Mathematical Foundations of ML – Function Spaces II

Def Hilbert space H : Complete metric vector space induced by an inner product $\langle , \rangle : H \times H \to \mathbb{C}$.
Properties of the inner product:

i. symmetric $\langle x, y\rangle = \overline{\langle y, x\rangle}$,
ii. linear $\langle \alpha x_1 + \beta x_2, y\rangle = \alpha \langle x_1, y\rangle + \beta \langle x_2, y\rangle$,
iii. Positive definite $\langle x, x\rangle \ge 0$, with $\langle x, x\rangle = 0 \Leftrightarrow x = 0$.

The natural norm $\|x\|_H := \langle x, x\rangle^{1/2}$ satisfies

(i) Cauchy-Schwartz

$$|\langle x, y\rangle| \le \|x\|_H \, \|y\|_H$$

(ii) Triangle inequality

$$\|x+y\|^2 = \|x\|^2 + 2\langle x, y\rangle + \|y\|^2 \le \|x\|^2 + 2\|x\|\|y\| + \|y\|^2 = \left(\|x\| + \|y\|\right)^2$$

So an Hilbert space is a Banach space.

Examples

(i) $l_2(\mathbb{Z})$

$$\langle \alpha, \beta\rangle_{l_2} := \sum_{i\in\mathbb{Z}} \alpha_i \overline{\beta}_i \ , \ \|\alpha\|_2 := \left(\sum_{i\in\mathbb{Z}} |\alpha_i|^2\right)^{1/2} .$$

(ii) $L^2(\Omega)$

$$f, g \text{ measurable} , \ \langle f, g\rangle := C_\Omega \int_\Omega f(x)\overline{g(x)}dx ,$$
$$\|f\|_{L_2(\Omega)} = \|f\|_2 = \langle f, f\rangle^{1/2} = \left(C_\Omega \int_\Omega |f(x)| dx\right)^{1/2} .$$

For $\Omega = \mathbb{R}^n, C_\Omega = 1$. For $\Omega = [-\pi, \pi]^n, C_\Omega = \frac{1}{(2\pi)^n}$.

Sobolev spaces

Multivariate derivatives: A partial derivative of order m

$$\alpha = (\alpha_1, \dots, \alpha_n) \in \mathbb{Z}_+^n, \quad D^\alpha f = \frac{\partial^m f}{\partial x_1^{\alpha_1} \cdots \partial x_n^{\alpha_n}} , \ |\alpha| := \sum_{i=1}^n \alpha_i = m .$$

$\underline{C^m(\Omega):}$

The space of all continuously differentiable functions of degree m in the classical sense.

$$\|f\|_{C^m(\Omega)} := \sum_{|\alpha|\le m} \left\|D^\alpha f\right\|_{L_\infty(\Omega)},$$

The ***semi-norm*** with the polynomials of degree m as a ***null-space***

$$|f|_{C^m(\Omega)} := \sum_{|\alpha|=m} \left\|D^\alpha f\right\|_\infty$$

Examples $C^m(\mathbb{R})$ Then $\|f\|_{C^m(\mathbb{R})} = \sum_{k=0}^{m} \left\|f^{(k)}\right\|_\infty$ is a norm $|f|_{C^m(\mathbb{R})} = \left\|f^{(m)}\right\|_\infty$ is a semi-norm with the polynomials as a null-space

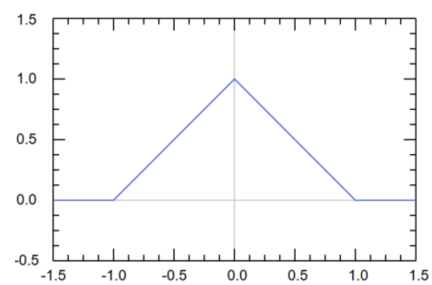
<u>Sobolev spaces</u> $W_p^r(\Omega)$, $1 \le p \le \infty$ <u>:</u>

Def I For $1 \le p < \infty$, completion in $L_p(\Omega)$ of $C^r(\Omega)$ with respect to the norm $\sum_{|\alpha|\le m} \left\|D^\alpha f\right\|_p$. One can also take closure of $C_0^r(\Omega)$.

