ \text{ en } \ X.$$

Apliquemos el teorema de Simons, Teorema 2.3, a las funciones convexas $f_1 - f, ..., f_n - f$, deduciendo la existencia de $\mathbf{s} \in \Delta_n$ tal que

$$\inf_{x \in X} \max_{i=1,\ldots,n} (f_i(x) - f(x)) = \inf_{x \in X} \sum_{i=1}^{n} s_i (f_i(x) - f(x)). \tag{2.1}$$

Analicemos cada miembro de esta igualdad. Empecemos por el de la izquierda:

$$\begin{aligned}
\text{(2.2)} \qquad \inf_{x\in X} \max_{i=1,\dots,n} (f_i(x) - f(x)) &= \inf_{x\in X} \left(\max_{i=1,\dots,n} f_i(x) - f(x) \right) \\
&\geq 0,
\end{aligned}$$

donde en la desigualdad se ha usado la hipótesis. Para el de la derecha, teniendo en cuenta que $\mathbf{s} \in \Delta_n$,

$$\begin{aligned}
\text{(2.3)} \qquad \inf_{x\in X} \sum_{i=1}^{n} s_i(f_i(x) - f(x)) &= \inf_{x\in X} \left(\sum_{i=1}^{n} s_i f_i(x) - \sum_{i=1}^{n} s_i f(x) \right) \\
&= \inf_{x\in X} \left(\sum_{i=1}^{n} s_i(f_i(x) - f(x)) \right).
\end{aligned}$$

Así pues, de (2.1), (2.2) y (2.3) se sigue

$$0 \leq \inf_{x\in X} \left(\sum_{i=1}^{n} s_i(f_i(x) - f(x)) \right),$$

esto es,

$$f \leq \sum_{i=1}^{n} s_i f_i \quad \text{en } X.$$

Comprobemos ahora que el teorema de Simons, Teorema 2.3, es consecuencia del Teorema 2.2. Sean pues E un espacio vectorial real, X un subconjunto no vacío y convexo de E, $n \geq 1$ y $f_1, \dots, f_n : X \longrightarrow \mathbb{R}$ funciones convexas dadas. Si

$$\inf_{x\in X} \max_{i=1,\dots,n} f_i(x) = -\infty$$

no hay nada que probar. En caso contrario, definimos la función cóncava $f : X \longrightarrow \mathbb{R}$ como

$$f(x) := \inf_{u\in X} \max_{i=1,\dots,n} f_i(u), \qquad (x \in X).$$

Para cada $x \in X$, $f(x) \leq \max_{i=1,...,n} f_i(x)$, luego estamos en condiciones de aplicar el Teorema 2.2, que nos garantiza la existencia de un $\mathbf{s} \in \Delta_n$ cumpliendo

$$f \leq \sum_{i=1}^{n} s_i f_i \quad \text{en } X,$$

es decir,

$$\inf_{x \in X} \max_{i=1,...,n} f_i(x) \leq \inf_{x \in X} \sum_{i=1}^{n} s_i f_i(x).$$

La desigualdad contraria siempre es cierta, luego tenemos que para cierto $\mathbf{s} \in \Delta_n$

$$\inf_{x \in X} \max_{i=1,...,n} f_i(x) = \inf_{x \in X} \sum_{i=1}^{n} s_i f_i(x).$$

■

Tal y como hemos anunciado al comienzo de esta sección, vamos a establecer una equivalencia entre el Teorema 2.2 y un teorema de la alternativa. Este tipo de resultados constituyen unas herramientas muy potentes cuando se trabaja con problemas de optimización ([6, 19, 29, 37, 28, 5, 10, 17, 7, 20, 9]). La mayoría de ellos son equivalentes al teorema de Hahn–Banach, aunque muchos aparecieron antes, como el teorema de Gordan, que data de 1873 ([21]), y el lema de Farkas ([13]), de 1902. Nos centramos en este punto en el teorema de Gordan, que establece que dados $N, n \geq 1$ y $\mathbf{x}_1, ..., \mathbf{x}_n \in \mathbb{R}^N$, entonces, o bien

$$\text{el sistema } \ \langle \mathbf{x}_i, \mathbf{x} \rangle > 0 \ \text{ admite solución } \ \mathbf{x} \in \mathbb{R}^N \ \text{ para } \ i = 1, ..., n,$$

o bien

$$\sum_{i=1}^{n} \lambda_i \mathbf{x}_i = \mathbf{0} \ \text{ tiene una solución no nula } \ 0 \leq \lambda_1, ..., \lambda_n \in \mathbb{R},$$

pero nunca se cumplen ambas afirmaciones a la vez. La interpretación geométrica del teorema se muestra en la siguiente figura, Figura 2.1: si el origen no

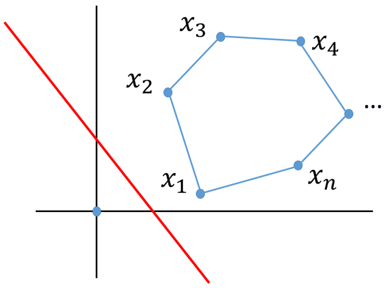


Figura 2.1: Teorema de la alternativa de Gordan

pertenece a la envolvente convexa de $\mathbf{x}_1, ..., \mathbf{x}_n$ podemos encontrar un hiperplano que separe a la envolvente convexa del origen. Una extensión interesante y muy aplicada del teorema de Gordan es el siguiente resultado de la alternativa para funciones convexas, debido a K. Fan, I. Glicksberg y A.J. Hoffman, [12, Theorem 1].

Teorema 2.5 *Si E es un espacio vectorial real, X es un subconjunto no vacío y convexo de E, $n \geq 1$ y $f_1, \ldots, f_n : X \longrightarrow \mathbb{R}$ son funciones convexas, entonces una, y solo una, de las siguientes desigualdades admite solución*

(i)
$$\textit{Existe } \ x \in X \ \ \textit{verificando} \ \ \max_{i=1,\ldots,n} f_i(x) < 0.$$

(ii)
$$\textit{Existe } \ \mathbf{s} \in \Delta_n \ \ \textit{tal que } \ 0 \leq \inf_{x \in X} \sum_{i=1}^{n} s_i f_i(x).$$

Este resultado admite la siguiente variante, ligeramente más general:

Teorema 2.6 *Sean E un espacio vectorial real, X un subconjunto no vacío y convexo de E, $n \geq 1$ y $f_1, \ldots, f_n : X \longrightarrow \mathbb{R}$ funciones convexas. Entonces para cada $\alpha \in \mathbb{R}$, una, y solo una, de las siguientes alternativas tiene solución*

(i)

$$\textit{Existe } \ x \in X \ \ \textit{tal que} \quad \max_{i=1,\ldots,n} f_i(x) < \alpha.$$

(ii)

$$\textit{Existe } \ \mathbf{s} \in \Delta_n \ \ \textit{cumpliendo} \ \ \alpha \leq \inf_{x \in X} \sum_{i=1}^{n} s_i f_i(x).$$

De hecho, este resultado es claramente equivalente al Teorema 2.5. En efecto; cuando se toma en él $\alpha = 0$ se recupera el Teorema 2.5. Y recíprocamente, solo hay que aplicar el Teorema 2.6 a las funciones convexas $f_1 - \alpha, ..., f_n - \alpha$.

Como hemos mencionado anteriormente, los teoremas de la alternativa no son más que reformulaciones del teorema de Hahn–Banach. Prueba de ello es la equivalencia entre la última generalización del teorema de Gordan, Teorema 2.6, y el teorema de Simons, Teorema 2.3:

Observación 2.7 Los Teoremas 2.6 y 2.3 son equivalentes.

Comprobemos que el Teorema 2.6 es consecuencia del teorema de Simons, Teorema 2.3. Sean entonces E un espacio vectorial real, X un subconjunto no vacío y convexo de E, $n \geq 1$ y $f_1, \ldots, f_n : X \longrightarrow \mathbb{R}$ funciones convexas.

Dado $\alpha \in \mathbb{R}$, es evidente que se da la alternativa

$$\begin{cases} \text{existe } \ x \in X \ \text{ tal que } \ \max\limits_{i=1,\ldots,n} f_i(x) < \alpha \\ \\ \text{o bien (excluyente)} \\ \\ \text{para todo } \ x \in X \ \text{ se verifica } \ \max\limits_{i=1,\ldots,n} f_i(x) \geq \alpha. \end{cases}$$

Con la idea de probar el Teorema 2.6, observemos que la segunda alternativa no es más que

$$\inf_{x \in X} \max_{i=1,\ldots,n} f_i(x) \geq \alpha. \tag{2.4}$$

Aplicamos el Teorema 2.3 a las funciones convexas $f_1, ..., f_n$, luego existe $\mathbf{s} \in \Delta_n$ tal que

$$\inf_{x \in X} \max_{i=1,\ldots,n} f_i(x) = \inf_{x \in X} \sum_{i=1}^{n} s_i f_i(x),$$

lo que combinado con (2.4) da lo que buscamos.

Probemos ahora que el teorema de Simons, Teorema 2.3, se sigue de la generalización del teorema de Gordan, Teorema 2.6. Sean pues E un espacio vectorial real, X un subconjunto no vacío y convexo de E, $n \geq 1$ y $f_1, \ldots, f_n : X \longrightarrow \mathbb{R}$ funciones convexas. Si tomamos $\alpha = \inf\limits_{x \in X} \max\limits_{i=1,\ldots,n} f_i(x)$ (podemos suponer, sin pérdida de generalidad, que es finito), entonces este escalar no verifica la alternativa (i) del Teorema 2.6, por lo que dicho resultado garantiza la existencia de $\mathbf{s} \in \Delta_n$ tal que

$$\alpha \leq \inf_{x \in X} \sum_{i=1}^{n} s_i f_i(x),$$

es decir,

$$\inf_{x \in X} \max_{i=1,\ldots,n} f_i(x) \leq \inf_{x \in X} \sum_{i=1}^{n} s_i f_i(x),$$

y como la desigualdad contraria siempre es válida, hemos concluido.

■

Así pues, hemos comprobado que la validez de cada uno de los siguientes resultados implica la de los demás: la generalización del teorema del máximo de König, Teorema 2.2, el teorema de Simons, Teorema 2.3, y la extensión del teorema de Gordan, Teorema 2.6.

2.2. Versión Óptima del Teorema de König

A continuación nos proponemos hacer la primera aportación no trivial de esta tesis, a saber, debilitar todo lo posible las hipótesis de la versión dada en la sección anterior del teorema del máximo de König, Teorema 2.2. En primer lugar, demostraremos que las condiciones de convexidad pueden ser reemplazadas por la infsup-convexidad de una determinada familia de funciones, que es un concepto más general, y seguidamente comprobaremos que el resultado así obtenido es óptimo.

Empezamos recordando la noción de familia infsup-convexa de funciones, ilustrándola además con algunos ejemplos.

Definición 2.8 Sean X un conjunto no vacío, $n \geq 1$ y $f_1, ..., f_n : X \longrightarrow \mathbb{R}$ funciones dadas. Decimos que la familia $f_1, ..., f_n$ es *infsup-convexa* en X si dados $m \geq 1, \mathbf{t} \in \Delta_m$ y $x_1, ..., x_m \in X$, se verifica

$$\inf_{x \in X} \max_{i=1,...,n} f_i(x) \leq \max_{i=1,...,n} \sum_{j=1}^{m} t_j f_i(x_j).$$

El concepto de infsup-convexidad (o *affine weakly convexlikeness*) surge en la teoría minimax ([54, 32, 47, 48]). En nuestro contexto aparece de manera natural buscando una versión óptima del teorema del máximo de König.

Ejemplos 2.9 1. La infsup-convexidad generaliza la convexidad en el sentido de Fan. Es fácil probar (véase, por ejemplo [47, Lemma 2.2]), que dados $n \geq 1$ y X no vacío, entonces la familia de funciones reales $f_1, ..., f_n$ definidas en X es convexa en el sentido de Fan si, y solo si, para cualesquiera $m \geq 1$, $\mathbf{t} \in \Delta_m$ y $x_1, ..., x_m \in X$ existe un x_0 en X tal que

$$i \in \{1, ..., n\} \;\Rightarrow\; f_i(x_0) \leq \sum_{j=1}^{m} t_j f_i(x_j)$$

–repárese en la analogía con la desigualdad de Jensen–, de donde se deduce

$$\max_{i=1,...,n} f_i(x_0) \leq \max_{i=1,...,n} \sum_{j=1}^{m} t_j f_i(x_j),$$

y consecuentemente

$$\inf_{x \in X} \max_{i=1,...,n} f_i(x) \leq \max_{i=1,...,n} \sum_{j=1}^{m} t_j f_i(x_j).$$

Para comprobar que esta generalización es estricta basta considerar las funciones $f_1, f_2 : [0,1] - \left\{\frac{1}{2}\right\} \longrightarrow \mathbb{R}$ definidas en cada $0 \leq x \leq 1$ como $f_1(x) := x$ y $f_2(x) := 1 - x$, y tomar $m = 2$, $t = \dfrac{1}{2}$, $x_1 = 0$ y $x_2 = 1$.

2. El concepto de infsup-convexidad es independiente del de quasiconvexidad. Para ello, basta considerar la familia de funciones $f_1, f_2 : [-1,1] \longrightarrow \mathbb{R}$ definidas por

$$f_1(x) := \begin{cases} 0 & \text{si} \quad -1 \leq x \leq 0 \\ \\ -x & \text{si} \quad 0 < x \leq 1 \end{cases}$$

y

$$f_2(x) := \begin{cases} x & \text{si} \quad -1 \leq x \leq 0 \\ \\ 0 & \text{si} \quad 0 < x \leq 1, \end{cases}$$

que claramente son quasiconvexas. Sin embargo esta familia no es infsup-convexa en $[-1,1]$. En efecto: por un lado, es claro que

$$\inf_{x \in X} \max_{i=1,2} f_i(x) = 0;$$

y por otro, si $m = 2$, $t_1 = t_2 = 0,5$, $x_1 = -1$ y $x_2 = 1$, entonces

$$\max_{i=1,2} \sum_{j=1}^{2} t_j f_i(x_j) = -0,5,$$

lo que muestra que la familia f_1, f_2 no es infsup-convexa en $[-1,1]$. En cambio, la familia $g_1, g_2 : [-1,1] \longrightarrow \mathbb{R}$ definida como

$$g_1(x) := -x^2 + 1, \quad (x \in [-1,1])$$

y

$$g_2(x) := 0, \quad (x \in [-1,1])$$

es infsup-convexa en $[-1,1]$ –basta realizar un sencillo cálculo–, aunque g_1 no es quasiconvexa, ya que el conjunto

$$\{y \in [-1,1] : \ g_1(y) \leq 0{,}5\}$$

no es convexo.

■

Aunque la infsup-convexidad extiende, en particular, el concepto de convexidad clásica, debemos insistir en este punto de que no ocurre así con muchas de las propiedades de esta. Por ejemplo, a diferencia de lo que ocurre con la convexidad clásica, la infsup-convexidad no es estable por sumas finitas. Por ejemplo, sean $f_1, f_2, g_1, g_2 : [-1,1] \longrightarrow \mathbb{R}$ las funciones definidas por

$$f_1(x) := 0, \qquad (x \in [-1,1]),$$

$$f_2(x) := \begin{cases} x & \text{si} \quad -1 \leq x \leq 0 \\ \\ 0 & \text{si} \quad 0 < x \leq 1 \end{cases},$$

$$g_1(x) := \begin{cases} 0 & \text{si} \quad -1 \leq x \leq 0 \\ \\ -x & \text{si} \quad 0 < x \leq 1 \end{cases}$$

y

$$g_2(x) := 0, \qquad (x \in [-1,1]).$$

Ambas familias, f_1, f_2 y g_1, g_2, son infsup-convexas en $[-1,1]$. En cambio, la familia suma

$$(f_1 + g_1)(x) = \begin{cases} 0 & \text{si} \quad -1 \leq x \leq 0 \\ \\ -x & \text{si} \quad 0 < x \leq 1 \end{cases}$$

y

$$(f_2 + g_2)(x) = \begin{cases} x & \text{si} \quad -1 \leq x \leq 0 \\ \\ 0 & \text{si} \quad 0 < x \leq 1, \end{cases}$$

no es infsup-convexa en $[-1,1]$, como se ha probado en el apartado 2 de Ejemplos 2.9.

Sin embargo, sí se mantiene la estabilidad por sumas de una constante, ya que si $n, m \geq 1$, $f_1, ..., f_n : X \longrightarrow \mathbb{R}$ es una familia infsup-convexa en X, $\mathbf{t} \in \Delta_m$, $x_1, ..., x_m \in X$ y α es un número real, entonces

$$\begin{aligned}
\operatorname*{ínf}_{x\in X} \operatorname*{máx}_{i=1,\ldots,n} (f_i(x) + \alpha) &= \left(\operatorname*{ínf}_{x\in X} \operatorname*{máx}_{i=1,\ldots,n} f_i(x)\right) + \alpha \\
&\leq \left(\operatorname*{máx}_{i=1,\ldots,n} \sum_{j=1}^{m} t_j f_i(x_j)\right) + \alpha \\
&= \operatorname*{máx}_{i=1,\ldots,n} \sum_{j=1}^{m} t_j (f_i(x_j) + \alpha).
\end{aligned}$$

Antes de ocuparnos de nuestra versión del teorema del máximo de König, establecemos el siguiente resultado técnico.

Lema 2.10 *Supongamos que* $\mathbf{s} \in \Delta_n$, *para cierto natural* n, X *es un conjunto no vacío,* $f, f_1, \ldots, f_n : X \longrightarrow \mathbb{R}$ *es una familia de funciones y consideremos el conjunto*

$$B := \{\mathbf{u} \in \mathbb{R}^n : \text{ existe } x \in X \text{ con } u_i \geq f_i(x) - f(x), \text{ para todo } i = 1, ..., n\}.$$

Entonces

$$\operatorname*{ínf}_{\mathbf{u}\in B} \sum_{i=1}^{n} s_i u_i = \operatorname*{ínf}_{x\in X} \sum_{i=1}^{n} s_i (f_i(x) - f(x)).$$

Demostración. Para probar la desigualdad " $\geq$ " elegimos un $\mathbf{u} \in B$, luego la desigualdad

$$\sum_{i=1}^{n} s_i u_i \geq \sum_{i=1}^{n} s_i (f_i(x_0) - f(x_0))$$

es válida para algún $x_0 \in X$. Como

$$\sum_{i=1}^{n} s_i (f_i(x_0) - f(x_0)) \geq \operatorname*{ínf}_{x\in X} \sum_{i=1}^{n} s_i (f_i(x) - f(x))$$

y $\mathbf{u} \in B$ es arbitrario, deducimos que

$$\inf_{\mathbf{u}\in B} \sum_{i=1}^{n} s_i u_i \geq \inf_{x\in X} \sum_{i=1}^{n} s_i(f_i(x) - f(x)).$$

Para demostrar la desigualdad contraria, consideremos el subconjunto de $\mathbb{R}^n$

$$D := \{(f_1(x) - f(x), ..., f_n(x) - f(x)) : \ x \in X\},$$

que claramente está contenido en B, luego

$$\begin{aligned} \inf_{\mathbf{u}\in B} \sum_{i=1}^{n} s_i u_i &\leq \inf_{\mathbf{u}\in D} \sum_{i=1}^{n} s_i u_i \\ &= \inf_{x\in X} \sum_{i=1}^{n} s_i(f_i(x) - f(x)). \end{aligned}$$

■

Ya estamos en condiciones de enunciar una versión más general del teorema del máximo de König que la dada en el Teorema 2.2 de la Sección 2.1. La hipótesis de infsup-convexidad permite prescindir de la convexidad y concavidad que se requieren en dicho resultado.

