A structural version of the theorem of Hahn–Banach.

Jan Brinkhuis

Econometrisch Instituut
Erasmus Universiteit Rotterdam
Postbus 1738
3000 DR Rotterdam
The Netherlands
Email: brinkhuis@few.eur.nl

Econometric Institute Report EI-2001-42

1 Introduction

We consider one of the basic results of functional analysis, the classical theorem of Hahn–Banach. This theorem gives the existence of a continuous linear functional on a given normed vectorspace extending a given continuous linear functional on a subspace with the same norm. In this paper we generalize this *existence* theorem to a result on the *structure* of the set of all these extensions. We establish a bijection between this set and the set of nonzero vectors in the conjugate of an explicit convex cone. Moreover we generalize the theorem of Hahn–Banach simultaneously in another direction. Instead of extending linear continuous functionals on normed vectorspaces we extend elements of conjugates of convex cones. These convex cones are allowed to lie in arbitrary vectorspaces; the only assumption which we have to make about these convex cones is that they possess a relatively internal point. Furthermore we solve some technical problems which arise because we do not restrict to pointed solid convex cones.

As an illustration of the new structural insight, let us look at the usual proof of the theorem of Hahn–Banach. We recall that it proceeds by repeated extension of a given functional to bigger and bigger linear subspaces, by way of jumps of codimension one. The main point of the proof is to show that this process of extension never gets stuck. Then it follows from Zorn's lemma that there exists an appropriate extension of the given functional to the whole space. Our structural result allows to consider arbitrary jumps; not only of dimension one. Moreover it gives a precise description of the freedom one has in each jump. For a one dimensional jump, the determination of this freedom amounts to the computation of the conjugate of a convex cone in the two–dimensional plane.

2 Standard results and definitions

Let V be a vectorspace (we consider only *real* vectorspaces). For a linear subspace W of V the quotientspace V/W can be characterized as follows: there is a pair $(V/W, \pi)$ consisting of a vector space V/W and a surjective linear mapping π from V to V/W with kernel W. The codimension of W in V, denoted by $\text{cod}\,_V W$, is defined to be the dimension of the quotient space V/W.

The dual vector space V' is the vector space consisting of all linear functionals on V. Taking duals preserves exactness. Explicitly, for each linear subspace W of a vector space V one can view $(V/W)'$ as a linear subspace of V' and the quotient space is isomorphic to W'; to be more precise, the inflation map inf from $(V/W)'$ to V' is injective, the restriction map res from V' to W' is surjective and the image of inf equals the kernel of res. Moreover the natural map i from V to its biconjugate $V'' = (V')'$ given by $i(v)(\varphi) = \varphi(v)\ \forall\ v \in V\ \forall \varphi \in V'$ is injective.

Let V be a vector space and let S be an arbitrary subset of V. The linear span of S, denoted by span S, is the smallest linear subspace of V containing S. An element s of S is a *relatively internal point* of S if

$$\forall d \in \text{span } (S)\ \exists \varepsilon > 0\ \forall \tau \in (0, \varepsilon] : s' + \tau d \in S.$$

The set of all relatively internal points of S is denoted by rint S. If S is moreover a convex set, then an element $s \in S$ belongs to rint S if and only if $\forall d \in \operatorname{span} S\ \exists \varepsilon > 0 : s + \varepsilon d \in S$.

The –algebraic– closure $\overline{S}$ of S is defined to be the set consisting of all vectors $v \in V$ such that

$$\exists t \in S \forall \alpha \in (0,1)\ \alpha v + (1-\alpha)s \in S.$$

The set S is called –algebraically– closed if $\overline{S} = S$. We observe that $\overline{S}$ is closed for each set S.

Let V be a vector space. A nonempty subset C of V is called a convex cone in V if it is closed under addition and under multiplication with nonnegative scalars. The convex cone C is called *pointed* if it does not contain any line through 0. One calls C *solid* if span $C = V$

Let V be a vector space and let C be a convex cone in V. The *conjugate cone* or *dual cone* $(C,V)'$, corresponding to this cone C, is the convex cone in the dual vector space V' which consists of all functionals $\varphi : V \to \mathbb{R}$ satisfying $\varphi(c) \geq 0\ \forall c \in C$.

Now we recall the result that each convex cone is the sum of a pointed convex cone and a linear subspace. We also show how this decomposition carries over to the conjugate cone.

