Lecture 1: Introduction & review of func. analysis

1.0: Welcome to Amath 563!

(No Scribe) This is an advanced course focusing on the mathematical aspects of data analysis & inference.

Much of the course is focused on Kernel methods & Gaussian processes as a versatile tool for various problems in data science & machine learning (ML).

Learning outcomes:

- ▷ understand fundamental theory of Reproducing Kernel Hilbert Spaces (RKHS)
- ▷ Be able to design & implement Kernel methods for various problems.
- ▷ Exposure to a wide range of problems in data science & ML.

Review Syllabus Doc

Why focus on theory?

Most ML/DS problems involve high. dim. data sets & complicated algorithms. Without abstract theory it is hard to tell what is going on? & which method is better/more efficient.

Why Kernel methods?

Kernel methods such as support vector machines (SVMs) were the state of the art until 2015! They are fairly simple & have alot of theoretical support. They also unify alot of existing methods under the same framework.

1.1 : Review of Functional Analysis

We will start with the very basics, ie, the def^n of a vector space & make our way through metric spaces, Banach spaces, & Hilbert spaces. Along the way I will recall some useful results without proof. It is helpful to have Kreyszig or other functional analysis books nearby.

Def^n: A set X is called a (real) vector space if it is closed under the operations of addition & scalar multiplication, ie,

$$x_1 + x_2 \in X \qquad \forall\, x_1, x_2 \in X$$
$$\alpha x \in X \qquad \forall\, x \in X \;\&\; \alpha \in \mathbb{R}$$

We would say X is a complex vector field if we allow $\alpha \in \mathbb{C}$ instead.

examples include:

- $\mathbb{R}^n$ equipped with the usual sum & scalar mult.
$$\underline{x} + \underline{y} = (x_1+y_1, x_2+y_2, \dots, x_n+y_n)$$
$$\alpha \underline{x} = (\alpha x_1, \alpha x_2, \dots, \alpha x_n)$$
- $C([a,b])$ equipped with
$$(x+y)(t) = x(t) + y(t) \qquad \forall t \in [a,b]$$
$$(\alpha x)(t) = \alpha \cdot x(t)$$

1.1 Banach Spaces

Def^n (Norm) A real valued function $\|\cdot\| : X \mapsto \mathbb{R}$ on a vector space X is called a norm if it satisfies the following four axioms

(1) $\|x\| \geq 0$ (iff)

(2) $\|x\| = 0 \iff x = 0$

(3) $\|\alpha x\| = |\alpha| \cdot \|x\|$, $\forall \alpha \in \mathbb{R}$

(4) $\|x + x'\| \leq \|x\| + \|x'\|$ (Triangle ineq.)

egs: The Euclidean norm on $\mathbb{R}^n$

$$\|\underline{x}\| := \left(\sum_{j=1}^{n} |x_j|^2 \right)^{1/2}$$

The L^2-norm on $C([a,b])$

$$\|x\| := \left(\int_a^b |x(t)|^2 dt \right)^{1/2}$$

The ℓ^p-norm on real sequences

$$\|x\| := \left(\sum_{j=1}^{\infty} |x_j|^p \right)^{1/p}$$

Def^n: A Banach space X is a Normed vector space that is also complete. That is, all Cauchy sequences have a limit in X.

Notation $(X, \|\cdot\|)$

Recall that $\{x_j\}_{j=1}^{\infty} \subset X$ is called a Cauchy sequence if $\forall \varepsilon > 0$, there exists $N > 0$ s.t. $\| x_j - x_k \| \leq \varepsilon$, for all $j, k > N$.

Completeness is not trivial to guarantee. For example, the space $C([a,b])$ equipped with the L^2-norm is not complete!

Hint: You can construct a Cauchy seq. that converges to a discontinuous function.

However, every normed vector space can be completed in a (essentially) unique way.

(see Thm 2.3-2 of Kreyszig). **Thm**: Let $(X, \|\cdot\|)$ be a normed space. Then there is a Banach space $\hat{X}$ & an isometry A from X to a subspace W of $\hat{X}$ which is dense in $\hat{X}$. The space $\hat{X}$ is unique up to isometries. i.e., if $\tilde{X}$ is another completion then $\hat{X}$ & $\tilde{X}$ are isometric!.

Todo: Look up the def^n of isometry & dense subsets.

- Isometries preserve distances:

$$\|x - x'\|_X = \|Ax - Ax'\|_{\hat{X}}$$

- $W \subset \hat{X}$ is dense if $\overline{W} = \hat{X}$.

eg. of Banach spaces.

- Euclidean space equipped with p-norms $p \geq 1$. $(\mathbb{R}^n, \|\cdot\|_p)$

- $L^2([a,b])$ space of (equivalence classes of) square integrable functions

$$L^2([a,b]) = \{f : [a,b] \to \mathbb{R} \mid \|f\|_{L^2([a,b])} < +\infty\}$$

$$\|f\|_{L^2([a,b])} := \left(\int_a^b |f(t)|^2 dt\right)^{1/2}$$

- $C([a,b])$ equipped with the sup norm

$$\|f\|_\infty := \sup_{t \in [a,b]} |f(t)|$$