# Recall

## Hahn-Banach Theorem(s)

*Dominated extension.* Let $Y$ be a subspace of a vector space $X$. Let $p$ be a positive homogeneous subadditive function on $X$. For every linear functional $f \in Y^\sharp$ with $f \leq p$ on $Y$, there exists $F \in X^\sharp$ extending $f$ and $F \leq p$ on $X$.

*Continuous extension.* Let $Y$ be a subspace of a normed space $X$. For every $f \in Y^*$, there exists $F \in X^*$ extending $f$ such that $\|F\| = \|f\|$.

*Existence of norm-attaining functional.* For every $x_0$ in a normed space $X$, there exists $f \in X^*$ such that $\|f\| = 1$ and $f(x_0) = \|x_0\|$.

*Chosure point checking.* Let $Y$ be a subspace of a normed space $X$. Then $x \in \overline{Y}$ if and only if for every $f \in X^*$ with $f = 0$ on $Y$, we have $f(x) = 0$.

*Hyperplane separation.* Let $C$ be a closed convex subset of a normed space $X$ and $x_0 \in X \setminus C$. Then there exists $f \in X^*$ such that $\sup_{y \in C} f(C) < f(x_0)$.

(Note that we restrict to normed space since the proof in [LN, Prop. 4.16] has used norm which can be avoided. But *hyperplane separation* holds for locally convex spaces.)

If the dual space $X^*$ is separable, then $X$ is separable.

Recall that to apply *dominated extension* in the proof of *hyperplane separation*, we have introduced the *Minkowski functional* $\mu_A$ defined for a set $A$. The properties of $A$ determine the behavior of $\mu_A$. The way of defining Minkowski functional is useful to construct natural functions from sets and reveals properties of the space.

Let $X, Y$ be normed spaces and $T \in B(X, Y)$. The adjoint operator $T^*\colon Y^* \to X^*$ is (formally) defined as, for $y^* \in Y^*, x \in X$,

$$T^* y^*(x) \coloneqq y^*(Tx).$$

Then $T^* \in B(Y^*, X^*)$ and $\|T^*\| = \|T\|$. (In symmetric notation, $\langle x, T^* y^* \rangle \coloneqq \langle Tx, y^* \rangle \coloneqq y^*(Tx)$.)

# Dual space of $C[a, b]$

Let $[a, b]$ be a closed bounded interval in $\mathbb{R}$. Let $C[a, b]$ be the space of $\mathbb{R}$-valued functions on $[a, b]$ with the sup-norm $\|\cdot\|_\infty$.

Let $\rho\colon [a, b] \to \mathbb{R}$ be a real-valued function and $P\colon \{a = x_0 < \cdots < x_n = b\}$ be a partition of $[a, b]$. Define the variation of $\rho$ with respect to $P$ by

$$V(\rho, P) \coloneqq \sum_{k=1}^{n} |\rho(x_k) - \rho(x_{k+1})|,$$

and the *total variation* by

$$V(\rho) \coloneqq \sup_{P \in \mathcal{P}} V(\rho, P)$$

where $\mathcal{P}$ denotes all the paritions of $[a, b]$. A function $\rho \colon [a, b] \to \mathbb{R}$ is called *bounded variation* if $V(\rho) < \infty$. Let $BV[a, b]$ denote the vector space of all the bounded variations.

Let $f \in C[a, b]$ and $\rho \in BV[a, b]$. Let $P : a = x_0 < \cdots < x_n = b$ with tags $t_k \in [x_{k-1}, x_k]$ be a tagged partition. Define the *Riemann-Stieltjes sum* with respect to $\rho$ and $P$ by

$$S(f, \rho, P) = \sum_{k=1}^{n} f(t_k)\Big(\rho(x_k) - \rho(x_{k-1})\Big).$$

Then the *Riemann-Stieltjes integral* is defined by

$$\int_a^b f(x)d\rho(x) \coloneqq \lim_{\|P\| \to 0} S(f, \rho, P).$$

where $\|P\|$ denotes the diameter of a partition. The Riemann-Stieljes integral exists by the uniform continuity of $f$ on $[a, b]$.

Observe that $V(\cdot)$ satisfies non-negativity, scaling property and the triangle inequality. However, $V(\cdot)$ is not non-degenerate since $V(\rho) = 0$ only implies that $\rho$ is constant on $[a, b]$. Hence we restrict to the following subspace (the notation may not be standard)

$$BV_0[a, b] = \{\rho \in BV[a, b] \colon \rho(a) = 0\}.$$

Then it is readily checked that $BV_0[a, b]$ is a Banach space under the norm $V(\cdot)$.