Teorema 2.11 *Sean X un conjunto no vacío, $n \geq 1$ y $f, f_1, \cdots, f_n$ funciones reales definidas en X tales que $f_1 - f, ..., f_n - f$ es una familia infsup-convexa y*

$$f \leq \max_{i=1,...,n} f_i \quad \text{en } X.$$

Entonces existe $\mathbf{s} \in \Delta_n$ de forma que

$$f \leq \sum_{i=1}^{n} s_i f_i \quad \text{en } X.$$

Demostración. Consideremos el funcional sublineal $H : \mathbb{R}^n \longrightarrow \mathbb{R}$ definido en cada $\mathbf{u} \in \mathbb{R}^n$ como

$$H(\mathbf{u}) := \max_{i=1,\dots,n} u_i.$$

Sea

$$B := \{\mathbf{u} \in \mathbb{R}^n : \text{ existe } \ x \in X \ \text{ con } \ u_i \geq f_i(x) - f(x), \text{ para todo } \ i = 1, ..., n\}$$

y apliquemos el teorema de Mazur–Orlicz, Teorema 1.13, al espacio vectorial $E = \mathbb{R}^n$, su subconjunto no vacío y convexo $C := \operatorname{co}(B)$ y al funcional sublineal $S = H$. Obtenemos entonces un funcional lineal $L \leq H$ de forma que

$$\inf_C L = \inf_C H. \tag{2.5}$$

El Lema 1.14 garantiza entonces la existencia de $\mathbf{s} \in \Delta_n$ tal que

$$L(\mathbf{u}) = \sum_{i=1}^n s_i u_i, \qquad (\mathbf{u} \in \mathbb{R}^n),$$

luego de este hecho y (2.5) deducimos que, para algún $\mathbf{s} \in \Delta_n$,

$$\inf_{\mathbf{u} \in C} \sum_{i=1}^n s_i u_i = \inf_{\mathbf{u} \in C} \max_{i=1,\dots,n} u_i. \tag{2.6}$$

Analicemos el segundo miembro de la igualdad. Sea $\mathbf{u} \in C$, i.e.,

$$\mathbf{u} = \sum_{j=1}^m t_j \mathbf{u}^\mathbf{j}$$

para ciertos $m \geq 1$, $\mathbf{t} \in \Delta_m$ y $\mathbf{u}^\mathbf{j} \in B$. Dado $j = 1, ..., m$, sea $x_j \in X$ con

$u_i^j \geq f_i(x^j) - f(x^j)$ para todo $i \in \{1, ..., n\}$. Entonces

$$\begin{aligned}
\max_{i=1,\dots,n} u_i &= \max_{i=1,\dots,n} \left(\sum_{j=1}^{m} t_j u^j \right)_i \\
&\geq \max_{i=1,\dots,n} \sum_{j=1}^{m} t_j (f_i(x^j) - f(x^j)) \\
&\geq \inf_{x \in X} \max_{i=1,\dots,n} (f_i(x) - f(x)) \\
&\geq 0,
\end{aligned}$$

donde hemos usado la infsup-convexidad de la familia $(f_i - f)_{i=1}^n$ y la hipótesis $f \leq \max_{i=1,\dots,n} f_i$ en X. Por tanto, la arbitrariedad de $\mathbf{u} \in C$ da

$$\inf_{\mathbf{u} \in C} \max_{i=1,\dots,n} u_i \geq 0. \tag{2.7}$$

Por otro lado, en virtud del Lema 1.9, podemos reescribir el primer miembro de la igualdad (2.6) como

$$\inf_{\mathbf{u} \in C} \sum_{i=1}^{n} s_i u_i = \inf_{\mathbf{u} \in B} \sum_{i=1}^{n} s_i u_i,$$

que a su vez coincide, de acuerdo con el Lema 2.10, con

$$\inf_{\mathbf{u} \in B} \sum_{i=1}^{n} s_i u_i = \inf_{x \in X} \sum_{i=1}^{n} s_i (f_i(x) - f(x)). \tag{2.8}$$

Combinando (2.6),(2.7) y (2.8)

$$\begin{aligned}
0 &\leq \inf_{\mathbf{u} \in C} \max_{i=1,\dots,n} u_i \\
&= \inf_{\mathbf{u} \in C} \sum_{i=1}^{n} s_i u_i \\
&= \inf_{x \in X} \sum_{i=1}^{n} s_i (f_i(x) - f(x)),
\end{aligned}$$

y en consecuencia,

$$0 \leq \inf_{x \in X} \sum_{i=1}^{n} s_i(f_i(x) - f(x)),$$

que es lo que había que demostrar.

$\blacksquare$

Podría pensarse que la hipótesis de infsup-convexidad debe exigirse sobre la familia $f_1, ..., f_n$; sin embargo, este extremo no es cierto:

Ejemplo 2.12 Sean $f_1, f_2 : [-1, 1] \longrightarrow \mathbb{R}$ las funciones definidas por

$$f_1(x) := \begin{cases} 2x, & \text{si } -1 \leq x \leq 0 \\ x, & \text{si } 0 < x \leq 1 \end{cases}$$

y

$$f_2(x) := \frac{x}{2}, \qquad (-1 \leq x \leq 1).$$

La familia f_1, f_2 es infsup-convexa en $[-1, 1]$, pues

$$\inf_{x \in [-1,1]} \max_{i=1,2} f_i(x) = -\frac{1}{2}$$

y para cada $m \geq 1$, $\mathbf{t} \in \Delta_m$ y $x_1, \ldots, x_m \in [-1, 1]$ se tiene que

$$\begin{aligned} -\frac{1}{2} &\leq \sum_{j=1}^{m} t_j f_2(x_j) \\ &\leq \max_{i=1,2} \sum_{j=1}^{m} f_i(x_j). \end{aligned}$$

Consideremos la función $f : [-1, 1] \longrightarrow \mathbb{R}$ definida por

$$f(x) := x, \qquad (x \in [-1, 1]).$$

Se tiene que $f \leq \max\{f_1, f_2\}$ en $[-1, 1]$, pero no es posible encontrar un $\mathbf{s} \in \Delta_n$ tal que

$$f \leq s_1 f_1 + (1 - s_1) f_2$$

en el intervalo $[-1, 1]$. Supongamos, por contra, la existencia de un tal $\mathbf{s} \in \Delta_n$. La desigualdad anterior en $[-1, 0]$ obliga a que $0 \leq s_1 \leq 1/3$, mientras que en $[0, 1]$ implica $s_1 = 1$, lo cual es contradictorio.

$\blacksquare$

Concluimos esta sección comprobando que el Teorema 2.11 –y más concretamente una reformulación equivalente suya– es óptima en términos de la hipótesis de infsup-convexidad. Para ello, necesitamos previamente el siguiente resultado elemental:

Lema 2.13 *Si X es un conjunto no vacío, $n \geq 1$, $\mathbf{s} \in \Delta_n$ y $f_1, ..., f_n : X \longrightarrow \mathbb{R}$ son funciones dadas, entonces*

$$\min_{j=1,...,m} \sum_{i=1}^{n} s_i f_i(x_j) \leq \max_{i=1,...,n} \sum_{j=1}^{m} t_j f_i(x_j),$$

para cualesquiera $m \geq 1$, $\mathbf{t} \in \Delta_m$ y $x_1, ..., x_m \in X$.

Demostración. Basta con tener en cuenta la siguiente cadena de desigualdades:

$$\begin{aligned}
\min_{j=1,...,m} \sum_{i=1}^{n} s_i f_i(x_j) &\leq \sum_{j=1}^{m} t_j \left(\sum_{i=1}^{n} s_i f_i(x_j) \right) \\
&= \sum_{i=1}^{n} s_i \left(\sum_{j=1}^{m} t_j f_i(x_j) \right) \\
&\leq \max_{i=1,...,n} \sum_{j=1}^{m} t_j f_i(x_j).
\end{aligned}$$

$\blacksquare$

Enunciamos ahora el resultado principal de esta sección, que es la versión óptima anunciada del teorema del máximo de König.

Teorema 2.14 *Sean X un conjunto no vacío, $n \geq 1$ y $f, f_1, \cdots, f_n : X \longrightarrow \mathbb{R}$ funciones dadas. Entonces la familia $f_1 - f, \cdots, f_n - f$ es infsup-convexa en X si, y solo si, para todo número real α, la validez de la desigualdad*

$$f + \alpha \leq \max_{i=1,\ldots,n} f_i \quad \text{en } X$$

implica la existencia de un $\mathbf{s} \in \Delta_n$ tal que

$$f + \alpha \leq \sum_{i=1}^{n} s_i f_i \quad \text{en } X.$$

Demostración. $\boxed{\Rightarrow}$ Sea $\alpha \in \mathbb{R}$. Supongamos que $f_1 - f, ..., f_n - f$ es una familia infsup-convexa en X. Definimos $\tilde{f} := f + \alpha$, con lo que la familia $f_1 - \tilde{f}, ..., f_n - \tilde{f}$ es infsup-convexa en X. Como para todo $x \in X$ es cierto que

$$\tilde{f}(x) \leq \max_{i=1,\ldots,n} f_i(x),$$

estamos en condiciones de aplicar el Teorema 2.11, asegurando así la existencia de $\mathbf{s} \in \Delta_n$ tal que para todo $x \in X$ se cumple

$$\tilde{f}(x) \leq \sum_{i=1}^{n} s_i f_i(x),$$

esto es,

$$x \in X \;\Rightarrow\; f(x) + \alpha \leq \sum_{i=1}^{n} s_i f_i(x).$$

$\boxed{\Leftarrow}$ Sea $\alpha := \inf\limits_{x\in X} \max\limits_{i=1,\ldots,n} (f_i(x) - f(x))$. Si $\alpha = -\infty$ no hay nada que probar. En caso contrario, como

$$\alpha \leq \max_{i=1,\ldots,n} (f_i - f) \ \text{ en } \ X,$$

entonces existe $\mathbf{s} \in \Delta_n$ tal que para cada x en X se tiene

$$f(x) + \alpha \leq \sum_{i=1}^{n} s_i f_i(x),$$

es decir,

$$x \in X \ \Rightarrow \ \alpha \leq \inf_{x\in X} \left(\sum_{i=1}^{n} s_i \left(f_i(x) - f(x)\right) \right). \tag{2.9}$$

Dados entonces cualesquiera $m \geq 1, \mathbf{t} \in \Delta_m$ y $x_1, ..., x_m \in X$, de acuerdo con el Lema 2.13 tenemos que

$$\min_{j=1,\ldots,m} \sum_{i=1}^{n} s_i(f_i(x_j) - f(x_j)) \leq \max_{i=1,\ldots,n} \sum_{j=1}^{m} t_j(f_i(x_j) - f(x_j)), \tag{2.10}$$

y como

$$\inf_{x\in X} \sum_{i=1}^{n} s_i(f_i(x) - f(x)) \leq \min_{j=1,\ldots,m} \sum_{i=1}^{n} s_i(f_i(x_j) - f(x_j)),$$

teniendo en cuenta (2.9) y (2.10), para todo $x \in X$ se verifica

$$\begin{aligned} \alpha &= \inf_{x\in X} \max_{i=1,\ldots,n} (f_i(x) - f(x)) \\ &\leq \max_{i=1,\ldots,n} \sum_{j=1}^{m} t_j(f_i(x_j) - f(x_j)), \end{aligned}$$

y por tanto la familia $f_1 - f, ..., f_n - f$ es infsup-convexa en X.

$\blacksquare$

Señalemos para concluir que no podemos prescindir de la variabilidad de α en el Teorema 2.14:

Ejemplo 2.15 Sean $X = [-1,1]$ y $f, f_1, f_2 : X \longrightarrow \mathbb{R}$ las funciones definidas por

$$f(x) := 0, \quad (x \in [-1,1]),$$

$$f_1(x) := -x + 1, \quad (x \in [-1,1])$$

y

$$f_2(x) := \begin{cases} x+1, & \text{si } -1 \leq x \leq 0 \\ 1, & \text{si } 0 < x \leq 1. \end{cases}$$

Como $f \leq \text{mín}\{f_1, f_2\}$ en $[-1,1]$, entonces obviamente $f \leq \text{máx}\{f_1, f_2\}$ en $[-1,1]$ y además, para todo $\mathbf{s} \in \Delta_n$ y $x \in [-1,1]$ se verifica

$$f(x) \leq \sum_{i=1}^{n} s_i f_i(x).$$

Pero la familia $f_1 - f = f_1$ y $f_2 - f = f_2$ no es infsup-convexa en $[-1,1]$ ya que

$$\inf_{x \in [-1,1]} \max_{i=1,2} f_i = 1,$$

y tomando $m = 2$, $x_1 = -1, x_2 = 1$, $s_1 = 0,4$ y $s_2 = 0,6$ se tiene que

$$\max_{i=1,2} \sum_{j=1}^{2} t_j f_i(x_j) = 0,8.$$

■

2.3. Aplicaciones a Programación no Lineal Semi-Infinita

En esta sección aplicaremos nuestra versión del teorema del máximo de König, Teorema 2.14, a problemas de programación no lineal semi-infinita con un número finito de restricciones tipo igualdad y desigualdad, para obtener generalizaciones de los teoremas de los multiplicadores de Lagrange, Karush–Kuhn–Tucker y Fritz John, cuya optimalidad depende de la infsup-convexidad de una conveniente familia de funciones, y suponen una extensión de los resultados clásicos.

Establecemos seguidamente algunos criterios de notación que usaremos a lo largo de esta sección. Sean $N, M \geq 0$, X un conjunto no vacío y $f, g_1, \ldots, g_N,\ h_1, \ldots, h_M : X \longrightarrow \mathbb{R}$ funciones dadas. En el caso de ser $N = 0$ (respectivamente, $M = 0$) entenderemos que no hay restricciones con desigualdades g (respectivamente, con las igualdades h) y que $\mathbb{R}^N$ o $\mathbb{R}^N_+$ (respectivamente, $\mathbb{R}^M$) desaparece del producto cartesiano donde se encuentre; y $\displaystyle\sum_{i=1}^{0} \cdots$ se entenderá como 0. Supongamos que el conjunto factible

$$X_0 := \{x \in X :\ g_1(x) \leq 0, ..., g_N(x) \leq 0, h_1(x) = 0, ..., h_M(x) = 0\}$$

es no vacío. Consideramos el problema de programación no lineal semi-infinita

$$\inf_{x \in X_0} f(x). \tag{2.11}$$

Los siguientes conceptos son bien conocidos, pero los recordamos por ser esenciales para exponer con precisión nuestros enunciados.

Definición 2.16 El *lagrangiano* asociado al problema de programación no lineal (2.11) es la función $\mathfrak{L} : X \times \mathbb{R}_+^N \times \mathbb{R}^M \longrightarrow \mathbb{R}$ que a cada $(x, \boldsymbol{\lambda}, \boldsymbol{\mu}) \in X \times \mathbb{R}_+^N \times \mathbb{R}^M$ le hace corresponder el escalar

$$\mathfrak{L}(x, \boldsymbol{\lambda}, \boldsymbol{\mu}) := f(x) + \sum_{i=1}^{N} \lambda_i g_i(x) + \sum_{j=1}^{M} \mu_j h_j(x).$$

Definición 2.17 Diremos que $(x^0, \boldsymbol{\lambda}^0, \boldsymbol{\mu}^0) \in X \times \mathbb{R}_+^N \times \mathbb{R}^M$ es un *punto de silla* para el lagrangiano $\mathfrak{L}$ asociado al problema de programación no lineal (2.11) si verifica que

$$\mathfrak{L}(x^0, \boldsymbol{\lambda}, \boldsymbol{\mu}) \leq \mathfrak{L}(x^0, \boldsymbol{\lambda}^0, \boldsymbol{\mu}^0) \leq \mathfrak{L}(x, \boldsymbol{\lambda}^0, \boldsymbol{\mu}^0) \tag{2.12}$$

para cualquier $(x, \boldsymbol{\lambda}, \boldsymbol{\mu}) \in X \times \mathbb{R}_+^N \times \mathbb{R}^M$.

El siguiente resultado estándar se incluye meramente por completitud expositiva.

Lema 2.18 *Sea $(x^0, \boldsymbol{\lambda}^0, \boldsymbol{\mu}^0) \in X \times \mathbb{R}_+^N \times \mathbb{R}^M$ un punto de silla del lagrangiano $\mathfrak{L}$ asociado al problema de programación no lineal (2.11). Entonces x^0 es una solución óptima para (2.11) y*

$$f(x^0) = \mathfrak{L}(x^0, \boldsymbol{\lambda}^0, \boldsymbol{\mu}^0).$$

Demostración. Como $(x^0, \boldsymbol{\lambda}^0, \boldsymbol{\mu}^0)$ es un punto de silla, en particular

$$(\boldsymbol{\lambda}, \boldsymbol{\mu}) \in \mathbb{R}_+^N \times \mathbb{R}^M \;\Rightarrow\; \mathfrak{L}(x^0, \boldsymbol{\lambda}, \boldsymbol{\mu}) \leq \mathfrak{L}(x^0, \boldsymbol{\lambda}^0, \boldsymbol{\mu}^0),$$

es decir, para todo $(\boldsymbol{\lambda}, \boldsymbol{\mu}) \in \mathbb{R}_+^N \times \mathbb{R}^M$ tenemos

$$f(x^0) + \sum_{i=1}^{N} \lambda_i g_i(x^0) + \sum_{j=1}^{M} \mu_j h_j(x^0) \leq f(x^0) + \sum_{i=1}^{N} \lambda_i^0 g_i(x^0) + \sum_{j=1}^{M} \mu_j^0 h_j(x^0). \tag{2.13}$$

Tomando $\boldsymbol{\mu} = \boldsymbol{\mu}^0$ obtenemos, que para cualquier $\boldsymbol{\lambda} \in \mathbb{R}^N_+$,

$$\sum_{i=1}^{N} \lambda_i g_i(x^0) \leq \sum_{i=1}^{N} \lambda_i^0 g_i(x^0). \tag{2.14}$$

Si elegimos ahora $\boldsymbol{\lambda} = (\rho, 0, ..., 0)$ con $\rho > 0$ arbitrario, deducimos

$$g_1(x^0) \leq 0.$$

Repitiendo el proceso en cada variable concluimos que $g_2(x^0) \leq 0, ..., g_N(x^0) \leq 0$. Volvemos ahora a (2.13) y tomando esta vez $\boldsymbol{\lambda} = \boldsymbol{\lambda}^0$ tenemos que para todo $\boldsymbol{\mu} \in \mathbb{R}^M$,

$$\sum_{j=1}^{M} \mu_j h_j(x^0) \leq \sum_{j=1}^{M} \mu_j^0 h_j(x^0).$$

Razonando como antes, con $\boldsymbol{\mu} = (\rho, 0, ..., 0)$ y $\rho \in \mathbb{R}$ llegamos a $h_1(x^0) = 0, ..., h_M(x^0) = 0$. Así pues x^0 pertenece a X_0.

Por otra parte, haciendo $\boldsymbol{\lambda} \searrow (0, ..., 0)$ en (2.14) se tiene que

$$0 \leq \sum_{i=1}^{N} \lambda_i^0 g_i(x^0),$$

y en vista de que $\boldsymbol{\lambda^0} \in \mathbb{R}^N_+$ y $g_i(x^0) \leq 0$, concluimos que

$$\sum_{i=1}^{N} \lambda_i^0 g_i(x^0) = 0,$$

de donde

$$f(x^0) = \mathfrak{L}(x^0, \boldsymbol{\lambda}^0, \boldsymbol{\mu}^0).$$

Finalmente, como para todo $x \in X$,

$$\mathfrak{L}(x^0, \boldsymbol{\lambda}^0, \boldsymbol{\mu}^0) \leq \mathfrak{L}(x, \boldsymbol{\lambda}^0, \boldsymbol{\mu}^0),$$

entonces para cada x en X_0 se verifica

$$\begin{aligned} f(x^0) &\leq f(x) + \sum_{i=1}^{N} \lambda_i^0 g_i(x) + \sum_{j=1}^{M} \mu_j^0 h_j(x) \\ &\leq f(x), \end{aligned}$$

luego,

$$f(x^0) = \inf_{x \in X_0} f(x).$$

■

Los teoremas de los multiplicadores de Lagrange, de Karush–Kuhn–Tucker y de Fritz John tienen como objeto transformar el problema de optimizar una función restringida a un dominio en otro donde la función objetivo no está sometida a ningún tipo de restricción. Para ello, se hace uso de diferentes vectores, como por ejemplo los multiplicadores de Lagrange o los de Karush–Kuhn–Tucker. Estos últimos se definen de la siguiente manera:

Definición 2.19 Un par $(\boldsymbol{\lambda}^0, \boldsymbol{\mu}^0) \in \mathbb{R}_+^N \times \mathbb{R}^M$ es un *vector Karush–Kuhn–Tucker* para el problema de programación no lineal (2.11) si verifica

$$\inf_{x \in X} \mathfrak{L}(x, \boldsymbol{\lambda}^0, \boldsymbol{\mu}^0) = \inf_{x \in X_0} f(x).$$

Como acabamos de probar en el Lema 2.18, dado un punto de silla tenemos una solución para el problema de programación no lineal (2.11). El siguiente resultado, [46, Theorem 28.3.], que se prueba sin dificultad, pone de manifiesto que bajo la existencia de vectores de tipo Karush–Kuhn–Tucker se satisface la implicación contraria.

