Functional Analysis: Quick reference

vishvAs vAsuki

April 16, 2012

Also see linear algebra ref.

Part I

Function spaces and operators

1 Function space

1.1 Notation

Consider a function class C defined on the input space X. Let p be a probability measure on X.

1.2 Euclidean Vector spaces

The definition of standard inner product and l_p norms can be extended to work with function spaces, when they are regarded as vector spaces with dimensionality equal to $|X|$; while also considering the p. These are described in the vector spaces survey.

1.3 Metric space using Disagreement probability

One can define a metric space using the norm $d(f, g) = Pr_p(f(x) \neq g(x))$.

1.4 Measuring size of function class C

If C is finite, we can simply use $|C|$. If C is not finite, we can consider the metric space defined using the 'probability of disagreement' metric. In this space, one can use covering and packing numbers to measure the size of C. These concepts are described in the topology ref.
If $dom(f), \forall f \in C$ is the boolean hypercube, we can use some other ways of measuring complexity of C, these are described in the boolean functions survey.

1.4.1 For smooth functions: just measure size of the input space

For functions which satisfy Holder continuity generalized to sth derivative: (see topology ref): $\left\|f^{(s)}(x) - f^{(s)}(y)\right\| \leq c\|x-y\|^a$: $\log(N(\epsilon, \mathrm{C}, \|\|)) \approx (\frac{1}{\epsilon})^{\frac{D}{s+a}}$.
[**Proof**]: : $\Box$ Take $\|f(x) - f(y)\| \leq c\|x-y\|$; ***now in input space, rather than in the function space!***; thence any ϵ covering of parameter space will define corresponding cover in the fn space.

1.4.2 Relationship with vcd of related binary classifiers

Let $G = \left\{g : R^D \to [0, B]\right\}$. Take set G_t of binary classifiers defined by super-graphs of g: see boolean fn ref; let $d(G_t)$ be its VCD. Let M be the packing number. $M(\epsilon, G, \|.\|_{L_p(v)}) \leq 3(2e(\frac{B}{\epsilon})^p \log(3e(\frac{B}{\epsilon})^p))^{d(G_t)}$. So, $M \approx O((\frac{1}{\epsilon})^{d(G_t)})$.
[**Find proof**]

2 Operators

2.1 Operator

A function: functions $\to$ functions. All functions are operators. Transform: an operator which simplifies some operations. Adjoint of operator T: T^* : $\langle Tu, v\rangle = \langle u, T^*v\rangle$.

2.1.1 Eigenfunctions

ev of D is e^x.

2.2 Differential operator

D or $D_x = \frac{d}{dx}$. A Linear operator.
Operators $T = \sum_k c_k D^k$. Any polynomial in D with function coefficients also a diff operator.

2.2.1 Vector and matrix calculus operators

See linear algebra ref.

2.2.2 Divergence operator

$divf := \nabla.f := \sum_i \frac{\partial f}{\partial x_i}$.

2.2.3 Laplacian and Elliptic operators

(Laplace) $\Delta = \frac{\partial^2}{\partial x_1^2} + \frac{\partial^2}{\partial x_2^2}$. Elliptic op: $L = div(a\nabla)$: a is scalar.

2.3 Integral operator

Aka Integral transform. $Tf(u) = \int_{t_1}^{t_2} K(t,u)f(t)dt$: changing the domain of the fn. K is the Kernel function or the nucleus of the transform. A Linear operator.
Symmetric kernels are indifferent to permutation of (t,u).

2.3.1 Inverse Kernel

$K^{-1}(u,t)$ yields inverse transform:
$f(t) = \int_{u_1}^{u^2} K^{-1}(u,t)(Tf(u))du.$

2.3.2 Convolution

$(f.g)(t) = \int_{-\infty}^{\infty} f(t)g(x-t)dt = \int_{-\infty}^{\infty} f(x-t)g(t)dt.$
You take a g, reflect it about 0, translate it by t. So, it is convoluted!
Properties: commutative, associative, distributive.

2.3.3 K(t, u) as a basis function

Maybe $\int_{t_1}^{t_2} abdt$ specifies an inner product in a function space. Maybe K(t, u) specifies the form of basis functions in that space: so for a fixed u, you get a fixed basis fn. Maybe you are trying to find the form of the component of f(t) along the basis fn K(t, u). The integral transform is the solution. Eg: Fourier transform.
Or maybe, you have the form of the component f(u) along a certain basis fn K(u, t), and ye want to reconstruct the fn form f(t) from it. Then the inverse integral operator is useful. Eg: Inverse Fourier transform.

2.4 Some vector differential operators

2.4.1 Jacobian matrix

Generalizes f'(x) for Vector map F.
Matrix $J = J_F(x_1 \dots) = \frac{\partial(y_1 \dots)}{\partial(x_1 \dots)}$: $J_{i,j} = \frac{\partial y_i}{\partial x_j}$.

2.4.2 Hessian matrix

Of scalar f(x) wrt vector x: $H_{i,j} = D_i D_j f(x)$: Always symmetric. Aka $\nabla^2 f(x)$. If $\nabla^2 f(x) > 0$ or +ve definite: f(x) convex, $\exists$ unique global minimum.

3 Analysis of functions

3.1 Polynomial fitting for continuous f(x)

$L^2[-1,1]$: Find closest n degree polynomial to f(t) in interval [u,v]: Project f(t) on that function sub-space: Let $A = [1x \dots x^n]$ (Continuous version of Vandermonde matrix), a [-1,1]*n matrix; $\perp$ in [-1,1]. Solve $A^*A\hat{y} = A^*f(t)$; A^*A a n*n (Hilbert) matrix with elements $\langle x^r, x^s\rangle$. Orthonormalize A to get $\hat{Q}$; Q is scalar multiple of Legendre polynomials: $1, x..$); use $P = \hat{Q}\hat{Q}^*$, a $[-1,1] \times [-1,1]$ matrix.