To justify the injectivity in our proof, we further remove the redundancy and modify the space to

$$BV_0^+[a, b] \coloneqq \left\{\rho \in BV_0[a, b] \colon \lim_{y \to x+} \rho(y) = \rho(x),\ \forall\, x \in (a, b)\right\}. \tag{1}$$

It can be checked that $BV_0^+[a, b]$ is closed in $BV_0[0, 1]$ since the right continuity is preserved by uniform convergence and $\|\cdot\|_\infty \le V(\cdot)$ on $BV_0[a, b]$.

*Remark.* The elements in (1) are defined explicitly. They can viewed as representatives of classes in a quotient space whose definition shares the same purpose to establish the injectivity. The details are given in the next section.

It follows from Jordan decomposition of $\rho \in BV[a, b]$ that $\rho = \rho_+ - \rho_-$ where $\rho_+$ and $\rho_-$ are increasing. Hence for $x \in (a, b)$,

$$\rho^* \coloneqq \lim_{y \to x+} \rho = \lim_{y \to x+} \rho_+(x) - \lim_{y \to x+} \rho_-(x) \tag{2}$$

is well defined, i.e., $\rho^* \in BV_0^+[a, b]$. Since $\rho_+$ and $\rho_-$ are monotone, $\rho^*$ and $\rho$ only differ on the at most countable discountinuities in $(a, b)$. Moreover, for $\rho \in BV_0[a, b]$,

$$V(\rho^*) \le V(\rho) \text{ and } \int_a^b f d\rho = \int_a^b f d\rho^* \text{ for all } f \in C[a, b] \tag{3}$$

and for $\rho_1, \rho_2 \in BV_0^+[a, b]$,

$$\text{if } \int_a^b f d\rho_1 = \int_a^b f d\rho_2 \text{ for all } f \in C[a, b], \text{ then } \rho_1 = \rho_2 \in BV_0^+[a, b]. \tag{4}$$

The proofs of (3) and (4) are given in Appendix. After the above modifications, we are ready to state the main theorem.

**Theorem 1.** *Under above notation, $C[a,b]^* = BV_0^+[a,b]$.*

*Proof.* We first introduce some convenient notation. For $f \in C[a,b]$ and $\rho \in BV_0[a,b]$, denote

$$\langle f, \rho \rangle := \int_a^b f d\rho.$$

Then $\langle \cdot, \cdot \rangle \colon C[a,b] \times BV_0[a,b] \to \mathbb{R}$ is well defined by the existence of Riemann-Stieltjes integral. It follows from the linearity of summation that for $\alpha \in \mathbb{R}$, $f, \tilde{f} \in C[a,b]$ and $\rho, \tilde{\rho} \in BV_0[a,b]$,

$$\langle \alpha f + \tilde{f}, \rho \rangle = \alpha \langle f, \rho \rangle + \langle \tilde{f}, \rho \rangle \text{ and } \langle f, \alpha \rho + \tilde{\rho} \rangle = \alpha \langle f, \rho \rangle + \langle f, \tilde{\rho} \rangle. \tag{5}$$

And since for any partition $P$, we have

$$\left| \sum_{k=1}^n f(t_k) \Big( \rho(x_k) - \rho(x_{k-1}) \Big) \right| \le \|f\|_\infty \sum_{k=1}^n |\rho(x_k) - \rho(x_{k-1})| = \|f\|_\infty V(\rho, P) \le \|f\|_\infty V(\rho).$$

Then take the limit $\|P\| \to 0$ on the LHS, we have

$$|\langle f, \rho \rangle| \le \|f\|_\infty V(\rho). \tag{6}$$

By (5) and (6), for any fixed $\rho \in BV_0^+[a,b]$, the map $\langle \cdot, \rho \rangle \colon C[a,b] \to \mathbb{R}$ is linear and bounded, i.e., $\langle \cdot, \rho \rangle \in C[a,b]^*$. To complete the proof, we will prove the map

$$\begin{aligned} T \colon BV_0^+[a,b] &\to C[a,b]^* \\ \rho &\mapsto \langle \cdot, \rho \rangle \end{aligned}$$

is an isometric isomorphism.

(i) (linear and injective) By (5), $T$ is linear. By (4), $T$ is injective.

(ii) (surjective) For any $\Lambda \in C[a,b]^*$, we will first find $\rho \in BV_0[a,b]$ such that $\Lambda f = \langle f, \rho \rangle$ for all $f \in C[a,b]$ and then modify $\rho$ to $\rho^* \in BV_0^+[a,b]$.

