Compactness arguments

The goal of these notes is to familiarize the reader with one kind of compactness arguments that are used quite frequently in several areas of mathematics. We will also describe them from several equivalent viewpoints. For people used to work with $\mathcal{F}$-limits this might be the most natural approach, but it is certainly useful to see different approaches too (at least in order to be able to understand such arguments, when I see them, and to be able to translate them to "my favourite" language.)

Basic facts and definitions

Nets and cluster point

A point $x \in X$ is a cluster point of a net $(x_d)_{d\in D}$ if and only if $\{d \in D; x_d \in U\}$ is cofinal in D for each neighborhood U of x. A point is a cluster point of some net if and only if it is limit of some subnet of this net.

It is also useful to notice that if $(x_d)_{d\in D}$ is a convergent net in a Hausdorff space, then the unique cluster point of this net is the limit.

We will mostly use cluster points of sequences. Notice that if x is a cluster point of some sequence, this does not imply that there exists a subsequence convergent to x. (We only know that there is a subnet, but not every subnet of a sequence is a subsequence.)

Remark 1. We will be working with the following definition of a subnet: Let $(x_d)_{d\in D}$ be a net. A net $(y_e)_{e\in E}$ is called a subnet of $(x_d)_{d\in D}$ if there exists a cofinal *monotone* map $h\colon E \to D$ such that $y_e = x_{h(e)}$. (Some authors use weaker condition instead of monotonicity of h.) Note that cofinality and monotonicity of h implies that

$$(\forall d \in D)(\exists e_0 \in E)(e \geq e_0 \Rightarrow h(e) \geq d),$$

which can be, in some sense, interpreted as $h(e) \to \infty$.

$\mathcal{F}$-limits

If (x_n) is a sequence in a compact space and $\mathcal{F}$ is an ultrafilter then $\mathcal{F}\text{-}\lim x_n$ exists. (In fact, this claim can be stated and proven for arbitrary domains, not just $\mathbb{N}$.)

Applications

Existence of invariant means/measures

In this paragraph, by a *measure* on $\mathbb{Z}$ we will understand a positive finitely additive probability measure $\mu\colon \mathcal{P}(\mathbb{Z}) \to [0, 1]$. *Invariant measure* is a measure which is shift-invariant, i.e.,

$$\mu(A + 1) = \mu(A)$$

for each $A \subseteq \mathbb{Z}$.

Proposition 1. *There exists an invariant measure on $\mathbb{Z}$.*

See e.g. [P, p.5]. The above proposition means that the discrete topological group $(\mathbb{Z},+)$ is *amenable.*

Proof. First, for each $n\in\mathbb{N}$ let us define

$$\mu_n(A)=\frac{|A\cap[-n,n]|}{2n+1}.$$

Each μ_n is a positive finitely additive probability measure on $\mathbb{N}$. If we work in the space of all finitely additive measure with the topology of pointwise convergence[1] then the set of all positive probability measures is a compact subset.

Note that

$$\mu_n(A+1)-\mu_n(A)=\frac{|(A+1)\cap[-n,n]|}{2n+1}-\frac{|A\cap[-n,n]|}{2n+1}=$$
$$=\frac{|(A)\cap[-n-1,n-1]|}{2n+1}-\frac{|A\cap[-n,n]|}{2n+1}=\frac{\chi_A(n)-\chi_A(-n-1)}{2n+1}$$

which implies

$$|\mu_n(A+1)-\mu_n(A)|\le\frac{2}{2n+1}.$$

Using subnets. By compactness, the sequence μ_n has a cluster point. Hence there exists a convergent subnet $(\mu_{n_d})_{d\in D}$. By Remark 1 we have $n_d\to\infty$. Let $\mu=\lim\mu_{n_d}$. Thus we get

$$|\mu(A+1)-\mu(A)|\le\lim_{d\in D}\frac{2}{2n_d+1}=0$$

and $\mu(A+1)=\mu(A)$, i.e., μ is shift-invariant.

It is easy to verify that μ is finitely additive and $\mu(\mathbb{Z})=1$.

Using filters. Let $\mathcal{F}$ be any free ultrafilter. Then for each $A\subseteq\mathbb{Z}$ put

$$\mu(A)=\mathcal{F}\text{-}\lim\mu_n(A).$$

Since $\mathcal{F}$ is ultrafilter and $\mu_n(A)\in[0,1]$ is a bounded sequence, this limit exists for each A.

The proof of the fact that $\mu(A)$ is a finitely additive probability measure on $\mathbb{Z}$ is straightforward. Let us check the shift-invariance. We have

$$|\mu(A+1)-\mu(A)|=|\mathcal{F}\text{-}\lim\mu_n(A+1)-\mu_n(A)|\le|\mathcal{F}\text{-}\lim\frac{2}{2n+1}|=0.$$

$\square$

Existence of Banach limits

Since there is a one-to-one correspondence between positive invariant means and positive finitely additive measures, this paragraph in fact proves the same results and the preceding one, only formulated using functionals instead of measures.

