Averaging in Banach spaces: Given a finite sequence \((x_i)_{i=1}^n\) in a Banach space \(X\) we want to study the average norm \(\|\sum_{i=1}^n \varepsilon_i x_i\|\) over all \(\varepsilon = (\varepsilon_i)_{i=1}^n \in \{-1,1\}^n\).

Notation: \(E(\|\sum_{i=1}^n \varepsilon_i x_i\|^p) = \frac{1}{2^n} \sum_{\varepsilon \in \{-1,1\}^n} \|\sum_{i=1}^n \varepsilon_i x_i\|^p,\) \(1 \leq p < \infty\)
i.e. \(E(\|\sum_{i=1}^n \varepsilon_i x_i\|^p)\) is the expectation of \(F: \{-1,1\}^n \to \mathbb{R},\)
\(F(\varepsilon) = \|\sum_{i=1}^n \varepsilon_i x_i\|^p,\) with the normalized counting measure.

Purpose: The value of \(E(\|\sum_{i=1}^n \varepsilon_i x_i\|^p)\) behaves differently depending on the underlying Banach space \(X\). We may thus use it as a tool to study whether a space \(X\) embeds in a space \(Y\).
Proposition: Let $x_1, \ldots, x_n$ be in a Banach space $X$. Then

$$\max_{1 \leq i \leq n} \|x_i\| \leq E\left(\|\sum_{i=1}^{n} \varepsilon_i x_i\|\right) \leq \sum_{i=1}^{n} \|x_i\|.$$

Proof: The second inequality is almost obvious. For the linear one, let $1 \leq i_0 \leq n$ and take $x^* \in S_{X^*}$ s.t. $x^*(x_{i_0}) = \|x_{i_0}\|$. Then

$$E(\|\sum_{i=1}^{n} \varepsilon_i x_i\|) = \frac{1}{2^n} \sum_{\varepsilon_i \in \{-1, 1\}} \|\sum_{i=1}^{n} \varepsilon_i x_i\| \geq \frac{1}{2^n} \sum_{\varepsilon_i \in \{-1, 1\}} \sum_{i=1}^{n} \varepsilon_{i_0} \varepsilon_i x^*_0(x_i)$$

$$= x^*_0(x_{i_0}) + \sum_{\varepsilon_{i_0} = 1}^{n} \varepsilon_{i_0} \sum_{i \neq i_0} \varepsilon_i x^*_0(x_i) =$$

$$\|x_{i_0}\| + \sum_{\varepsilon_{i_0} = 1}^{n} \sum_{i \neq i_0} \varepsilon_i x^*_0(x_i) - \sum_{\varepsilon_{i_0} = -1}^{n} \sum_{i \neq i_0} \varepsilon_i x^*_0(x_i)$$

To finish the proof it remains to show that

$$\sum_{\varepsilon_{i_0} = 1}^{n} \sum_{i \neq i_0} \varepsilon_i x^*_0(x_i) = \sum_{\varepsilon_{i_0} = 1}^{n} \sum_{i \neq i_0} \varepsilon_i x^*_0(x_i) .$$

Indeed, if we define $\Phi: \{ \varepsilon: \varepsilon_{i_0} = 1 \} \to \{ \varepsilon: \varepsilon_{i_0} = -1 \}$ by $\Phi(\varepsilon_i)_{i=1}^{n} = (\varepsilon'_i)_{i=1}^{n}$ with $\varepsilon_i = \varepsilon'_i$ for $i \neq i_0$ then $\Phi$ is a bijection, thus

$$\sum_{\varepsilon_{i_0} = 1}^{n} \sum_{i \neq i_0} \varepsilon_i x^*_0(x_i) = \sum_{\varepsilon_{i_0} = 1}^{n} \sum_{i \neq i_0} \varepsilon'_i x^*_0(x_i) = \sum_{\varepsilon_{i_0} = -1}^{n} \sum_{i \neq i_0} \varepsilon_i x^*_0(x_i) .$$
More notation: Given a B space $X$ we define $L_{simple}(X)$ to be the vector space of (X-valued) simple functions $f : [0,1] \to X$. That is, $f \in L_{simple}(X)$ means that there is a partition of $[0,1]$ into measurable sets $A_1, \ldots, A_n$ and $x_1, \ldots, x_n \in X$ s.t. $f(t) = x_i$ iff $t \in A_i$.