(Inspired by the 'formal' argument that $\rho(x) - \rho(a) = \int_a^x d\rho = \langle \chi_{[a,x]}, \rho \rangle \approx \Lambda \chi_{[a,x]}$. But we can NOT apply $\Lambda$ directly to $\chi_{[a,x]}$, which is where *Hahn-Banach* comes into the stage.) Observing that $C[a,b]$ is a subspace in the normed space $B[a,b]$ of bounded functions, we apply *Hahn-Banach* to extend $\Lambda$ to $\widetilde{\Lambda} \in B[a,b]^*$ with $\|\widetilde{\Lambda}\| = \|\Lambda\|$. Hence we are able to define $\rho(x) := \widetilde{\Lambda} \chi_{[a,x]}$ for $x \in (a,b]$ and $\rho(0) := 0$.

First we check $\rho \in BV_0[a,b]$. For any partition $P$, write $\theta_k = \operatorname{Sgn}(\rho(x_k) - \rho(x_{k-1}))$. Then by the linearity and $\|\widetilde{\Lambda}\| = \|\Lambda\|$,

$$\begin{aligned} \sum_{k=1}^n |\rho(x_k) - \rho(x_{k-1})| &= \sum_{k=1}^n \theta_k \left( \rho(x_k) - \rho(x_{k-1}) \right) \\ &= \theta_1 \widetilde{\Lambda} \chi_{[a,x_1]} + \sum_{k=2}^n \theta_k \left( \widetilde{\Lambda} \chi_{[a,x_k]} - \widetilde{\Lambda} \chi_{[a,x_{k-1}]} \right) \\ &= \widetilde{\Lambda} \Big( \theta_1 \chi_{[a,x_1]} + \sum_{k=2}^n \theta_k (\chi_{[a,x_k]} - \chi_{[a,x_{k-1}]}) \Big) \\ &\le \|\widetilde{\Lambda}\| \|\theta_1 \chi_{[a,x_1]} + \sum_{k=2}^n \theta_k \chi_{(x_{k-1},x_k]} \|_\infty \\ &= \|\Lambda\|, \end{aligned} \tag{7}$$

where in the last equality we used that the function in $\|\cdot\|_\infty$ is bounded by 1. Take supremum over partition $P$ on LHS to obtain $V(\rho) \le \|\Lambda\|$. Hence $\rho \in BV_0[a,b]$.

Next we check $\Lambda f = \langle f, \rho\rangle$ for all $f \in C[a,b]$.

(Inspired by the facts that Riemann-Stieljes integral is contuinous w.r.t. $\|\cdot\|_\infty$ and $f$ can by uniformly approximated by step functions.) Let $\varepsilon > 0$. By the uniform continuity of $f$, there exists $\delta_1$ such that for any partition $P$ with $\|P\| \le \delta_1$, $\sup_{x\in[x_{k-1},x_k]}|f(x) - f(x_k)| \le \varepsilon$ for $1 \le k \le n$. Then define the step function $\tilde{f} = f(x_1)\chi_{[a,x_1]} + \sum_{k=2}^n f(x_k)\chi_{(x_{k-1},x_k]}$

$$\|f - \tilde{f}\|_\infty \le \varepsilon \tag{8}$$

By the definition of Riemann-Stieltjes integral, there exists $\delta_2 > 0$, such that for any partition $P$ with $\|P\| \le \delta_2$ and the tags $t_k = x_k, 1 \le k \le n$, we have

$$|\langle f, \rho\rangle - S(f, \rho, P)| \le \varepsilon. \tag{9}$$

It follows from a similar check in (7) that $S(f, \rho, P) = \widetilde{\Lambda}\tilde{f}$. Hence when $\|P\| < \min\{\delta_1, \delta_2\}$, by (8) and (9),

$$\begin{aligned}
|\langle f, \rho\rangle - \Lambda f| &= \left|\langle f, \rho\rangle - \widetilde{\Lambda} f\right| \\
&\le \left|\langle f, \rho\rangle - \widetilde{\Lambda}\tilde{f}\right| + |\widetilde{\Lambda}\tilde{f} - \widetilde{\Lambda} f| \\
&\le |\langle f, \rho\rangle - S(f, \rho, P)| + \|\widetilde{\Lambda}\|\|f - \tilde{f}\|_\infty \\
&\le (1 + \|\Lambda\|)\varepsilon.
\end{aligned}$$

Letting $\varepsilon \to 0$, we have $\langle f, \rho\rangle = \Lambda f$.

Replace $\rho$ with $\rho^*$ defined in (2). It follows from (3) that $\langle f, \rho^*\rangle = \langle f, \rho\rangle = \Lambda f$ and $V(\rho^*) \le V(\rho) \le \|\Lambda\|$.