Proposición 2.20 *Sea $\mathfrak{L}$ el lagrangiano del problema de programación no lineal (2.11). Entonces $(x^0, \boldsymbol{\lambda}^0, \boldsymbol{\mu}^0) \in X \times \mathbb{R}_+^N \times \mathbb{R}^M$ es un punto de silla de $\mathfrak{L}$ si, y solo si,*

$$\begin{cases} x^0 \text{ es una solución del problema (2.11)} \\ y \\ (\boldsymbol{\lambda}^0, \boldsymbol{\mu}^0) \text{ es un vector Karush–Kuhn–Tucker para el problema (2.11).} \end{cases}$$

Además, en caso afirmativo,

$$\inf_{x \in X_0} f(x) = f(x^0) = \mathfrak{L}(x^0, \boldsymbol{\lambda}^0, \boldsymbol{\mu}^0) = \inf_{x \in X} \mathfrak{L}(x, \boldsymbol{\lambda}^0, \boldsymbol{\mu}^0).$$

Nos disponemos a continuación a establecer el resultado principal de esta sección. Reemplazamos la existencia de un vector Karush–Kuhn–Tucker, que no es una hipótesis útil desde un punto de vista práctico, por otra comprobable en gran parte de los problemas concretos de forma automática: la condición de Slater. Para ello necesitaremos del siguiente par de resultados técnicos:

Lema 2.21 *Dados X un conjunto no vacío, $N \geq 1$, $f, g_1, ..., g_N, h_1, ..., h_M : X \longrightarrow \mathbb{R}$ unas funciones tales que el conjunto*

$$X_0 = \{x \in X : g_1(x) \leq 0, ..., g_N(x), h_1(x) = 0, ..., h_M(x) = 0\}$$

es no vacío y $x^0 \in X$ una solución del problema (2.11), entonces

$$\inf_{x \in X} \max\{g_1(x), ..., g_N(x), \pm h_1(x), ..., \pm h_M(x), f(x) - f(x^0)\} = 0.$$

Demostración. Basta reescribir X_0 como

$$X_0 = \{x \in X : g_1(x) \leq 0, ..., g_N(x) \leq 0, \pm h_1(x) \leq 0, ..., \pm h_M(x) \leq 0\},$$

y aplicar [49, Lemma 3.1].

■

Lema 2.22 *Si X es un conjunto no vacío, n un natural, $\mathbf{s} \in \Delta_n$, y $\varphi_1, ..., \varphi_n$ es una familia de funciones reales definidas en X tales que*

$$\inf_{x \in X} \max_{i=1,...,n} \varphi_i(x) \leq \inf_{x \in X} \sum_{i=1}^{n} s_i \varphi_i(x),$$

entonces la familia $(\varphi_i)_{i=1}^n$ es infsup-convexa en X.

Demostración. Para cualesquiera $m \geq 1$, $\mathbf{t} \in \Delta_m$ y $x_1, ..., x_m \in X$, a la luz del Lema 2.13 tenemos que

$$\begin{aligned} \inf_{x \in X} \max_{i=1,...,n} \varphi_i(x) &\leq \inf_{x \in X} \sum_{i=1}^{n} s_i \varphi_i(x) \\ &\leq \min_{j=1,...,m} \sum_{i=1}^{n} s_i \varphi_i(x_j) \\ &\leq \max_{i=1,...,n} \sum_{j=1}^{m} t_j \varphi_i(x_j), \end{aligned}$$

y en consecuencia, la familia $(\varphi_i)_{i=1}^n$ es infsup-convexa en X como habíamos anunciado.

■

El siguiente resultado supone una extensión de [49, Proposition 3.2].

Proposición 2.23 *Sean X un conjunto no vacío, $N, M \geq 0$ y f, $g_1, ..., g_N$, $h_1, ..., h_M : X \longrightarrow \mathbb{R}$ funciones dadas. Supongamos que el conjunto*

$$X_0 = \{x \in X : \ g_1(x) \leq 0, ..., g_N(x) \leq 0, h_1(x) = 0, ..., h_M(x) = 0\}$$

es no vacío y que el lagrangiano del problema de programación no lineal (2.11) admite un punto de silla $(x^0, \boldsymbol{\lambda}^0, \boldsymbol{\mu}^0) \in X \times \mathbb{R}_+^N \times \mathbb{R}^M$. *Entonces la familia* $g_1, ..., g_N, \pm h_1, ..., \pm h_M, f - f(x^0)$ *es infsup-convexa en* X.

Demostración. Por un lado, de acuerdo con el Lema 2.21 ,

$$\inf_{x \in X} \max\{g_1(x), ..., g_N(x), \pm h_1(x), ..., \pm h_M(x), f(x) - f(x^0)\} = 0. \tag{2.15}$$

Por otro, como $(x^0, \boldsymbol{\lambda}^0, \boldsymbol{\mu}^0)$ es un punto de silla del lagrangiano, entonces todo $x \in X$

$$f(x^0) = \mathfrak{L}(x^0, \boldsymbol{\lambda}^0, \boldsymbol{\mu}^0) \leq \mathfrak{L}(x, \boldsymbol{\lambda}^0, \boldsymbol{\mu}^0),$$

y en consecuencia,

$$0 \leq \inf_{x \in X} \left(\sum_{i=1}^{N} \lambda_i^0 g_i(x) + \sum_{j=1}^{M} \mu_j^0 h_j(x) + (f(x) - f(x^0)) \right),$$

equivalentemente

$$0 \leq \inf_{x \in X} \left(\sum_{i=1}^{N} \lambda_i^0 g_i(x) + \sum_{j=1}^{M} (\mu_j^0)_+ h_j(x) + \sum_{j=1}^{M} (\mu_j^0)_- (-h_j(x)) + (f(x) - f(x^0)) \right), \tag{2.16}$$

donde

$$(\mu_j^0)_+ := \max\{\mu_j^0, 0\} \quad \text{y} \quad (\mu_j^0)_- := \max\{-\mu_j^0, 0\}.$$

Para concluir, basta definir $\mathbf{s} \in \Delta_{N+2M+1}$ como

$$\mathbf{s} := \frac{1}{\rho}(\lambda_1^0, ..., \lambda_N^0, (\lambda_1^0)_+, ..., (\mu_M^0)_+, (\mu_1^0)_-, ..., (\mu_M^0)_-, 1)$$

siendo

$$\rho := 1 + \sum_{i=1}^{N} \lambda_i^0 + \sum_{j=1}^{M} ((\mu_j^0)_+ + (\mu_j^0)_-),$$

ya que en vista de (2.15), (2.16) y el Lema 2.22, la familia $g_1, ..., g_N, \pm h_1, ..., \pm h_M, f - f(x^0)$ es infsup-convexa en X.

■

Enunciemos el resultado anunciado, una extensión óptima del teorema de los multiplicadores de Lagrange, que generaliza el resultado clásico ([46, Corollary 28.3.1] y [57, Theorem 2]). En él mostramos que, bajo cierta condición de Slater, la infsup-convexidad nos permite establecer una dualidad perfecta entre los puntos de silla y las soluciones del problema de programación no lineal (2.11).

Teorema 2.24 *Sean X un conjunto no vacío y $f, g_1, ..., g_N, h_1, ..., h_M$ funciones reales definidas en X de forma que el conjunto factible*

$$X_0 = \{x \in X : \ g_1(x) \leq 0, ..., g_N(x) \leq 0, h_1(x) = 0, ..., h_M(x) = 0\}$$

es no vacío. Supongamos además que $x^0 \in X$ es una solución del problema de programación semi-infinita (2.11) y se cumple la siguiente condición de Slater: existe $x^1 \in X$ tal que

$$g_1(x^1) < 0, ..., g_N(x^1) < 0, h_1(x^1) = 0, ..., h_M(x^1) = 0, \tag{2.17}$$

y dados $\boldsymbol{\alpha}, \boldsymbol{\beta} \in \mathbb{R}^M_+$ alguno distinto de cero, encontramos $x^2 \in X$ de forma que

$$\sum_{j=1}^{M} \alpha_j h_j(x^2) < \sum_{j=1}^{M} \beta_j h_j(x^2). \tag{2.18}$$

Entonces existe $(\boldsymbol{\lambda}^0, \boldsymbol{\mu}^0) \in \mathbb{R}^N_+ \times \mathbb{R}^M$ tal que $(x^0, \boldsymbol{\lambda}^0, \boldsymbol{\mu}^0)$ es un punto de silla del lagrangiano asociado a (2.11) si, y solo si, la familia $g_1, ..., g_N, \pm h_1, ..., \pm h_M$, $f - f(x^0)$ es infsup-convexa en X.

Demostración. Una de las direcciones es obvia gracias a la Proposición 2.23. Para la otra, dado que x^0 es solución del problema (2.11), podemos aplicar el Lema 2.21, deduciendo que

$$\inf_{x\in X} \max\{g_1(x), ..., g_N(x), \pm h_1(x), ..., \pm h_M(x), f(x) - f(x^0)\} = 0.$$

Definimos la función $\tilde{f} : X \longrightarrow \mathbb{R}$ como

$$\tilde{f}(x) := 0, \quad (x \in X),$$

y en virtud del Teorema 2.14, existe $\mathbf{s} \in \Delta_{N+2M+1}$ tal que para todo $x \in X$

$$\begin{aligned} 0 \quad &\leq \quad \sum_{i=1}^{N} s_i g_i(x) + \sum_{i=N+1}^{N+M} s_i^{'} h_{i-N}(x) \\ &+ \quad \sum_{i=N+M+1}^{N+2M} s_i^{''}(-h_{i-N-M}(x)) + s_{N+2M+1}(f(x) - f(x^0)) \\ &= \quad \sum_{i=1}^{N} s_i g_i(x) + \sum_{i=N+1}^{N+2M} \tilde{s}_i h_{i-N}(x) + s_{N+2M+1}(f(x) - f(x^0)), \end{aligned}$$

donde hemos renombrado $s_i^{'} - s_i^{''}$ como $\tilde{s}_i$. Sabemos que $s_{N+2M+1} > 0$, ya que si $s_{N+2M+1} = 0$, por la condición de Slater (2.17) y (2.18) tendríamos que $\mathbf{s} = 0$, lo cual es absurdo. Tomando

$$\lambda_i^0 := \frac{s_i}{s_{N+2M+1}} \quad \text{y} \quad \mu_i^0 := \frac{\tilde{s}_i}{s_{N+2M+1}},$$

queda que para cada $x \in X$

$$f(x^0) \leq f(x) + \sum_{i=1}^{N} \lambda_i^0 g_i(x) + \sum_{i=N+1}^{N+M} \mu_i^0 h_i(x), \tag{2.19}$$

y eligiendo $x = x^0$,

$$\sum_{i=1}^{N} \lambda_i^0 g_i(x^0) = 0. \tag{2.20}$$

Por tanto, notando por $\mathfrak{L}$ el lagrangiano asociado al programa (2.11), dado $(\boldsymbol{\lambda}, \boldsymbol{\mu}) \in \mathbb{R}_+^N \times \mathbb{R}^M$ se tiene que

$$\begin{aligned}
\mathfrak{L}(x^0, \boldsymbol{\lambda}, \boldsymbol{\mu}) &= f(x^0) + \sum_{i=1}^{N} \lambda_i g_i(x^0) + \sum_{j=1}^{M} \mu_j h_j(x^0) \\
&\leq f(x^0) \\
&= \mathfrak{L}(x^0, \boldsymbol{\lambda}^0, \boldsymbol{\mu}^0),
\end{aligned}$$

y como además, por (2.19) y (2.20), para todo x en X

$$\begin{aligned}
\mathfrak{L}(x^0, \boldsymbol{\lambda}^0, \boldsymbol{\mu}^0) &= f(x^0) \\
&\leq \mathfrak{L}(x, \boldsymbol{\lambda}^0, \boldsymbol{\mu}^0),
\end{aligned}$$

entonces $(x^0, \boldsymbol{\lambda}^0, \boldsymbol{\mu}^0)$ es un punto de silla para $\mathfrak{L}$.

■

Observemos que las restricciones tipo igualdad han sido tratadas como desigualdades de una función y su opuesta. De igual manera, podríamos haber expresado las restricciones de este tipo como una desigualdad del valor absoluto de las mismas.

Merece la pena mencionar que los teoremas de Karush–Kuhn–Tucker (véase [31, 36] y, por ejemplo, [4]) y de Fritz John ([30, 27, 16]) para el programa semi-infinito (2.11) pueden obtenerse a partir del Teorema 2.14. Para el primero es fácil demostrar que $(x^0, \boldsymbol{\lambda}^0, \boldsymbol{\mu}^0) \in X \times \mathbb{R}_+^N \times \mathbb{R}^M$ es un punto de silla del lagrangiano $\mathfrak{L}$ del problema de programación no lineal (2.11) si, y solo si, bajo cierta hipótesis tipo Slater se cumplen unas adecuadas condiciones de Karush–Kuhn–Tucker. En consecuencia, podemos obtener una versión más general del teorema de Karush–Kuhn–Tucker –el caso $M = 0$ está demostrado en [49, Corollary 4.1]– que es equivalente al Teorema 2.24.

Teorema 2.25 *Sean X un conjunto no vacío, $x^0 \in X$, $N, M \geq 0$ y f, $g_1, \ldots, g_N$, $h_1, \ldots, h_M : X \longrightarrow \mathbb{R}$ unas funciones tales que el conjunto*

$$X_0 = \{x \in X : \ g_1(x) \leq 0, \ldots, g_N(x) \leq 0, h_1(x) = 0, \ldots, h_M(x) = 0\}$$

es no vacío. Supongamos que la familia $g_1, \ldots, g_N, \pm h_1, \ldots, \pm h_M, f - f(x^0)$ es infsup-convexa en X, y que además se cumple la siguiente condición de Slater: existe x^1 en X de forma que

$$g_1(x^1) < 0, \ldots, g_N(x^1) < 0, h_1(x^1) = 0, \ldots, h_M(x^1) = 0,$$

y para cualesquiera $\boldsymbol{\alpha}, \boldsymbol{\beta} \in \mathbb{R}^M_+$ alguno no nulo, podemos encontrar un $x^2 \in X$ tal que

$$\sum_{j=1}^{M} \alpha_j h_j(x^2) < \sum_{j=1}^{M} \beta_j h_j(x^2).$$

Entonces x^0 es solución del problema de programación no lineal (2.11) si, y solo si, existe $(\boldsymbol{\lambda^0}, \boldsymbol{\mu^0}) \in \mathbb{R}^N_+ \times \mathbb{R}^M$ verificando las siguientes condiciones tipo Karush–Kuhn–Tucker:

$$f + \sum_{i=1}^{N} \lambda_i^0 g_i + \sum_{j=1}^{M} \mu_j^0 h_j \text{ alcanza su ínfimo en } X \text{ en } x^0,$$

$$i = 1, \ldots, N \ \Rightarrow \ g_i(x^0) \leq 0 \quad y \quad \lambda_i^0 g_i(x^0) = 0$$

y

$$j = 1, \ldots, M \ \Rightarrow \ h_j(x^0) = 0.$$

Demostración. Tal y como acabamos de mencionar, a la luz del Teorema 2.24, la demostración se reduce a probar que la existencia de un punto de silla para el lagrangiano $\mathfrak{L}$ del programa (2.11) equivale a que las condiciones de Karush–Kuhn–Tucker se satisfacen. Supongamos en primer lugar que existe $(\boldsymbol{\lambda^0}, \boldsymbol{\mu^0}) \in \mathbb{R}^N_+ \times \mathbb{R}^M$ tal que $(x^0, \boldsymbol{\lambda^0}, \boldsymbol{\mu^0})$ es un punto de silla. En particular,

$$x \in X \ \Rightarrow \ \mathfrak{L}(x^0, \boldsymbol{\lambda}^0, \boldsymbol{\mu}^0) \leq \mathfrak{L}(x, \boldsymbol{\lambda}^0, \boldsymbol{\mu}^0),$$

esto es, para cada $x \in X$,

$$f(x^0) + \sum_{i=1}^{N} \lambda_i^0 g_i(x^0) + \sum_{j=1}^{M} \mu_j^0 h_j(x^0) \leq f(x) + \sum_{i=1}^{N} \lambda_i^0 g_i(x) + \sum_{j=1}^{M} \mu_j^0 h_j(x),$$

o dicho con otras palabras,

$$f + \sum_{i=1}^{N} \lambda_i^0 g_i + \sum_{j=1}^{M} \mu_j^0 h_j$$

alcanza su ínfimo en x^0. Usando nuevamente el hecho de que $(\boldsymbol{\lambda^0}, \boldsymbol{\mu^0})$ es un punto de silla de $\mathfrak{L}$, y en concreto,

$$(\boldsymbol{\lambda}, \boldsymbol{\mu}) \in \mathbb{R}_+^N \times \mathbb{R}^M \quad \Rightarrow \quad \mathfrak{L}(x^0, \boldsymbol{\lambda}, \boldsymbol{\mu}) \leq \mathfrak{L}(x^0, \boldsymbol{\lambda}^0, \boldsymbol{\mu}^0),$$

tenemos que para todo $(\boldsymbol{\lambda}, \boldsymbol{\mu}) \in \mathbb{R}_+^N \times \mathbb{R}^M$

$$\sum_{i=1}^{N} \lambda_i g_i(x^0) + \sum_{j=1}^{M} \mu_j h_j(x^0) \leq \sum_{i=1}^{N} \lambda_i^0 g_i(x^0) + \sum_{j=1}^{M} \mu_j^0 h_j(x^0).$$

Procediendo como en la prueba del Lema 2.18, deducimos de forma inmediata las restantes condiciones de Karush–Kuhn–Tucker:

$$g_i(x^0) \leq 0 \quad \text{y} \quad \lambda_i^0 g_i(x^0) = 0$$

para cada $i = 1, ..., N$ y

$$h_j(x^0) = 0$$

con $j = 1, ..., M$.

Para la otra implicación, puesto que la función

$$f + \sum_{i=1}^{N} \lambda_i^0 g_i + \sum_{j=1}^{M} \mu_j^0 h_j$$

alcanza su ínfimo en x^0, tenemos que para cada x en X

$$f(x^0)+\sum_{i=1}^{N}\lambda_i^0 g_i(x^0)+\sum_{j=1}^{M}\mu_j^0 h_j(x^0) \leq f(x)+\sum_{i=1}^{N}\lambda_i^0 g_i(x)+\sum_{j=1}^{M}\mu_j^0 h_j(x). \tag{2.21}$$

Gracias a las restantes condiciones de Karush–Kuhn–Tucker podemos concluir que para todo $(\boldsymbol{\lambda},\boldsymbol{\mu}) \in \mathbb{R}_+^N \times \mathbb{R}^M$ se cumple

$$\begin{aligned} f(x^0)+\sum_{i=1}^{N}\lambda_i g_i(x^0)+\sum_{j=1}^{M}\mu_j h_j(x^0) &\leq f(x^0) \\ &= f(x^0)+\sum_{i=1}^{N}\lambda_i^0 g_i(x^0)+\sum_{j=1}^{M}\mu_j^0 h_j(x^0), \end{aligned}$$

lo que junto a (2.21) prueba que $(x^0,\boldsymbol{\lambda^0},\boldsymbol{\mu^0})$ es un punto de silla.

■

Al igual que el teorema tipo Karush–Kuhn–Tucker que acabamos de establecer, nos disponemos a probar otro resultado óptimo en términos de infsup-convexidad, un teorema de Fritz John, esta vez sin exigir condición de Slater alguna:

Teorema 2.26 *Sean X un conjunto no vacío, $M, N \geq 0$, f, $g_1,\ldots,g_N$, $h_1,\ldots,h_M : X \longrightarrow \mathbb{R}$ funciones tales que el conjunto*

$$X_0 = \{x \in X : \; g_1(x) \leq 0,\ldots,g_N(x) \leq 0, h_1(x) = 0,\ldots,h_M(x) = 0\}$$

es no vacío y $x^0 \in X$ una solución del programa semi-infinito (2.11). Entonces la familia $f - f(x^0), g_1,...,g_N, \pm h_1,...,\pm h_M$ es infsup-convexa en X si, solo si, existen $y \geq 0$, $\boldsymbol{\lambda^0} \in \mathbb{R}_+^N$, $\mathbf{u^0},\mathbf{v^0} \in \mathbb{R}_+^M$, con

$$y+\sum_{i=1}^{N}\lambda_i^0+\sum_{j=1}^{M}(u_j^0+v_j^0) \neq 0,$$

verificando las siguientes condiciones Fritz John:

$$yf + \sum_{i=1}^{N} \lambda_i^0 g_i + \sum_{j=1}^{M} (u_j^0 - v_j^0) h_j \text{ alcanza su ínfimo en } X \text{ en } x^0 \tag{2.22}$$

y

$$i = 1, \ldots, N \quad \Rightarrow \ \lambda_i^0 g_i(x^0) = 0. \tag{2.23}$$

Demostración. Supongamos que $f - f(x^0), g_1, \ldots, g_N, \pm h_1, \ldots, \pm h_M$ es una familia infsup-convexa en X. Como x^0 es solución del problema de programación no lineal (2.11), el Lema 2.21 nos asegura que

$$\inf_{x \in X} \max\{f(x) - f(x^0), g_1(x), \ldots, g_N(x), \pm h_1(x), \ldots, \pm h_M(x)\} = 0,$$

y aplicando el Teorema 2.14, como en la demostración del Teorema 2.24, garantizamos la existencia de un $\mathbf{s} \in \Delta_{1+N+2M}$ tal que

$$\begin{aligned} &0 \le \inf_{x \in X} \langle \mathbf{s}, \\ &\quad (f(x) - f(x^0), g_1(x), \ldots, g_N(x), h_1(x), \ldots, h_M(x), -h_1(x), \ldots, -h_M(x)) \rangle, \end{aligned} \tag{2.24}$$

donde $< \cdot, \cdot >$ denota el producto escalar usual en $\mathbb{R}^{1+N+2M}$. Comprobemos que se cumplen las condiciones de Fritz John (2.22) y (2.23), sin más que elegir

$$y := s_1, \quad \boldsymbol{\lambda} := (s_2, \ldots, s_{N+1})$$

y

$$\mathbf{u^0} := (s_{N+2}, \ldots, s_{N+M+1}), \quad \mathbf{v^0} := (s_{N+M+2}, \ldots, s_{1+N+2M}).$$

En efecto: claramente $y \ge 0$, $\boldsymbol{\lambda^0} \in \mathbb{R}_+^N$ y $\mathbf{u^0}, \mathbf{v^0} \in \mathbb{R}_+^M$, y como consecuencia de (2.24),

$$x \in X \ \Rightarrow \ yf(x^0) \le yf(x) + \sum_{i=1}^{N} \lambda_i^0 g_i(x) + \sum_{j=1}^{M} (u_j^0 - v_j^0) h_j(x). \tag{2.25}$$

Tomando $x = x^0$ en esta desigualdad, se tiene que

$$\sum_{i=1}^{N} \lambda_i^0 g_i(x^0) = 0, \tag{2.26}$$

y de (2.25) y (2.26) deducimos (2.22) y (2.23).

