Escuela Politécnica Nacional

Nombre: Daniel Lara **Fecha:** Enero, 2024
Tema: Teoría de funciones convexas conjugadas **Materia:** -

Las presentes notas corresponden a un desarrollo de los detalles para la demostración del Teorema de Fenchel-Rockafellar con base en [1]. Se agradece al lector dirigir cualquier observación sobre este documento al correo `daniel.lara@alephsub0.org`

Definición 1: –Mapa convexo–

Sean E un espacio vectorial, $C \subseteq E$ no vacío y convexo y $\phi\colon C \subseteq E \to \mathbb{R}$, decimos que el mapa ϕ es *convexo* si

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

Definición 2

Sean E un espacio vectorial real y $\phi\colon E \to (-\infty, +\infty]$.

- Al conjunto:
  $$D(\phi) = \{x \in E : \phi(x) < \infty\},$$
  le llamamos *dominio* de ϕ y escribimos $\phi \not\equiv +\infty$, cuando $D(\phi) \neq \varnothing$.
- El *epígrafo* de ϕ es el conjunto:
  $$\operatorname{epi}\phi = \{(x, \lambda) \in E \times \mathbb{R} : \phi(x) \leq \lambda\}.$$
- Para cada $\lambda \in \mathbb{R}$ escribimos:
  $$[\phi \leq \lambda] = \{x \in E : \phi(x) \leq \lambda\}.$$

Lema 1. Sea $C \subseteq E$ un conjunto convexo, entonces $\operatorname{Int} C$ es convexo. Además, si $\operatorname{Int} C \neq \varnothing$, entonces

$$\overline{C} = \overline{\operatorname{Int} C}.$$

Lema 2. Sean E un espacio vectorial real normado, $A, B \subseteq E$ no vacíos, convexos y disjuntos con alguno de ellos abierto. Entonces existe un hiperplano cerrado que separa $\overline{A}$ y B.

Lema 3. Sean E un espacio vectorial normado, $f\colon E \to \mathbb{R}$ continua en $a \in E$ y $c \in \mathbb{R}$ tal que $f(a) < c$, entonces existe $r > 0$ tal que

$$\forall x \in B_r(a) \quad f(x) < c.$$

Definición 3: –Función conjugada–

Sea $\phi\colon E \to (-\infty, +\infty]$ una función tal que $\phi \not\equiv +\infty$, i.e. $(\operatorname{Dom}(\phi) \neq \varnothing)$. Se define la función conjugada de ϕ como:

$$\begin{aligned} \phi^*\colon E^* &\longrightarrow (-\infty, +\infty] \\ f &\longmapsto \sup_{x \in E} \{\langle f, x\rangle - \phi(x)\}. \end{aligned}$$

Observación 1.

1. Note que ϕ^* es convexa y semi continua inferiormente en E^*.
2. Se tiene la siguiente desigualdad
   $$\langle f, x\rangle \leq \phi(x) + \phi^*(f) \quad \forall x \in E, \quad \forall f \in E^*.$$

3. En algunos contextos se denomina transformada de Legendre a la función conjugada ϕ^*.

Teorema 4: –Fenchel-Rockafellar–
Sean E un espacio vectorial normado y $\phi, \psi \colon E \to (-\infty, +\infty]$ mapas convexos. Si existe $x_0 \in D(\phi) \cap D(\psi)$ tal que ϕ es continuo en x_0, entonces

$$\begin{aligned}
\inf_{x\in E} \{\phi(x) + \psi(x)\} &= \sup_{f\in E^*} \{-\phi^*(-f) - \psi^*(f)\} \\
&= \max_{f\in E^*} \{-\phi^*(-f) - \psi^*(f)\} \\
&= -\min_{f\in E^*} \{\phi^*(-f) + \psi^*(f)\}.
\end{aligned}$$

Demostración. Definamos:

$$a = \inf_{x\in E} \{\phi(x) + \psi(x)\} \qquad \text{y} \qquad b = \sup_{f\in E^*} \{-\phi^*(-f) - \psi^*(f)\}.$$