(iii) (isometric) Let $\rho \in BV_0^+[a,b]$. It follows from (6), that $\|T\rho\| \le V(\rho)$. By (ii), there exists $\tilde{\rho} \in BV_0^+[a,b]$ with $V(\tilde{\rho}) \le \|T\rho\|$ and $\langle f, \tilde{\rho}\rangle = \langle f, \rho\rangle$. By (i), $\rho = \tilde{\rho} \in BV_0^+[a,b]$. Hence $V(\rho) \le \|T\rho\|$, thus $V(\rho) = \|T\rho\|$.

□

*Remark.* A similar proof shows Theorem 1 also holds when the scalar field is $\mathbb{C}$. These results are the special cases of Riesz representation of $C_0(X)^*$ via Borel regular measures when $X$ is locally compact Hausdorff.

It's good to stop here.

*Remark.* Actually the explicit candidate in (1) is found in a 'cheated' way. We can reason as following, for $\Lambda \in C[a,b]^*$, by the general Riesz representation, there exists a unique Borel regular measure $\mu \in M[a,b]$ such that $\Lambda f = \int_a^b f d\mu$ (the integral is defined in Lebesgue way). Then the cumulative distribution function $F_\mu(t) := \mu[a,t]$ is right countinuous (by the continuity of measure) and of bounded variation. Moreover, $\mu$ is the measure extension of the premeasure induced by $F_\mu$ on semiring $\{a, \emptyset, (c,d], a \le c < d \le b\}$. However, notice that for the Dirac measure $\delta_a$ (representing the evaluation $\Lambda f = f(a)$), the Riemann-Stieljes integral w.r.t. $F_{\delta_a} = \chi_{[a,b]}$

identically vanish on $C[a,b]$ ! Then we realize the Riemann-Stieltjies integral will 'forget' the jump at $a$ if the $\rho$ is right-continuous at $a$. Hence we made the following modification for $\mu \in M[a,b]$,

$$\Lambda f = \int_a^b f d\mu = \mu\{a\} f(a) + \int_a^b f dF_\mu = \int_a^b f d(F_\mu + G_{\mu\{a\}})$$

where $G_{\mu\{a\}}(a) \coloneqq -\mu\{a\}\chi_{\{a\}}$. Hence $\widetilde{F_\mu} \coloneqq F_\mu + G_{\mu\{a\}} \in BV_0^+[a,b]$.

Then there comes a natural follow-up question that why we don't have to modify the distribution function in the 'Stieljes' integral defined in probability theory (e.g. MATH3280). One reason is that $F(-\infty) = 0$ and the integration is on the whole real line.

# A quotient space perspective

Instead of explicitly finding the representatives like (1), another natural way to achieve the injectivity is to define the quotient space. Define a subspace of $BV_0[a,b]$ as

$$H \coloneqq \{\rho \in BV_0[a,b] \colon \langle f, \rho \rangle = 0,\ \forall\, f \in C[a,b]\}. \tag{10}$$

By (6), $H$ is closed as the intersection of the kernels of continuous function $\langle f, \cdot \rangle$. Explicitly, $H$ is exactly the subspace of $BV_0[a,b]$ consisting of the functions differing from 0 only on at most countable points on $(a,b)$. Hence $BV_0[a,b]/H$ is well-defined. Let $\pi$ be the natural projection. Define $T \colon BV_0[a,b]/H \to C[a,b]^*$ by $T(\pi(\rho)) = \langle \cdot, \rho \rangle$. Recall for any $\pi(\rho) \in BV_0[a,b]$, the quotient norm $\|\pi(\rho)\| \le \|\rho\|$ and for any $h \in H$, $|\langle f, \rho \rangle| = |\langle f, \rho + h \rangle| \le \|f\|_\infty \|\rho + h\|$, we have $\|T\rho\| \le \|\pi(\rho)\|$. The linear and injectivity follows as we expected. The surjectivity is obtained in the same way in the proof of Theorem 1. Thus $C[a,b]^* = BV_0[a,b]/H$ also holds.

# Appendix

In this Appendix, we will establish the intuition that with respect to $\langle f, \cdot \rangle, \forall\, f \in C[a,b]$, a change at the countable **interior** discountinuites of $\rho \in BV_0[a,b]$ doesn't matter.