Lemma 2.1. *Let V be a vector space and C a convex cone in V. Then the following statements hold true.*

- **(i)** *There are linear subspaces W_1, W_2, W_3 of V and there exists a solid, pointed convex cone D in W_2 such that $V = W_1 \oplus W_2 \oplus W_3$ and $C = W_1 + D$. Here W_1 (resp. $W_1 \oplus W_2$) is uniquely determined as the maximal (resp. minimal) linear subspace of V which is contained in C (resp. which contains C).*

- **(ii)** *$(\overline{C}, V)$ equals (C,V).*

- **(iii)** *Assume $\overline{C} = C$ and let W_1, W_2, W_3 and D be as in (i). Then $V' = W_1' \oplus W_2' \oplus W_3'$ and $(C,V)' = W_3' + (D, W_2)'$; moreover the convex cone $(D, W_2)'$ is pointed.*

This result suggests that for the analysis of the conjugate cone it is useful to view the set of its nonzero elements $(C,V)' \setminus \{0\}$ as the union of two disjoint subsets: its *singular part* $S(C,V$ defined by

$$S(C,V) := \{\varphi \in V' \setminus \{0\} \mid \varphi(0) = 0\ \forall c \in C\}$$

and its *regular part* $R(C, V)$ defined by

$$R(C, V := (C, V)' \setminus S(C, V) \cup \{0\}.$$

The connection between these definition and the lemma above is as follows: $S(C, V) \cup \{0\}$ is the maximal linear subspace of V' which is contained in $(C, V)'$.

The conjugate cone and its regular and singular elements allow the following geometrical interpretation. The rays of the conjugate cone (C, V) correspond bijectively to the linear subspaces in V of codimension 1 with C on one of its two sides together with a choice of side which contains C: for each $\varphi \in (C, V) \backslash \{0\}$ we associate to the ray $\mathbb{R}^+\varphi$ the linear subspace $\{v \in V | \varphi(v) = 0\}$ and its side $\{v \in V | \varphi(v) \geq 0\}$. For each $\varphi \in (C, V) \backslash \{0\}$ one has $\varphi \in S(C, C)$ precisely if $C \subseteq \ker \varphi$. If $\operatorname{rint} C \neq \varnothing$ one has $\varphi \in R(C, V)$ precisely if $\ker \varphi \cap \operatorname{rint} C = \varnothing$.

Now we are going to recall the well-known classification of closed convex cones in a two-dimensional vector space. Moreover we recall the explicit description of their conjugates. It is convenient to do this for the complex plane $\mathbb{C}$ viewed as a two-dimensional vector space. We define for $\phi_1, \phi_2 \in \mathbb{R}$ with $0 \leq \phi_2 - \phi_1 \leq \pi$ the convex cone C_{ϕ_1,ϕ_2} in $\mathbb{C}$ by $C_{\phi_1,\phi_2} = \{re^{i\phi} | r \in \mathbb{R}^+$ and $\phi_1 \leq \phi \leq \phi_2\}$.

We identify the dual vector space $(\mathbb{C})'$ with $\mathbb{C}$ by letting $T \in (\mathbb{C})'$ and $W \in \mathbb{C}$ correspond if and only if $T(z) = Re(w\overline{z})$ $\forall w \in W$. Here $^-$ denotes complex conjugate and Re denotes 'real part'.

Lemma 2.2. **i** *Each closed convex cone C in $\mathbb{C}$ with $C \neq \operatorname{span} C$ is of the form C_{ϕ_1,ϕ_2} for suitable $\phi_1, \phi_2 \in \mathbb{R}$ with $0 \leq \phi_2 - \phi_1 \leq \pi$.*

ii *Let $\phi_1, \phi_2 \in \mathbb{R}$ with $0 \leq \phi_2 - \phi_1 \leq \pi$, then the conjugate cone $(C_{\phi_1,\phi_2}, \mathbb{C})'$ equals $C_{\phi_2 - \frac{1}{2}\pi, \phi_1 + \frac{1}{2}\pi}$*

Furthermore we record the following easy fact.

Lemma 2.3. *Let V be a vector space and C a convex cone in V with $\operatorname{rint} C \neq \varnothing$. Then $C + \operatorname{rint} C \subseteq C$.*

3 The main result

Let V be a vectorspace and C a convex cone in V which has a relatively internal point, rint $C \neq \phi$. We will analyze the structure of the conjugate cone $(C, V)'$ in a number of steps, leading up to the main theorem (3.9). To begin with, it is useful to view the set of nonzero elements in the dual cone, $(C, V)' \backslash \{0\}$, as the union of its *singular* part $S(C, V)$ and its regular part $R(C, V)$. The structure of $S(C, V)$ is clarified by the following observation.