`P.D. Vamos a probar que` $a \leq b$ `y` $b \leq a$.
`P.D. Primero probemos que` $a \leq b$.
Idea: Utilizar las propiedades de ínfimo y supremo.
Dado que

$$\begin{aligned}
-\phi^*(-f) - \psi^*(f) &= -\sup_{x\in E} \{-\langle f, x\rangle - \phi(x)\} - \sup_{x\in E} \{\langle f, x\rangle - \psi(x)\} && \text{Definición de función conjugada} \\
&= \inf_{x\in E} \{f(x) + \phi(x)\} + \inf_{x\in E} \{-f(x) + \psi(x)\} && \text{Propiedades del ínfimo y supremo} \\
&\leq \inf_{x\in E} \{\phi(x) + \psi(x)\}.
\end{aligned}$$

Por lo tanto

$$-\phi^*(-f) - \psi^*(f) \leq \inf_{x\in E} \{\phi(x) + \psi(x)\},$$

con lo cual

$$\sup_{f\in E^*} \{-\phi^*(-f) - \psi^*(f)\} \leq \inf_{x\in E} \{\phi(x) + \psi(x)\} \quad \Longrightarrow \quad b \leq a.$$

`P.D. Con lo anterior, mostremos que lo requerido se cumple trivialmente para el caso` $a = -\infty$.
Por otra parte, si $a = -\infty$, entonces, puesto que $b \leq a$, $b = -\infty$, es decir,

$$\inf_{x\in E} \{\phi(x) + \psi(x)\} = \sup_{f\in E^*} \{-\phi^*(-f) - \psi^*(f)\}.$$

Además, puesto que se «alcanza» el supremo y por las propiedades de máximo

$$\begin{aligned}
\sup_{f\in E^*} \{-\phi^*(-f) - \psi^*(f)\} &= \max_{f\in E^*} \{-\phi^*(-f) - \psi^*(f)\} \\
&= \max_{f\in E^*} \{-(\phi^*(-f) + \psi^*(f))\} \\
&= -\min_{f\in E^*} \{\phi^*(-f) + \psi^*(f)\}.
\end{aligned}$$

`P.D. Por lo anterior, podemos suponer que` $a \in \mathbb{R}$. `Así, mostremos que` $a \leq b$.
Así, supongamos que $a \in \mathbb{R}$ y tomemos,

$$C = \text{epi}\,(\phi) = \{(x, \lambda) \in E \times \mathbb{R} : \phi(x) \leq \lambda\}.$$

`P.D. Mostremos que Int(C) es no vacío.`
Como $x_0 \in D(\phi)$, por hipótesis, entonces $\phi(x_0) \in \mathbb{R}$ y por lo tanto $(x_0, \phi(x_0)+1) \in C$; más aún, como ϕ es continua en x_0, entonces existe $r^* > 0$ tal que

$$\phi(x) < \phi(x_0) + 1 \quad \forall x \in B_{r^*}(x_0),$$

así puesto que $B_{r^*}(x_0) \times]\phi(x_0), \phi(x_0)+1[$ es un abierto, entonces por construcción se tiene que

$$(x_0, \phi(x_0)+1) \in B_{r^*}(x_0) \times]\phi(x_0), \phi(x_0)+1[\subseteq C,$$

por lo tanto, $\operatorname{Int} C \neq \varnothing$. Por otro lado, puesto que ϕ es convexo, entonces C es convexo y por el *Lema 1* $A = \operatorname{Int}(C)$ es convexo.
Ahora, definamos