3.2 Fourier basis in the function space

3.2.1 Fourier basis for 2 pi periodic functions

3.2.1.1 Inner product

$\langle f, g\rangle = (2\pi)^{-1}\int_X f(x)g(\bar{x})dx$.

3.2.1.2 Orthogonal basis

For f having an interval or period 2π : $X = [0, 2\pi]$ or $X = [-\pi, \pi]$ or $X = [-\infty, \infty]$, for distinct $n, m \in Z^+$: $\cos nx \perp \cos mx \perp \sin nx \perp \sin mx$.
Also, for distinct $n, m \in Z : e^{inx} \perp e^{imx}$.

3.2.2 Fourier basis for 2L periodic functions

3.2.2.1 Inner product

$\langle f, g\rangle = (2L)^{-1}\int_X f(x)g(\bar{x})dx$. If this were not scaled, f(x) series could've been scaled.

3.2.2.2 Orthogonal basis

In general, if f has a period $X = [-L, L]$, for distinct $n, m \in Z : e^{inx\frac{\pi}{L}} \perp e^{imx\frac{\pi}{L}}$.
Frequency spectrum: $\frac{n\pi}{L}\forall n \in Z$.

3.2.3 Fourier basis for N dimensional space

3.2.3.1 Inner product

Same inner product as in the case of 2L periodic functions. For different inner product space see boolean functions ref.

3.2.3.2 Orthogonal basis

f defined on Z_N. In general, if f has a period $X = [0, N-1]$, for distinct $n, m \in Z_N : e^{inx\frac{2\pi}{N}} \perp e^{imx\frac{2\pi}{N}}$.

3.2.4 Fourier Series

Project f(x) on $1, \sin nx, \cos nx$: $f(x) = a + s_1 \sin x + c_1 \cos x + s_2 \sin 2x \ldots$. $s_n = \frac{\langle f(x), \sin nx \rangle}{\| \sin nx \|}$.
Also, $f(x) = \sum_{n=-\infty}^{\infty} \hat{f}(n) e^{inx}$. $\hat{f}(n) = \langle f, e^{inx} \rangle$.

3.2.5 Fourier basis for non periodic functions

Use $X = [-\infty, \infty]$; compare with $X = [-L, L]$ case. Frequency spectrum becomes R. In this limiting case, coefficients F(w) correspond to Fourier transform. [**Find proof**]Inverse Fourier transform: $f(x) = \int_{x=-\infty}^{\infty} F(w) e^{iwx} dx$. [**Find proof**]

3.3 Fourier transforms

3.3.1 Fourier Transform (FT)

Use $X = [-\infty, \infty]$; Let $\int_X |f(x)| dx < M$. Fourier transform of f(x) is $F(w) = \frac{1}{2\pi} \int_X f(x) e^{-iwx} dx$. Similarly inverse FT.
A transformation from time domain fn to frequency domain fn. DFT used for practical purposes.
Fourier series of fn is the evaluation of Inverse FT for a particular fn.

3.3.2 Discrete Fourier Transform (DFT)

Use $X = [0, N-1]$ as a period of length N. See analysis for Fourier basis for 2L periodic fn. x measured at N points, so deal with N dim space: find the N Fourier components. Or see this as approximating f by taking only n frequencies.
$F(n) = \sum_{x=0}^{N-1} f(x) e^{\frac{-2\pi}{N} inx}$. Inverse DFT: $f(x) = N^{-1} \sum_{n=0}^{N-1} F(n) e^{inx}$: an avg over many frequencies.
Use in frequency analysis and synthesis of signals.
Let $w_n = e^{-i\frac{2\pi}{n}} = 1^{-1/n}$; make Fourier matrix elements $F_{j,k} = w^{jk}$. F is a Vandermonde type matrix. Get N long sample vector c of f(x) in the interval X. Then, do Fc=y to find coefficients y.
Inverse DFT accomplished similarly with matrix $\bar{F}$.
$F\overline{F} = nI$: $w_n^{jk} w_n^{-jk} = 1$, $\sum_l w_n^{(j-k)l} = 0$.

3.3.3 Fast Fourier Transform (FFT) alg

Find $F_n c = y$ when $n = 2^l$ (Naively n^2 mults).

3.3.3.1 Butterfly diagram

Important in understanding FFT. Input: n numbers c, Output n numbers y. Compute all these in l steps, parallelly. 2 point FFT butterfly diagram. Combine them to get 4 point FFT diagram etc. : see top down view for decomposition.

3.3.3.2 Top down view

For $m = n/2$, $c' = c_i$ for even i, $c'' = c_i$ for odd i, Find $Fc' = y'$ and $Fc'' = y''$, $y_j = \sum_k w_n^{2kj} c_{2k} + \sum_k w_n^{(2k+1)j} c_{2k+1} = y'_j + w_n^j y''_j$; . Do recursively until you hit the 2 point fourier transform.
FFT is a l-time factorization of Fc.
Extend to other prime factors of highly composite n. Divide and conquer! Opcount: $O(n \log n)$.
Inverse FFT is similar.

3.4 Fourier analysis of Boolean fns

See Boolean fns ref.

3.5 Approximation of functions, interpolation

See Approximation theory ref.