Let us denote $C_n\colon\ell_\infty\to\mathbb{R}$,

$$C_n(x)=\frac{x_1+\dots+x_n}{n}.$$

Clearly, each C_n belongs to ℓ_∞^* and $\|C_n\|=1$. Each C_n is also positive.

[1] i.e. a net $(\mu_d)_{d\in D}$ converges to μ in this space if and only if $\mu_d(A)\to\mu(A)$ for each A

Proof using ultrafilters. Let us define $f(x) = \mathcal{F}\text{-}\lim C_n(x)$. It is easy to show that $f(x)$ is a linear positive functional and $\|f\| = 1$. It is also shift-invariant, since

$$f(Sx - x) = \mathcal{F}\text{-}\lim \frac{x_{n+1} - x_1}{n} = 0.$$

□

Proof using subnets. All C_n belong to unit ball of ℓ_∞^*, which is compact by Banach-Alaoglu theorem. There exists a convergent subnet $(C_{n_d})_{d \in D}$ with $n_d \to \infty$. If we put

$$f = \lim_{d \in D} C_{n_d}$$

then the functional $f \in \ell_\infty^*$ has norm 1 and it is shift-invariant since

$$f(Sx - x) = \lim_{d \in D} \frac{x_{n_d+1} - x_1}{n_d} = 0.$$

□

Krylov-Bogolyubov Theorem

In this part we would like to show the existence of invariant measure for a continuous transformation of a compact metric space. See [AB, Theorem 16.48], [BS, Theorem 4.6.1], [S, p.8, Theorem 1.1], [W, p.152, Corollary 6.9.1].[2]

By Riesz representation theorem, regular Borel measures on X can be identified with functionals from $C^*(X)$. If we work with weak*-topology, then the set of all probability measures can be considered as a closed subset of unit ball of $C^*(X)$, which is weak*-compact by Banach-Alaoglu theorem.

The T-invariant measures correspond to T-invariant functionals. This can be shown using change of variables [AB, Theorem 13.46].

Theorem 1 (Change of Variables Theorem I)**.** *Let Σ_X and Σ_Y be σ-algebras of subsets of X and Y respectively, and let $T\colon (X, \Sigma_X) \to (Y, \Sigma_Y)$ be a measurable transformation. Assume that μ is a measure on Σ_X and let $\nu = \mu T^{-1}$ be the measure induced from μ by T on Σ_Y. For e function $f\colon Y \to \mathbb{R}$ we have*

1. *If f is ν-integrable, then $f \circ T$ is μ-integrable and*

$$\int_Y f \, \mathrm{d}\nu = \int_Y f \, \mathrm{d}\mu T^{-1} = \int_X f \circ T \, \mathrm{d}\mu.$$

2. *If ν is σ-finite, f is ν-measurable, and $f \circ T \in L_1(\mu)$, then $f \in L_1(\nu)$ and*

$$\int_Y f \, \mathrm{d}\nu = \int_X f \circ T \, \mathrm{d}\mu.$$

Lemma 1. *A measure μ is T-invariant if and only if the corresponding functional*

$$\varphi(f) = \int_X f \, \mathrm{d}\mu$$

is T-invariant, i.e. it fulfills the condition $\varphi(f \circ T) = \varphi(f)$ for each continuous $f\colon X \to \mathbb{R}$.

[2] `http://mathoverflow.net/questions/66669/proof-of-krylov-bogoliubov-theorem`

Proof. $\boxed{\Leftarrow}$ $\mu(T^{-1}A) = \int_X \chi_{T^{-1}A}\ \mathrm{d}\mu = \int_X \chi_A \circ T\ \mathrm{d}\mu = \varphi(\chi_A \circ T) = \varphi(\chi_A) = \int_X \chi_A\ \mathrm{d} = \mu(A)$

$\boxed{\Rightarrow}$ T-invariance of the measure μ means $\mu = \mu T^{-1}$. Thus

$$\varphi(f \circ T) = \int_X f \circ T\ \mathrm{d}\mu = \int_X f\ \mathrm{d}\mu T^{-1} = \int_X f\ \mathrm{d}\mu = \varphi(f).$$

□

Let us note also that μ is positive if and only if the corresponding functional is positive and μ is a probability measure if and only if the corresponding functional is normalized.

Another fact needed in the following proof is that every finite Borel measure on a completely metrizable space is regular [AB, Theorem 12.7]. [3]

Theorem 2 (Krylov-Bogolyubov)**.** *Let X be a compact metrizable topological space and $T\colon X \to X$ be a continuous function. Then there exists a T-invariant Borel measure on X, i.e. a σ-additive positive measure μ defined on the σ-algebra $\mathcal{B}$ of all Borel sets such that*

$$\mu(T^{-1}A) = \mu(A)$$

holds for each $A \in \mathcal{B}$.

Proof. From the above remarks if follows that it is sufficient to show the existence of $\varphi \in C^*(X)$ such that

$$\varphi(f \circ T) = \varphi(f)$$

for each $f \in C(X)$. As usually, we will work with the weak* topology on $C^*(X)$.