Lemma 3.1. *The vectorspace $S(C, V) \cup \{0\}$ is naturally isomorphic to $(V/\text{span } C)'$, the dual of the quotient space $V/\text{span } C$. In particular, $S(C, V) \neq \varnothing$ if and only if span $C \neq V$.*

Proof. It is seen from the definition that $S(C, V) \cup \{0\}$ equals the kernel of the restriction map from V' to $(\text{span } C)'$ (see section 2). This kernel is isomorphic to $(V/\text{span } C)'$ (see section 2). This dual space is nonzero if and only if span $C \neq V$ (see section 2). □

Lemma 3.2. *Let a linear subspace U of V be given with $U \not\subseteq \text{span } C$ and $U \cap \text{rint } C \neq \varnothing$. Then the following statements hold.*

(i) $U \cap \text{span } C = \text{span } (U \cap C)$.

(ii) $S(C \cap U, U) \neq \varnothing$.

(iii) Each $\varphi \in S(C \cap U, U)$ can be extended to an element $\psi \in S(C, V)$.

(iv) For each $\varphi \in S(C \cap U, U)$ there is a natural bijection between the set of elements in $S(C, V)$ which extend φ and the set $(V/U)' \backslash \{0\}$.

Proof. (i) We will check the inclusion $\subseteq$; the inclusion $\supseteq$ is obvious. Choose $c_0 \in U \cap \text{rint } C$. Then for each $v \in U \cap \text{span } C$ there exists $\varepsilon > 0$ with $c_0 + \varepsilon v \in U \cap C$; therefore $v = \varepsilon^{-1}((c_0 + \varepsilon v) - c_0)$ lies in span $(U \cap C)$, as desired.

(ii) We have span $(U \cap C) \neq U$ by statement (i) and our assumption $U \not\subseteq \text{span } C$. Therefore, by Lemma (3.1), we get

$$S(C \cap U, U) \neq \varnothing.$$

(iii) By (i) the canonical mapping from $U/\text{span } (U \cap C)$ to $V/\text{span } C$ is injective.
Therefore the restriction map from $(V/\text{span } C)'$ to $(U/\text{span } (U \cap C))'$ is surjective. Hence, by Lemma (3.1), the restriction map from $S(C,V) \cup \{0\}$ to $S(C \cap U, U) \cup \{0\}$ is surjective. Thus each $\varphi \in S(C \cap U, U)$ can be extended to an element $\psi \in S(C,V)$.

(iv) By Lemma (3.1) it suffices to prove that the kernel of the restriction map from $(V/\text{span } C)'$ to $(U/\text{span } (C \cap U))'$ is isomorphic to $(V/U)'$. This follows from the injectivity of the canonical mapping from $U(\text{span } (U \cap C)$ to $V/\text{span } C$ (see section 2).

□

Now we turn to the regular part $R(C,V)$.

Lemma 3.3. *Assume* $C \neq \text{span } C$ *and* $\dim\ V = 2$. *Then* $R(C,V) \neq \varnothing$.

Proof. The statement of this lemma is a consequence of Lemma 2.2. □

Lemma 3.4. *Let* $\varphi \in (C,V)'$ *be given. Then* $\varphi \in R(C,V)$ *if and only if there exists* $c_0 \in \text{rint } C$ *with* $\varphi(c_0) > 0$. *Moreover if* $\varphi \in R(C,V)$ *then* $\varphi(c) > 0$ *for all* $c \in \text{rint } C$.

Proof. Both statements of this lemma are a consequence of the following observation: the existence of an element $d_0 \in \text{rint } C$ with $\varphi(d_0) = 0$ implies that $\varphi \in S(C,V)$. □

Lemma 3.5. $C + \text{rint } C \subseteq \text{rint } C$.