Def II We define the space of ***test-functions*** $C_0^r(\Omega)$ - continuously differentiable with compact support in Ω. Let $f \in L_p(\Omega) \cap L_1(\Omega)$. Now for $\alpha \in \mathbb{Z}_+^d$, $|\alpha| = r$, $g := D^\alpha f$ is the ***distributional (generalized) derivative*** of f if for all $\phi \in C_0^r(\Omega)$

$$\int_\Omega g\phi = (-1)^{|\alpha|} \int_\Omega f D^\alpha \phi \ .$$

Assignment: For $H(x) := \begin{cases} x+1, & -1 \le x < 0, \\ 1-x, & 0 \le x \le 1, \\ 0, & else. \end{cases}$,



prove that $H'(x) = g(x) = \begin{cases} 1, & -1 \le x < 0, \\ -1, & 0 \le x \le 1, \\ 0, & else. \end{cases}$.

So, in this sense $H \in W_p^1(\mathbb{R})$.

The Sobolev norm and semi-norm. We require that the distributional derivatives exist as functions(!) and

$$\|f\|_{W_p^r(\Omega)} := \sum_{|\alpha|\le r} \left\|D^\alpha f\right\|_{L_p(\Omega)} < \infty \qquad |f|_{W_p^r(\Omega)} := \sum_{|\alpha|=r} \left\|D^\alpha f\right\|_{L_p(\Omega)}.$$

Theorem W_p^r is a Banach space

Theorem For $f \in W_p^r\left(\mathbb{R}^n\right)$ and $0 \le j \le r$, $\varepsilon > 0$

$$|f|_{j,p} \le c\left(\varepsilon |f|_{r,p} + \varepsilon^{-j/(r-j)} \|f\|_p\right),$$
$$\|f\|_{j,p} \le c\left(\varepsilon \|f\|_{r,p} + \varepsilon^{-j/(r-j)} \|f\|_p\right),$$
$$\|f\|_{j,p} \le c\|f\|_{r,p}^{j/r} \|f\|_p^{(r-j)/r}$$

Remarks

(i) Sometimes one sees $\|f\|_{W_p^r(\Omega)} := \|f\|_{L_p(\Omega)} + |f|_{W_p^r(\Omega)}$, since by the theorem the two definitions are equivalent.

(ii) This is also true for 'nice' domains and the constants depend on the 'smoothness' of the boundary.

Approximation using uniform piecewise constants (numerical integration)

The B-Spline of order one (degree zero, smoothness -1) $N_1(x) = \mathbf{1}_{[0,1]}(x)$.

Let $\Omega = \mathbb{R}$ or $\Omega = [a,b]$. We approximate from the space

$$S(N_1)^h := \left\{ \sum_{k \in \mathbb{Z}} c_k N_1\left(h^{-1}x - k\right) \right\} = \left\{ \sum_{k \in \mathbb{Z}} c_k \mathbf{1}_{[kh,(k+1)h]}(x) \right\}.$$

Theorem Let $f \in W_p^1(\mathbb{R})$, $1 \le p \le \infty$. Then

$$E\left(f, S(N_1)^h\right)_{L_p(\mathbb{R})} := \inf_{g \in S(N_1)^h} \|f - g\|_{L_p(\mathbb{R})} \le h|f|_{W_p^1(\mathbb{R})}.$$

Proof First assume $f \in C^1(\mathbb{R}) \cap W_p^1(\mathbb{R})$. Let's take the interval $[kh,(k+1)h]$. Then, for $p = \infty$

$$|f(x) - f(kh)| = \left| \int_{kh}^{x} f'(u)\,du \right| \le h \sup_u |f'(u)|.$$

So select $c_k := f(kh)$ and you get the theorem for $p = \infty$. For $1 < p < \infty$ we do something similar

$$|f(x) - f(kh)|^p \le \left(\int_{kh}^{(k+1)h} |f'(u)|\,du \right)^p.$$

Then

$$\int_{kh}^{(k+1)h} |f(x)-f(kh)|^p\,dx \le h\left(\int_{kh}^{(k+1)h} |f'(u)|\,du\right)^p$$

$$\le h\left(\|f'\|_{L_p([kh,(k+1)h])}\|1\|_{L_{p'}([kh,(k+1)h])}\right)^p$$

$$= hh^{p/p'}\|f'\|^p_{L_p([kh,(k+1)h])}$$

$$= h^p\|f'\|^p_{L_p([kh,(k+1)h])}.$$

$$1+\frac{p}{p'} = 1+p\left(1-\frac{1}{p}\right) = 1+p-1 = p$$

Therefore, with $g(x) := \sum_k f(kh) N_1(h^{-1}x-k)$, we get

$$\|f-g\|_p^p = \int_{-\infty}^{\infty} |f(x)-g(x)|^p\,dx = \sum_k \int_{kh}^{(k+1)h} |f(x)-f(kh)|^p\,dx \le \sum_k h^p\|f'\|^p_{L_p([kh,(k+1)h])} = h^p\|f'\|_p^p.$$