We start with an arbitrary $g \in C^*(X)$ such that $\|g\| = 1$ and g is positive. (We can choose, for example, $g\colon f \mapsto f(x_0)$ for some $x_0 \in X$.) Now define $\varphi_n = \frac{g+g\circ T+\cdots+g\circ T^{n-1}}{n}$.

Ultrafilters. Let $\varphi(f) = \mathcal{F}\text{-}\lim \varphi_n(f) = \mathcal{F}\text{-}\lim g\left(\frac{f+Tf+\cdots+T^{n-1}f}{n}\right)$. It is easy to see that φ is a linear positive functional on $C^*(X)$. Moreover

$$|\varphi(f\circ T)(x) - \varphi(f)(x)| = |\mathcal{F}\text{-}\lim g\left(\frac{T^n f - f}{n}\right)| \leq \limsup_{n\to\infty} \frac{\|g\|.2\|f\|}{n} = 0,$$

hence $\varphi(f \circ T) = \varphi(f)$ and φ is T-invariant.

Convergent subnet. If $\varphi = \lim_{d\in D} \varphi_{n_d}$, then

$$|\varphi(f\circ T) - \varphi(f)| = \lim_{d\in D} |g\left(\frac{T^{n_d} f - f}{n_d}\right)| \leq \lim_{d\in D} \frac{\|g\|.2\|f\|}{n_d} = 0.$$

□

Maybe it is worth mentioning that the proof given in [AB, Theorem 16.48] utilizes existence of Banach limits. For any Banach limit L and any $x \in X$ the formula $\varphi(f) = L(T^n f(x))$ defines a T-invariant positive functional $\varphi \in C^*(X)$. The same proof is given in [K, p.2].

[3] To be honest, I still do not see, where exactly in the proof the metrizability of X is used.

Banach density and maximal value of shift-invariant mean

Recall that $\overline{u}(A) = \lim_{n\to\infty} \sup_k \frac{A(k,k+n-1)}{n}$.

Proposition 2. *For every $A_0 \subseteq \mathbb{N}$ there exists a shift-invariant positive finitely additive measure μ on $\mathbb{N}$ such that $\mu(A) = \overline{u}(A)$.*

Proof. From the definition of $\overline{u}(A)$ we get that there exists a sequence I_n of intervals in $\mathbb{N}$ such that $\lim_{n\to\infty} |I_n| = +\infty$ and

$$\lim_{n\to\infty} \frac{|A \cap I_n|}{|I_n|} = \overline{u}(A).$$

Define μ_n by $\mu_n(B) = \frac{|B\cap I_n|}{|I_n|}$. Clearly, each μ_n is a finitely additive probability measure.

Ultrafilters. If we put

$$\mu(B) = \mathcal{F}\text{-}\lim \mu_n(B)$$

for $B \subseteq \mathbb{N}$, then it is easy to show that μ is a finitely additive probability measure. Moreover,

$$|\mu(B+1) - \mu(B)| = |\mathcal{F}\text{-}\lim \frac{|(B+1)\cap I_n| - |B \cap I_n|}{|I_n|}| \le \lim_{n\to\infty} \frac{2}{|I_n|} = 0$$

holds for any $B \subseteq \mathbb{N}$.

For the set A we have

$$\mu(A) = \mathcal{F}\text{-}\lim \mu_n(A) = \lim_{n\to\infty} \frac{|A\cap I_n|}{|I_n|} = \overline{u}(A).$$

Convergent subnet. We put $\mu(B) = \lim_{n\in D} \mu_{n_d}(B)$.

$$|\mu(B+1) - \mu(B)| \le \lim_{d\in D} \frac{2}{|I_{n_d}|} = 0$$

$$\mu(A) = \lim_{d\in D} \frac{|A\cap I_{n_d}|}{|I_{n_d}|} = \lim_{n\to\infty} \frac{|A \cap I_n|}{|I_n|} = \overline{u}(A),$$

since $\left(\frac{|A\cap I_{n_d}|}{|I_{n_d}|}\right)_{d\in D}$ is a subnet of a convergent sequence. □

References

[AB] Charalambos D. Aliprantis and Kim C. Border. *Infinite Dimentional Analysis, A Hitchhiker's Guide.* Springer, Berlin, 3rd edition, 2006.

[BS] Michael Brin and Garrett Stuck. *Introduction to dynamical systems.* Cambridge University Press, Cambridge, 2003.

[K] Ultrich Krengel. *Ergodic Theorems.* Walter de Gruyter, Berlin, New York, 1985.

[P] Alan L. T. Paterson. *Amenability.* Amer. Math. Soc., Providence, 1988.

[S] Ya. G. Sinai. *Dynamical Systems II.* Springer-Verlag, Berlin.

[W] Peter Walters. *An Introduction to Ergodic Theory.* Springer-Verlag, New York, 1982. Graduate texts in mathematics 79.