Y recíprocamente; como se verifican (2.22) y (2.23) para algún $y \geq 0$, $\boldsymbol{\lambda}^{\mathbf{0}} \in \mathbb{R}_+^N$ y $\mathbf{u}^{\mathbf{0}}, \mathbf{v}^{\mathbf{0}} \in \mathbb{R}_+^M$, podemos definir

$$\mathbf{s} := \left(\frac{y}{\rho}, \frac{\lambda_1^0}{\rho}, \ldots, \frac{\lambda_N^0}{\rho}, \frac{u_1^0}{\rho}, \ldots, \frac{u_M^0}{\rho}, \frac{v_1^0}{\rho}, \ldots, \frac{v_M^0}{\rho} \right) \in \Delta_{1+N+2M},$$

con

$$\rho := y + \sum_{i=1}^{N} \lambda_i^0 + \sum_{j=1}^{M} (u_j^0 + v_j^0) > 0,$$

con lo que

$$\begin{aligned} &0 \leq \operatorname{ínf}_{x \in X} \langle \mathbf{s}, \\ &\quad (f(x) - f(x^0), g_1(x), \ldots, g_N(x), h_1(x), \ldots, h_M(x), -h_1(x), \ldots, -h_M(x)) \rangle, \end{aligned}$$

Para concluir solo hay que observar que los Lemas 2.21 y 2.22 nos dan la infsup-convexidad de la familia $f - f(x^0), g_1, \ldots, g_N, \pm h_1, \ldots, \pm h_M$ en X.

■

Mencionemos para concluir que la caracterización dada en el Teorema 2.26 es una extensión de [49, Proposition 4.2].

Capítulo 3

Programas Infinitos sin Igualdades

Nuestro punto de partida en este capítulo es el teorema del supremo de König, una generalización del teorema del máximo para el caso de una familia arbitraria de funciones, y que es equivalente al teorema de Hahn–Banach. Al igual que con el teorema del máximo de König, obtenemos una versión óptima del teorema del supremo de König que depende de la infsup-convexidad de una conveniente familia de funciones. La aplicación de este resultado a programación infinita nos permite refinar los clásicos teoremas de los multiplicadores de Lagrange, Karush–Kuhn–Tucker y Fritz John para un número no necesariamente finito de restricciones tipo desigualdad.

3.1. Teorema del Supremo de König

En esta sección obtendremos una versión óptima del teorema del supremo de König, [34, Erweiterter Maximumssatz p. 501], que como acabamos de comentar es una reformulación equivalente del teorema de Hahn–Banach (los detalles pueden encontrarse en [41, Proposition 2.3]), y que a diferencia del teorema del máximo de König, se establece para un número arbitrario de funciones y en lugar del máximo de estas aparece su supremo. Con tal fin, nos ocupamos de recordar algunos conceptos y propiedades del espacio de Banach $\ell^\infty(\Lambda)$, que en este contexto más general desempeña el papel del espacio euclídeo finito dimensional en el teorema del máximo de König. De igual forma precisamos del concepto de infsup-convexidad para familias con un número arbitrario de funciones. Posteriormente presentamos un ejemplo de familia infinita de funciones infsup-convexa, que será importante para elegir la condición de Slater en este nuevo ambiente. En último lugar probamos una generalización del teorema del supremo de König, para dar posteriormente su versión óptima.

Tal y como acabamos de mencionar, recordamos algunos hechos elementales sobre el espacio de Banach $\ell^\infty(\Lambda)$.

Dado un conjunto Λ no vacío, diremos que la función $x \in \ell^\infty(\Lambda)$ es *positiva* si verifica $x(\lambda) \geq 0$ para todo $\lambda \in \Lambda$, y notaremos por $\ell^\infty(\Lambda)_+$ al cono de todas las funciones positivas de $\ell^\infty(\Lambda)$. Escribimos el dual topológico de $\ell^\infty(\Lambda)$ como $\ell^\infty(\Lambda)^*$, i.e.,

$$\ell^\infty(\Lambda)^* := \{L : \ell^\infty(\Lambda) \longrightarrow \mathbb{R} \ : \ L \text{ es lineal y continuo}\},$$

y por último, denotaremos por $\ell^\infty(\Lambda)^*_+$ el cono de los funcionales *positivos* de $\ell^\infty(\Lambda)^*$, esto es,

$$\ell^\infty(\Lambda)^*_+ := \{L \in \ell^\infty(\Lambda)^* : x \in \ell^\infty(\Lambda)_+ \Rightarrow L(x) \geq 0\}.$$

Recordemos que un funcional $L \in \ell^\infty(\Lambda)^*$ es *monótono* si

$$x \leq y \Rightarrow L(x) \leq L(y), \quad (x, y \in \ell^\infty(\Lambda)),$$

lo que equivale de forma inmediata a la positividad de dicho L.

En lo que sigue, $S : \ell^\infty(\Lambda) \longrightarrow \mathbb{R}$ es el funcional definido como

$$S(x) := \sup_{\lambda \in \Lambda} x(\lambda), \qquad (x \in \ell^\infty(\Lambda)), \tag{3.1}$$

que claramente es sublineal.

El siguiente resultado, que caracteriza de forma cómoda los funcionales lineales de $\ell^\infty(\Lambda)$ mayorados por el funcional sublineal S, es conocido, pero se incluye por completitud. Estos funcionales juegan el papel de los elementos del n-símplex de probabilidad, ahora en un contexto no necesariamente finito dimensional. Antes de probar el resultado, notaremos por $\mathbb{1}$ el elemento de $\ell^\infty(\Lambda)$ definido por

$$\mathbb{1}(\lambda) := 1, \qquad (\lambda \in \Lambda).$$

Lema 3.1 *Sean Λ un conjunto no vacío y $L : \ell^\infty(\Lambda) \longrightarrow \mathbb{R}$ un funcional lineal. Entonces*

$$L \leq S \Leftrightarrow \begin{cases} L(\mathbb{1}) = 1 \\ y \\ L \text{ es positivo.} \end{cases}$$

Demostración. $\boxed{\Leftarrow}$ Dado x en $\ell^\infty(\Lambda)$, como

$$0 \leq S(x)\mathbb{1} - x,$$

de la linealidad de L y las hipótesis se sigue que

$$\begin{aligned} 0 &\leq L(S(x)\mathbb{1} - x) \\ &= S(x)L(\mathbb{1}) - L(x) \\ &= S(x) - L(x). \end{aligned}$$

Así pues,

$$L(x) \leq S(x),$$

y la arbitrariedad de $x \in \ell^\infty(\Lambda)$ da $L \leq S$.

$\boxed{\Rightarrow}$ Empezaremos probando que $L(\mathbb{1}) = 1$. Para ello basta observar que

$$\begin{aligned} L(-\mathbb{1}) &\leq S(-\mathbb{1}) \\ &= \sup_{\lambda\in\Lambda} - \mathbb{1}(\lambda) \\ &= -1, \end{aligned} \tag{3.2}$$

mientras que

$$\begin{aligned} L(\mathbb{1}) &\leq S(\mathbb{1}) \\ &= \sup_{\lambda\in\Lambda} \mathbb{1}(\lambda) \\ &= 1. \end{aligned} \tag{3.3}$$

Por la linealidad de L, (3.2) y (3.3), se deduce que $L(\mathbb{1}) = 1$. Pasemos a comprobar que L es positivo. Es claro que, por ser $L \leq S$, entonces $L \in \ell^\infty(\Lambda)^*$. Además, si $x \in \ell^\infty(\Lambda)_+$, es decir,

$$\sup_{\lambda\in\Lambda}(-x)(\lambda) \leq 0,$$

luego, a luz de la hipótesis,

$$\begin{aligned} L(-x) &\leq S(-x) \\ &= \sup_{\lambda\in\Lambda}(-x)(\lambda) \\ &\leq 0, \end{aligned}$$

y como L es lineal, tenemos que $0 \leq L(x)$.

$\blacksquare$

Pasemos a definir el conjunto formado por los funcionales en $\ell^\infty(\Lambda)$ que, en vista del Lema 3.1, desempeñarán el papel de los elementos del n-símplex en dimensión arbitraria.

Definición 3.2 Dado Λ un conjunto no vacío, notamos por Δ_Λ el subconjunto de $\ell^\infty(\Lambda)^*$

$$\Delta_\Lambda := \{L \in \ell^\infty(\Lambda)^* : \ L \leq S\}.$$

Gracias al Lema 3.1, tenemos que

$$\Delta_\Lambda = \{L \in \ell^\infty(\Lambda)^* : \ L(\mathbb{1}) = 1 \ \text{ y L es positivo}\}.$$

En caso de que Λ sea finito, Δ_Λ coincide con el n-símplex de probabilidad. Concretamente, si el cardinal de Λ es n, se tiene que $\ell^\infty(\Lambda) \equiv \mathbb{R}^n$ y $\Delta_\Lambda = \Delta_n$.

Recogemos el concepto de infsup-convexidad dado en el capítulo anterior, ahora para un número arbitrario de funciones.

Definición 3.3 Sean X y Λ conjuntos no vacíos y $\{f_\lambda\}_{\lambda\in\Lambda}$ una familia de funciones reales definidas en X. La familia $\{f_\lambda\}_{\lambda\in\Lambda}$ es *infsup-convexa* en X si para todo $m \geq 1, \mathbf{t} \in \Delta_m$ y $x_1, ..., x_m \in X$ se verifica

$$\inf_{x\in X} \sup_{\lambda\in\Lambda} f_\lambda(x) \leq \sup_{\lambda\in\Lambda} \sum_{j=1}^m t_j f_\lambda(x_j).$$

A continuación ilustramos este concepto –conocido en la literatura también como *affine weakly convexlikeness*– con un ejemplo que usaremos más adelante.

Ejemplo 3.4 Sean $X = \mathbb{R}$, $\Lambda = \mathbb{N}$, $f : \mathbb{R} \longrightarrow \mathbb{R}$ la función definida por

$$f(x) := x, \qquad (x \in \mathbb{R}),$$

y $g_n : \mathbb{R} \longrightarrow \mathbb{R}$, con $n \geq 1$, las funciones dadas por

$$g_n(x) := -\frac{x^3}{n}, \qquad (x \in \mathbb{R}).$$

Para comprobar que la familia $\{g_n\}_{n\in\mathbb{N}}, f$ es infsup-convexa en $X = \mathbb{R}$ debemos probar que

$$\inf_{x\in\mathbb{R}} \left(f(x) \vee \sup_{n\geq 1} g_n(x) \right) \leq \sum_{j=1}^{m} t_j f(x_j) \vee \sup_{n\geq 1} \sum_{j=1}^{m} t_j g_n(x_j), \tag{3.4}$$

cualesquiera sean $m \geq 1, \mathbf{t} \in \Delta_m$ y $x_1, ..., x_m \in X$. Empecemos analizando el miembro izquierdo de la desigualdad (3.4):

$$\begin{aligned} \inf_{x\in\mathbb{R}} \left(f(x) \vee \sup_{n\geq 1} g_n(x) \right) &= \inf_{x\in\mathbb{R}_+} \left(x \vee \sup_{n\geq 1} -\frac{x^3}{n} \right) \wedge \inf_{x\in\mathbb{R}_-} \left(x \vee \sup_{n\geq 1} -\frac{x^3}{n} \right) \\ &= \left(\inf_{x\in\mathbb{R}_+} x \right) \wedge \left(\inf_{x\in\mathbb{R}_-} -x^3 \right) \\ &= 0 \wedge 0 \\ &= 0. \end{aligned}$$

Para el miembro derecho de (3.4),

$$\begin{aligned} \sum_{j=1}^{m} t_j f(x_j) \vee \sup_{n\geq 1} \sum_{j=1}^{m} t_j g_n(x_j) &= \sum_{j=1}^{m} t_j x_j \vee \sup_{n\geq 1} \left(\sum_{j=1}^{m} t_j x_j^3 \right) \left(-\frac{1}{n} \right) \\ &=: \alpha. \end{aligned}$$

Pero, si suponemos

$$\sum_{j=1}^{m} t_j x_j^3 \geq 0,$$

entonces

$$\sup_{n\geq 1} \left(\sum_{j=1}^{m} t_j x_j^3\right)\left(-\frac{1}{n}\right) = 0,$$

luego $\alpha \geq 0$. Si por el contrario

$$\sum_{j=1}^{m} t_j x_j^3 < 0,$$

se tiene que

$$\sup_{n\geq 1} \left(\sum_{j=1}^{m} t_j x_j^3\right)\left(-\frac{1}{n}\right) \geq 0.$$

En ambos casos $\alpha \geq 0$, y por tanto la familia es infsup-convexa en $\mathbb{R}$.

■

Enunciamos seguidamente una primera versión del teorema del supremo de König, que constituye una generalización del Teorema 2.11.

Teorema 3.5 *Supongamos que X y Λ son conjuntos no vacíos, $f : X \longrightarrow \mathbb{R}$ es una función y $\{f_\lambda\}_{\lambda\in\Lambda}$ es una familia de funciones reales definidas en X de forma que $\{f_\lambda - f\}_{\lambda\in\Lambda}$ es una familia infsup-convexa en X, $\{f_\lambda(x)\}_{\lambda\in\Lambda} \in \ell^\infty(\Lambda)$ para cada x en X y*

$$x \in X \;\Rightarrow\; f(x) \leq \sup_{\lambda\in\Lambda} f_\lambda(x).$$

Entonces existe $L \in \Delta_\Lambda$ tal que

$$x \in X \;\Rightarrow\; f(x) \leq L(\{f_\lambda(x)\}_{\lambda\in\Lambda}).$$

Demostración. Sea C el subconjunto no vacío de $\ell^\infty(\Lambda)$

$$C := \operatorname{co}\left\{\{f_\lambda(x) - f(x)\}_{\lambda\in\Lambda} : x \in X\right\}.$$

Aplicamos el teorema de Mazur–Orlicz, Teorema 1.13, al espacio vectorial $\ell^\infty(\Lambda)$, al funcional sublineal S definido en (3.1) y al subconjunto convexo C de $\ell^\infty(\Lambda)$, para obtener un $L \in \Delta_\Lambda$ tal que

$$\inf_C L = \inf_C S. \tag{3.5}$$

Analicemos con detalle esta igualdad. Por el Lema 1.9, el miembro izquierdo no es más que

$$\inf_{\varphi\in C} L(\varphi) = \inf_{x\in X} L(\{f_\lambda(x) - f(x)\}_{\lambda\in\Lambda}), \tag{3.6}$$

mientras que para el segundo se tiene

$$\begin{aligned} \inf_{\varphi\in C}\sup_{\lambda\in\Lambda}\varphi(\lambda) &= \inf_{\substack{m\geq 1, \mathbf{t}\in\Delta_m\\ x_1,\ldots,x_m\in X}}\ \sup_{\lambda\in\Lambda}\sum_{j=1}^m t_j(f_\lambda(x_j) - f(x_j)) \\ &\geq \inf_{x\in X}\sup_{\lambda\in\Lambda}(f_\lambda(x) - f(x)) \\ &\geq 0, \end{aligned} \tag{3.7}$$

donde hemos usado que $\{f_\lambda - f\}_{\lambda\in\Lambda}$ es una familia infsup-convexa y $f \leq \sup_{\lambda\in\Lambda}\{f_\lambda\}_{\lambda\in\Lambda}$ en X. Combinando (3.5), (3.6) y (3.7) llegamos a

$$\inf_{x\in X} L(\{f_\lambda(x) - f(x)\}_{\lambda\in\Lambda}) \geq 0,$$

y como $L(\mathbb{1}) = 1$ y L es lineal, tenemos finalmente que

$$x \in X \;\Rightarrow\; f(x) \leq L(\{f_\lambda(x)\}_{\lambda\in\Lambda}).$$

■

Concluimos esta sección refinando el resultado anterior, deduciendo la versión óptima previamente anunciada del teorema del supremo de König, lo que supone una generalización del Teorema 2.14.

Teorema 3.6 *Sean X y Λ conjuntos no vacíos, $\{f_\lambda\}_{\lambda\in\Lambda}$ una familia de funciones reales definidas en X tal que para cada x en X, $\{f_\lambda(x)\}_{\lambda\in\Lambda} \in \ell^\infty(\Lambda)$ y sea $f : X \longrightarrow \mathbb{R}$ una función dada. Entonces $\{f_\lambda - f\}_{\lambda\in\Lambda}$ es infsup-convexa en X si, y solo si, para todo α real haciendo cierta la desigualdad*

$$x \in X \ \Rightarrow \ f(x) + \alpha \leq \sup_{\lambda\in\Lambda} \{f_\lambda(x)\}_{\lambda\in\Lambda},$$

se tiene que, para algún $L \in \Delta_\Lambda$,

$$x \in X \ \Rightarrow \ f(x) + \alpha \leq L(\{f_\lambda(x)\}_{\lambda\in\Lambda}).$$

Demostración. Supongamos inicialmente que $\{f_\lambda - f\}_{\lambda\in\Lambda}$ es una familia infsup-convexa en X. Dado $\alpha \in \mathbb{R}$, $\{f_\lambda - (f + \alpha)\}_{\lambda\in\Lambda}$ también es una familia infsup-convexa y, como por hipótesis tenemos que

$$x \in X \ \Rightarrow \ f(x) + \alpha \leq \sup_{\lambda\in\Lambda} \{f_\lambda(x)\},$$

solo queda aplicar el Teorema 3.5.

Con objeto de demostrar la otra implicación, definimos α como

$$\alpha := \inf_{x\in X} \sup_{\lambda\in\Lambda} (f_\lambda(x) - f(x)).$$

Partimos de que α es finito, ya que en caso contrario no hay nada que probar. Al ser $\alpha \leq \sup_{\lambda\in\Lambda}(f_\lambda(x) - f(x))$, cualquiera sea $x \in X$, por hipótesis existe $L \in \Delta_\Lambda$ tal que

$$x \in X \ \Rightarrow \ f(x) + \alpha \leq L(\{f_\lambda(x)\}_{\lambda\in\Lambda}),$$

esto es,

$$\operatorname*{ínf}_{x\in X}\sup_{\lambda\in\Lambda}(f_\lambda(x)-f(x))\leq \operatorname*{ínf}_{x\in X} L(\{f_\lambda(x)-f(x)\}_{\lambda\in\Lambda}).$$

Para probar que la familia $\{f_\lambda - f\}_{\lambda\in\Lambda}$ es infsup-convexa en X, fijemos $m\geq 1, \mathbf{t}\in\Delta_m$ y $x_1,...,x_m\in X$. Por tanto,

$$\begin{aligned}
\operatorname*{ínf}_{x\in X}\sup_{\lambda\in\Lambda}(f_\lambda(x)-f(x)) &\leq \operatorname*{ínf}_{x\in X} L(\{f_\lambda(x)-f(x)\}_{\lambda\in\Lambda}) \\
&\leq \operatorname*{mín}_{j=1,...,m} L(\{f_\lambda(x_j)-f(x_j)\}_{\lambda\in\Lambda}) \\
&\leq \sum_{j=1}^m t_j L(\{(f_\lambda(x_j)-f(x_j))\}_{\lambda\in\Lambda}) \\
&\leq \sum_{j=1}^m t_j \sup_{\lambda\in\Lambda}(f_\lambda(x_j)-f(x_j)) \\
&\leq \sup_{\lambda\in\Lambda}\sum_{j=1}^m t_j(f_\lambda(x_j)-f(x_j)),
\end{aligned}$$

y la arbitrariedad de m, $\mathbf{t}\in\Delta_m$ y $x_1,...,x_m\in X$ da la infsup-convexidad de la familia $\{f_\lambda - f\}_{\lambda\in\Lambda}$ en X.