For $1 \leq p < \infty$ we define the $L^p(X)$-norm on $L_{simple}(X)$ as:

$$
||f||_{L^p(X)} = \left( \int ||f(t)||^p d\lambda(t) \right)^{1/p}, \text{ i.e., if } f \text{ is as above then } ||f||_{L^p(X)} = \left( \frac{1}{n} \sum_{i=1}^{n} ||x_i||^p \lambda(A_i) \right)^{1/p}.
$$

Then, $L^p(X)$ is the completion of $(L_{simple}(X), || \cdot ||_{L^p(X)})$ (we will not use/study this space).

Notation: for a simple $f : [0,1] \to \mathbb{R}$ and $x \in X$ we define $f x \in L_{simple}(X)$ by $(f x)(t) = f(t) x$.

The Rademacher sequence

Recall: For $n \in \{0\} \cup \mathbb{N}$ define $D_n = \left\{ \left[ \frac{i-1}{2^n}, \frac{i}{2^n} \right) : 1 \leq i \leq 2^n \right\}$.

Let $I_k = 2^n + 1 \leq k \leq 2^{n+1}$.

We define for $n \in \mathbb{N}$

$$
F_n = \sum_{k=2^n+1}^{2^{n+1}} h_k = \sum_{k=2^n+1}^{2^{n+1}} (-1)^{k+1} 1_{I_k}.
$$

That is, $F_n$ takes its constant on each $I \in D_n$, it obtains the values 1, -1, and it alternates between these values successively.

Remark: For $1 \leq p < \infty$, $(F_n)_n$ is a normalized sequence.

For $p = 2$, $(F_n)_n$ is orthonormal.

Exercise: Prove that in $L^2$ $(F_n)_n$ is isometrically equivalent to the unit ball of $l^2$.

Proposition: Let $x_1, \ldots, x_n$ be vectors in a B-space $X$. Then,

$$
E (\sum_{i=1}^{n} \varepsilon_i x_i)^p = \left( \sum_{i=1}^{n} \left( E (x_i)^p \right) \right)^{1/p}, \text{ for } 1 \leq p < \infty.
$$
Proof: For \( n = 1, 2, \ldots \) we define a bijection \( i: \{-1, 1\}^n \rightarrow \mathbb{D}_n \) as follows: \( i(-1) = (1, 1/2) = \mathbb{I}_2 \), \( i(1) = (1/2, 1) = \mathbb{I}_2 \).

If \( i: \{-1, 1\}^{n-1} \rightarrow \mathbb{D}_{n-1} \) has been defined, for \( \varepsilon = (\varepsilon_i)_{i=1}^n \in \{-1, 1\}^n \) we define \( i(\varepsilon) \) as follows.

Set \( \kappa = i((\varepsilon_i)_{i=1}^{n-1}) \) and \( \varepsilon_n = \begin{cases} 1 & \text{if } \kappa = 1, \\ -1 & \text{otherwise}. \end{cases} \)

Schematically:

\[ \xymatrix{ \{-1, 1\}^{n-1} & \mathbb{D}_{n-1} \ar[l] \ar[r] & \mathbb{D}_n \} \]

The map \( i: \{-1, 1\}^n \rightarrow \mathbb{D}_n \) is a bijection (by an easy induction).

Then, for any \( x_1, \ldots, x_n \in \mathbb{X} \) & \( \varepsilon = (\varepsilon_i)_{i=1}^n \in \{-1, 1\}^n \) we have that \( \forall t \in \mathbb{T}_i(\varepsilon) \), \( \sum_{i=1}^n \varepsilon_i x_i = \sum_{i=1}^n r_i(t) x_i \).

Indeed, if \( n = 1 \), then \( x_1 = r_1(t) x_1 \forall t \in (0, 1/2) \).

- \( x_1 = r_1(t) x_1 \forall t \in (C, 1) \). If the statement holds for \( n \) and \( (\varepsilon_i)_{i=1}^{n+1} \), if \( k = i((\varepsilon_i)_{i=1}^n) \), then if \( \varepsilon_{n+1} = 1 \implies i((\varepsilon_i)_{i=1}^{n+1}) = 2k-1 \).

\[ \sum_{i=1}^{n+1} r_i(t) x_i = \sum_{i=1}^{n} r_i(t) x_i + r_{n+1}(t) x_{n+1} = \sum_{i=1}^{n} \varepsilon_i x_i + x_{n+1} \]

Similarly, if \( \varepsilon_{n+1} = -1 \), \( \sum_{i=1}^{n+1} r_i(t) x_i = \sum_{i=1}^{n} \varepsilon_i x_i - x_{n+1} \) on \( \mathbb{T}_{2k-1} \).