Proof. Immediate from the definitions. □

Let a linear subspace U of V with $U \cap \text{rint } C \neq \varnothing$ be given. Choose once and for all $c_0 \in U \cap \text{rint } C$. Let moreover an element $\varphi \in R(C \cap U, U)$ be given. Our next aim is to give a description of all extensions of φ to elements of $R(C,V)$. Let H denote the kernel of φ and j the natural map from V to the quotientspace V/H. Furthermore we introduce the following notations.

$$R_n(C,V) = \{\psi \in R(C,V) \mid \psi(c_0) = \varphi(c_0)\}$$

$$R_n(C \cap U, U) = \{\psi \in R(C \cap U, U) \mid \psi(c_0) = \varphi(c_0)\}$$

$$R_n\Big(j(C), j(V)\Big) = \{\psi \in R\Big(j(C), j(V)\Big) \mid \psi\Big(j(c_0)\Big) = \varphi(c_0)\}$$

(the 'n' in the notation refers to normalisation).

Finally let res denote the restriction map from $R_n(C, V)$ to $R_n(C \cap U, U)$ and inf the inflation map from $R_n(j(C),\ j(V))$ to $R_n(C, V)$, given by composition with j; this inflation map is injective, clearly. Now we are ready to give the promised description of the set of all extensions of φ to elements of $R(C, V)$.

The following lemma will play a key role.

Lemma 3.6. *Let $\psi \in R_n(C, V)$ be given. Then res $\psi = \varphi$ if and only if $\psi =$inf ρ for some $\rho \in R_n(j(C),\ j(V))$.*

Proof. Assume that $\psi =$inf ρ for some $\rho \in R_n(j(C),\ j(V))$. We have to prove that res $\psi = \varphi$. As $U = H \oplus Rc_0$ ('direct sum') it suffices to verify that res ψ and φ are equal in H and on c_0. Well, res $\psi\mid_H = \rho \circ \gamma \mid_H$ by the definitions and this is equal to 0 as $H =$ Ker j; moreover $\varphi \mid_H = 0$ as $H =$ Ker φ. Furthermore res $\psi(c_0) = \rho(j(c_0))$ and this equals $\varphi(c_0)$ by the definition of $R_n(j(C),\ j(V))$, to which $\rho \circ j$ belongs. This finishes the proof that res $\psi = \varphi$.

Now assume conversely that res $\psi = \varphi$. We have to prove that $\psi =$ inf ρ for some $\rho \in R_n(j(C),\ j(V))$.

As $\psi \mid_H = \varphi \mid_H = 0$, using $H =$ Ker φ, we conclude that ψ factorizes over j, using that $H =$ Ker j and that j is surjective, say $\psi = \rho \circ j$ for some linear functional ρ on $j(V)$. Now we are going to verify that $\rho \in (j(C),\ j(V))'$.

We have $\rho(j(C)) = \psi(C)$, which is contained in $[0, \infty)$ as $\psi \in (C, V)'$ (using $\psi \in R(C, V)$). Therefore $\rho \in (j(C),\ j(V))'$. It remains to verify that $\rho \in R_n(j(C),\ j(V))$. We have $\rho(j(c_0)) = \psi(c_0)$ and this equals $\varphi(c_0)$ by the definition of $R_n(C, V)$, to which ψ belongs. Therefore $\rho \in R_n(j(C),\ j(V))$. $\square$

Warning. Lemma (3.6) gives no information about the *existence* of elements $\psi \in R(C, V)$ extending a given element $\varphi \in R(C \cap U, U)$. It does give a bijection between the set of these elements ψ and the set $R_n(j(C),\ j(V))$. However we do not yet know at this point that $R_n(j(C),\ j(V))$ is always nonempty. The next lemma is a first step towards proving this.

Lemma 3.7. $j(C) \neq \text{span}\, j(C)$

Proof. We have $(\text{rint } C) \cap H = \varnothing$: this follows from lemma (3.4) as $H = \text{Ker } \varphi$ and $\varphi \in R(C, V)$. Combining this with lemma (3.5) gives $(C + \text{rint } C) \cap H = \varnothing$. This can be formulated in terms of the natural map j from V to V/H as follows $j(C) \cap j(-\text{rint } C) = \varnothing$. Combining this with the assumption $\text{rint } C \neq \varnothing$ and the observation $j(-\text{rint } C) \subseteq \text{span } j(C)$ we arrive at the following conclusion:

$$j(C) \neq \text{span } j(C).$$

$\square$

Lemma 3.8. *Let W be a vectorspace and let D a convex cone in W with $\text{rint } D \neq \varnothing$. Then there exist regular elements, $R(D, W) \neq \varnothing$, if and only if $D \neq \text{span } D$.*