■

3.2. Programación Infinita con Restricciones Tipo Desigualdad

En esta última sección extendemos los teoremas clásicos de los multiplicadores de Lagrange, de Karush–Kuhn–Tucker y de Fritz John cuando el número de restricciones es arbitrario –no necesariamente finito–, utilizando para ello la versión del teorema del supremo de König dada en el Teorema 3.6.

Comenzamos presentando nuestro problema de programación infinita. Sean X y Λ conjuntos no vacíos, $f : X \longrightarrow \mathbb{R}$ una función y $\{g_\lambda\}_{\lambda\in\Lambda}$ una familia de funciones reales definidas en X tal que para cualquier $x \in X$, $\{g_\lambda(x)\}_{\lambda\in\Lambda} \in \ell^\infty(\Lambda)$. Supongamos además que el conjunto factible

$$X_0 := \left\{ x \in X : \sup_{\lambda\in\Lambda} g_\lambda(x) \leq 0 \right\}$$

es no vacío. El problema de programación infinita a estudiar es

$$\inf_{x\in X_0} f(x). \tag{3.8}$$

Seguidamente, recordamos un par de conceptos básicos bien conocidos, que generalizan los análogos presentados en la Sección 2.3 cuando allí solo se considera un número finito de restricciones de desigualdades.

Definición 3.7 Dado el programa infinito (3.8), su *lagrangiano* es la función $\mathfrak{L} : X \times \ell^\infty(\Lambda)^*_+ \longrightarrow \mathbb{R}$ dada por

$$\mathfrak{L}(x,\phi) := f(x) + \phi(\{g_\lambda(x)\}_{\lambda\in\Lambda}), \qquad \left((x,\phi) \in X \times \ell^\infty(\Lambda)^*_+\right). \tag{3.9}$$

Definición 3.8 Diremos que $(x^0, \phi_0) \in X \times \ell^\infty(\Lambda)^*_+$ es un *punto de silla* del lagrangiano $\mathfrak{L}$ del programa (3.8) si

$$\mathfrak{L}(x^0, \phi) \leq \mathfrak{L}(x^0, \phi_0) \leq \mathfrak{L}(x, \phi_0), \tag{3.10}$$

cualquiera sea $(x, \phi) \in X \times \ell^\infty(\Lambda)^*_+$.

En primer lugar damos un resultado técnico elemental.

Lema 3.9 *Sea Λ un conjunto no vacío y $x \in \ell^\infty(\Lambda)$ una función verificando $\phi(x) \leq 0$ para todo funcional $\phi \in \ell^\infty(\Lambda)^*_+$. Entonces*

$$-x \in \ell^\infty(\Lambda)_+.$$

Demostración. Supongamos que $-x \notin \ell^\infty(\Lambda)_+$, luego existe

$$\lambda_0 \in \Lambda \text{ tal que } x(\lambda_0) > 0.$$

Definimos el funcional $\phi : \ell^\infty(\Lambda) \longrightarrow \mathbb{R}$ como

$$\phi(y) := y(\lambda_0), \quad (y \in \ell^\infty(\Lambda)),$$

que claramente es lineal y además continuo, pues si notamos por $B_{\ell^\infty(\Lambda)}$ la bola unidad cerrada de $\ell^\infty(\Lambda)$, entonces

$$\begin{aligned} \sup_{y \in B_{\ell^\infty(\Lambda)}} \phi(y) &= \sup_{y \in B_{\ell^\infty(\Lambda)}} y(\lambda_0) \\ &\leq 1. \end{aligned}$$

Más aún, $\phi \in \ell^\infty(\Lambda)^*_+$, porque

$$y \in \ell^\infty(\Lambda)_+ \Rightarrow \phi(y) = y(\lambda_0) \geq 0.$$

Pero

$$\begin{aligned} \phi(x) &= x(\lambda_0) \\ &> 0, \end{aligned}$$

lo que contradice nuestra hipótesis.

■

Para el siguiente resultado, sobradamente conocido, incluimos su prueba, no solo por complitud, sino porque algunas de las ideas que aparecen en ella se usarán más adelante.

Lema 3.10 *Si* $(x^0, \phi_0) \in X \times \ell^\infty(\Lambda)^*_+$ *es un punto de silla del lagrangiano* $\mathfrak{L}$ *del problema de programación no lineal* (3.8), *entonces* x^0 *es una solución óptima de dicho problema y además*

$$f(x^0) = \mathfrak{L}(x^0, \phi_0).$$

Demostración. Como (x^0, ϕ_0) es un punto de silla, dado $(x, \phi) \in X \times \ell^\infty(\Lambda)^*_+$ se tiene que

$$\mathfrak{L}(x^0, \phi) \leq \mathfrak{L}(x^0, \phi_0) \leq \mathfrak{L}(x, \phi_0). \tag{3.11}$$

La primera desigualdad asegura que para todo $\phi \in \ell^\infty(\Lambda)^*_+$ se tiene

$$\phi(\{g_\lambda(x^0)\}_{\lambda\in\Lambda}) \leq \phi_0(\{g_\lambda(x^0)\}_{\lambda\in\Lambda}).$$

Así pues, fijado un tal ϕ, para cualquier $\rho > 0$ se verifica

$$\rho\phi(\{g_\lambda(x^0)\}_{\lambda\in\Lambda}) \leq \phi_0(\{g_\lambda(x^0)\}_{\lambda\in\Lambda}), \tag{3.12}$$

y haciendo tender ρ a infinito llegamos a $\phi(\{g_\lambda(x^0)\}_{\lambda\in\Lambda}) \leq 0$. Por tanto, el Lema 3.9 da

$$-\{g_\lambda(x^0)\}_{\lambda\in\Lambda} \in \ell^\infty(\Lambda)_+,$$

es decir, $x^0 \in X_0$.

Para concluir, esto es, comprobar que f alcanza su mínimo en X_0 en el punto x^0, hagamos $\rho \longrightarrow 0$ en (3.12), deduciendo que $\phi_0(\{g_\lambda(x^0)\}_{\lambda\in\Lambda}) \geq 0$. Pero $\phi_0 \in \ell^\infty(\Lambda)^*_+$ y $-\{g_\lambda(x^0)\}_{\lambda\in\Lambda} \in \ell^\infty(\Lambda)_+$, luego

$$\phi_0(\{g_\lambda(x^0)\}_{\lambda\in\Lambda}) = 0, \tag{3.13}$$

de donde

$$f(x^0) = \mathfrak{L}(x^0, \phi_0).$$

A partir de la segunda desigualdad de (3.11), y teniendo en cuenta (3.13), llegamos a

$$\begin{aligned} f(x^0) &\leq f(x) + \phi_0(\{g_\lambda(x)\}_{\lambda\in\Lambda}) \\ &\leq f(x) \end{aligned}$$

para cualquier x en X_0, y por tanto x^0 es solución de (3.8).

■

Presentamos ahora una adaptación del Lema 2.21 a nuestro nuevo contexto:

Lema 3.11 *Sean X y Λ conjuntos no vacíos, $\{f_\lambda\}_{\lambda\in\Lambda}$ una familia de funciones reales definidas en X tal que $\{f_\lambda(x)\}_{\lambda\in\Lambda} \in \ell^\infty(\Lambda)$ para cualquier $x \in X$*

y $f : X \longrightarrow \mathbb{R}$ *una función. Si el conjunto*

$$X_0 = \left\{ x \in X : \sup_{\lambda \in \Lambda} g_\lambda(x) \leq 0 \right\}$$

es no vacío y x^0 *es una solución del problema* (3.8), *entonces*

$$\inf_{x \in X} \max \left\{ \sup_{\lambda \in \Lambda} g_\lambda(x), f(x) - f(x^0) \right\} = 0.$$

Demostración. Como x^0 es una solución del programa infinito (3.8),

$$0 = \inf_{x \in X_0} (f(x) - f(x^0)),$$

y por tanto,

$$0 \leq \inf_{x \in X_0} \max \left\{ \sup_{\lambda \in \Lambda} g_\lambda(x), f(x) - f(x^0) \right\}. \tag{3.14}$$

En el caso de que $x \in X \setminus X_0$, existe un $\lambda \in \Lambda$ tal que $g_\lambda(x) > 0$, luego

$$0 \leq \inf_{x \in X \setminus X_0} \max \left\{ \sup_{\lambda \in \Lambda} g_\lambda(x), f(x) - f(x^0) \right\}. \tag{3.15}$$

De (3.14) y (3.15) se deduce

$$0 \leq \inf_{x \in X} \max \left\{ \sup_{\lambda \in \Lambda} g_\lambda(x), f(x) - f(x^0) \right\},$$

y tomando $x = x^0$ obtenemos la igualdad.

■

Enunciamos el primer resultado importante de esta sección, un teorema de los multiplicadores de Lagrange, que de forma análoga a como se hizo en el capítulo anterior, vuelve a poner de manifiesto que la dualidad entre las soluciones del problema de programación infinita (3.8) y la existencia de puntos de silla del lagrangiano se da justamente bajo la infsup-convexidad de una conveniente familia de funciones.

Teorema 3.12 *Sean X y Λ conjuntos no vacíos, $f : X \longrightarrow \mathbb{R}$ una función y $\{g_\lambda\}_{\lambda\in\Lambda}$ una familia de funciones reales definidas en X con $\{g_\lambda(x)\}_{\lambda\in\Lambda} \in \ell^\infty(\Lambda)$ para todo x en X y tal que el conjunto factible*

$$X_0 = \left\{x \in X : \sup_{\lambda\in\Lambda} g_\lambda(x) \leq 0\right\}$$

es no vacío. Supongamos además que

(i) *x^0 es solución del problema* (3.8) *y*

(ii) *se cumple la siguiente condición de Slater: existe $x^1 \in X$ tal que*

$$\sup_{\lambda\in\Lambda} g_\lambda(x^1) < 0.$$

Entonces el lagrangiano $\mathfrak{L}$ asociado al problema de programación infinita (3.8) *admite un punto de silla $(x^0, \phi_0) \in X \times \ell^\infty(\Lambda)^*_+$ si, y solo si, la familia $\{g_\lambda\}_{\lambda\in\Lambda}, f - f(x^0)$ es infsup-convexa en X.*

Demostración. $\boxed{\Rightarrow}$ Dado que (x^0, ϕ_0) es un punto de silla del lagrangiano $\mathfrak{L}$ tenemos, gracias al Lema 3.10 y al Lema 3.11, que

$$\inf_{x\in X} \max\left\{\sup_{\lambda\in\Lambda} g_\lambda(x), f(x) - f(x^0)\right\} = 0. \tag{3.16}$$

Haciendo uso de la segunda desigualdad en (3.10), para cada x en X tenemos que

$$f(x^0) = \mathfrak{L}(x^0, \phi_0) \leq \mathfrak{L}(x, \phi_0),$$

es decir,

$$f(x^0) \leq f(x) + \phi_0(\{g_\lambda(x)\}_{\lambda\in\Lambda}),$$

luego

$$0 \leq \phi_0(\{g_\lambda(x)\}_{\lambda\in\Lambda}) + (f(x) - f(x^0)),$$

y en consecuencia, dada la arbitrariedad de $x \in X$,

$$0 \leq \inf_{x\in X}(\phi_0(\{g_\lambda(x)\}_{\lambda\in\Lambda}) + (f(x) - f(x^0))). \tag{3.17}$$

Consideremos a continuación la siguiente notación: sea $\mu \notin \Lambda$ y sea $\Lambda_\mu := \Lambda \cup \{\mu\}$, con lo que

$$\ell^\infty(\Lambda_\mu) \equiv \mathbb{R} \oplus \ell^\infty(\Lambda)$$

y

$$\ell^\infty(\Lambda_\mu)^* \equiv \mathbb{R} \oplus \ell^\infty(\Lambda)^*.$$

Sea entonces $\overline{L} \in \ell^\infty(\Lambda_\mu)^*_+$ el funcional

$$\overline{L} := (L, \rho),$$

donde

$$L := \frac{\phi_0}{\phi_0(\mathbb{1}) + 1}$$

y

$$\rho := \frac{1}{\phi_0(\mathbb{1}) + 1}.$$

Comprobemos que $\overline{L} \in \Delta_{\Lambda_\mu}$, o lo que es igual por ser $\overline{L}$ positivo, que su valor en $\mathbb{1}$ es 1. En efecto:

$$\begin{aligned} \overline{L}(\mathbb{1}) &= L(\mathbb{1}) + \rho \\ &= \tfrac{\phi_0(\mathbb{1})}{\phi_0(\mathbb{1})+1} + \tfrac{1}{\phi_0(\mathbb{1})+1} \\ &= 1. \end{aligned}$$

Reescribimos pues (3.17) como

$$0 \leq \inf_{x\in X} \overline{L}(\{g_\lambda(x)\}_{\lambda\in\Lambda}, f(x) - f(x^0)).$$

Para terminar, dados $m \geq 1$, $\mathbf{t} \in \Delta_m$ y $x_1, ..., x_m \in X$ se tiene

$$\begin{aligned} 0 &\leq \inf_{x \in X} \overline{L}\left(\{g_\lambda(x)\}_{\lambda\in\Lambda}, f(x) - f(x^0)\right) \\ &\leq \min_{j=1,\ldots,m} \overline{L}\left(\{g_\lambda(x_j)\}_{\lambda\in\Lambda}, f(x_j) - f(x^0)\right) \\ &\leq \max\left\{\sup_{\lambda\in\Lambda} \sum_{j=1}^m t_j g_\lambda(x_j), \sum_{j=1}^m t_j(f(x_j) - f(x^0))\right\}, \end{aligned}$$

lo que junto con (3.16) nos da la infsup-convexidad en X de la familia $\{g_\lambda\}_{\lambda\in\Lambda}, f - f(x^0)$.

$\boxed{\Leftarrow}$ Como x^0 es una solución del problema de programación infinita (3.8), en virtud del Lema 3.11, se cumple

$$\inf_{x\in X} \max\left\{\sup_{\lambda\in\Lambda} g_\lambda(x), f(x) - f(x^0)\right\} = 0.$$

Definimos la función $\tilde{f} : X \longrightarrow \mathbb{R}$ como

$$\tilde{f}(x) := 0, \qquad (x \in X),$$

y por el Teorema 3.5, existe $\overline{L} \in \Delta_{\Lambda_\mu}$ tal que

$$x \in X \;\Rightarrow\; 0 \leq \overline{L}(\{g_\lambda(x)\}_{\lambda\in\Lambda}, f(x) - f(x^0)),$$

es decir,

$$x \in X \;\Rightarrow\; 0 \leq L(\{g_\lambda(x)\}_{\lambda\in\Lambda}) + \rho(f(x) - f(x^0)), \tag{3.18}$$

siendo $L \in \ell^\infty(\Lambda)^*_+$ y $\rho \in \mathbb{R}$ con $L(\mathbb{1}) + \rho = 1$. La condición de Slater garantiza que ρ es distinto de cero. Efectivamente, supongamos que $\rho = 0$, luego

$$\begin{aligned} \inf_{x\in X} L(\{g_\lambda(x)\}_{\lambda\in\Lambda}) &\leq L(\{g_\lambda(x^1)\}_{\lambda\in\Lambda}) \\ &\leq \sup_{\lambda\in\Lambda} g_\lambda(x^1) \\ &< 0, \end{aligned}$$

llegando a una contradicción. Si $\phi_0 := \dfrac{L}{\rho} \in \ell^\infty(\Lambda)^*_+$, entonces (3.18) queda como

$$x \in X \;\Rightarrow\; f(x^0) \le f(x) + \phi_0(\{g_\lambda(x)\}_{\lambda\in\Lambda}). \tag{3.19}$$

Eligiendo $x = x^0$ tenemos que $\phi_0(\{g_\lambda(x^0)\}_{\lambda\in\Lambda}) \ge 0$, pero como x^0 es solución del programa (3.8) y $\phi_0 \in \ell^\infty(\Lambda)^*_+$, $\phi_0(\{g_\lambda(x^0)\}_{\lambda\in\Lambda}) \le 0$, y por tanto

$$\phi_0(\big\{g_\lambda(x^0)\big\}_{\lambda\in\Lambda}) = 0. \tag{3.20}$$

En consecuencia,

$$\begin{array}{rcl} f(x^0) + \phi(\{g_\lambda(x^0)\}_{\lambda\in\Lambda}) & \le & f(x^0) \\ & = & f(x^0) + \phi_0(\{g_\lambda(x^0)\}_{\lambda\in\Lambda}), \end{array}$$

que combinado con (3.19) nos permite afirmar que $(x^0, \phi_0) \in X \times \ell^\infty(\Lambda)^*_+$ es un punto de silla asociado al lagrangiano $\mathfrak{L}$ del programa (3.8).

■

Se podría pensar que una condición de Slater válida para el teorema anterior es, de forma más débil, la existencia de un $x^1 \in X$ de forma que para todo λ en Λ, $g_\lambda(x^1) < 0$. Como demuestra el siguiente ejemplo, no es suficiente.

Ejemplo 3.13 Sean $X = \mathbb{R}$, $\Lambda = \mathbb{N}$, $f : \mathbb{R} \longrightarrow \mathbb{R}$ dada por

$$f(x) := x, \;\; (x \in \mathbb{R}),$$

y $g_n : \mathbb{R} \longrightarrow \mathbb{R}$ con $n \ge 1$ definida como

$$g_n(x) := -\frac{x^3}{n}, \;\; (x \in \mathbb{R}).$$

En el Ejemplo 3.4 probamos que esta familia es infsup-convexa en $\mathbb{R}$. Si consideramos el problema de programación infinita (3.8) para estos datos, entonces $X_0 = \mathbb{R}_+$ y falla la condición de Slater exigida en el Teorema 3.12, pues para cualquier $x^1 \in \mathbb{R}_+$ se tiene que

$$\sup_{n\geq 1} g_n(x^1) = 0,$$

aunque existe $x^1 \in X$ tal que $g_\lambda(x^1) < 0$ para todo λ en Λ: basta tomar $x^1 = 1$. Vamos a demostrar que aunque el programa (3.8) tiene solución, ($x_0 = 0$), no existe $\phi_0 \in (\ell^\infty)^*_+$ tal que (x^0, ϕ_0) es un punto de silla del lagrangiano $\mathfrak{L}$. Supongamos, procediendo por reducción al absurdo, que lo admite. Dado que $\phi_0 \in (\ell^\infty)^*_+$, el teorema de Dixmier, Teorema 1.15, garantiza la existencia de $y_0 \in \ell^1$ y $\varphi_0 \in c_0^\perp$ tales que

$$\phi_0 = y_0 + \varphi_0.$$

Observemos que $y_0(n) \geq 0$ para cada n natural, ya que como $\phi_0 \in (\ell^\infty)^*_+$, entonces, al ser $e_n \in (\ell^\infty)_+$ (e_n es la sucesión nula salvo en la posición n-ésima, en la que vale 1) se tiene

$$\begin{array}{rcl} 0 & \leq & \phi_0(e_n) \\ & = & y_0(e_n) + \varphi_0(e_n) \\ & = & y_0(n). \end{array}$$

Como hemos supuesto que (x^0, ϕ_0) es un punto de silla del lagrangiano $\mathfrak{L}$, en particular se satisface la segunda desigualdad de (3.10), esto es,

$$x \in X \ \Rightarrow\ f(x^0) + \phi_0\left(\{g_n(x^0)\}_{n\geq 1}\right) \leq f(x) + \phi_0\left(\{g_n(x)\}_{n\geq 1}\right).$$

Teniendo en cuenta que $\{g_n(x)\}_{n\geq 1} \in c_0$, para cada $x \in \mathbb{R}$ se verifica

$$0 \leq x - x^3 \sum_{n=1}^{\infty} \frac{y_0(n)}{n},$$

lo que es absurdo, porque si $\sum_{n=1}^{\infty} \frac{y_0(n)}{n} = 0$, entonces basta tomar $x < 0$ para llegar a una contradicción; y en el caso de que $\sum_{n=1}^{\infty} \frac{y_0(n)}{n} \neq 0$ elegimos un $x > 0$ suficientemente grande obteniendo

$$x - x^3 \sum_{n=1}^{\infty} \frac{y_0(n)}{n} < 0.$$

■

Para terminar este capítulo, establecemos las versiones óptimas de los teoremas de Karush–Kuhn–Tucker y Fritz John para el problema de programación infinita (3.8).