$$B = \{(x, \lambda) \in E \times \mathbb{R};\, \lambda \leq a - \psi(x)\}.$$

`P.D. Mostremos que B es no vacío.`
Como $x_0 \in D(\psi)$, por hipótesis, entonces $\psi(x_0) \in \mathbb{R}$ y por lo tanto $(x_0, a - \psi(x_0)) \in B$, con lo que se sigue que $B \neq \varnothing$.
`P.D. Mostremos que A y B son disjuntos.`
Tomemos $(x_0, \lambda_0) \in A$, arbitrario pero fijo, por definición del epígrafo, $\lambda_0 > \phi(x_0)$; además, por la definición del ínfimo $\phi(x_0) + \psi(x_0) \geq a$, y así

$$\lambda_0 > \phi(x_0) \geq a - \psi(x_0) \quad \Longrightarrow \quad \lambda_0 > a - \psi(x_0),$$

de esta manera $(x_0, \lambda_0) \notin B$; con lo que, como $(x_0, \lambda_0) \in A$ es arbitrario, se sigue que $A \cap B = \varnothing$.
`P.D. Mostremos que B es convexo.`
Ahora, notemos que B es convexo; en efecto, tomemos (x, α) y $(y, \beta) \in B$, vamos a probar

$$(x, \alpha)t + (1-t)(y, \beta) \in B \quad \forall t \in (0,1). \tag{1}$$

Notemos que para $t = 0$ y $t = 1$ se tiene el resultado. Así sea $t \in]0,1[$, cualquiera; notemos que

$$t\alpha \leq ta - t\psi(x) \qquad \text{y} \qquad (1-t)\beta \leq (1-t)a - (1-t)\psi(y).$$

Por tanto

$$t\alpha + (1-t)\beta \leq ta - t\psi(x) + (1-t)a - (1-t)\psi(y) \leq a - (t\psi(x) + (1-t)\psi(y))$$

y como ψ es convexa, se tiene que

$$\psi(tx + (1-t)y) \leq t\psi(x) + (1-t)\psi(y) \quad \Longrightarrow \quad -(t\psi(x) + (1-t)\psi(y)) \leq -\psi(tx + (1-t)y),$$

por lo tanto

$$t\alpha + (1-t)\beta \leq a - (t\psi(x) + (1-t)\psi(y)) \leq a - \psi(tx + (1-t)y).$$

Así, como t es arbitrario, hemos probado que B es convexo.
`P.D. Existe un hiperplano cerrado que separa a C y B, y encontrar una forma explícita para el mismo.`
Usando el teorema de Hanh-Banach, primera forma geométrica, existe un hiperplano $\mathscr{H} = [h = \alpha]$ que separa A y B, más aún por el *Lema 2*, se tiene que $\mathscr{H}$ también separa a $\overline{A}$ y B y dado que $C \subseteq \overline{C} = \overline{A}$, por el *Lema 1*. Entonces $\mathscr{H}$ también separa a C y B. Más aún, por la forma del espacio[1], se sigue que existen $f \in E^*$ y $k \in \mathbb{R}$ tales que h tiene la siguiente forma:

$$h(x, \lambda) = \langle f, x \rangle + k\lambda \quad \forall (x, \lambda) \in E \times \mathbb{R}.$$

[1]Usando el hecho de que los espacios $E^* \times E^*$ y $(E \times E)^*$ son isomorfos.

Por lo tanto, tenemos que

$$\langle f,\, x\rangle + k\lambda \geq \alpha \quad \forall (x,\lambda) \in C, \tag{2}$$
$$\langle f,\, x\rangle + k\lambda \leq \alpha \quad \forall (x,\lambda) \in B. \tag{3}$$

`P.D. Mostremos que` $k \geq 0$.
Tomando $x = x_0$ en (2), se sigue que para $n \in \mathbb{N}$,

$$\langle f,\, x_0\rangle + k\,(n\phi(x_0)+1)) \geq \alpha \quad \iff \quad k \geq \frac{\alpha - \langle f,\, x_0\rangle}{n\phi(x_0)+1} \quad \forall n \in \mathbb{N}.$$

y cuando $n \to \infty$ tenemos que $k \geq 0$.
`P.D. Mostremos que` $k > 0$.
En efecto, por contradicción, supongamos que $k = 0$, entonces como el mapa asociado al hiperplano $\mathscr{H}$ no puede ser idénticamente nulo, por hipótesis, se sigue que

$$\|f\| \neq 0. \tag{4}$$

Puesto que $k = 0$, se tiene que por (2) y (3),

$$\langle f,\, x\rangle \geq \alpha \quad \forall x \in D(\phi),$$
$$\langle f,\, x\rangle \leq \alpha \quad \forall x \in D(\psi).$$