To conclude, \( \| \sum_{i=1}^n r_i(t)x_i \|_p \leq \sum_{i=1}^n \| r_i(t) x_i \|_p \).

\[ = \sum_{i=1}^n \int_{\mathbb{T}_{2k-1}} \| r_i(t) x_i \|_p d\lambda = \sum_{i=1}^n \int_{\mathbb{X}_{2k-1}} \| r_i(t) x_i \|_p d\lambda = \int_{\mathbb{X}_{2k-1}} \sum_{i=1}^n \| r_i(t) x_i \|_p d\lambda \]

\[ = \frac{1}{2^n} \sum_{i=1}^n \lambda(\mathbb{I}_{2k-1}) \| x_i \|_p = \frac{1}{2^n} \sum_{i=1}^n \lambda(\mathbb{I}_{2k-1}) \| x_i \|_p. \]
Recall: Let \((X, \mathcal{A}, P)\) be a probability space

- A collection \((A_i)_{i \in I}\), \(A_i \in \mathcal{A}\), \(i \in I\), is called independent if for any \(i_1, \ldots, i_n \in I\), pairwise different, \(P(\bigcap_{j=1}^n A_{i_j}) = \prod_{j=1}^n P(A_{i_j})\).
- A collection of \(\sigma\)-algebras \((\mathcal{A}_i)_{i \in I}\) is called independent if for any \(A_i \in \mathcal{A}_{i_1}, \ldots, A_i \in \mathcal{A}_i\), \((A_i)_{i \in I}\) is independent.
- Given a measurable \(f : X \to \mathbb{R}\) we define \(s(f) = \{ f^{-1}(A) : A \in \mathcal{B}(\mathbb{R})\}\).
- A collection of functions \(f_i : X \to \mathbb{R}, i \in I\), is called independent if \((s(f_i))_{i \in I}\) is independent.

Example: Let \(R_n\) be the nth Rademacher function. Then, \(s(R_n) = \{ \emptyset, [0, 1], \cup_{n \text{ odd}} \left[\frac{j-1}{2^n}, \frac{j}{2^n}\right], \cup_{n \text{ even}} \left[\frac{j-1}{2^n}, \frac{j}{2^n}\right]\} = \{ f^*(\emptyset), f^*(\{0\}), f^*(\{1\}), f^*(\{-1\})\}\).

Proposition: The sequence \((R_n)\) is independent.

Sketch of proof: By induction on \(k\) show that for \(n \leq \cdots \leq n_k \in \mathbb{N}\) and \(A_i \in s(R_{n_k})\) with \(\lambda(A_i) = \frac{1}{2^n}\) for \(i = 1, \ldots, k\), we have that \(\cap_{i=1}^k A_i\) consists of \(2^{n_k-k}\) intervals of the form \(\left[\frac{i-1}{2^{n_k}}, \frac{i}{2^{n_k}}\right]\), thus \(\lambda(\cap_{i=1}^k A_i) = \frac{1}{2^k} = \prod_{i=1}^k \lambda(A_i)\).

Remark: Given a probability space \((X, \mathcal{A}, P)\), any seq. \((\epsilon_n)\) of independent random variables on \(X\) with \(P(\epsilon_n = 1) = P(\epsilon_n = -1) = \frac{1}{2}\) is called a Rademacher sequence.

Lemma (for later use): Let \((A_i)_{i \in \mathcal{I}}\) be independent \(\sigma\)-algebras.
- \(I = I_1, I_2, \ldots, I_m\) be disjoint subsets of \(I\). Then, \(\mathcal{A} = \bigoplus_{i \in I} \mathcal{A}_i\), \(i \in I\), are independent.

Sketch of proof for \(n = 2\): fix \(A_1, A_2 \subseteq \mathcal{A}\) and \(A_i \in \mathcal{A}_i\). Show that \(\mathcal{C}_0 = \{ C \in \mathcal{C} : C \text{ independent of } A_0 \}\) is a Dynkin system. Observe that \(\mathcal{E} = \{ \bigcap_{i \in I_1} (A_{i_1} \cap A_i) \cup (A_{i_2} \cup A_{i_1}) \}\) is closed under finite intersection. Thus, by Dynkin's theorem \(\mathcal{B}_1 = \mathcal{C}\) \(\mathcal{C}_0\). Reverse the argument to show that for \(B \in \mathcal{B}_1\), \(B \in \mathcal{C}_0\). Then, \(B \in \mathcal{B}_2\) are independent.