Teorema 3.14 *Sean X y Λ conjuntos no vacíos, $f : X \longrightarrow \mathbb{R}$ una función y $\{g_\lambda\}_{\lambda\in\Lambda}$ una familia de funciones reales definidas en X tal que $\{g_\lambda(x)\}_{\lambda\in\Lambda} \in \ell^\infty(\Lambda)$ para cada x en X y de forma que el conjunto factible*

$$X_0 = \left\{ x \in X : \sup_{\lambda\in\Lambda} g_\lambda(x) \leq 0 \right\}$$

es no vacío. Supongamos además que $\{g_\lambda\}_{\lambda\in\Lambda}, f - f(x^0)$ es una familia infsup-convexa en X y que se verifica la siguiente condición de Slater: existe $x^1 \in X$ tal que

$$\sup_{\lambda\in\Lambda} g_\lambda(x^1) < 0.$$

Entonces x^0 es una solución del problema de programación infinita (3.8) *si, y solo si, existe $\phi_0 \in \ell^\infty(\Lambda)^*_+$ verificando las condiciones tipo Karush–Kuhn–Tucker*

(KKT1)

$$f+\phi_0(\{g_\lambda\}_{\lambda\in\Lambda}) \text{ alcanza su ínfimo en } X \text{ en } x^0$$

y

(KKT2)

$$\phi_0(\{g_\lambda(x^0)\}_{\lambda\in\Lambda})=0 \quad y \quad \sup_{\lambda\in\Lambda} g(x^0)\leq 0.$$

Demostración. Al igual que sucedía en el capítulo anterior, el cumplimiento de las condiciones de Karush–Kuhn–Tucker equivale a la existencia de un $\phi_0\in\ell^\infty(\Lambda)^*_+$ tal que (x^0,ϕ_0) es un punto de silla del lagrangiano $\mathfrak{L}$ del programa (3.8). En efecto; supongamos inicialmente que (x^0,ϕ_0) es un punto de silla del lagrangiano $\mathfrak{L}$. La segunda desigualdad de (3.10) garantiza entonces que

$$x\in X \;\Rightarrow\; f(x^0)+\phi_0(\{g_\lambda(x^0)\}_{\lambda\in\Lambda})\leq f(x)+\phi_0(\{g_\lambda(x)\}_{\lambda\in\Lambda}),$$

que es precisamente la condición (KKT1). Usando otra vez que (x^0,ϕ_0) es un punto de silla del lagrangiano $\mathfrak{L}$,

$$\phi_0\in\ell^\infty(\Lambda)^*_+ \;\Rightarrow\; \phi(\{g_\lambda(x^0)\}_{\lambda\in\Lambda})\leq \phi_0(\{g_\lambda(x^0)\}_{\lambda\in\Lambda}).$$

Razonando como en la prueba del Lema 3.10 obtenemos la validez de las condiciones (KKT2).

Para la otra implicación, ya hemos señalado que (KKT1) no es más que una de las desigualdades de (3.10). Por otro lado, haciendo uso de (KKT2),

$$\sup_{\lambda\in\Lambda} g_\lambda(x^0)\leq 0, \tag{3.21}$$

lo que implica que para cada $\phi\in\ell^\infty(\Lambda)^*_+$ se satisface

$$\phi(\{g_\lambda(x^0)\}_{\lambda\in\Lambda})\leq 0,$$

por lo que podemos asegurar que para todo $\phi_0 \in \ell^\infty(\Lambda)^*_+$ se tiene que

$$\begin{aligned} f(x^0) + \phi(\{g_\lambda(x^0)\}_{\lambda\in\Lambda}) &\leq f(x^0) \\ &= f(x^0) + \phi_0(\{g_\lambda(x^0)\}_{\lambda\in\Lambda}), \end{aligned}$$

donde hemos usado que $\phi_0(\{g_\lambda(x^0)\}_{\lambda\in\Lambda}) = 0$. Así pues, (x^0, ϕ_0) es un punto de silla para el lagrangiano $\mathfrak{L}$ del problema de programación no lineal (3.8).

■

Teorema 3.15 *Sean X y Λ conjuntos no vacíos, $f : X \longrightarrow \mathbb{R}$ una función y $\{g_\lambda\}_{\lambda\in\Lambda}$ una familia de funciones reales definidas en X tales que $\{g_\lambda(x)\}_{\lambda\in\Lambda} \in \ell^\infty(\Lambda)$ para cualquier x en X. Si el conjunto*

$$X_0 = \left\{x \in X : \sup_{\lambda\in\Lambda} g_\lambda(x) \leq 0\right\}$$

es no vacío y $x^0 \in X$ es una solución del problema de programación infinita (3.8), entonces existen $\rho \geq 0$ y $\phi_0 \in \ell^\infty(\Lambda)^*_+$ tales que $\rho + \phi_0(\mathbb{1}) = 1$ y verificando las condiciones tipo Fritz John

(FJ1)

$$\rho f + \phi_0(\{g_\lambda\}_{\lambda\in\Lambda}) \text{ alcanza su ínfimo en } X \text{ en } x^0$$

y

(FJ2)

$$\phi_0(\{g_\lambda(x^0)\}_{\lambda\in\Lambda}) = 0,$$

si, y solo si, la familia $\{g_\lambda\}_{\lambda\in\Lambda}, f - f(x^0)$ es infsup-convexa en X.

Demostración. Empecemos suponiendo que $\{g_\lambda\}_{\lambda\in\Lambda}, f-f(x^0)$ es una familia infsup-convexa en X. Razonando como en la prueba del Teorema 3.12, podemos asegurar la existencia de $\phi_0 \in \ell^\infty(\Lambda)_+^*$ y $\rho \geq 0$ verificando $\phi_0(\mathbb{1})+\rho = 1$ y la desigualdad

$$(3.22) \qquad x \in X \;\Rightarrow\; 0 \leq \phi_0(\{g_\lambda(x)\}_{\lambda\in\Lambda}) + \rho(f(x) - f(x^0)),$$

i.e.,

$$(3.23) \qquad x \in X \;\Rightarrow\; \rho f(x^0) \leq \rho f(x) + \phi_0(\{g_\lambda(x)\}_{\lambda\in\Lambda}).$$

Haciendo $x = x^0$, deducimos que $\phi_0(\{g_\lambda(x^0)\}_{\lambda\in\Lambda}) \geq 0$, pero como x^0 es solución del programa (3.8), $\phi_0(\{g_\lambda(x^0)\}_{\lambda\in\Lambda}) \leq 0$, y por lo tanto se cumple (FJ2). En particular, podemos reescribir (3.23) como

$$x \in X \;\Rightarrow\; \rho f(x^0) + \phi_0(\{g_\lambda(x^0)\}_{\lambda\in\Lambda}) \leq \rho f(x) + \phi_0(\{g_\lambda(x)\}_{\lambda\in\Lambda}),$$

que no es más que la condición (FJ1).

Y recíprocamente; usando las condiciones (FJ1) y (FJ2), para cada x en X se cumple

$$\rho f(x^0) \leq \rho f(x) + \phi_0(\{g_\lambda(x)\}_{\lambda\in\Lambda}),$$

es decir,

$$(3.24) \qquad x \in X \;\Rightarrow\; 0 \leq \phi_0(\{g_\lambda(x)\}_{\lambda\in\Lambda}) + \rho(f(x) - f(x^0)).$$

Retomando la notación introducida en la prueba del Teorema 3.12, sean

$$\alpha := \rho + \phi_0(\mathbb{1}),$$

y

$$L := \left(\frac{\phi_0}{\alpha}, \frac{\rho}{\alpha}\right) \in \ell^\infty(\Lambda_\mu)_+^*.$$

Es inmediato comprobar que $L \in \Delta_{\Lambda_\mu}$ y que podemos reescribir (3.24) como

$$0 \leq \inf_{x \in X} L(\{g_\lambda(x)\}_{\lambda \in \Lambda}, f(x) - f(x^0)).$$

Aplicando el Lema 3.11, y razonando como en la demostración del Teorema 3.6, tenemos que $\{g_\lambda\}_{\lambda \in \lambda}, f - f(x^0)$ es una familia infsup-convexa en X.

$\blacksquare$

Capítulo 4

Programas Infinitos Arbitrarios

Este capítulo es el último del presente trabajo. En el primero obtuvimos una generalización del teorema del máximo de König que, aplicado a programación no lineal semi-infinita, nos llevó a obtener versiones óptimas de los teoremas de los multiplicadores de Lagrange, Karush–Kuhn–Tucker y Fritz John. En el segundo capítulo hicimos otro tanto con el teorema del supremo de König, pero en el contexto de programación infinita con restricciones tipo desigualdad. Ahora permitimos que dichas restricciones admitan además un número arbitrario de igualdades, siendo esta una generalización no trivial. Sin embargo, a diferencia de los dos capítulos anteriores, nuestro punto de partida no será el teorema de Mazur–Orlicz, Teorema 1.13, sino el teorema de Hahn–Banach en su formulación geométrica, Teorema 1.12, pues hemos tenido en cuenta que este resultado es más conocido en optimización. También destacamos que a diferencia de los demás capítulos, deduciremos en primer lugar la optimalidad del resultado de tipo Karush–Kuhn–Tucker, y a partir de él, el teorema los multiplicadores de Lagrange. Por último, presentaremos una nueva generalización para el teorema de Fritz John.

Mencionemos así mismo que mediante un ejemplo pondremos de manifiesto que los resultados aquí presentados mejoran estrictamente los ya conocidos para nociones más débiles de convexidad, quasi-convexidad y convexidad en el sentido de Fan, destacando una vez más la importancia del concepto de infsup-convexidad, del cual depende la optimalidad de nuestros resultados.

4.1. Un Teorema Tipo Hahn–Banach en $\ell^\infty(\Lambda)$

En esta sección obtenemos la herramienta clave para la demostración de la versión del teorema de Karush–Kuhn–Tucker: un caso particular del teorema de Mazur-Orlicz deducido directamente a partir del teorema de separación de convexos de Hahn–Banach. Antes, vamos a probar un resultado técnico fácil.

Lema 4.1 *Si Λ es un conjunto no vacío, entonces*

$$A := \left\{\Phi \in \ell^\infty(\Lambda) : \sup_{\lambda \in \Lambda} \Phi(\lambda) < 0\right\}$$

es un subconjunto abierto de $\ell^\infty(\Lambda)$.

Demostración. Sean $\Phi \in A$ y $\varphi \in \ell^\infty(\Lambda)$. Supongamos que

$$\|\Phi - \varphi\| < -\frac{1}{2} \sup_{\lambda \in \Lambda} \Phi(\lambda),$$

entonces,

$$\begin{aligned} \sup_{\lambda \in \Lambda} \varphi(\lambda) &\leq \frac{1}{2} \sup_{\lambda \in \Lambda} \Phi(\lambda) \\ &< 0, \end{aligned}$$

y por tanto, $\varphi \in A$. Como Φ es cualquier elemento de A, hemos probado que dicho subconjunto de $\ell^\infty(\Lambda)$ es abierto.

∎

Pasamos a demostrar directamente de la versión geométrica del teorema Hahn–Banach el caso particular del teorema de Mazur–Orlicz anunciado, con el espacio vectorial $\ell^\infty(\Lambda)$ y el funcional sublineal supremo.

Lema 4.2 *Sea Λ un conjunto no vacío. Dado C un subconjunto no vacío y convexo de $\ell^\infty(\Lambda)$, existe $L \in \Delta_\Lambda$ tal que*

$$\inf_{\Phi \in C} L(\Phi) = \inf_{\Phi \in C} \sup_{\lambda \in \Lambda} \Phi(\lambda).$$

Demostración. Antes de comenzar la demostración observemos que si $L \in \Delta_\Lambda$, entonces

$$\Phi \in \ell^\infty(\Lambda) \;\Rightarrow\; L(\Phi) \leq \sup_{\lambda \in \Lambda} \Phi(\lambda),$$

luego

$$\inf_{\Phi \in C} L(\Phi) \leq \inf_{\Phi \in C} \sup_{\lambda \in \Lambda} \Phi(\lambda). \tag{4.1}$$

En consecuencia, y sin que ello suponga pérdida de generalidad alguna, suponemos que

$$r := \inf_{\Phi\in C} \sup_{\lambda\in\Lambda} \Phi(\lambda) > -\infty.$$

Consideremos ahora el subconjunto K de $\ell^\infty(\Lambda)$

$$K := C - r\mathbb{1},$$

que es convexo y no vacío. Por otro lado, como

$$\inf_{\Phi\in K} \sup_{\lambda\in\Lambda} \Phi(\lambda) = 0,$$

se tiene

$$K \cap A = \emptyset,$$

siendo

$$A := \left\{ \Phi \in \ell^\infty(\Lambda) : \sup_{\lambda\in\Lambda} \Phi(\lambda) < 0 \right\}.$$

El Lema 4.1 garantiza que el conjunto A es abierto, lo cual nos permite aplicar el teorema de Hahn–Banach en su versión geométrica, Teorema 1.12, que nos asegura la existencia de $\tilde{L} \in \ell^\infty(\Lambda)^*\backslash\{0\}$ y $\alpha \in \mathbb{R}$ verificando

$$\left.\begin{array}{r} \Phi \in K \\ \varphi \in A \end{array}\right\} \Rightarrow \tilde{L}(\varphi) < \alpha \leq \tilde{L}(\Phi). \tag{4.2}$$

Dados $\varphi \in A$ y $\rho > 0$, es obvio que $\rho\varphi \in A$, y en consecuencia

$$\rho\tilde{L}(\varphi) < \alpha.$$

La arbitrariedad de $\rho > 0$ implica que

$$0 \leq \alpha \tag{4.3}$$

y

$$\tilde{L}(\varphi) \leq 0.$$

Como para todo $\varphi \in A$

$$0 < \inf_{\lambda\in\Lambda}(-\varphi(\lambda)),$$

entonces para cualquier $\delta > 0$ se tiene

$$0 < \inf_{\lambda\in\Lambda}(-\varphi + \delta\mathbb{1})(\lambda),$$

con lo que

$$0 \leq \tilde{L}(-\varphi) + \delta\tilde{L}(\mathbb{1}).$$

Como la desigualdad es cierta para cualquier $\delta > 0$, se deduce que

$$\tilde{L}(-\varphi) \geq 0,$$

y al ser φ cualquier función de A, concluimos que $\tilde{L}$ es positivo. Comprobemos que $\tilde{L}(\mathbb{1}) > 0$, o equivalentemente, $\tilde{L}(\mathbb{1}) \neq 0$, pues $\tilde{L}$ es positivo. Supongamos que $\tilde{L}(\mathbb{1}) = 0$ y sea $\Phi_0 \in C$ (recordemos que C es no vacío). Entonces llegamos a la siguiente cadena contradictoria de desigualdades:

$$\begin{aligned}\tilde{L}\left(\Phi_0 - \left(1 + \sup_{\lambda\in\Lambda}\Phi_0(\lambda)\right)\mathbb{1}\right) &< \alpha \\ &\leq \tilde{L}(\Phi_0 - r\mathbb{1}) \\ &= \tilde{L}\left(\Phi_0 - \left(1 + \sup_{\lambda\in\Lambda}\Phi_0(\lambda)\right)\mathbb{1}\right),\end{aligned}$$

donde hemos usado que $\left\{\Phi_0 - \left(1 + \sup_{\lambda\in\Lambda}\Phi_0(\lambda)\right)\mathbb{1}\right\} \in A$, $\Phi_0 - r\mathbb{1} \in K$ y (4.2). En último lugar definimos $L \in \ell^\infty(\Lambda)^*_+$ como

$$L := \frac{\tilde{L}}{\tilde{L}(\mathbb{1})}.$$

En vista del Lema 3.1, $L \in \Delta_\Lambda$, y usando además dicho lema, (4.2) y (4.3), deducimos

$$\inf_{\Phi\in C}\sup_{\lambda\in\Lambda}\Phi(\lambda) \leq \inf_{\Phi\in C} L(\Phi),$$

lo que junto a (4.1) da la igualdad buscada:

$$\operatorname*{ínf}_{\Phi \in C} \sup_{\lambda \in \Lambda} \Phi(\lambda) = \operatorname*{ínf}_{\Phi \in C} L(\Phi).$$

■

4.2. Programación Infinita

Centramos nuestros esfuerzos ahora en deducir del Lema 4.2 versiones óptimas de los teoremas clásicos de programación no lineal, generalizando simultáneamente los resultados de los dos capítulos anteriores. A diferencia de la estructura seguida en estos, mostraremos como el resultado de Lagrange se puede probar a partir del de Karush–Kuhn–Tucker. A continuación, presentaremos un ejemplo que pone de manifiesto el papel esencial de la infsup-convexidad en nuestros resultados y su mejora respecto a otros que también debilitan la convexidad, como se comentó en la introducción del capítulo.

Describimos inicialmente el problema de optimización: sean X, Λ y Ω conjuntos con X no vacío (adoptamos un convenio análogo al del caso finito del Capítulo 2, contemplando la posibilidad de que Λ u Ω sean vacíos), $f : X \longrightarrow \mathbb{R}$ una función y $\{g_\lambda\}_{\lambda \in \Lambda}$ y $\{h_\omega\}_{\omega \in \Omega}$ familias de funciones reales definidas en X tales que para todo $x \in X$, $\{g_\lambda(x)\}_{\lambda\in\Lambda} \in \ell^\infty(\Lambda)$ y $\{h_\omega(x)\}_{\omega\in\Omega} \in \ell^\infty(\Omega)$. Supongamos no vacío el conjunto factible

$$X_0 := \left\{x \in X : \sup_{\lambda\in\Lambda} g_\lambda(x) \leq 0, \sup_{\omega\in\Omega} \mid h_\omega(x) \mid = 0\right\}.$$

Consideramos entonces el problema de programación infinita

$$\inf_{x\in X_0} f(x). \tag{4.4}$$

Tal y como acabamos de mencionar, en primer lugar vamos a probar el Teorema de Karush–Kuhn–Tucker mediante el Lema 4.2.

Teorema 4.3 *Sean X, Λ y Ω conjuntos con X no vacío, $f : X \longrightarrow \mathbb{R}$ una función y $\{g_\lambda\}_{\lambda\in\Lambda}$ y $\{h_\omega\}_{\omega\in\Omega}$ familias de funciones reales definidas en X cumpliendo $\{g_\lambda(x)\}_{\lambda\in\Lambda} \in \ell^\infty(\Lambda)$ y $\{h_\omega(x)\}_{\omega\in\Omega} \in \ell^\infty(\Omega)$ para todo $x \in X$. Supongamos que el conjunto*

$$X_0 = \left\{x \in X : \sup_{\lambda\in\Lambda} g_\lambda(x) \leq 0, \sup_{\omega\in\Omega} \mid h_\omega(x) \mid = 0\right\}$$

es no vacío, x^0 es un punto de X, la familia $\{g_\lambda\}_{\lambda\in\Lambda}, \{h_\omega\}_{\omega\in\Omega}, \{-h_\omega\}_{\omega\in\Omega}, f - f(x^0)$ es infsup-convexa en X y se verifica la siguiente condición de Slater: existe $x^1 \in X$ tal que

$$\sup_{\lambda\in\Lambda} g_\lambda(x^1) < 0 \quad y \quad \sup_{\omega\in\Omega} \mid h_\omega(x^1) \mid = 0,$$

y para cualesquiera $\varphi', \varphi'' \in \ell^\infty(\Omega)^* \setminus \{0\}$ podemos encontrar un $x^2 \in X$ tal que

$$\varphi''(\{h_\omega(x^2)\}_{\omega\in\Omega}) < \varphi'(\{h_\omega(x^2)\}_{\omega\in\Omega}).$$

Entonces x^0 es solución del problema de programación infinita (4.4) si, y solo si, existen $\Phi_0 \in \ell^\infty(\Lambda)^*_+$ y $\varphi_0 \in \ell^\infty(\Omega)^*$ satisfaciendo las condiciones tipo Karush–Kuhn–Tucker

(KKT1)

$$f + \Phi_0(\{g_\lambda\}_{\lambda\in\Lambda}) + \varphi_0(\{h_\omega\}_{\omega\in\Omega}) \text{ alcanza su ínfimo en } X \text{ en } x^0$$

y

(KKT2)

$$\Phi_0(\{g_\lambda(x^0)\}_{\lambda\in\Lambda}) = 0, \{h_\omega(x^0)\}_{\omega\in\Omega} = 0 \quad y \quad \sup_{\lambda\in\Lambda} g_\lambda(x^0) \le 0.$$

Demostración. Partamos en primer lugar de que x^0 es solución del problema de programación infinita (4.4). Definimos el conjunto $\Theta := \Lambda \cup \Omega \cup \Omega_0 \cup \{\mu\}$ donde Ω y Ω_0 son conjuntos biyectivos, $\mu \notin \Lambda \cup \Omega \cup \Omega_0$ y Λ, Ω y Ω_0 son mutuamente disjuntos. Aplicamos el Lema 4.2 a

$$C := \mathrm{co}\{(\{g_\lambda(x)\}_{\lambda\in\Lambda}, \{h_\omega(x)\}_{\omega\in\Omega}, \{-h_\omega(x)\}_{\omega\in\Omega}, f(x) - f(x^0)) : x \in X\},$$

que es un subconjunto convexo y no vacío de $\ell^\infty(\Theta)$, para obtener un $L \in \Delta_\Theta$ cumpliendo

$$\inf_{\Phi\in C} L(\Phi) = \inf_{\Phi\in C} \sup_{\theta\in\Theta} \Phi(\theta). \tag{4.5}$$