We then use a density argument to go from $C^1(\mathbb{R})\cap W_p^1(\mathbb{R})$ to $W_p^1(\mathbb{R})$.

<u>Modulus of smoothness</u>

Def The ***difference operator*** Δ_h^r. For $h\in\mathbb{R}^d$ we define $\Delta_h(f,x)=f(x+h)-f(x)$. For general $r\ge 1$ we define

$$\Delta_h^r(f,x) = \underbrace{\Delta_h\circ\cdots\Delta_h}_{r}(f,x) = \sum_{k=0}^{r}\binom{r}{k}(-1)^{r-k} f(x+kh).$$

Remarks

1. For $\Omega\subset\mathbb{R}^n$, we in fact modify to $\Delta_h^r(f,x) := \Delta_h^r(f,x,\Omega)$, where $\Delta_h^r(f,x,\Omega)=0$, in the case $[x,x+rh]\not\subset\Omega$. So for $\Omega=[a,b]$, $\Delta_h^r(f,x)=0$ on $[b-rh,b]$, for any function.
2. As an operator on $L_p(\Omega)$, $1\le p\le\infty$, we have that $\|\Delta_h^r\|_{L_p\to L_p}\le 2^r$. Assume $\Omega=\mathbb{R}^n$, then

$$\|\Delta_h^r(f,\bullet)\|_p \le \sum_{k=0}^{r}\binom{r}{k}\|f(\bullet+kh)\|_p = \sum_{k=0}^{r}\binom{r}{k}\|f\|_p = 2^r\|f\|_p$$

Def The ***modulus of smoothness*** of order r of a function $f\in L_p(\Omega)$, $0<p\le\infty$, at the parameter $t>0$

$$\omega_r(f,t)_p := \sup_{|h|\le t}\|\Delta_h^r(f,x)\|_{L_p(\Omega)}.$$

For $r=1$ the modulus of smoothness is called the ***modulus of continuity***.

Example of non continuous functions. Let $\Omega=[-1,1]$. $f(x)=\begin{cases}0 & x<0\\ 1 & 0\le x\end{cases}$.

Let's compute $\omega_r(f,t)_{L_p([-1,1])}$.

$$\Delta_h(f,x)=\begin{cases} 0 & -1\le x\le -h \\ 1 & -h<x\le 0 \\ 0 & 0<x\le 1 \end{cases}$$

For $p=\infty$ we get $\omega_1(f,t)_{L_\infty([-1,1])}=1$.

For $p\neq\infty$ we get $\omega_1(f,t)_{L_p([-1,1])}=t^{1/p}$.

$$\Delta_h^2(f,x)=\Delta_h(\Delta_h f,x)=\begin{cases} 0 & -1\le x\le -2h \\ 1 & -2h<x\le -h \\ -1 & -h<x\le 0 \\ 0 & 0\le x\le 1 \end{cases}$$

We get $\omega_2(f,t)_{L_p([-1,1])}=(2t)^{1/p}$.

In general, we'll get $\omega_r(f,t)_{L_p([-1,1])}\le C(r,p)t^{1/p}$.

Quick jump into the "future" (Generalized Lipschitz / Besov smoothness)… for $\alpha<1/\tau$, $r=\lfloor\alpha\rfloor+1$,

$$|f|_{B^\alpha_{\tau,\infty}}:=\sup_{t>0}t^{-\alpha}\omega_r(f,t)_\tau\le\sup_{0<t\le 2}t^{-\alpha}\omega_r(f,t)_\tau\le c\sup_{0<t\le 2}t^{1/\tau-\alpha}<\infty.$$

We then say that f has α (weak-type) smoothness. Observe that in this example α can be arbitrarily large as long as the integration takes place with τ sufficiently small.