Por otra parte, dado que $x_0 \in D(\phi)$ y ϕ es continua en x_0, entonces existe un ε_0 tal que $B_{\varepsilon_0}(x_0) \subseteq D(\phi)$ y por lo tanto como $B_{\varepsilon_0}(x_0) = x_0 + \varepsilon_0 z$ con $z \in B_1(0)$, entonces

$$\langle f,\, x_0 + \varepsilon_0 z\rangle \geq \alpha \quad \forall z \in B_1(0).$$

Puesto que $z \in B_1(0)$, entonces $-z \in B_1(0)$ y por lo tanto

$$\begin{aligned}
\langle f,\, x_0 - \varepsilon_0 z\rangle \geq \alpha \quad &\Longrightarrow \quad \langle f,\, x_0\rangle + \langle f,\, -\varepsilon_0 z\rangle \geq \alpha && \forall z \in B_1(0)\\
&\Longrightarrow \quad \langle f,\, x_0\rangle - \varepsilon_0 \langle f,\, z\rangle \geq \alpha && \forall z \in B_1(0)\\
&\Longrightarrow \quad \langle f,\, x_0\rangle - \alpha \geq \varepsilon_0 \langle f,\, z\rangle && \forall z \in B_1(0)\\
&\Longrightarrow \quad \frac{\langle f,\, x_0\rangle - \alpha}{\varepsilon} \geq \langle f,\, z\rangle && \forall z \in B_1(0).
\end{aligned}$$

Por lo tanto, se tiene que

$$\|f\| = \sup_{z \in B_1(0)} \langle f,\, z\rangle \leq \frac{\langle f,\, x_0\rangle - \alpha}{\varepsilon_0}$$

Además, dado que $x_0 \in D(\psi)$, entonces se sigue que $0 \geq \alpha - \langle f,\, x_0\rangle$ y así

$$\|f\| = 0.$$

pero esto contradice (4). Así, se concluye que $k > 0$.
Comentario: Con lo probado anteriormente, vamos a continuar con la demostración del teorema.
De esta manera, por (2), tenemos que

$$\begin{aligned}
\langle f,\, x\rangle + k\lambda \geq \alpha \quad &\Longrightarrow \quad \frac{\langle f,\, x\rangle}{k} + \lambda \geq \frac{\alpha}{k} && \forall (x,\lambda) \in C\\
&\Longrightarrow \quad \left\langle -\frac{f}{k},\, x\right\rangle - \lambda \leq -\frac{\alpha}{k} && \forall (x,\lambda) \in C.
\end{aligned}$$

Puesto que $-\frac{\alpha}{k}$ es una cota superior para $\left\langle -\frac{f}{k}, x \right\rangle - \phi(x)$, entonces

$$\left\langle -\frac{f}{k}, x \right\rangle - \phi(x) \leq -\frac{\alpha}{k} \quad \forall x \in E \quad \Longrightarrow \quad \sup_{x\in E}\left\{ \left\langle -\frac{f}{k}, x \right\rangle - \phi(x) \right\} \leq -\frac{\alpha}{k}$$

y así, por definición de función conjugada, tenemos que

$$\phi^*\left(-\frac{f}{k}\right) \leq -\frac{\alpha}{k}. \tag{5}$$

Por otra parte, de (3), se sigue que

$$\langle f, x\rangle + \lambda k \leq \alpha \quad \Longrightarrow \quad \frac{\langle f, x\rangle}{k} + \lambda \leq \frac{\alpha}{k} \quad \forall (x, \lambda) \in B$$

así, para $x \in D(\psi)$, tenemos que $(x, a - \psi(x)) \in B$ y por lo tanto $\lambda = a - \psi(x)$, se sigue que

$$\left\langle \frac{f}{k}, x \right\rangle + a - \psi(x) \leq \frac{\alpha}{k} \quad \Longrightarrow \quad \left\langle \frac{f}{k}, x \right\rangle - \psi(x) \leq \frac{\alpha}{k} - a \quad \forall x \in D(\psi).$$