En la siguiente cadena de desigualdades usamos, en este orden, el Lema 3.11, la infsup-convexidad de la familia $\{g_\lambda\}_{\lambda\in\Lambda}$, $\{h_\omega\}_{\omega\in\Omega}$, $\{-h_\omega\}_{\omega\in\Omega}$, $f - f(x^0)$ en X y la existencia de $L \in \Delta_\Theta$ verificando la igualdad (4.5):

$$\begin{aligned}
0 &= \inf_{x\in X} \max\left\{\sup_{\lambda\in\Lambda} g_\lambda(x), \sup_{\omega\in\Omega} h_\omega(x), \sup_{\omega\in\Omega}(-h_\omega(x)), f(x) - f(x^0)\right\} \\
&\le \inf_{\substack{m\ge1,\mathbf{t}\in\Delta_m \\ x_1,\dots,x_m\in X}} \max\left\{\sup_{\lambda\in\Lambda} \sum_{j=1}^m t_j g_\lambda(x_j), \sup_{\omega\in\Omega} \sum_{j=1}^m t_j h_\omega(x_j),\right. \\
&\qquad \left.\sup_{\omega\in\Omega} \sum_{j=1}^m t_j(-h_\omega)(x_j), \sum_{j=1}^m t_j(f(x_j) - f(x))\right\} \\
&= \inf_{\phi\in C} \sup_{\theta\in\Theta} \phi(\Theta) \\
&= \inf_{\phi\in C} L(\phi),
\end{aligned}$$

es decir,

$$0 \leq \inf_{x \in X} L\left(\{g_\lambda(x)\}_{\lambda\in\Lambda}, \{h_\omega(x)\}_{\omega\in\Omega}, \{-h_\omega(x)\}_{\omega\in\Omega}, (f(x)-f(x^0))\right). \tag{4.6}$$

Sabemos que $L = (L^*, L^{**}, L^{***}, \rho)$, donde $L^* \in \ell^\infty(\Lambda)^*_+$, $L^{**} \in \ell^\infty(\Omega)^*_+$, $L^{***} \in \ell^\infty(\Omega)^*_+$ y $\rho \in \mathbb{R}_+$, y gracias a la condición de Slater tenemos que $\rho \neq 0$, hecho que nos permite definir los siguientes funcionales: $\Phi_0 := \dfrac{L^*}{\rho} \in \ell^\infty(\Lambda)^*_+$ y $\varphi_0 := \dfrac{L^{**} - L^{***}}{\rho} \in \ell^\infty(\Omega)^*$. Con ellos, reescribimos (4.6) como

$$x \in X \;\Rightarrow\; f(x^0) \leq f(x) + \Phi_0\left(\{g_\lambda(x)\}_{\lambda\in\Lambda}\right) + \varphi_0\left(\{h_\omega(x)\}_{\omega\in\Omega}\right). \tag{4.7}$$

Dado que x^0 es solución del problema de programación infinita (4.4) y que $\Phi_0 \in \ell^\infty(\Lambda)^*_+$,

$$\Phi_0(\left\{g_\lambda(x^0)\right\}_{\lambda\in\Lambda}) \leq 0,$$

y tomando $x = x^0$ en (4.7) tenemos la igualdad

$$\Phi_0(\left\{g_\lambda(x^0)\right\}_{\lambda\in\Lambda}) = 0.$$

Este hecho y la definición de X_0 implican la veracidad de (KKT2). Solo queda comprobar (KKT1), pero como x^0 es solución del problema no lineal (4.4) y φ_0 es lineal tenemos que

$$\varphi_0(\left\{h_\omega(x^0)\right\}_{\omega\in\Omega}) = 0,$$

y por lo tanto, para cada x en X_0 se cumple

$$\begin{aligned} f(x^0) + \Phi_0(\{g_\lambda(x^0)\}_{\lambda\in\Lambda}) + \varphi_0(\{h_\omega(x^0)\}_{\omega\in\Omega}) \;\;\leq\;\; & f(x) + \Phi_0(\{g_\lambda(x)\}_{\lambda\in\Lambda}) \\ & + \varphi_0(\{h_\omega(x)\}_{\omega\in\Omega}). \end{aligned}$$

Para probar la otra implicación supongamos que existen $\Phi_0 \in \ell^\infty(\Lambda)^*_+$, $\varphi_0 \in \ell^\infty(\Omega)^*$ y $x^0 \in X$ verificando (KKT1) y (KKT2). Que x^0 pertenece a

X_0 aparece explícitamente en las condiciones dadas en (KKT2). Con la idea de comprobar que x^0 es solución del problema de programación infinita (4.4), escribimos (KKT1), esto es, para cada x en X tenemos que

$$\begin{aligned} f(x^0) + \Phi_0(\{g_\lambda(x^0)\}_{\lambda\in\Lambda}) + \varphi_0(\{h_\omega(x^0)\}_{\omega\in\Omega}) \quad \leq \quad & f(x) + \Phi_0(\{g_\lambda(x)\}_{\lambda\in\Lambda}) \\ & + \varphi_0(\{h_\omega(x)\}_{\omega\in\Omega}), \end{aligned}$$

y de (KKT2) se deduce

$$x \in X \quad \Rightarrow \quad f(x^0) \leq f(x) + \Phi_0(\{g_\lambda(x)\}_{\lambda\in\Lambda}) + \varphi_0(\{h_\omega(x)\}_{\omega\in\Omega}). \tag{4.8}$$

Por otro lado, dado $x \in X_0$ se tiene que

$$f(x) + \Phi_0(\{g_\lambda(x)\}_{\lambda\in\Lambda}) + \varphi_0(\{h_\omega(x)\}_{\omega\in\Omega}) \leq f(x). \tag{4.9}$$

Así pues, combinando (4.8) y (4.9), concluimos que

$$x \in X_0 \;\Rightarrow\; f(x^0) \leq f(x),$$

es decir, x^0 es solución del programa infinito (4.4).

■

Para enunciar el teorema de los multiplicadores de Lagrange, lógicamente debemos extender los conceptos de lagrangiano y de punto de silla dados en los capítulos anteriores:

Definición 4.4 El *lagrangiano* asociado al problema de programación infinita (4.4) es la función $\mathfrak{L} : X \times \ell^\infty(\Lambda)^*_+ \times \ell^\infty(\Omega)^* \longrightarrow \mathbb{R}$ que a cada $(x, \Phi, \varphi) \in X \times \ell^\infty(\Lambda)^*_+ \times \ell^\infty(\Omega)^*$ le hace corresponder

$$\mathfrak{L}(x, \Phi, \varphi) := f(x) + \Phi(\{g_\lambda(x)\}_{\lambda\in\Lambda}) + \varphi(\{h_\omega(x)\}_{\omega\in\Omega}). \tag{4.10}$$

Definición 4.5 Una terna $(x^0, \Phi_0, \varphi_0) \in X \times \ell^\infty(\Lambda)^*_+ \times \ell^\infty(\Omega)^*$ es un *punto de silla* del lagrangiano $\mathfrak{L}$ asociado al programa infinito (4.4), si para todo $(x, \Phi, \varphi) \in X \times \ell^\infty(\Lambda)^*_+ \times \ell^\infty(\Omega)^*$ se cumple que

$$\mathfrak{L}(x^0, \Phi, \varphi) \leq \mathfrak{L}(x^0, \Phi_0, \varphi_0) \leq \mathfrak{L}(x, \Phi_0, \varphi_0). \tag{4.11}$$

En la línea de los casos particulares recogidos en los dos capítulos anteriores, probamos la relación que se da entre puntos de silla y soluciones óptimas:

Lema 4.6 *Sea* $(x^0, \Phi_0, \varphi_0) \in X \times \ell^\infty(\Lambda)^*_+ \times \ell^\infty(\Omega)^*$ *un punto de silla del lagrangiano* $\mathfrak{L}$ *del problema de programación infinita* (4.4). *Entonces* x^0 *es una solución óptima del programa* (4.4) *y*

$$f(x^0) = \mathfrak{L}(x^0, \Phi_0, \varphi_0).$$

Demostración. Como (x^0, Φ_0, φ_0) es un punto de silla de $\mathfrak{L}$, entonces para cada $(\phi, \varphi) \in \ell^\infty(\Lambda)^*_+ \times \ell^\infty(\Omega)^*$, la siguiente desigualdad es cierta:

$$\phi(\{g_\lambda(x^0)\}_{\lambda\in\Lambda}) + \varphi(\{h_\omega(x^0)\}_{\omega\in\Omega}) \leq \phi_0(\{g_\lambda(x^0)\}_{\lambda\in\Lambda}) + \varphi_0(\{h_\omega(x^0)\}_{\omega\in\Omega}). \tag{4.12}$$

Tomando $\varphi = \varphi_0$ y razonando como en el Lema 3.10, tenemos que

$$-\{g_\lambda(x^0)\}_{\lambda\in\Lambda} \in \ell^\infty(\Lambda)_+,$$

y que

$$\phi_0(\{g_\lambda(x^0)\}_{\lambda\in\Lambda}) = 0, \tag{4.13}$$

mientras que si $\phi = \phi_0$ en (4.12), obtenemos

$$\varphi \in \ell^\infty(\Omega)^* \Rightarrow \varphi(\{h_\omega(x^0)\}_{\omega\in\Omega}) \leq \varphi_0(\{h_\omega(x^0)\}_{\omega\in\Omega}).$$

Fijado $\varphi \in \ell^\infty(\Omega)^*$, para cada ρ real se cumple

$$\rho\varphi(\{h_\omega(x^0)\}_{\omega\in\Omega}) \leq \varphi_0(\{h_\omega(x^0)\}_{\omega\in\Omega}). \tag{4.14}$$

Haciendo $\rho \longrightarrow +\infty$ tenemos que $\varphi(\{h_\omega(x^0)\}_{\omega\in\Omega}) \leq 0$, pero tomando $\rho \longrightarrow -\infty$ obtenemos $\varphi(\{h_\omega(x^0)\}_{\omega\in\Omega}) \geq 0$, luego para todo $\varphi \in \ell^\infty(\Omega)^*$,

$$\varphi(\{h_\omega(x^0)\}_{\omega\in\Omega}) = 0. \tag{4.15}$$

Así pues, $\{h_\omega(x^0)\}_{\omega\in\Omega} = 0$, lo que nos permite concluir que $x^0 \in X_0$.

Usando nuevamente que (x^0, Φ_0, φ_0) es un punto de silla, para cada x en X_0 se verifica la siguiente cadena de desigualdades:

$$\begin{aligned} f(x^0) &\leq f(x) + \phi_0(\{g_\lambda(x)\}_{\lambda\in\Lambda}) + \varphi_0(\{h_\omega(x)\}_{\omega\in\Omega}) \\ &\leq f(x), \end{aligned}$$

y por tanto x^0 es solución del problema de programación infinita (4.4).

Que $f(x^0) = \mathfrak{L}(x^0, \Phi_0, \varphi_0)$, es consecuencia directa de (4.13) y (4.15).

■

Bajo las hipótesis del Teorema 4.3, las condiciones de Karush–Kuhn–Tucker equivalen a encontrar un punto de silla para el lagrangiano asociado al problema (4.4).

Teorema 4.7 *Consideremos X un conjunto no vacío y sean $f : X \longrightarrow \mathbb{R}$ una función, Λ y Ω conjuntos cualesquiera y $\{g_\lambda\}_{\lambda\in\Lambda}$ y $\{h_\omega\}_{\omega\in\Omega}$ dos familias de funciones reales definidas en X tales que para cada x en X, $\{g_\lambda(x)\}_{\lambda\in\Lambda} \in \ell^\infty(\Lambda)$ y $\{h_\omega(x)\}_{\omega\in\Omega} \in \ell^\infty(\Omega)$. Supongamos que el conjunto factible*

$$X_0 = \left\{x \in X : \sup_{\lambda\in\Lambda} g_\lambda(x) \leq 0,\ \sup_{\omega\in\Omega} \mid h_\omega(x) \mid = 0\right\}$$

es no vacío, x^0 es un punto de X, la familia $\{g_\lambda\}_{\lambda\in\Lambda}, \{h_\omega\}_{\omega\in\Omega}, \{-h_\omega\}_{\omega\in\Omega}, f - f(x^0)$ es infsup-convexa en X y la siguiente condición de Slater es cierta: existe $x^1 \in X$ satisfaciendo

$$\sup_{\lambda\in\Lambda} g_\lambda(x^1) < 0 \quad y \quad \sup_{\omega\in\Omega} \mid h_\omega(x^1) \mid = 0$$

y para cualesquiera $\varphi^{'}, \varphi^{''} \in \ell^\infty(\Omega)^* \setminus \{0\}$ encontramos $x^2 \in X$ tal que

$$\varphi^{''}(\big\{h_\omega(x^2)\big\}_{\omega\in\Omega}) < \varphi^{'}(\big\{h_\omega(x^2)\big\}_{\omega\in\Omega}).$$

Si x^0 es solución del problema de programación infinita (4.4), entonces existen $\Phi_0 \in \ell^\infty(\Lambda)^*_+$ y $\varphi_0 \in \ell^\infty(\Omega)^*$ verificando las condiciones de Karush–Kuhn–Tucker

(KKT1)

$$f + \Phi_0(\{g_\lambda\}_{\lambda\in\Lambda}) + \varphi_0(\{h_\omega\}_{\omega\in\Omega}) \ \textit{alcanza su ínfimo en } X \textit{ en } x^0$$

y

(KKT2)

$$\Phi_0(\big\{g_\lambda(x^0)\big\}_{\lambda\in\Lambda}) = 0, \quad \{h_\omega(x^0)\}_{\omega\in\Omega} = 0 \quad y \quad \sup_{\lambda\in\Lambda} g_\lambda(x^0) \leq 0,$$

si, y solo si, (x^0, Φ_0, φ_0) es un punto de silla para el lagrangiano $\mathfrak{L}$ asociado al problema (4.4).

Demostración. Si (x^0, Φ_0, φ_0) es un punto de silla, en particular, para cada x en X,

$$\begin{aligned} f(x^0) + \Phi_0(\{g_\lambda(x^0)\}_{\lambda\in\Lambda}) + \varphi_0(\{h_\omega(x^0)\}_{\omega\in\Omega}) \;\; &\leq \;\; f(x) + \Phi_0(\{g_\lambda(x)\}_{\lambda\in\Lambda}) \\ &\quad\; +\varphi_0(\{h_\omega(x)\}_{\omega\in\Omega}), \end{aligned}$$

es decir, se verifica (KKT1). Pero gracias al Lema 4.6, x^0 es solución del problema de programación infinita (4.4), lo que junto a la linealidad de $\Phi_0 \in \ell^\infty(\Lambda)^*_+$ nos da (KKT2).

Recíprocamente: la segunda desigualdad que aparece en la definición de punto de silla es inmediata por (KKT1). Como $\Phi \in \ell^\infty(\Lambda)^*_+$ y x^0 es solución del programa infinito (4.4), es evidente que

$$\Phi(\{g_\lambda(x^0)\}_{\lambda\in\Lambda}) \leq 0.$$

Solo resta usar (KKT2) y la linealidad de $\varphi \in \ell^\infty(\Omega)^*$ para comprobar la validez de la primera desigualdad de punto de silla.

$\blacksquare$

Pasamos a probar la versión más general del teorema de los multiplicadores de Lagrange dada en esta memoria.

Teorema 4.8 *Sean X, Λ y Ω conjuntos tales que X es no vacío, $f : X \longrightarrow \mathbb{R}$ una función y $\{g_\lambda\}_{\lambda\in\Lambda}, \{h_\omega\}_{\omega\in\Omega}$ familias de funciones reales definidas en X tales que para todo x en X se tiene $\{g_\lambda(x)\}_{\lambda\in\Lambda} \in \ell^\infty(\Lambda)$ y $\{h_\omega(x)\}_{\omega\in\Omega} \in \ell^\infty(\Omega)$. Supongamos además que el conjunto*

$$X_0 = \left\{x \in X : \sup_{\lambda\in\Lambda} g_\lambda(x) \leq 0, \ \sup_{\omega\in\Omega} \mid h_\omega(x) \mid = 0\right\}$$

es no vacío, x^0 es un punto de X y que se cumple la siguiente condición de Slater: existe $x^1 \in X$ tal que

$$\sup_{\lambda\in\Lambda} g_\lambda(x^1) < 0 \quad y \quad \sup_{\omega\in\Omega} \mid h_\omega(x^1) \mid = 0,$$

y para cualesquiera $\varphi^{'}, \varphi^{''} \in \ell^\infty(\Omega)^* \setminus \{0\}$ *podemos encontrar un* $x^2 \in X$ *tal que*

$$\varphi^{''}(\{h_\omega(x^2)\}_{\omega\in\Omega}) < \varphi^{'}(\{h_\omega(x^2)\}_{\omega\in\Omega}).$$

Entonces x^0 *es solución del problema de programación infinita* (4.4) *y la familia* $\{g_\lambda\}_{\lambda\in\Lambda}, \{h_\omega\}_{\omega\in\Omega}, \{-h_\omega\}_{\omega\in\Omega}, f - f(x^0)$ *es infsup-convexa en* X *si, y solo si, existen* $\Phi_0 \in \ell^\infty(\Lambda)^*_+$ *y* $\varphi_0 \in \ell^\infty(\Omega)^*$ *tales que* (x^0, Φ_0, φ_0) *es un punto de silla para el lagrangiano* $\mathfrak{L}$ *asociado al programa* (4.4)*.*

Demostración. $\boxed{\Leftarrow}$ Sea (x^0, Φ_0, φ_0) un punto de silla asociado al lagrangiano $\mathfrak{L}$. Solo tenemos que comprobar que la familia de funciones en X $\{g_\lambda\}_{\lambda\in\Lambda}, \{h_\omega\}_{\omega\in\Omega}$, $\{-h_\omega\}_{\omega\in\Omega}$, $f - f(x^0)$ es infsup-convexa, ya que x^0 es solución del problema de programación infinita (4.4) gracias al Lema 4.6. Utilizando la segunda desigualdad de la definición de punto de silla, y teniendo en cuenta que x^0 es solución del problema de programación infinita (4.4), para cada x en X se tiene

$$f(x^0) \leq f(x) + \Phi_0(\{g_\lambda(x)\}_{\lambda\in\Lambda}) + \varphi_0(\{h_\omega(x)\}_{\omega\in\Omega}),$$

es decir,

$$0 \leq \Phi_0(\{g_\lambda(x)\}_{\lambda\in\Lambda}) + \varphi_0(\{h_\omega(x)\}_{\omega\in\Omega}) + f(x) - f(x^0). \tag{4.16}$$

En virtud del teorema de Hahn–Banach (los detalles pueden consultarse en [25, Lemma, p. 203]), podemos encontrar una descomposición de φ_0 en partes positiva y negativa, esto es, existen $\varphi_1, \varphi_2 \in \ell^\infty(\Omega)^*_+$ tales que $\varphi_0 = \varphi_1 - \varphi_2$, lo que nos permite reescribir (4.16) como

$$0 \leq \Phi_0(\{g_\lambda(x)\}_{\lambda\in\Lambda}) + \varphi_1(\{h_\omega(x)\}_{\omega\in\Omega}) + \varphi_2(\{-h_\omega(x)\}_{\omega\in\Omega}) + f(x) - f(x^0). \tag{4.17}$$

El funcional $L \in \ell^\infty(\Theta)^*_+$, siendo Θ el conjunto definido en la prueba del

Teorema (4.3), dado por $L := \dfrac{1}{\Phi_0(\mathbb{1}) + \varphi_1(\mathbb{1}) + \varphi_2(\mathbb{1}) + 1}(\Phi_0, \varphi_1, \varphi_2, 1)$ pertenece a Δ_Θ. Volviendo a (4.16) podemos escribir

$$0 \leq L(\{g_\lambda(x)\}_{\lambda\in\Lambda}, \{h_\omega(x)\}_{\omega\in\Omega}, -\{h_\omega(x)\}_{\omega\in\Omega}, f(x) - f(x^0)),$$

y como por el Lema 3.11

$$0 = \inf_{x\in X} \max\left\{\sup_{\lambda\in\Lambda} g_\lambda(x), \sup_{\omega\in\Omega} h_\omega(x), \sup_{\omega\in\Omega}(-h_\omega(x)), f(x) - f(x^0)\right\},$$

solo hay que razonar como en la prueba del Teorema 3.6 para deducir la infsup-convexidad de la familia $\{g_\lambda\}_{\lambda\in\Lambda}, \{h_\omega\}_{\omega\in\Omega}, \{-h_\omega\}_{\omega\in\Omega}, f - f(x^0)$ en X.