Proof. If $D = \text{span } D$, then $R(D, W) = \varnothing$, by the definitions. Now assume $D \neq \text{span } D$. Consider the collection of pairs (U, φ) consisting of a linear subspace U of W with $U \cap \text{rint } D \neq \varnothing$ and an element $\varphi \in R(D \cap U, U)$. We define a partial ordering on this collection in the following way: $(U_1, \varphi_1) \preceq (U_2, \varphi_2)$ if and only if $U_1 \subseteq U_2$ and moreover φ_2 extends φ_1. By Zorn's lemma, this collection has a maximal element, say (U, φ). It remains to prove that U equals the whole space W. Let us argue by contradiction. Assume $U \neq W$. Choose a linear subspace T of W with $U \subseteq T$ and $\dim(T/U) = 1$. Now we apply lemma (3.6) with $V = T$, $C = D \cap T$, $U = U$. It follows that there is a bijection between the set of extensions of φ to an element of $R(D \cap T, T)$ and the set $R(j(D \cap T), j(T))$ where j is the canonical mapping from T to the quotient space $T/\text{Ker } \varphi$. By lemma (3.7) we have $j(D \cap T) \neq \text{span } j(D \cap T)$. Therefore, as $\dim j(T) = 2$, we conclude by Lemma (3.2) that $R(j(D \cap T), j(T)) \neq \varnothing$. This finishes the proof that φ can be extended to an element of $R(D \cap T, T)$. However this contradicts the maximality of φ. $\square$

Now we collect everything we have proved about the dual cone $(C, V)'$, its singular part $S(C, V)$ and its regular part $R(C, V)$; moreover we include some additional statements, all of which follow immediately from the lemmas above.

Theorem 3.9 (Structural theorem of Hahn–Banach). *Let V be a vectorspace and C a convex cone in V with $\text{rint } C \neq \varnothing$.*

(i) *The conjugate cone* $(C,V)'$ *has non–zero elements ,* $(C,V)' \neq 0$*, if and only if* $C \neq V$.

(ii) *The conjugate cone* $(C,V)'$ *has singular elements,* $S(C,V) \neq \varnothing$*, if and only if* $\mathrm{span}\, C \neq V$.

(iii) *The conjugate cone* $(C,V)'$ *has regular elements,* $R(C,V) \neq \varnothing$*, if and only if* $C \neq \mathrm{span}\, C$.

(iv) *The vectorspace* $S(C,V) \cup \{0\}$ *is naturally isomorphic to* $(V/\mathrm{span}\, C)'$*, the dual of the quotientspace* $V/\mathrm{span}\, C$.

Let moreover U *be a linear subspace of* V *with* $U \cap \mathrm{rint}\, C \neq \varnothing$.

(v) $R(C \cap U, U) \neq \varnothing$*; moreover* $S(C \cap U, U) \neq \varnothing$ *if* $U \nsubseteq \mathrm{span}\, C$.
Let furthermore an element $\varphi \in (C \cap U, U)'$ *be given.*

(vi) *There exists an extension of* φ *to an element* $\psi \in (C,V)'$*. Moreover if* φ *is singular (resp. regular), then each such extension* ψ *is also singular (resp. regular).*

(vii) *Assume* $U \nsubseteq \mathrm{span}\, C$*. If* φ *is singular,* $\varphi \in S(C \cap U, U)$*, then there is a natural bijection from the set of* $\psi \in S(C,V)$ *extending* φ *to the set* $(V/U)' \backslash \{0\}$.

(viii) *Choose* $c_0 \in U \cap \mathrm{rint}\, C$*. If* φ *is regular,* $\varphi \in R(C \cap U, U)$*, then there is a natural bijection from the set of* $\psi \in R(C,V)$ *extending* φ*, to the set* $R_n(j(C),\ j(V))$*. Here we let* $\psi \in R(C,V)$ *with* $\psi\,|_U = \varphi$ *correspond to* $\rho \in R_n(j(C),\ j(V))$ *if and only if* $\psi = \rho \circ j$.

Remark 3.10. Statement (viii) of the theorem above is especially noteworthy. It establishes a bijection between the inverse image of $\varphi \in R(C \cap U, U)$ under the restriction map from $R(C,V)$ to $R(C \cap U, U)$ and the subset of $R((C + \mathrm{Ker}\, \varphi)/\mathrm{Ker}\, \varphi,\ V/\mathrm{Ker}\, \varphi)$ consisting of all elements which take value $\varphi(c_0)$ on the element $c_0 + \mathrm{Ker}\, \varphi$.