Más aún, si ahora $x \notin D(\psi)$, entonces $\psi(x) = +\infty$ y por lo tanto

$$\left\langle \frac{f}{k}, x \right\rangle - \psi(x) \leq \frac{\alpha}{k} - a \quad \forall x \in E.$$

Así, puesto que $\frac{\alpha}{k} - a$ es una cota superior para $\left\langle \frac{f}{k}, x \right\rangle - \psi(x)$, entonces

$$\left\langle \frac{f}{k}, x \right\rangle - \psi(x) \leq \frac{\alpha}{k} - a \quad \forall (x, \lambda) \in B \quad \Longrightarrow \quad \sup_{x\in E}\left\{ \left\langle \frac{f}{k}, x \right\rangle - \psi(x) \right\} \leq \frac{\alpha}{k} - a,$$

y así, en consecuencia

$$\psi^*\left(\frac{f}{k}\right) \leq \frac{\alpha}{k} - a. \tag{6}$$

De esta manera, por (5) y (6) tenemos que

$$\phi^*\left(-\frac{f}{k}\right) \leq -\frac{\alpha}{k} \qquad \text{y} \qquad \psi^*\left(\frac{f}{k}\right) \leq \frac{\alpha}{k} - a.$$

De donde, uniendo las desigualdades anteriores se sigue que

$$\phi^*\left(-\frac{f}{k}\right) + \psi^*\left(\frac{f}{k}\right) \leq -a \quad \Longrightarrow \quad -\phi^*\left(-\frac{f}{k}\right) - \psi^*\left(\frac{f}{k}\right) \geq a.$$

Por otra parte, por la definición de b, tenemos que

$$b \geq \phi^*\left(-\frac{f}{k}\right) - \psi^*\left(\frac{f}{k}\right) \geq a,$$

y como habíamos probado que $b \leq a$, entonces se sigue que $a = b$, es decir:

$$\inf_{x\in E}\{\phi(x) + \psi(x)\} = \sup_{f\in E^*}\{-\phi^*(x)(-f) - \psi^*(f)\}.$$

`P.D. Mostremos las equivalencias restantes.`

Además, puesto que la imagen de ϕ y ψ incluye a $+\infty$, entonces el máximo se alcanza y se tiene que

$$\sup_{f\in E^*}\{-\phi^*(x)(-f) - \psi^*(f)\} = \max_{f\in E^*}\{-\phi^*(x)(-f) - \psi^*(f)\};$$

de donde, por propiedades del máximo, tenemos que

$$\max_{f\in E^*}\{-\phi^*(x)(-f)-\psi^*(f)\}=-\min_{f\in E^*}\{-\phi^*(x)(-f)-\psi^*(f)\}\,.$$

Así, hemos probado que

$$\begin{aligned}\inf_{x\in E}\{\phi(x)+\psi(x)\} &= \sup_{f\in E^*}\{-\phi^*(x)(-f)-\psi^*(f)\}\\ &= \max_{f\in E^*}\{-\phi^*(x)(-f)-\psi^*(f)\}\\ &= -\min_{f\in E^*}\{-\phi^*(x)(-f)-\psi^*(f)\}\,,\end{aligned}$$

como se quería. □

Referencias

[1] Haim Brezis. *Functional Analysis, Sobolev Spaces and Partial Differential Equations*. Springer New York, 2010. DOI: 10.1007/978-0-387-70914-7.