$\boxed{\Rightarrow}$ Para probar esta implicación basta aplicar el Teorema 4.3 para obtener $\Phi_0 \in \ell^\infty(\Lambda)^*_+$ y $\varphi_0 \in \ell^\infty(\Omega)^*$ verificando las condiciones Karush–Kuhn–Tucker, y en virtud del Teorema 4.7, (x^0, Φ_0, φ_0) es un punto de silla del lagrangiano $\mathfrak{L}$.

■

Como hemos señalado en la introducción del capítulo, estos teoremas suponen una generalización estricta en relación a los conocidos para otros conceptos más débiles que el de convexidad, la quasiconvexidad y la convexidad tipo Fan ([1, Corollary 3.1], [24, Theorems 1 y 2] y [26, 59, 56]). Concretamente, comprobemos que se puede encontrar un programa no lineal verificando las hipótesis de nuestros resultados, pero no los que exigen hipótesis de quasiconvexidad o convexidad en sentido de Fan.

Ejemplo 4.9 Sea $X := \mathbb{R}$; definimos las funciones $f, g_1 : \mathbb{R} \longrightarrow \mathbb{R}$ como

$$f(x) := \mid x(x - \sqrt{2}) \mid, \qquad (x \in \mathbb{R})$$

y

$$g_1(x) := \begin{cases} 1 & \text{si} \quad x \in \mathbb{Q} \\ -1 & \text{si} \quad x \in \mathbb{R} \setminus \mathbb{Q} \end{cases},$$

y para cada $n \geq 2$, $g_n : \mathbb{R} \longrightarrow \mathbb{R}$ por

$$g_n(x) := x - \sqrt{2} - \frac{1}{n}, \qquad (x \in \mathbb{R}).$$

Consideramos el problema de programación infinita

$$\inf_{x \in X_0} f(x), \tag{4.18}$$

siendo X_0 el conjunto

$$X_0 = \{x \in X : \sup_{n \geq 1} g_n(x) \leq 0\}.$$

Resulta inmediato probar que

$$X_0 = (\mathbb{R} \setminus \mathbb{Q}) \cap (-\infty, \sqrt{2}] \neq \emptyset$$

y que $x^0 = \sqrt{2} \in X_0$ es solución del problema de programación infinita (4.18). Por otro lado, $x^1 = -\sqrt{3}$ cumple la condición de Slater. Comprobemos en primer lugar que la familia $\{g_n\}_{n \geq 1}, f - f(x^0)$ es infsup-convexa en $\mathbb{R}$. A la vista del Lema 3.11 tenemos que

$$\inf_{x \in X} \max \left\{ \sup_{n \geq 1} g_n(x), f(x) - f(x^0) \right\} = 0$$

y por otro lado, la siguiente cadena de desigualdades es cierta –recuérdese que x^0 es solución del problema de programación infinita (4.18)– para cualesquiera $m \geq 1, \mathbf{t} \in \Delta_m$ y $x_1, ..., x_m \in X$:

$$\begin{aligned} 0 &\leq \sum_{j=1}^{m} t_j \left(f(x_j) - f(x^0) \right) \\ &\leq \max \left\{ \sup_{n \geq 1} \sum_{j=1}^{m} t_j g_n(x_j), \sum_{j=1}^{m} t_j \left(f(x_j) - f(x^0) \right) \right\}, \end{aligned}$$

y por tanto la familia es infsup-convexa en $\mathbb{R}$.

Ya que $X_0 = (\mathbb{R} \setminus \mathbb{Q}) \cap (-\infty, \sqrt{2}]$ no es un conjunto convexo, no tiene sentido hablar de los resultados clásicos para funciones quasiconvexas. Además, la familia $\{g_n\}_{n\geq 1}, f$ no es convexa en el sentido de Fan. Basta tomar $t = \frac{1}{2}, x_1 = 0$ y $x_2 = \sqrt{2}$ para llegar a la conclusión de que no existe $y_0 \in \mathbb{R}$ verificando simultáneamente las siguientes desigualdades

$$f(y_0) \leq \frac{1}{2}f(0) + \frac{1}{2}f(\sqrt{2}), \tag{4.19}$$

$$g_1(y_0) \leq \frac{1}{2}g_1(0) + \frac{1}{2}g_1(\sqrt{2}), \tag{4.20}$$

y

$$g_n(y_0) \leq \frac{1}{2}g_n(0) + \frac{1}{2}g_n(\sqrt{2}), \quad (n \geq 2). \tag{4.21}$$

Para que se cumpla (4.19), necesariamente $y_0 = 0$ o $y_0 = \sqrt{2}$. Que $y_0 = 0$ imposibilita el cumplimiento de (4.20) y en caso de que $y_0 = \sqrt{2}$, no se cumpliría la desigualdad (4.21).

■

Concluimos presentando la versión más general del teorema de Fritz John:

Teorema 4.10 *Consideremos X, Λ y Ω conjuntos tales que X es no vacío, $f : X \longrightarrow \mathbb{R}$ una función, $\{g_\lambda\}_{\lambda\in\Lambda}$ y $\{h_\omega\}_{\omega\in\Omega}$ dos familias de funciones reales definidas en X tales que para x en X, $\{g_\lambda(x)\}_{\lambda\in\Lambda} \in \ell^\infty(\Lambda)$ y $\{h_\omega(x)\}_{\omega\in\Omega} \in \ell^\infty(\Omega)$. Si el conjunto*

$$X_0 = \left\{ x \in X : \sup_{\lambda\in\Lambda} g_\lambda(x) \leq 0, \ \sup_{\omega\in\Omega} \mid h_\omega(x) \mid= 0 \right\}$$

es no vacío y x^0 *es una solución del problema de programación infinita* (4.4)*, entonces existen* $\rho \geq 0$, $\phi_0 \in \ell^\infty(\Lambda)^*_+$, $\varphi_1 \in \ell^\infty(\Omega)^*_+$ *y* $\varphi_2 \in \ell^\infty(\Omega)^*_+$ *tales que* $\rho + \phi_0(\mathbb{1}) + \varphi_1(\mathbb{1}) + \varphi_2(\mathbb{1}) = 1$ *y cumpliendo las condiciones Fritz John*

(FJ1)

$$\rho f + \phi_0(\{g_\lambda\}_{\lambda\in\Lambda}) + (\varphi_1 - \varphi_2)(\{h_\omega\}_{\omega\in\Omega}) \ \textit{alcanza su ínfimo en } X \textit{ en } x^0$$

y

(FJ2)

$$\phi_0(\{g_\lambda(x^0)\}_{\lambda\in\Lambda}) = 0 \quad \textit{y} \quad (\varphi_1 - \varphi_2)(\{h_\omega(x^0)\}_{\omega\in\Omega}) = 0,$$

si, y solo si, $\{g_\lambda\}_{\lambda\in\Lambda}, \{h_\omega\}_{\omega\in\Omega}, \{-h_\omega\}_{\omega\in\Omega}, f - f(x^0)$ *es una familia infsup-convexa en* X*.*

Demostración. $\boxed{\Rightarrow}$ De (FJ1) y (FJ2) deducimos que para cada $x \in X$ se cumple que

$$0 \leq \phi_0(\{g_\lambda(x)\}_{\lambda\in\Lambda}) + \varphi_1(\{h_\omega(x)\}_{\omega\in\Omega}) + \varphi_2(\{-h_\omega(x)\}_{\omega\in\Omega}) + \rho(f(x) - f(x^0)).$$

Para terminar esta implicación basta considerar el siguiente funcional $L \in \Delta_\Theta$

$$L := \left(\frac{\phi_0}{\alpha}, \frac{\varphi_1}{\alpha}, \frac{\varphi_2}{\alpha}, \frac{\rho}{\alpha}\right),$$

con

$$\alpha := \rho + \phi_0(\mathbb{1}) + \varphi_1(\mathbb{1}) + \varphi_2(\mathbb{1}),$$

y proceder como en la prueba de Teorema 4.8.

$\boxed{\Leftarrow}$ Razonando como en la demostración del Teorema 4.3, se garantiza la existencia de $L^* \in \ell^\infty(\Lambda)^*_+$, $L^{**} \in \ell^\infty(\Omega)^*_+$, $L^{***} \in \ell^\infty(\Omega)^*_+$ y $\rho \in \mathbb{R}_+$ tales que para todo x en X se satisface la desigualdad

$$\rho f(x^0) \leq \rho f(x) + L^*\left(\{g_\lambda(x)\}_{\lambda\in\Lambda}\right) + (L^{**} - L^{***})\left(\{h_\omega(x)\}_{\omega\in\Omega}\right).$$

Para probar (FJ1) y (FJ2) basta con reescribir $\Phi_0 := L^*$, $\varphi_1 := L^{**}$ y $\varphi_2 := L^{***}$ en la desigual anterior y tener en cuenta que $x^0 \in X$.

■

Capítulo 5

Conclusions and Open Problems

The original results contained in this thesis are framed in the context of two closely related fields: convex analysis, with its results of the Hahn–Banach type, and optimization. For them, the framework where they are optimal has been determined, a framework that depends on the infsup-convexity of an adequate family of functions, a concept that arises in minimax theory.

Starting from the Mazur–Orlicz theorem, we have established generalized versions of the theorems of the maximum and the supremum of König, which are sharp, in the sense that the validity of these statements is equivalent to the infsup-convexity of a certain family of functions. Both results have allowed us to prove, in the context of nonlinear programming, theorems of Lagrange multiplier, Karush–Kuhn–Tucker and Fritz John. In particular, this type of result has been studied for semi-infinite programs –arbitrary domain and finite number of constraints of the form of equality and inequality– and infinite with an arbitrary number of constraints. For an arbitrary infinite program, on the other hand, we have decided to prove the aforementioned theorems about optimization, equally optimal, directly from the theorem of separation of convex, since it is of common use in optimization.

With regard to open problems, let us begin with those involving only the Hahn–Banach theorem and its equivalent versions of König. Bearing in mind its application to programs that allow infinite values:

Problem I Is the maximum theorem of König valid when the functions involved take values in $\mathbb{R} \cup \{+\infty\}$? And that of the supremum?

The question is not trivial, taking into account that the tool used to establish our König's theorems is the Mazur–Orlicz theorem, which is not valid at full generality for sublinear functionals with infinite values, and it requires additional hypotheses.

On the other hand, the optimality of our mentioned above König results ultimately depends on the concept of infsup-convexity, whose origin lies in the minimax theory. If we also realize that many of the minimax inequalities, classic and recent, are equivalent reformulations of the Hahn–Banach theorem, we raise the question:

Problem II Is there any relation between our König theorems and any known minimax inequality?

Or even:

Problem III Can we derive a new minimax inequality from our König theorems?

Regarding the results on nonlinear programming, it is possible to consider, in view of Problem I:

Problem IV Do the theorems of Lagrange multiplier, Karush–Kuhn–Tucker or Fritz John admit infinite-valued versions?

Our future research could be determined by this other question, which is beyond our initial purposes, but which implies a natural extension of the results developed in this thesis:

Problem V Is there any kind of vector optimization problem for which it is possible to obtain optimal results along the lines of those established in this thesis?

Obviously, first we should deal –probably in the context of vector equilibrium problems– with generalizing the notion of infsup-convexity.

Bibliografía

[1] K.J. Arrow, A.C. Enthoven, *Quasi-concave programming*, Econometrica **29** (1961), 779–800.

[2] S. Boyd, L. Vandenberghe, *Convex optimization*, Cambridge University Press, Cambridge, 2008.

[3] H. Brezis, *Functional analysis, Sobolev spaces and partial differential equations*, Universitext, Springer, 2011.

[4] O. Brezhnevaa, A.A. Tret'yakov, *An elementary proof of the Karush–Kuhn–Tucker theorem in normed linear spaces for problems with a finite number of inequality constraints*, Optimization **60** (2011), 613–618.

[5] T.D. Chuong, *L-invex-infine functions and applications*, Nonlinear Analysis **75** (2012), 5044–5052.

[6] A. Dax, V.P. Sreedharan, *Theorems of the alternative and duality*, Journal of Optimization Theory and Applications **94** (1997), 561–590.

[7] N. Dinh, V. Jeyakumar, *Farkas' lemma: three decades of generalizations for mathematical optimization*, Top **22** (2014), 1–22.

[8] N. Dihn, M.A. Goberna, M.A. López, T.H. Mo, *From the Farkas lemma to the Hahn–Banach theorem*, SIAM Journal on Optimization **24** (2014), 678–701.

[9] N. Dinh, T.H. Mo, *Farkas lemma for convex systems revisited and applications to sublinear-convex optimization problems*, Vietnam Journal of Mathematics **43** (2015), 297–321.

[10] A.R. Doagooei, *Farkas-type theorems for positively homogeneous systems in ordered topological vector spaces*, Nonlinear Analysis **75** (2012), 5541–5548.

[11] K. Fan, *Minimax theorems*, Proceedings of the National Academy of Sciences of the United States of America **39** (1953), 42–47.

[12] K. Fan, I. Glicksberg, A. J. Hoffman *Systems of inequalities involving convex functions*, The American University, University of Notre Dame and National Bureau of Standards, 1957.

[13] J. Farkas, *Theorie der einfachen ungleichungen*, Journal für die reine und angewandte Mathematik **124** (1902), 1–27.

[14] W. Fenchel, D.W. Blackett, *Convex cones, sets and functions*, Princeton University, Department of Mathematics, Princeton 1953.

[15] B. de Finetti, *Sulle stratificazioni convesse*, Annali di Matematica Pura ed Applicata **30** (1949), 173–183.

[16] F. Flores-Bazán, *Fritz John necessary optimality conditions of the alternative-type*, Journal of Optimization Theory and Applications **161** (2014), 807–818.

[17] F. Flores-Bazán, C. Vera, *Gordan-type alternative theorems and vector optimization revisited*. Recent Developments in Vector Optimization, 29–59, Vector Optimization **1**, Springer, Berlin, 2012.

[18] A.I. Garralda-Guillem, M. Ruiz Galán, *Mixed variational formulations in locally convex spaces*, Journal of Mathematical Analysis and Applications **414** (2014), 825-849.

[19] G. Giorgi, A. Guerraggio, A., J. Thierfelder, *Mathematics of optimization: smooth and nonsmooth case*. Elsevier Science B.V., Amsterdam (2004).

[20] G. Giorgi, T.H. Kjeldsen (Eds.), *Traces and emergence of nonlinear programming*, Birkhäuser/Springer Basel AG, Basel, 2014.

[21] P. Gordan, *Uber die auflösung linearer gleichungen mit reelen coefficienten*, Mathematische Annalen **6** (1873), 23–28.

[22] J. Grzybowski, H. Przybycień, R. Urbański, *On Simon's version of Hahn–Banach–Lagrange theorem*, Function Spaces X, 99–104, Banach Center Publications **102**, Polish Academy of Sciences, Institute of Mathematics, Warsaw, 2014.

[23] A. Guerraggio, E. Molho, *The origins of quasi-concavity: a development between mathematics and economics*, Historia Mathematica **31** (2004), 62–75.

[24] M. Hayasi, H. Komiya, *Perfect duality for convexlike programs*, Journal of Optimization Theory and Applications **38** (1982), 179–189.

[25] R.B. Holmes, *Geometric functional analysis and its applications*, Graduate Texts in Mathematics **24**, Springer–Velag, New York, 1975.

[26] T. Illés, G. Kassay, *Theorems of the alternative and optimality conditions for convexlike and general convexlike programming*, Journal of Optimization Theory and Applications **101** (1999), 243–257.

[27] K. Ito, K. Kunisch, *Karush–Kuhn–Tucker conditions for nonsmooth mathematical programming problems in function spaces*, SIAM Journal on Control and Optimization **49** (2011), 2133–2154.

[28] V. Jeyakumar, G.M. Lee, G. Li, *Global optimality conditions for classes of non-convex multi-objective quadratic optimization problems.* Variational analysis and generalized differentiation in optimization and control, 177–186, Springer Optimization and its Applications **47**, Springer, New York, 2010.

[29] Y. Jin, B. Kalantari, *A procedure of Chvátal for testing feasibility in linear programming and matrix scaling*, Linear Algebra and its Applications **416** (2006), 795–798.

[30] F. John, *Extremum problems with inequalities as subsidiary conditions*, Studies and Essays Presented to R. Courant on his 60th Birthday, 1948, 187–204. Interscience Publishers, Inc., New York, 1948.

[31] W. Karush, *Minima of functions of several variables with inequalities as side conditions*, MSc Thesis, Department of Mathematics, University of Chicago, 1939.

[32] G. Kassay, J. Kolumbán, *On a generalized sup-inf problem*, Journal of Optimization Theory and Applications **91** (1996), 651-670.

[33] H. König, *Über das von Neumannsche minimax-theorem*, Archiv der Mathematik **19** (1968), 482–487.

[34] H. König, *Sublineare funktionale*, Archiv der Mathematik **23** (1972), 500–508.

[35] H. König, *Sublinear functionals and conical measures*, Archiv der Mathematik **77** (2001), 56-64.

[36] H.W. Kuhn, A.W. Tucker, *Nonlinear programming*, Proceedings of the Second Berkeley Symposium on Mathematical Statistics and Probability, 1950, pp. 481–492. University of California Press, Berkeley and Los Angeles, 1951.

[37] J.B. Lasserre, *Duality and a Farkas lemma for integer programs*. Structure and Applications, 15–39, Springer Optimization and Its Applications **32**, Springer, New York, 2009.

[38] F.-C. Liu, *Mazur–Orlicz equality*, Studia Mathematica **189** (2008), 53-63.

[39] S. Mazur, W. Orlicz, *Sur les spaces metriques lineaires*, Studia Mathematica **13** (1953), 137-179.

[40] P. Montiel López, M. Ruiz Galán, *Nonlinear programming via König's maximum theorem*, Journal of Optimization Theory and Applications **170** (2016), 838-852.

[41] P. Montiel López, M. Ruiz Galán, *Revisiting the Hahn–Banach theorem and nonlinear infinite programming*, Journal of Mathematical Analysis and Applications, **455** (2017), 1037-1050.

[42] P. Montiel López, M. Ruiz Galán, *Infinite programming and theorems of the alternative*, preprint.

[43] P. Montiel López, M. Ruiz Galán, *Recent convex tools for nonlinear programming*, Porceedings of the 17$^{\text{th}}$ International Conference on Computational and Mathematical Methods in Science and Engineering, Rota (Cádiz), July 4$^{\text{th}}$–8$^{\text{th}}$ 2017, ISBN 978–84–617–8694–7.

[44] V. Pták, *On a theorem of Mazur and Orlicz*, Studia Mathematica **15** (1956), 365–366.

[45] P.J. Rabier, *Quasiconvexity and density topology*, Canadian Mathematical Bulletin **57** (2014), 178–187.

[46] R.T. Rockafellar, *Convex analysis*, Princeton University Press, Princeton, 1970.

[47] M. Ruiz Galán, *An intrisic notion of convexity for minimax*, Journal of Convex Analysis **21** (2014), 1105-1139.

[48] M. Ruiz Galán, *The Gordan theorem and its implications for minimax theory*, Journal of Nonlinear and Convex Analysis **17** (2016), 2385-2405.

[49] M. Ruiz Galán, *A sharp Lagrange multiplier theorem for nonlinear programs*, Journal of Global Optimization **65** (2016), 513-530.

[50] S. Simons, *Minimal sublinear functionals*, Studia Mathematica **37** (1970), 37–56.

[51] S. Simons, *Maximinimax, minimax, and antiminimax theorems and a result of R. C. James*, Pacific Journal of Mathematics **Vol. 40, No. 3**, (1972), 709-718.

[52] S. Simons, *The Hahn–Banach–Lagrange theorem*, Optimization **56** (2007), 149–169.

[53] S. Simons, *Bootstrapping the Mazur–Orlicz–König theorem*, Journal of Convex Analysis **25** (2018), to appear.

[54] A. Stefanescu, *A theorem of the alternative and a two-function minimax theorem*, Journal of Applied Mathematics **2004:2** (2004), 167-177.

[55] C. Sun, *The Mazur–Orlicz theorem for convex functionals*, Journal of Convex Analysis **24** (2017), to appear.

[56] S. Suzuki, D. Kuroiwa, *Optimality conditions and the basic constraint qualification for quasiconvex programming*, Nonlinear Analysis **74** (2011), 1279–1285.

[57] H. Uzawa, *The Kuhn–Tucker theorem in concave programming*. In: K.J. Arrow, L. Hurwicz, H. Uzawa (Eds.), *Studies in linear and nonlinear programming*, Stanford University Press, Stanford, pp.: 32–37, 1958.

[58] J. von Neumann, *Zur theorie der gesellschaftsspiele*, Mathematische Annalen **100** (1928), 295–320.

[59] R. Zeng, *A general Gordan alternative theorem with weakened convexity and its application*, Optimization **51** (2002), 709–717.