**Lemma 2.** *Let $c \in (a,b)$ and $\alpha \in \mathbb{K}$. The Riemann-Stieljes integral $\int_a^b f d(\alpha \chi_{\{c\}}) = 0$ for all $f \in C[a,b]$.*

*Proof.* Let $f \in C[a,b]$ and $P$ be any tagged partition of $[a,b]$. If $c$ is in the interior of some $[x_k, x_{k-1}]$, then $S(f, \chi_{\{c\}}, P) = 0$. If $c = x_k$ for some $x_k \in (a,b)$, then by choosing the tag $c = x_k$ at both $[x_{k-1}, x_k]$ and $[x_k, x_{k+1}]$, we have $S(f, \chi_{\{c\}}, P) = \alpha f(c) - \alpha f(c) = 0$. Hence $\int_a^b f d(\alpha \chi_{\{c\}}) = 0$. $\square$

**Lemma 3.** *Let $\rho \in BV_0[a,b]$. Denote $(c_n)_{n=1}^\infty$ the discountinuous points of $\rho$ in $(a,b)$ and $(\alpha_n)_{n=1}^\infty$ the oscillations of $\rho$, more precisely, $\alpha_n = |\lim_{y \to c_n -} \rho(y) - \lim_{y \to c_n +} \rho(y)|$. For any sequence $(\beta_n)_{n=1}^\infty$ such that $|\beta_n| \le \alpha_n$ for all $n \in \mathbb{N}$, define $\eta = \sum_{n=1}^\infty \beta_n \chi_{\{c_n\}}$. Then $\eta \in BV_0[a,b]$ and $\int_a^b f d\eta = 0$ for all $f \in C[a,b]$.*

*If $\rho$ has only finitely many discontinuities, Lemma 2 finishes the proof.*

*Proof.* Since $\rho \in BV_0[a,b]$, $\sum_{n=1}^{\infty}|\beta_n| \leq \sum_{n=1}^{\infty} \alpha_n \leq V(\rho) < \infty$. Hence for any $\varepsilon > 0$, there exists $N \in \mathbb{N}$ such that $\sum_{n=N+1}^{\infty}|\beta_n| \leq \varepsilon/2$. Then

$$V(\sum_{n=N+1}^{\infty} \beta_n \chi_{\{c_n\}}) \leq 2 \sum_{n=N+1}^{\infty} |\beta_n| \leq \varepsilon.$$

Let $f \in C[a,b]$. By Lemma 2 and (6),

$$|\langle f, \eta \rangle| = \left| \langle f, \sum_{n=1}^{N} \beta_n \chi_{\{c_n\}} \rangle + \langle f, \sum_{n=N+1}^{\infty} \beta_n \chi_{\{c_n\}} \rangle \right| \leq 0 + \|f\|_\infty \varepsilon.$$

Letting $\varepsilon \to 0$, we have $\langle f, \eta \rangle = 0$. □

*Proof of* (3). Let $\rho \in BV_0[a,b]$. Define $\rho_n := \begin{cases} \rho(x+1/n) & \text{if } x \in [a, b-1/n] \\ \rho(b) & \text{if } x \in (b-1/n, b]. \end{cases}$ Then it is readily checked that $V(\rho_n) \leq V(\rho)$ and $\rho^* = \lim_{n\to\infty} \rho_n$. By the lower semi-continuity of $V(\cdot)$ (see e.g. [Royden-Fitzpatrick Real Analysis, Sec 6.3 Problem 33]), $V(\rho^*) \leq V(\rho)$.

By the definition of $\rho^*$, we have $\rho^* - \rho = \sum_{n=1}^{\infty} \beta_n \chi_{\{c_n\}}$ for some sequence $(\beta_n)_{n=1}^{\infty}$ satisfying the condition in Lemma 3. Hence $\langle f, \rho^* \rangle = \langle f, \rho \rangle$ for all $f \in C[a,b]$. □

*Proof of* (4). It suffices to prove that if $\rho \in BV_0^+[a,b]$ and $\langle f, \rho \rangle = 0$ for all $f \in C[a,b]$, then $\rho = 0$. Let $\mu$ be the measure extended from by $\rho^*$. Then for any $c \in (a,b]$, $\rho^*(c) = \mu[a,c] = \lim_{n\to\infty} \int f_n d\mu = \lim_{n\to\infty} \left( \rho^*(a) f_n(a) + \int_a^b f_n d\rho + \int_a^b f_n d\rho^*(a)\chi_{\{a\}} \right) = 0$ where the second equality follows from Lebesgue dominated convergence theorem for sequence $f_n(x) := \begin{cases} 1 & [a,c] \\ \text{linear} & (c, c+1/n] \\ 0 & (c+1/n, b]. \end{cases}$ Hence $\rho^* = 0$ on $(a,b]$ and $\rho = 0$ on $[a,b]$. □